

THE WORLD BANK GROUP ARCHIVES

PUBLIC DISCLOSURE AUTHORIZED

Folder Title: Shourie, Arun - Articles and Speeches (1970 - 1973) - 1v

Folder ID: 1654606

Fonds: Records of Office of External Affairs (WB IBRD/IDA EXT)

Digitized: February 27, 2014

To cite materials from this archival folder, please follow the following format:
[Descriptive name of item], [Folder Title], Folder ID [Folder ID], World Bank Group Archives, Washington, D.C., United States.

The records in this folder were created or received by The World Bank in the course of its business.

The records that were created by the staff of The World Bank are subject to the Bank's copyright.

Please refer to <http://www.worldbank.org/terms-of-use-earchives> for full copyright terms of use and disclaimers.



THE WORLD BANK
Washington, D.C.

© 2012 International Bank for Reconstruction and Development / International Development Association or
The World Bank
1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

PUBLIC DISCLOSURE AUTHORIZED

Shourie, Arun - Articles and Speeches (1970-1973)



1654606

A1992-007 Other #: 18

212063B

Shourie, Arun - Articles and Speeches (1970 - 1973) - 1v

DECLASSIFIED
WBG Archives

REMARKS

xi I can't think of anything to do with this except file it in Shourie speeches. Mr. Lind says paper presented at Copenhagen Seminar, and will become a Working Paper.

From

Doris

Mr. Gamstein:

Another top-down

paper -

will become WP

LL

*(Hawkins
another)*

THE USE OF MACROECONOMIC REGRESSION MODELS FOR FORECASTS AND
POLICY PRESCRIPTION: SOME REFLECTIONS ON CURRENT PRACTICE

August, 1970

Arun Shourie
Economics Department
International Bank
for Reconstruction
and Development

THE USE OF MACROECONOMIC REGRESSION MODELS FOR FORECASTS AND
POLICY PRESCRIPTION: SOME REFLECTIONS ON CURRENT PRACTICE

I. Introduction

The need for long-term forecasting of various magnitudes in under-developed economies is increasingly recognized. Lending agencies that make loans with maturities ranging from ten to fifty years and that must make concrete decisions about the appropriate volume and terms of lending naturally want to know all they can about likely developments in these economies for as many years into the future as possible. An international agency like UNCTAD which has learned that it takes a good many years to wrest action from a hundred-odd countries must plan now for steps it would want taken in the mid- or late 1970's, and to do so it needs projections relating to those distant periods. National planning agencies often want to examine the problems of their economies in a long-term context and, just as often, they must put together plan documents that give a long-term perspective to current developments, their programs and their requests for external assistance.

It is a common practice in making these forecasts whether medium- or long-term, and for policy prescription in general, to use econometric models - models that were originally developed for short-term forecasting of relationships in developed economies. Thus UNCTAD has a large project under way of constructing some 30-odd models of various economies; ECAFE recently

I am grateful to P.D. Henderson, B.S. Minhas, E.P. Holland, N.G. Carter and R. Cheetham for many useful comments and to S. Malik and E. Ramiscal for their help in the computations. None of them is responsible for any errors that remain. The author is a member of the Economics Department of the World Bank. The views expressed in this note are his personal views and not those of the World Bank group.

published econometric models for ten countries; other international organizations, such as the IMF and the World Bank, and national donor agencies, such as the U.S. AID, are making increasing use of them; and consultants and research organizations working on problems of the third world often base their projections and advice on econometric models.

This paper is based on a review of about 45 such models that have been proposed for various developing countries. All of these models use statistically estimated parameters to provide a quantitative description of certain macro-relationships in developing economies. They might, for short, be described as aggregative macro-economic regression models.

Sections II to VI review the models and their relationships from five points of view: the reliability of the data from which they are estimated; the significance of the coefficients of their equations; the functional form of their equations; their propensity to use certain statistical tests in a mechanical way; and their use for making forecasts and prescribing policies for years far into the future. Section VII summarizes the principal conclusions and contains some general remarks about the manner in which econometric models are being used at present.

The discussion in Sections II to VI is illustrated by cases from various models that have been proposed in the last five years and from a conventional two-gap model of the Kenya economy put together by the author. Many of the illustrations are taken from a recent UNCTAD study that contains

models for 18 countries.^{1/} Raul Prebisch in his preface to the study states that "the study provides a solid basis for the elaboration of new national and international targets for the United Nations Development Decade of the 1970's."^{2/}

II. The Data Base

It may appear trite or even superfluous to begin yet another article on econometric models by drawing attention to the data from which the models are estimated. But the fact is that reliable and ample data is critical to an econometric model: if the underlying data is seriously deficient then our care in adopting optimal methods of estimation is of little avail. Every economist constructing a model is conscious of this fact and acknowledges the problems connected with the basic data by some ritualistic disclaimers: the data are of uncertain quality but this is the best that is available; the alternative to using it is to abandon all hopes of quantifying our policy recommendations; in any case the exercise is worthwhile for it highlights the particular areas in which the data-collection

1/ UNCTAD, Trade Prospects and Capital Needs of Developing Countries, United Nations, New York, 1968, (henceforth, UNCTAD, 1968). A similar volume has been published by the Economic Commission for Asia and the Far East: ECAFE, Feasible Growth and Trade Gap Projections in the ECAFE Region, Bangkok, 1968 (henceforth, ECAFE, 1968). This volume contains models for nine countries. The model for Thailand, however, is the same as the one contained in UNCTAD, 1968. All subsequent references to ECAFE, 1968 relate to the eight models other than the model for Thailand.

2/ UNCTAD, 1968, op. cit., p. (iii).

efforts need to be strengthened.^{1/} Having protected his flanks with these disarming assertions the model-builder often proceeds to act in precisely the way he would have acted had the data been abundant and reliable.

The data of most developing countries presents three problems to the econometrician.

First, the series for most macro-variables cover only a few years. Consequently, from an econometric point of view the models operate on the subsistence margin in that their relationships are estimated from no more than 5 to 16 observations.^{2/}

Second, the figures are of varying and often very low reliability. Estimates of many macro-variables are based on guesses or on rules of thumbs derived from scanty empirical evidence. Investigators who have had the patience to examine the data carefully have often concluded that "we may have to admit that up to now we really are not in a position to tell how the main economic variables are changing in large areas of the third world."^{3/} Trade data, in fact, constitute the only set of figures for these countries in

^{1/} J.T. Scott and Earl O. Heady, "Econometricians and the Data Gap; Reply," American Journal of Agricultural Economics, Vol. 51, No. 1, February 1969, p. 188. UNCTAD, 1968 notes "the paucity, inadequacy and unreliability of the relevant statistical data (for, in this case, Nigeria). . . Suffice it to say that in many cases the estimates are rather crude and impressionistic with very little firm basis. . .," p. 150. It goes on to note that the national accounts in Nigeria are intimately tied up with population estimates as for many items the aggregate magnitude is obtained by multiplying an estimated per capita figure by the population estimate and that the 1963 census has shown population to be much higher than the figures assumed in the national accounts. "Nevertheless, use of available data is probably better than not using any data at all and provides a useful check on more qualitative approaches," ibid.

^{2/} The frequency distribution of the models in UNCTAD, 1968 and ECAFE, 1968 by the number of observations on which they are based is as follows:

<u>Number of observations:</u>		<u>5-7</u>	<u>8-10</u>	<u>11-13</u>	<u>14-16</u>
<u>Number of countries:</u>	(i) UNCTAD, 1968	2	7	3	2
	(ii) ECAFE, 1968	4	1	1	2

Equations in the four models of UNCTAD, 1968 that have not been included above are based on varying numbers of observations: Argentina (12 to 16), Ghana (10 to 16), Nigeria (10 to 13) and Philippines (9 to 14).

^{3/} L. Whitehead, "Basic data in poor countries: the Bolivian case," Oxford University Institute of Economics and Statistics, Bulletin, Vol. 31, No. 3, August 1969, p. 226.

which most economists are willing to place a good deal of confidence. But even this information when it is examined systematically - for example, when it is compared to the returns of the country's trading partners - reveals very substantial discrepancies.^{1/}

The underlying data are such that many refinements in estimation procedures, such as multiple stage least squares, seem to be somewhat pointless. In fact, one does not quite know how to assess even the more basic refinements. Consider the fact that most models in use base their estimates on constant price data.^{2/} Relationships are estimated at constant prices so as to separate the income and price effects. More often than not the model builder adopts the national accounts data for a country without examining the deflators used to arrive at the constant price series. To cite but one instance: UNCTAD, 1968 gives projections for Ceylon at constant 1960 prices.^{3/} It is interesting to examine some of the deflators used in Ceylonese national accounts. Production in the agricultural sector consists of a few identifiable crops: as the product is homogenous over space and time output indices are used and we do, in fact, obtain a constant price series that is meaningful. For all the other sectors value added at current prices is adjusted by the following odd assortment of deflators:

^{1/} As a recent and relevant illustration, cf., S. Naya and T. Morgan, The Accuracy of International Trade Data: The Case of Southeast Asian Countries, University of Wisconsin, mimeographed, August, 1967.

^{2/} Fourteen of the eighteen models in UNCTAD, 1968 use constant price series.

^{3/} UNCTAD, 1968, p. 75.

Sector	Deflator
Mining and Quarrying	Wage index for workers in industry and commerce.
Manufacturing	
(a) Tea, rubber, coconut processing	Indices of physical output.
(b) Factory and cottage industries excluding (a)	Domestic group of the Colombo cost of living index.
Construction	Domestic group of the Colombo cost of living index.
Electricity, water and sanitary services	Wage index of Government technical and clerical employees.
Transport, storage and communication	A mixture of indices of wages of Central Government employees, workers in industry and commerce and the Colombo cost of living index.
Wholesale and retail trade	Domestic group of the Colombo cost of living index.
Banking, insurance and real estate	Wage index for employees in industry and commerce.
Ownership of dwellings	Domestic group of the Colombo cost of living index.
Services	Colombo cost of living index.

Weights for the more widely used indices are based on surveys conducted many years ago when the structure of the economy was very different from what it is now. The most widely used deflator - the Colombo cost of living index - for instance, is based on a survey conducted in 1952.^{1/} The problems are

^{1/} Some studies explicitly acknowledge that for their countries current price series are more reliable than available constant price series: cf. for Thailand, UNCTAD, 1968, pp. 388, 390; for Tanzania, ibid., p. 190; Kenya, ibid., p. 107; Uganda, ibid., pp. 233-234.

compounded for economies that in addition to having deflators of somewhat limited reliability have had multiple or oft-changing exchange rates.^{1/}

For a similar reason one must be wary of reading too much into high coefficients of determination and t-ratios before examining the manner in which the relevant data were estimated in the first place. If, for instance, we fit an equation relating imports of construction materials to value added by construction in Ceylon we obtain a very satisfactory fit and a highly significant t-ratio. The reason for this is simply that value added by the bulk of the construction sector is itself estimated by multiplying the imports of construction materials by (4.25) - after allowing for a three-month lag. The \bar{R}^2 and the t-ratios in this case would tell us no more than the national accounts themselves. The multiplier of (4.25) is based on a survey conducted some years ago and unless we examine the survey itself we cannot attach significance to the impressive \bar{R}^2 and t-ratios.^{2/}

^{1/} For an instructive illustration of the gyrations one must go through to take account of this problem see B.A. de Vries and J.C. Liu, "An Econometric Analysis of Inflation and Growth in Brazil," Washington, D.C., 1969, (mimeographed), presented at the New York meetings of the Econometric Society, December 1969.

^{2/} These remarks are suggested by some equations in a recently published model of the Colombian economy: K. Marwah, "An Econometric Model of Colombia: A Prototype Devaluation View," *Econometrica*, April 1969, pp. 228-251. We are told that "there seems to be a strong complementary relation between the imports of capital goods and investment. A marginal increment of one billion pesos worth of imports of capital goods seems to be accompanied by an addition of 1.49 billion pesos of total investment," *ibid.*, p. 236. The t-ratio is more than 16 and the \bar{R}^2 is 0.981. I understand that in Colombia fixed investment is estimated by applying some multipliers (which have remained unchanged for some time) to the imports of capital goods.

The third source of difficulties is that the national accounts of many of the countries for which econometric models are being constructed at present are in their infancy and are still being revised periodically. Therefore, each of the models is "dated" in the sense of being tied to the national income series that were in use at the time when it was estimated. A rough check indicates that of the 37 countries for which UNCTAD, 1968 provides trade gap estimates the accounts for as many as 14 have been revised since 1965 (the last year for the sample period of UNCTAD, 1968) in ways that would materially affect the estimates of some important coefficients. The national accounts of Kenya, for instance, have been revised often and very considerably over the last few years. One has only to compare the data used by Paul Clark in his models for the East African economies,^{1/} that used by Faaland and Dahl in their model of Kenya,^{2/} and the data in use now, to see that series like gross domestic product at factor cost and gross investment have been revised upward by 15 to 20 percent.

Clearly, national accounts have to be revised as and when new data or new information becomes available. Unfortunately, one cannot be equally philosophic about the consequences these revisions have for the confidence one can retain in models estimated at so early a stage in the country's history of preparing national accounts. Consider the consequences of revising capital formation figures upwards by 15 or 20 percent: the savings figures in Kenya, as in many other countries, are derived as a

^{1/} Paul G. Clark, Development Planning in East Africa, East African Publishing House, 1965, especially, pp. 34-35.

^{2/} UNCTAD, 1968, pp. 104-148.

residual using the capital formation and balance of payments accounts. Now, the capital formation figures have been revised upwards by one-eighth to one-fifth but the balance of payments account has not undergone a comprehensive revision of any comparable magnitude. Therefore, for any given period one gets very different savings coefficients and a very different picture of the accelerations or decelerations in the country's savings efforts today than one would have if one were using the old series on capital formation. The same holds true of functions linking imports of goods to GDP at factor cost or to fixed investment. And there is no assurance that no further revisions will be needed.

These revisions affect not only individual coefficients but also the overall conclusions that emerge from the models. The Faaland-Dahl model for Kenya, for example, projects the savings gap to be the dominant one and proposes that Kenya should aim at increasing its marginal savings ratio to 25% in the 1970's.^{1/} The revised accounts being used now indicate, however, that the marginal savings rate in the period 1960-1968 itself was 25.2% in contrast to the 11.3% that emerged from the accounts Faaland and Dahl used. Mechanical projections from the author's Kenya model - a model that in its specification is quite similar to the Faaland-Dahl model - indicate that savings are not likely to be an important issue in the 1970's and that the trade gap rather than the savings gap is liable to be the dominant one.

The equations contained in the kind of models we are now considering do provide a compact description of large amounts of data. As such they may sometimes highlight its peculiarities and may dramatize the fact that the assumptions on which the data has been put together need to be reexamined.

^{1/} UNCTAD, 1968, pp. 126, 137-138.

Sometimes one obtains such insights as by-products of the model building exercise. But clearly they cannot by themselves be a sufficient justification for constructing elaborate models out of shaky data. Of course, in cases where the underlying relationships are "mechanical" or "biological" and where they have been extensively documented for a number of similar systems, models can be used to generate data from only a few apparently unrelated bits of information. Demographic models of human and animal populations exemplify the use of models for this purpose. But given our limited understanding of change in developing economies and the even more limited extent to which their experience has been quantified, one can scarcely hope to use these simple and highly aggregated models to generate data just yet.

These deficiencies in the underlying data cannot be allowed for in any straightforward way. Methods that have been developed for handling errors in variables require a good deal of information about the nature and extent of the errors. Furthermore, these methods require that the errors in each variable be independent of the variable, of other variables that figure in the equation and of the errors in other variables. Given the way in which the national accounts of many countries are put together these requirements are seldom met. Estimates for the series of an individual aggregate variable are derived from diverse sources of which some sketchy information about other variables is certainly an important one. Not only is one estimate derived or 'built up' from the others but in the final stages they are all adjusted and doctored for securing consistency.

III. The Significance of Individual Coefficients

Even if the data was entirely reliable and covered many years, neither of which is the case, the coefficients of the regression equations

of these models would have only a very limited significance. This is often overlooked as evidenced by the fact that in addition to their use for forecasting these highly aggregated models are almost invariably used for purposes of control - i.e., for making policy recommendations. The coefficients of a fitted regression equation are treated not merely as coefficients that have emerged from a particular system of simultaneous equations, which is in fact all that they are, but are too readily identified with the marginal propensities and elasticities to which the economist is accustomed and which have traditionally had overtones of causality around them.

The abuse of regression analysis in this particular form is almost universal and a very large number of examples can be listed. A few will have to suffice.

Thus, for example, a model for Argentina fits private consumption expenditures to wage and non-wage income and gets the equation:

$$\begin{aligned} \text{Private consumption expenditure} &= -106.04 + 1.006 \text{ Wage Income} \\ &\quad (0.12) \quad (3.574) \\ &\quad + 0.655 \text{ Non-Wage Income} \\ &\quad (6.625) \end{aligned}$$

with an \bar{R}^2 of 0.93. Figures in parentheses are t-ratios. It goes on to assert that "as may be seen from the above equations, there exists a markedly different consumption behavior between wage and non-wage income groups. Thus, changes in the distribution of income are bound to have a significant effect on the level of total consumption."^{1/} The coefficients seem "reasonable" only because we are accustomed to thinking that the marginal propensity to consume out of "wage incomes" is much higher than out of "non-wage incomes" primarily because we associate the former with the poorer and the latter with the richer sections of the population. In fact, in this model most of the income originating in the agricultural sector is treated as "non-wage income": a fact that must bring the per capita incomes of the wage and non-wage groups closer together. But far more important from our point of view is the fact that the two independent variables are highly collinear.^{2/} It is only fortuitous that the coefficient for "non-wage income" has turned out to be smaller than that for "wage income".

1/ UNCTAD, 1968, p. 408.

2/ The simple correlation between wage and non-wage incomes at current prices is 0.9982 (cf., data in UNCTAD, 1968, p. 408). The equation is based on incomes at constant prices. It has not been possible to obtain the correlation coefficient for these as the report does not furnish the data on an income at constant prices.

Similarly, a recently published model of the Colombian economy designed specifically to analyze the consequences of changes in exchange rate policies treats each of the coefficients as economic propensities and elasticities - even when some important explanatory variables that are left out are highly collinear with the ones that have been included.^{1/}

B.A. de Vries and J.C. Liu in their recent paper on Brazil^{2/} provide a very explicit illustration of economists attributing causal significance to each of their coefficients. At each step this instructive paper carefully stresses the number of significant explanatory variables that have not been included in each equation (many of these are probably collinear with the ones that have been included), the poor quality of data, the partial specification of the model itself and many other difficulties; and yet it interprets the resulting coefficients as if they were the conventional elasticities and propensities to which the economist is accustomed. The paper proceeds to derive trade-offs between inflation and growth.

For a large number of reasons we are not justified in using regression equations and their coefficients for control purposes. We shall discuss two of these reasons. First, there is the fact that in a very large number of cases the explanatory variables that have been included in an equation are highly collinear and, second, the likelihood that the equations are underspecified and that some of the excluded variables may well be collinear with the variables that have been included in the equation.

(a) Collinearity Among Regressors^{3/}

We know that if the regressors are orthogonal, if the equation is specified correctly and if the data is reliable an unambiguous meaning can be attached to the estimated coefficients. At the other extreme, one

^{1/} K. Marwah, op. cit., in particular, the discussion of the consumption equations, pp. 234-235.

^{2/} B.A. de Vries and J.C. Liu, op. cit.

^{3/} The following is by no means a comprehensive treatment of the problems that arise from multicollinearity in regression analysis and their solutions. A brief account is given in J. Johnston, Econometric Methods, McGraw Hill, New York, 1963, pp. 201-207. For two recent statements see D.C. Farrar and R.R. Glauber "Multicollinearity in Regression Analysis: The Problem Revisited," Review of Economics and Statistics, Vol. XLIX, February 1967, and Y. Haitovsky, "Multicollinearity in Regression Analysis: Comment," Review of Economics and Statistics, LI, November, 1969.

cannot estimate the coefficients if the regressors are perfectly collinear. In practice the situation one encounters is an intermediate one: over short periods of five to fifteen years a surprisingly large number of aggregative time series have strong trend components and are, therefore, highly though not perfectly collinear.^{1/} The use of such variables as independent regressors in multiple regression equations raises difficult problems of interpretation.

Individual coefficients have little significance if the regressors in a multiple regression equation are collinear. In particular if the regressors X_1 and X_2 are not strictly independent it is misleading to assume that a coefficient, say b_2 , in the equation ($Y = b_1X_1 + b_2X_2$) indicates the extent to which Y will change as a consequence of a unit change in X_2 'other things remaining the same'. To begin with, the high correlation between X_1 and X_2 over the sample period may imply that there is some organic dependence between the two regressors so that when x_2 is altered 'other things' - in this case, X_1 - may not 'remain the same'. The coefficient b_2 in the single equation model ($Y = b_1X_1 + b_2X_2$) does not take account of these indirect

^{1/} Of the 630-odd estimated equations that are listed in UNCTAD, 1968 about 168 have more than one regressors. The data given in UNCTAD, 1968 enabled us to calculate the intercorrelations among the regressors in 118 of these equations. A total of 168 intercorrelations among different regressors were involved. The distribution of these intercorrelations is as follows:

	<u>Value of the Correlation Coefficient</u>								
	<u>0.0 to 0.6</u>	<u>>0.6</u> <u>≤0.65</u>	<u>>0.65</u> <u>≤0.7</u>	<u>>0.7</u> <u>≤0.75</u>	<u>>0.75</u> <u>≤0.8</u>	<u>>0.8</u> <u>≤0.85</u>	<u>>0.85</u> <u>≤0.9</u>	<u>>0.9</u> <u>≤0.95</u>	<u>>0.95</u> <u>≤1.0</u>
<u>Number</u>	12	6	6	12	6	4	6	13	103

In other words, about 60% of the intercorrelations among supposedly 'independent' variables are higher than 0.95 and about 70% are higher than 0.90.

effects of X_2 on Y via the changes it induces in X_1 . Even if one has reason to believe that there is no organic dependence between X_1 and X_2 and that they 'just happened to have been collinear in the sample period' there are various reasons why we should be wary of using the coefficients as if they were the partial derivatives of calculus.

First, if the correlations between regressors are indeed as high as (0.95) - as they most often are in the models under review - or (0.9985) as in Tanzanian equation cited below the values one obtains for the coefficients are very sensitive to sample coverage, data errors, specification of the equation and even rounding-off errors. The argument can be stated compactly as follows. Let R be the matrix of coefficients of correlation (r_{ij}) between the regressor variables; let b^* denote the vector of standard partial regression coefficients so that the regression coefficient b_j is equal to $(b_j^* S_y / S_{x_j})$ and let g be the vector of coefficients of correlation between the dependent variable and the regressors. Then,

$$Rb^* = g$$

Therefore,

$$b^* = R^{-1}g$$

Now, R^{-1} is $[\bar{R}_{ij} / |R|]'$ where $[\bar{R}_{ij}]'$ is the adjoint matrix of R consisting of cofactors and $|R|$ is the determinant obtained from R . The closer a regressor is to being the linear function of other regressors the closer the value of the determinant $|R|$ will be to zero. With a divisor close to zero, the values of elements in the inverse matrix R^{-1} (from which the vector b^* and ultimately b and the standard errors of the individual coefficients are estimated) become extremely volatile: even small changes in

the r_{ij} arising from sample coverage, or data errors or from the effects of rounding-off can alter the value of $|R|$ greatly. In order to gauge the extent of this volatility it is sufficient to examine the simple correlation between the regressors when the equation has only two regressors. But if there are more than two regressors the value of R can be close to zero even if none of the correlation coefficients is individually close to 1.0. The reason for this is that an individual regressor X_1 though not very highly correlated with any one of the other regressors may yet be a linear function of all or some (i.e., more than one) of the other regressors. In this case an inspection of the simple correlation coefficients will not suffice; one must look at the value of the determinant $|R|$ itself.^{1/}

^{1/} The diagonal elements of the symmetric matrix (R) are 1.0; the off-diagonal elements can vary between -1.0 and 1.0. Thus the value of the determinant $|R|$ can vary between 1.0 and zero. In the two-variable case if the two regressors are completely independent the matrix (R) is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ so that the determinant is 1.0. If they are perfectly correlated, (R) is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and so the determinant is zero. As the value of the determinant is a function solely of the coefficient of correlation between the two regressors it is sufficient to examine the value of the coefficient. But when more than two regressors are involved the individual correlation coefficients are not always adequate guides to the value of the determinant. Consider the following array of four regressors:

	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>
X_1	2.0	0.0	0.0	3.5	0.0	0.0	4.0	0.0	0.0
X_2	0.0	2.5	0.0	0.0	3.0	0.0	0.0	3.5	0.0
X_3	0.0	0.0	3.0	0.0	0.0	2.5	0.0	0.0	4.0
X_4	2.0	2.5	3.0	3.5	3.0	2.5	4.0	3.5	4.0

As X_4 is a linear combination (in this case just the sum) of the other three the value of the determinant to be used in the regression computations is zero. Yet the individual correlation coefficients are as follows: $r_{12} = 0.0$; $r_{13} = 0.0$; $r_{14} = 0.59531$; $r_{23} = 0.0$; $r_{24} = 0.54972$ and $r_{34} = 0.58600$. Compared to the correlation coefficients one encounters in the models under consideration none of these is alarmingly high. In fact, as there are only nine observations one could even maintain that none of the coefficients is significantly different from zero at, say, the 5% significance level.

Second, the standard error of the estimate of a coefficient increases as the degree of collinearity increases. When the correlations among regressors are as high as 0.9985 or 0.95 the standard errors are often so high as to make the point estimate of the coefficient virtually useless for policy purposes. In this context two conventions regarding standard errors and t-ratios deserve comment. First, the customary rule of thumb that is used in evaluating the t-ratios is to regard them as acceptable if they are greater than two. This rule is a carry-over from the days when the primary interest was in verifying if the b-coefficient was significantly different from zero; that is, whether - in a qualitative sense - the independent and dependent variables were related. But if as in the models we are considering the equations and their coefficients are being used to make forecasts and to prescribe policy then one would expect that the model builders would look for t-ratios that are much higher than two. A t-ratio of two, after all, means that the standard error of the coefficient is one-half as large as the coefficient itself. Consequently, the forecasts and the policy conclusions one can draw from it will indeed cover a very wide band. Second, the fact that the standard errors increase with increasing collinearity has been taken to imply that one need not worry about multicollinearity as long as the standard errors do not increase so much as to render the coefficients insignificantly different from zero; as long, say, as the t-ratios are not below the magic figure of two. This is a mistaken deduction. High collinearity does not always result in high or 'unacceptably high' standard errors, though it always affects the meaningfulness of the estimated coefficients. This point of view needs some elaboration.

Consider an equation with two regressors: $Y = b_1X_1 + b_2X_2 + u$.
 The standard error for b_1 is $S_{b_1} = \Sigma y^2(1 - R_{y.12}^2)/(n - k - 1)\Sigma x_1^2(1 - r_{12}^2)$
 where $y^2 = (Y - \bar{Y})^2$, $x_i = (X_i - \bar{X}_i)$, $R_{y.12}$ is the coefficient of multiple
 determination and r_{12} is the coefficient of correlation between X_1 and X_2 .^{1/}
 The greater the linear dependence between X_1 and X_2 - i.e. the closer r_{12} is
 to 1.0 - the closer the entire denominator is to zero and, hence, given other
 things, higher is the standard error. But in a very large proportion of mul-
 tiple regression equations reported, for instance, in UNCTAD, 1968 while the
 intercorrelation among regressors is 0.95 or 0.9985 the R^2 term in the numerator
 term is also close to 1.0. Therefore, the term $(1 - R_{y.12}^2)/(1 - r_{12}^2)$ is almost
 unity and the effects of extremely high collinearity are - depending on one's
 point of view - 'absorbed', 'submerged' or 'neutralized' by the equally high $R_{y.12}^2$.

^{1/} The expression used in the text can be derived as follows in the two
 variable case. We know that

$$\begin{bmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \Sigma x_1 y \\ \Sigma x_2 y \end{bmatrix}$$

where $y = (Y_i - \bar{Y})$, $x_1 = (X_1 - \bar{X}_1)$ and $x_2 = (X_2 - \bar{X}_2)$.

Thus

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \Sigma x_1^2 & \Sigma x_1 x_2 \\ \Sigma x_1 x_2 & \Sigma x_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma x_1 y \\ \Sigma x_2 y \end{bmatrix}$$

The inverse matrix is:

$$1/(\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1 x_2)^2) \begin{bmatrix} \Sigma x_2^2 & -\Sigma x_1 x_2 \\ -\Sigma x_1 x_2 & \Sigma x_1^2 \end{bmatrix}$$

The standard error for b_1 is

$$S_{b_1} = \text{MSE}[\Sigma x_2^2 / (\Sigma x_1^2 \Sigma x_2^2 - (\Sigma x_1 x_2)^2)]$$

where MSE is the Mean Square Error.

$$\begin{aligned} \text{Thus, } S_{b_1} &= \text{MSE}[1/\Sigma x_1^2(1 - (\Sigma x_1 x_2)^2/\Sigma x_1^2 \Sigma x_2^2)] \\ &= \text{MSE}[1/\Sigma x_1^2(1 - r_{12}^2)] \\ &= \Sigma y^2(1 - R_{y.12}^2)/(n - k - 1) \Sigma x_1^2(1 - r_{12}^2) \end{aligned}$$

Technically the estimated coefficients are unbiased and among the class of linear estimators they are the best. Moreover, as the relatively low standard errors testify, they are also relatively 'precise'. A coefficient may have all these properties and yet may completely fail to reflect the actual economic relationships. A sample of observations, even when it yields the best linear unbiased estimates of coefficients that have high t-ratios, may yet not be rich enough to be able to disentangle the separate effects of the regressors in an economically meaningful sense.

A set of equations reported in UNCTAD, 1968 provide an instructive illustration. Faaland and Dahl fit the following regression to account for the value of agricultural output in Tanzania:^{1/}

$$P_A = 34.322 + 0.402(\text{GDP}) \quad \bar{R}^2 = 0.94 \quad \text{DW} = 1.59 \\ (0.036)$$

where P_A is the value of agriculture and the expression in parenthesis is the standard error of the coefficient. To take account of the fact that demand patterns may be changing because of the growth of per capita income and the diminishing importance of the subsistence sector they fit another regression:

$$P_A = -8.001 + 5.698(\text{GDP}/N) \quad \bar{R}^2 = 0.96 \quad \text{DW} = 1.57 \\ (0.397)$$

Then, to see how the relationship between agricultural output and GDP is being modified by changing income levels they fit an equation containing both GDP and GDP per head as explanatory variables:

$$P_A = -131.716 - 1.226(\text{GDP}) + 22.816(\text{GDP}/N) \quad \bar{R}^2 = 0.99 \quad \text{DW} = 2.13 \\ (0.239) \quad (3.345)$$

The coefficient of correlation between (GDP) and (GDP/N) is 0.9985 and, as the authors point out, the negative derivative for GDP (-1.226 + 22.816/N) makes little economic sense. Yet the t-ratios are highly significant and, if we assume that the third equation is correctly specified, the estimated coefficients are the best linear unbiased estimates.

When collinearities among regressors lead to such patently implausible results one is at once made aware of their distorting influence.

^{1/} UNCTAD, 1968, p. 204.

But this influence is present even when the resulting coefficients do not appear to be obviously or dramatically unreasonable. After all the negative and significant coefficient for (GDP) in the third equation cited above is only an extreme illustration of the effect of collinearity among regressors. The problem here is that too little attention is paid to formulating refutable and precise hypotheses before sending the data down to the computers. As the initial hypotheses are not precisely formulated often one has little reason to scrutinize or doubt the results that one obtains. Consequently, the text accompanying the equations just sets out either to rationalize the values that emerge from the calculations or to reject the results in a perfunctory manner.

From current practice it would seem that for many relationships it is possible to rationalize a wide range of values of coefficients including positive as well as negative signs. Examples can be cited about the direction and extent of the relationship between inflation and consumption, the determinants of Government consumption expenditures, changes in consumption expenditures and their impact on investments via their effect on expectations and so on. A single example may suffice. Consider the relationship of the existing capital stock to current or future gross investment. A negative sign or a very low value of the coefficient could reflect the depressing effects of a larger existing stock on net investment - a hypothesis that is not altogether unreasonable in the kind of capital-stock-adjustment or accelerator models in which demand considerations dominate. On the other hand, a positive and numerically significant coefficient could reflect, first, the stimulus that a more adequate infrastructure provides for investment and, second, the purely arithmetical fact that larger capital stock

requires larger replacement outlays. As one seldom has reliable estimates for replacement outlays and as, very often, one does not even have sectoral investment detail one has to make do with treatments such as the following:

"The role of the capital stock in an investment function has been of concern to many researchers. Various alternative specifications of the function have been tried and the coefficient invariably bears a positive sign. Others working with developed economies have found it to be negative. Of course, in a developing economy the positive sign may be valid. However, the capital stock variable has not been included in the relevant equation in the model."^{1/}

The scope for rationalizations is even greater when what is at issue is not the sign but the numerical magnitude of a coefficient. Ought we to expect constant returns to scale when fitting aggregate production functions for developing economies? What is a reasonable income elasticity of demand for imported consumer durables? 'Common sense' or one's 'feel for figures' is a poor guide on these issues: its only criterion is that the result under consideration should not look too different from past results. And in many cases the past results in their turn are only based on a more or less unsystematic examination (for instance, with the aid of some earlier inadequately specified regression model) of more or less unsatisfactory data.

We see, therefore, that a very large proportion of the regressors that are used are highly collinear, that the resulting coefficients may have little economic meaning and that we may not always be able to detect this either by using statistical tests like the t-test or by appealing to our intuition. Given these facts, how do the model builders tackle the problems posed by collinear regressors? The method that is most often mentioned in econometrics texts as a way of dealing with collinearities is to use some

^{1/} M. Dutta and V. Su, "An Econometric Model of Puerto Rico," Review of Economic Studies, Vol. XXXVI (3); No. 107, July 1969, p. 326.

external or a priori information to supply estimates for the coefficients of all but one of the collinear variables. In practice this 'solution' is seldom adopted - reflecting no doubt the difficulties of obtaining a priori estimates for the coefficients. Nor do the models extract the principal components of the regressors and use these instead of the regressors in the equations. Perhaps the reason is that the model builders are interested in the regressors themselves rather than in entities called their 'principal components' which they cannot very readily manipulate for policy or projection purposes. The 'solution' that is most frequently adopted in practice is to drop one or more of the collinear regressors from the estimating equation. Once again a number of examples could be cited, of which two may suffice.

So as to find a function determining Government consumption expenditures, de Vries and Liu try a number of explanatory variables and report three equations:^{1/}

$$\begin{array}{ll} C_G = -337.8 + 8.96N & \bar{R}^2 = 0.95 \quad d = 1.34 \\ & (14.89) \\ = 31.6 + 0.62T & \bar{R}^2 = 0.94 \quad d = 1.42 \\ & (14.4) \\ = -2290 - 0.30T + 48.0 N - 61.7t - 5.87P_3 & \bar{R}^2 = 0.97 \quad d = 1.34 \\ & (2.08) \quad (4.82) \quad (3.79) \quad (3.54) \end{array}$$

where N is population, T is taxes, t is time and P_3 is a price index. The entries in parenthesis are t-ratios. The authors choose the first equation "because it seems more important in expressing the economic reason behind an increase in government consumption." The third function is rejected because "though showing significant parameters for all four variables, the sign of T is reverse to the expectation and others are due to multicollinearity," (sic).^{2/} Notice the extent to which the coefficient for N in the equation that is selected (8.96) differs from that in the third (48.0).

On six occasions authors in UNCTAD - 1968^{3/} postulate that output in the agricultural sector and in "other sectors, mainly manufactures" is a function of total GDP and of GDP per capita. When the output of either of the sectors is regressed on GDP alone or on GDP per capita alone it yields highly 'significant' coefficients. When the two regressors are used

^{1/} B.A. de Vries and J.C. Liu, op. cit., p. 11.

^{2/} Ibid.

^{3/} UNCTAD, 1968, pp. 123-124, 204-205, 248-249.

together the values of the coefficients and their standard errors alter dramatically. Some of the coefficients are no longer "significantly different from zero," though some are - as in the equation for Tanzania's agricultural output cited above. But in each of the six cases after a perfunctory discussion of the standard errors being 'relatively high' or of a sign being 'improbable' on a priori grounds the second regressor - per capita GDP - is dropped.

The authors seldom pause to notice that by dropping a regressor or a group of regressors to avoid problems of collinearity they are altering the reduced form of their models very considerably and hence the policy conclusions that can be drawn from them. Moreover, and this leads us to the next sub-section, by dropping these variables the authors introduce serious specification errors into their equations.

(b) Specification Errors: The Choice of Explanatory Variables

The second and in practice an equally important group of reasons for which it may be misleading to treat individual coefficients as the partial derivatives of calculus has to do with the set of explanatory variables that have not been included in an equation. The standard argument for least squares proceeds as follows: let X and Z be matrices and Y , b , β and θ be vectors; assume that the true model is

$$Y = X\beta + u \quad (i)$$

Least squares procedures estimate β by b where

$$b = (X'X)^{-1} X'Y \quad (ii)$$

Substituting for Y from (i) and taking expectations we get

$$E(b) = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'E(u) \quad (iii)$$

But as $(X'X)^{-1} X'X = I$ and $E(u) = 0$

$$E(b) = \beta \quad (iv)$$

That is, b is an unbiased estimator of β . If however, the correct model is

$$Y = X\beta + Z\theta + w \quad (v)$$

while the model used for estimation is

$$Y = X\beta + v$$

That is, let us assume that the analyst has left out some of the explanatory variables because, as we saw in the preceding section, he wants to avoid problems arising from collinearities among regressors, because he has misconstrued the causal relationships, because he cannot obtain satisfactory data about them or for some other reason then we obtain biased estimates of β . As before,

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

Substituting for Y - but this time from (v) - and taking expectations we get

$$E(b) = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'Z\theta + (X'X)^{-1} X'E(w) \quad (\text{vi})$$

As $(X'X)^{-1} X'X = I$ and $E(w) = 0$ we have

$$E(b) = \beta + A\theta \quad (\text{vii})$$

where A is $(X'X)^{-1} X'Z$ and is known as the alias matrix. Thus the omission of a variable that has some causal links with Y means not only that in the end we do not have an estimate of the coefficient linking the excluded variable with Y but also that the estimates of coefficients linking all the other X_i to Y may be vitiated. Incidentally, if the Z are included when in fact they do not have any links with Y , or when even if they are linked to Y they are not correlated with X , they will not bias the estimates of β . If $\theta = 0$ or if $\delta = 0$, then $b = \beta$ in (vi) or (vii) and we have an unbiased estimator of β .

If one is only interested in forecasting and if one has reason to believe that the future relationship between (X) and (Z) will be the same as in the past, there is nothing wrong in using an abbreviated model

$\underline{Y} = f(X)$ and having the (X) do the work for themselves as well as for the (Z) .

But if we are going to use the coefficients for policy prescription then this is precisely the situation when leaving out (Z) as in the abbreviated model

will give us highly misleading estimates of the coefficients linking (Y) and (X). Instead of leaving (Z) out the model-builder should search out the parent variables that explain (Z) and (X) themselves - if such variables exist - and substitute those variables into the equation. If the search does not yield the ultimate explanation for the collinearity of (X) and (Z) one can indeed argue that both sets of variables, (X) and (Z), should be included even if the t-ratios of one or both are not significant so that mechanical regression procedures do not attribute to one of them the variance of the dependent variable that the two together help explain. From this negative point of view a variable may be very important even if it is "statistically insignificant".

We have already stressed that many econometric models are being used for control purposes also. The question that arises then is: how often is it that in practice explanatory variables that have a quantitatively significant influence on the dependent variables get left out from one's estimating equations? One has only to scan the text (and perhaps the footnotes) accompanying the equations of a model to realize that - for one reason or another - every author has had to leave out quite a few variables from most equations that he would have liked to include.^{1/} Apart from the fact that for many of the important explanatory variables the relevant data is not available, and that some of them are very difficult to quantify anyhow, there is the pervasive difficulty that economic theory offers insufficient guidance as to what constitutes an adequate list of explanatory variables. Some models assume very simple functional relationships; others

^{1/} One repeatedly comes across passages like the following: "The suitable explanatory variable for raw material imports seems to be the level of output in secondary industry (manufacturing, construction and electricity). However, in the absence of necessary data, the following two alternative forms (linking imports of raw materials to fixed investment and to GDP) were obtained," UNCTAD, 1968, p. 360. "The model fails to include all the relevant variables affecting inflation and growth either because of the lack of statistical data or because of the inability to identify them properly," B.A. de Vries and J.C. Liu, *op. cit.*, p. 1; "Since the data about the availability of tractors and consumption of agricultural products are not available, the former is excluded and the latter is substituted by a proxy variable, aggregate consumption," *ibid.*, p. 7. And so on.

are more venturesome and imaginative, making the birth rate a function of per capita consumer expenditure and time; rates of exchange a function of capital transfers into the country in the preceding period, the ratio of the official exchange rate to the GDP deflator and the "capacity to import" which in turn is to be a function of the price index of the country's imports; and investment a function of earnings of the country from tourism and net credit extended to the private sector.

Now, there is no logical reason why any of these or more esoteric variables should not have been included; but the list does suggest that when one opts for one or a very small number of explanatory variables one has possibly not exhausted the field for potentially relevant candidates. Thus the risk is always present that the estimates of the coefficients are misleading. In a sense, the very fact that in devising a model the model-builder almost invariably has a specific purpose in mind (for instance, to project the two gaps, to analyze the consequences of different exchange rate policies, to estimate the "trade-offs" between various rates of growth and inflation) itself greatly influences his choice of explanatory variables. This becomes apparent when one compares three or four models for the same economy. The sets of equations in them differ much more in their list of explanatory variables than they do in their lists of dependent variables: a model aimed at analyzing the choices between inflation and growth attempts to include price variables at as many points as possible; another shows a marked preference for the exchange rate; a third for ratios depicting structural changes in the economy. As the purpose influences one's selection of explanatory variables, it may indirectly influence the values and signs of the coefficients derived and, thus, the numerical conclusions that are

drawn from the model.^{1/}

Each of the functional equations in a model can be specified in a variety of ways. When we use the model as a whole - when, for example, we focus on the import-export gap (a magnitude that depends on a large number of equations) - the uncertain specifications of each structural equation are communicated to the final result: one cannot be certain that the biases will cancel and not reinforce each other. This is especially important because often, as in the case of the import-export gap, we are interested in what are relatively small differences between fairly large variables so that even small errors in the latter are liable to affect the former in a major way; in fact, there often are other variables in the system - a country's debt servicing obligations are a case in point - that depend on the cumulative values of these small differences. Moreover, just as there is no unique way of specifying a single equation so is there no unique way of specifying the model as a whole. To cite but one example: the conventional two-gap models are designed to project the investment-savings gap and the import-export gap independently and are by their very nature overdetermined. There are many ways of reconciling the two gaps and the outcome in terms of the final projections and policy prescriptions will differ depending on which particular reconciliation procedure is adopted. Considering the numerous possibilities open to him in specifying a model and its individual relationships the model builder may find it very difficult to be sure that his

^{1/} The exceptions would consist of two types of cases: the somewhat uninteresting one in which all the explanatory variables are assumed to change at the same or very similar rates of growth; and the very special case in which they change in such a manner as to exactly offset each other's impact.

projections and policy prescriptions are anything more than a mechanical result of the particular way in which he has specified his model.

IV. Functional Form of the Equations

One way in which a model-builder may misspecify the relationship between a dependent variable and a set of independent variables is to assume that the relationship is, say, linear when in fact it is, say, quadratic. This is a particular case of the more general class of specification errors considered above. A few additional comments about it, however, may be worthwhile.

For a variety of reasons the choice between one functional form and another is not a matter of indifference.

First, each form implies a very specific view of underlying economic relationships. There is nothing in economic theory or in the empirical research conducted thus far which tells us that the most reasonable assumption about the future is that the marginal rate will remain unchanged (the assumption one makes when one uses a linear equation of the $Y = a + bX$ variety). It may be just as reasonable to assume that both the average and marginal rates will remain constant and use the relationship ($Y = bX$); or to assume that both will vary but that the elasticity will remain unchanged and use ($\log Y = \log a + b \log X$); or to assume that the elasticity will fall as Y rises and use a semi-log relationship, ($Y = a + b \log X$); or that the elasticity will fall as X increases using, then, the log-inverse function ($\log Y = a - b/X$); or that the average and marginal rates as well as the elasticity will vary without any particular restriction as, say, in the quadratic equation, ($Y = a + bX + cX^2$). In dealing with a few problems one can invoke some consistency conditions (for

instance, the homogeneity conditions in demand analysis) to limit the choices open to the model-builder or to constrain the coefficients in some particular way. But for most of the relationships that highly aggregated models of the kind considered here try to quantify our knowledge of economic relationships is just not specific enough to unambiguously suggest that we should use one form and not another.

Second, one cannot just abdicate one's responsibility and say that one will choose that particular form which fits the sample period data best. For in many cases various alternative forms fit this data equally well. But, one may ask, if different forms fit the sample period data equally well, does it make much practical difference as to which of those one uses for making projections or drawing policy conclusions? Indeed it does. Consider the following equations which have been computed from Kenyan data for 1960-1968.

		DW	F-Ratio	\bar{R}^2
$(DIRTAX)_t$	$= 0.04847 (GDPFAC)_t$ (.00191)	0.378	36.997	0.822
$(DIRTAX)_t$	$= -9.03007 + 0.07468 (GDPFAC)_t$ (0.39871) (0.00704)	1.124	112.502	0.933
$\log(DIRTAX)_t$	$= -2.54325 + 1.48174 \log(GDPFAC)_t$ (0.00969) (0.13391)	1.129	122.441	0.936
$(DIRTAX)_t$	$= 34.15114 - 0.18211 (GDPFAC)_t + 0.00037 (GDPFAC)_t^2$ (0.19761) (0.05425) (0.00008)	3.069	240.242	0.984
$(MCAPTL)_t$	$= .27497 (INVEST)_t$ (.01945)	0.612	34.920	0.814
$(MCAPTL)_t$	$= -5.28821 + 0.36031 (INVEST)_t$ (1.02273) (0.05305)	0.942	46.125	0.849
$\log(MCAPTL)_t$	$= -1.37567 + 1.44553 \log(INVEST)_t$ (0.03465) (0.25389)	0.866	32.417	0.797
$(MCAPTL)_t$	$= -27.87714 + 1.14340 (INVEST)_t - 0.00604 (INVEST)_t^2$ (0.96160) (0.56760) (0.00436)	1.446	27.047	0.867
$\log(MCAPTL)_t$	$= 1.74490 - 28.71198 \sqrt{1/(INVEST)_t}$ (0.05169) (2.49690)	1.864	132.228	0.943

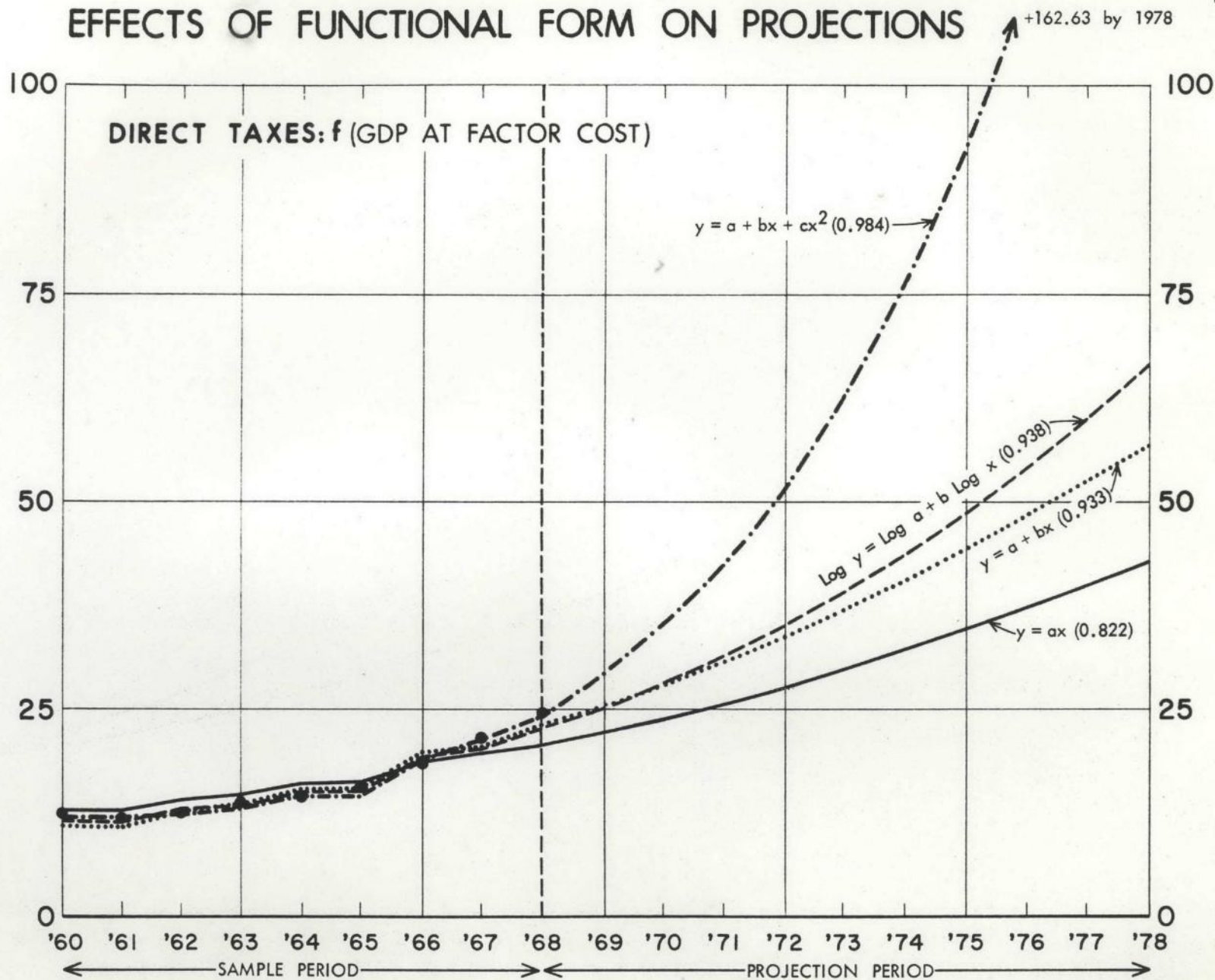
In the first set of four equations receipts from direct taxes are made a function of GDP at factor cost and in the second set of five equations imports of capital goods are made a function of gross fixed investment. Figures in parenthesis are standard errors. Each of the equations yields reasonably good fits for the sample period and the t-ratios are satisfactory for all of them. In Figures 1 and 2 the different equations are used to project the dependent variables into the future. In each figure identical values of the independent variable are used for all the equations. The diagrams show very clearly that the different forms yield widely divergent projections for the future.

The same sort of differences arise when one uses the first derivatives of the reduced form equations to assess the consequences of a change in some of the independent variables. While the first derivatives of equations like $(Y = aX)$ or $(Y = a + bX)$ are constant and unambiguous, those of equations like $(Y = aX^b)$ or $(Y = a + bX + cX^2)$ vary with the value of X . Therefore, their absolute values and, in many cases, the rankings of, say, individual industries according to their effect on the import-export gap, are very different depending on whether one takes X to be equal to the mean of the sample period or assigns to it the value it is projected to have at the end of the projection period. As the values of the derivatives vary, so must one's policy prescription.^{1/}

^{1/} If the data has some non-linearities in it non-linear equations like the quadratic one are bound to explode when extended far enough into the future. The point being made in the text is not that quadratic equations should always be used for long-term projections even when they yield patently absurd results; rather that the data often contains significant accelerations or decelerations, and if so these should be examined instead of being mechanically suppressed in a linear equation.

FIGURE 1

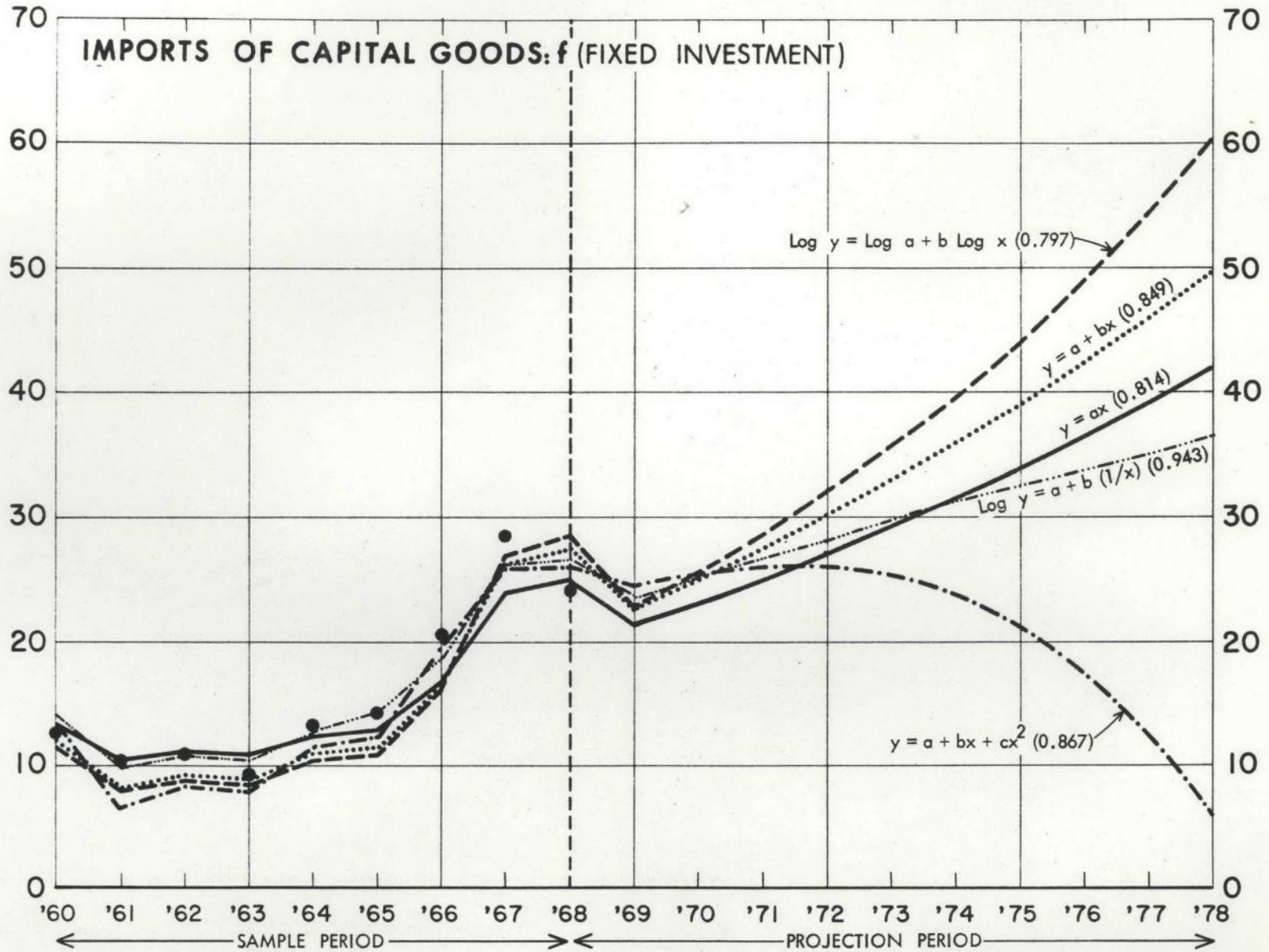
EFFECTS OF FUNCTIONAL FORM ON PROJECTIONS



1. Actual values in the sample period are represented by (●).
2. Values of R^2 are given in parenthesis.

EFFECTS OF FUNCTIONAL FORM ON PROJECTIONS

FIGURE 2



1. Actual values in the sample period are represented by (●).
2. Values of R^2 are given in parenthesis.

One does not have to seek far for the reasons that account for these differences. Least squares estimates of coefficients are fairly sensitive to extreme values. As the sample consists of only five to fifteen observations even a few aberrant values can affect the estimates very considerably. Moreover, as the equations are used to project variables many time periods beyond the sample period or to gauge the impact of very substantial changes in the 'policy variables' even small differences in the marginal rates of change implicit in the equations get magnified in a very substantial manner.

Finally, the choice of a particular functional form is not without its consequences for the efficacy of one's initial estimates of the coefficients. It can be shown - and the case is just a particular variety of specification errors considered above - that in general a linear equation will not yield an unbiased estimate of the true marginal rate even close to the center of gravity of the observed point set when the true underlying relationship is, say, quadratic or log-linear. Biases arise both because of non-linearity in the true relationship and because of an erroneous assessment of disturbances.^{1/}

V. Mechanical Use of Certain Statistical Tests

Sections III and IV have sketched some of the difficulties that arise from collinearities and different types of specification errors. The following question naturally comes to one's mind: do the standard statistical tests not alert us to the sorts of deficiencies that have been illustrated above? For example, if some important explanatory variables have been left

^{1/} Cf., H. Theil, "Specification Errors and the Estimation of Economic Relationships," Revue de l'Institut International de Statistique, Vol. 25, No. 1/3, 1957, pp. 41-51.

out of one's equation could we not assume that the value of the coefficient of determination will be extremely low or that the residuals would be highly auto-correlated and that the standard tests would draw our attention to the excluded variables? Similarly, if the independent variables are highly collinear will the low values of the t-ratios not warn us about their consequences? These questions need to be considered at some length for if the standard statistical tests used by the models are indeed sufficient to alert one to the sorts of problems considered in this paper then one can assume that the model-builders are justified in relying heavily on their equations and coefficients provided these pass the tests. The answers unfortunately are far from reassuring.

Most models display statistics relevant to three statistical tests: the coefficient of determination, the t-test for individual coefficients and the Durbin-Watson test for auto-correlation in the residuals. Among these the coefficient of determination receives the greatest attention. In fact, many models display an almost obsessive concern with the \bar{R}^2 .

Consider some extracts from a paper projecting employment patterns over 20 years:

"Group 0. For professional workers, the semi-logarithmic form gave the best overall fit and appeared to cope with elements of non-linearity in the relationship. The equation chosen was:

$$S_0 = -12.05 + 2.94 \ln Y + e, \\ (.21)$$

with R^2 equal to .843. Only per capita income shows a regression coefficient significant at the five percent level, although the linear form (with smaller R^2 of .811) showed population size significant as well. . . .

"Group 2. A semi-logarithmic form was shown to fit the data on clerical workers quite well ($R^2 = .784$), with no evidence against the linearity of this relationship. Although growth rate as well as income was significant in the simple linear fit, R^2 was lower (.766). Hence, the equation chosen was:

$$S_2 = -13.12 + 3.22 \ln Y + e.$$

"Group 4. For farmers and related occupations, the best fit is found with the semi-logarithmic relationship:

$$S_4 = 153.33 - 19.34 \ln Y + e, \\ (1.40)$$

with R^2 of .838. The scatter diagram revealed no problems with the linearity of this fit. The simple linear form shows population and growth rate significant (at the 10 percent confidence level) as well as income, but R^2 was much lower at .677. . . .

"Group 9. As with sales workers, neither equation form is very successful in explaining the variance. The semi-logarithmic form, however, had the higher R^2 (.278, significant by F-test), and will be used despite the appearance of non-linear elements in the corresponding scatter diagram (sic.). The equation used is:

$$S_9 = 5.55 + 1.63 \ln Y - 0.78 \ln P + e. \\ (.53) \quad (.31) \quad \dots \text{"1/}$$

Similarly, the Colombia model referred to above cites two equations explaining investment in construction:

$$\frac{I_{hH}}{P_{cnt}} = -0.9595 + 0.1312 X + 1.208 \left(\frac{I_{mcnt}}{P_{icnt}} \right)^{-1} \\ (0.0068) \quad (0.187) \\ \bar{R}^2 = 0.977, d = 2.327$$

and

$$\frac{I_{hH}}{P_{cnt}} = -0.311945 + 0.39590 \left[X - X_{-1} \right] + 0.5920 \left(\frac{I_{mcnt}}{P_{icnt}} \right)^{-1} \\ + 0.94580 \left(\frac{I_{hH}}{P_{cnt}} \right)^{-1} \\ (0.11380) \quad (0.3550) \\ \bar{R}^2 = 0.912, d = 1.590$$

The coefficients for X and (I_{mcnt}/P_{icnt}) are very different in the two equations and this should have provided an important clue to the "reliability" of the coefficients. However, the text merely notes that the former equation "was accepted in the final analysis on the basis of its predictive value."^{2/}

1/ J.S. Scoville, "The Occupational Structure of Employment, 1960-1980" in Sectoral Aspects of Projections for the World Economy, First International Seminar on Long-Term Economic Projections, United Nations, New York, 1969, Vol. 3, pp. 77-109; especially, pp. 86-90.

2/ K. Marwah, op. cit., pp. 233 and 236-237. (X) is GNP at 1958 prices; (I_{mcnt}) is value of imports of construction materials and (P_{cnt}) and (P_{icnt}) are indices of prices of construction materials in pesos and dollars.

In a superficial sense this preoccupation with the \bar{R}^2 is well rewarded. The models are fairly simple and highly aggregated representations of the economies of developing countries.^{1/} Few of them incorporate any institutional detail or any information about production relationships in the economy. Yet if one considers their structural equations individually one is struck by their very impressive coefficients of determination. It is often and rightly said that the process of growth and change in developing countries is a complex and in many respects an incompletely understood phenomenon. The equations of these models, on the other hand, seem to suggest that the broad features of growth and change in a host of dissimilar countries over the past 5 to 15 years can be represented by extremely simple equations with no more than one or two, or at the most four or five, explanatory variables.^{1/}

^{1/} The 18 models in UNCTAD, 1968 utilize about 370 estimated equations (though, as noted elsewhere, about 630 estimated equations are reported in the text). Their distribution by the number of regressors used in the equation is as follows.

Total number of equations	Equations with 1 regressor	Equations with 2 regressors	Equations with 3 regressors	Equations with 4 or more regressors
370	281	67	21	1

^{2/} On a rough count the 18 models in UNCTAD, 1968 between them use about 370 estimated equations and the 8 models of ECAFE, 1968 list about 120 computed equations. The distribution of these equations by their coefficients of determination is as follows:

Coefficient of determination:	<0.3	>0.3 & <0.4	>0.4 & <0.5	>0.5 & <0.6	>0.6 & <0.7	>0.7 & <0.8	>0.8 & <0.9	>0.9
% of Total Equations								
(i) UNCTAD, 1968	2.2	1.3	6.1	6.1	11.2	11.7	23.9	37.5
(ii) ECAFE, 1968	2.5	5.0	4.2	3.4	4.20	11.8	18.5	50.4

Even though the point has been made that complex social phenomena can often be represented by fairly simple mathematical expressions,^{1/} the high degree of success of the models and their simple equations, in that they seem to explain a wide range of dependent variables, comes as a considerable surprise. However, this success of the models in explaining developments in particular economies is almost entirely deceptive. And the models pay a substantial price for this apparent success. The foregoing extracts, for example, illustrate one consequence of relying so heavily on the coefficient of determination. As already noted, by mechanically adopting the equation with the higher \bar{R}^2 the authors in each case exclude a number of explanatory variables from the particular segment of their model. The reduced form and, perhaps also the dynamics of the model and hence the policy conclusions that can be derived from it are all altered drastically. Now, it so happens that often the coefficients of determination can be affected by so many extraneous factors which have nothing to do with functional or causal relationships between variables that their use can be very misleading - especially when the models or equations are to be used as a basis for policy prescription. Three sets of considerations will be used to illustrate some common situations in which reliance on the \bar{R}^2 can mislead the model builder.

By far the most frequently encountered situation is the one that Yule documented 45 years ago:^{2/} two series may display "nonsense correlations" because, even though unrelated in a causal sense, they have pronounced trend

1/ M.G. Kendall, "Natural Law in the Social Sciences," Journal of the Royal Statistical Society, Part 1, 1961, pp. 1-16, en passant.

2/ G. Udny Yule, "Why Do We Sometimes Get Nonsense Correlations Between Time Series?", Journal of the Royal Statistical Society, January 1926, Vol. LXXXIX, Part 1, pp. 1-64.

or cyclical components. Consider, for instance, Kenya's exports and imports of services other than tourism. These consist mainly of freight, transportation, insurance and similar items. Let us postulate that most of these arise from the movement of goods in foreign trade and, therefore, let us regress them against the single regressor: the total value of goods imported into and exported from Kenya. For different equation forms the \bar{R}^2 for the 1960-1968 data for the exports of services are between 0.912 and 0.950 and for the imports of services between 0.715 and 0.779. The F-ratios are highly significant as are the t-ratios. The fact, however, is that Kenya's foreign exchange earnings from "exports of services" and expenditures on "imports of services" are related only in an incidental way to its imports and exports of goods: first, Kenya has no large shipping lines in which these goods may be ferried and, second, imports of goods in the merchandise account and as used in computing the equations are recorded c.i.f. so that whatever impact they have on freight and transportation costs is already taken account of elsewhere. Kenya's earnings from freight and transport arise more from the imports and exports of Uganda and Tanzania rather than its own imports and exports. The incidental way in which its own imports and exports of goods enter the picture is that a small part of Kenya's earnings from exports of services consist of port-charges and part of these are collected from the ships that ferry goods specifically meant for or specifically originating in Kenya. In this case, therefore, the \bar{R}^2 would certainly have misled us if we abandoned our search for explanatory variables once we had hit upon the total value of goods imported into and exported from Kenya. But, one may argue, what is wrong in using the equations for projections or policy prescription? After all, the variables have moved together in the past. The difficulty

is that the less certain one is about the existence of a causal connection between the dependent and the explanatory variables, the less confident one is about asserting that "as they have moved together in the past it is reasonable to assume that they will move together in the future." Taking the 'exports of services' as an example, one would, in fact, be on firmer ground if one asserted that these are likely to grow at a rate lower than the growth of the total value of goods imported into and exported from Kenya: Uganda's exports and imports together are expected to grow at a rate about one-half of the rate being currently projected for Kenya's imports and exports of goods and Tanzania has definite plans of developing Dar-es-Salaam which will mean that it will be relying somewhat less on Mombasa in the future. To avoid pitfalls of this kind models of developed economies almost invariably estimate their relationships from data that has been transformed - for instance, by taking its first differences - to avoid the effects of serial correlation in the series. But hardly any of the newer models for developing countries is estimated in terms of first differences or similarly transformed data.^{1/}

^{1/} Not one among the 18 models of UNCTAD, 1968 estimates its relationships using data that has been transformed to remove the effects of autocorrelation in the original series. The argument here is not that models should always be calculated from first differences: in fact, one can think of situations in which taking first differences may itself introduce spurious correlations into the data. The argument simply is that in many of the equations of the aggregative models considered here the high \bar{R}^2 may well be due to factors like trend components in the series. The difficulties outlined in the text are not confined to time series models. Just as trends often result in spurious correlations in time series models so may factors like scale mislead us in models based on cross section data, liberty is not the only thing that demands eternal vigilance.

The second illustration of how one may obtain a spuriously high \bar{R}^2 relates not to a failure to transform data when it should have been transformed but to transforming it in such a way as to generate once again the kind of situation that Yule warned us against:^{1/} a situation in which one has "conjunct series with conjunct differences." This illustration is provided by the output-investment functions used in ECAFE, 1968 and UNCTAD, 1968.^{2/} Fourteen of the eighteen models of UNCTAD, 1968 and seven of the eight models of ECAFE, 1968 use the following output function:

$$Y_t = f \left[\sum_{i=0}^{t-1} I_i \right]$$

where Y is GDP at factor cost and I is gross fixed investment. The function yields extremely impressive fits: the \bar{R}^2 being almost invariably higher than 0.95. The two series Y and I have pronounced trend components in almost all cases; by cumulating I we are generating a new series which not only has high serial auto-correlation but the first differences of which are also auto-correlated. The nature of the effect on the coefficients of determination is perhaps best demonstrated by a simple example. Consider three series: Y_t , X_t and Z_t where Z_t is $\sum_{i=0}^{t-1} X_i$. Y_t is growing at 10% and X_t fluctuates around a mean of about 11.5.

1/ G. Udny Yule, op. cit.

2/ The same function is used in Kanta Marwah, op. cit. The function is used for 14 of the 18 countries for which UNCTAD, 1968 contains models. For Argentina and India imports of capital and intermediate goods have been included in addition to cumulative investment. Among the remaining four countries the function was fitted for Kenya, Uganda and Tanzania also but the regressions revealed a negative correlation between investment and GDP for Kenya and no significant correlation for Uganda and Tanzania and so the authors assumed a capital-output ratio of 3:1 for Kenya and 2.5:1 for the latter two (UNCTAD, 1968, pp. 124-125, 205 and 249). For Thailand there is no separate function, the relevant magnitudes being derived from an identity.

Y_t	X_t	$Z_t = \sum_{i=0}^{t-1} X_i$
1.10000	12	0
1.21000	11	12
1.33100	10	23
1.46410	14	33
1.61051	9	47
1.77156	10	56
1.94871	12	66
2.14358	11	78
2.35794	10	89
2.59374	12	99
2.85311	11	111
3.13842	13	122
3.45226	9	135
3.79749	14	144
4.17724	8	158

In fitting Y_t to X_t we get:

$$Y_t = 3.00572 - 0.06106 X_t$$

(0.407)

R^2	F-ratio
0.0125	0.165

The figure in parentheses is the t-ratio. But when we cumulate X_t and fit Y_t to Z_t as in the UNCTAD, 1968 and ECAFE, 1968 output functions we get:

$$Y_t = 0.81828 + 0.01933Z_t$$

(20.740)

.970	430.172
------	---------

The very impressive coefficient of determination and the significant t-ratio are due solely to the fact that we have cumulated X_t and do not in any way imply a reliable, stable or significant relationship between X_t and Y_t . The reader can indeed experiment with any set of positive random numbers and obtain similar results.

The final illustration of how \bar{R}^2 may be misleading involves regressing one ratio on another. We know that gross fixed investment in a sector $(\text{INVEST})_i$ consists of net investment $(\text{INVNET})_i$ and depreciation $(\text{INVDEP})_i$:

$$(\text{INVEST})_i = (\text{INVNET})_i + (\text{INVDEP})_i$$

Assuming that a given proportion (d) of the capital stock is replaced every year and that the capital stock bears a fixed relation to the value added in a sector

$$(\text{INVDEP})_i = d(\text{KSTOCK}) = d \cdot k(\text{GDPFAC})_i = b(\text{GDPFAC})_i$$

Therefore

$$(\text{INVEST})_i = (\text{INVNET})_i + b(\text{GDPFAC})_i$$

Dividing both sides by $\Delta(\text{GDPFAC})_i$ we get

$$\frac{(\text{INVEST})_i}{\Delta(\text{GDPFAC})_i} = \frac{(\text{INVNET})_i}{\Delta(\text{GDPFAC})_i} + \frac{b(\text{GDPFAC})_i}{\Delta(\text{GDPFAC})_i} \quad (1)$$

$$k_{gi} = k_{ni} + b \frac{1}{r_i}$$

where k_{gi} is the gross capital-output ratio in sector (i), k_{ni} the net capital-output ratio and r_i the sectoral rate of growth. On fitting equation (1) to the data we obtain the relevant gross-capital output ratios^{1/} Given the sectoral growth rates (r_i) we obtain numerical values of the capital-output ratios that can be used for determining, say, the amounts of investment required to assure a certain specified growth of value added in the sectors. For the sample period we may estimate the (r_i) by fitting the trend equation $\sqrt{Y_t} = Y_0(1+r)^t$ or $\log Y_t = \log Y_0 + t \log(1+r)$ to the sample

^{1/} The preceding formulation of arriving at capital-output ratios is taken from A.M. Strout, "Savings, Imports and Capital Productivity in Developing Countries," First World Congress of the Econometric Society, Rome, September, 1965. A somewhat similar reasoning underlies the investment functions in J. Vanek, Estimating Foreign Resource Needs for Economic Development, McGraw Hill, New York, 1967. Strout uses the function for cross section data. Vanek's estimates of coefficients are not derived from regression techniques.

period data. The sectoral growth rates and the gross capital-output ratios estimated from the Kenyan data for 1960-1968 along with the coefficients of determination are given below.

(1) Sector	(2) Rates of Growth in the Sample Period	(3) R^2 for Col. 2	(4) Capital- Output Ratios for the Sample Period	(5) \bar{R}^2 for Equations Estimating Capital- Output Ratios
Agriculture, Forestry and Fishing	5.39	0.801	1.58	0.979
Mining and Quarrying	1.84	0.000	5.20	0.624
Manufacturing and Repairing	9.27	0.966	2.87	0.987
Building and Construction	3.80	0.134	3.63	0.961
Electricity and Water	8.79	0.966	4.22	0.986
Transport, Storage and Communications	8.84	0.981	4.85	0.751
Other Services	6.86	0.969	2.61	0.970
TOTAL: When Estimated Directly	<u>6.56</u>	<u>0.964</u>	<u>2.69</u>	<u>0.999</u>

The coefficients of determination for the trend equations and for the capital-output ratios are satisfactory for all the sectors other than mining and quarrying. In addition, the capital-output ratios do not seem unreasonable. Would we be justified in presuming that the satisfactory R^2 and the reasonable values of the ratios indicate that we now have a satisfactory basis for forecasting sectoral gross investment?

Notice that the denominator, $(\Delta G D P F A C)_{it}$, is the same on both sides of equation (1). In fact, it turns out that the variance of $(\Delta G D P F A C)_{it}$ is

much greater than the variance of either $(\text{INVEST})_{it}$ or $(\text{GDPFAC})_{it}$. So that the reason we get the very impressive \bar{R}^2 for the capital-output ratios is largely that the variance of $(\Delta\text{GDPFAC})_{it}$ is just explaining itself. The literature contains formal tests relevant to this situation^{1/} but they are seldom employed in practice. Whether one is going to regard the very high values of \bar{R}^2 as "spurious" or not depends to a large extent on one's predilections; for philosophically the situation is an ambiguous one.^{2/}

Perhaps the following practical consideration will be more persuasive than lengthy arguments about whether or not the correlations are properly regarded as spurious. The sectoral capital-output ratios reported above were used along with actual year to year changes in sectoral (GDPFAC) to estimate fixed investment in the sample period. The following χ^2 result from a comparison of forecasted and actual investment in 1961-1968.

^{1/} For formal tests relevant to this situation cf.: E. Kuh and J. Meyer, "Correlation and Regression Estimates when the Data are Ratios," Econometrica, Vol. 23, 1955, pp. 400-416; and A. Madansky, "Spurious Correlation Due to Deflating Variables," Econometrica, Vol. 32, 1964, pp. 652-655.

^{2/} "The question of spurious correlation quite obviously does not arise when the hypothesis to be tested has initially been formulated in terms of ratios, for instance in problems involving relative prices. Similarly, when a series such as money value of output is divided by a price index to obtain a 'constant dollar' estimate of output, no question of spurious correlation need arise. Thus, spurious correlation can only exist when a hypothesis pertains to undeflated variables and the data have been divided through by another series for reasons extraneous to but not in conflict with the hypothesis framed as an exact, i.e., non-stochastic, relation," E. Kuh and J. Meyer, op. cit., pp. 401-402.

Comparisons of Actual and Forecasted Gross Fixed
Investment in Kenya, 1961-1968

Sector (1)	χ^2 : Total 1961-68 (2)	χ^2 : Years 1961-68 taken individually (3)
Agriculture, Livestock, Forestry and Fishing	0.002	347.635
Mining and Quarrying	0.0368	141.052
Manufacturing and Repairing	0.255	18.071
Building and Construction	0.098	268.240
Electricity and Water	0.026	8.395
Transport, Storage and Communications	0.505	14.011
Other Sectors	0.186	40.502
<u>All Sectors</u>	<u>1.109</u>	<u>837.906</u>

Column 2 indicates that the forecasts of cumulated investment for the period as a whole are indeed satisfactory; the χ^2 for all sectors in column 2 is less than the permissible limit with 6 degrees of freedom. But the cumulated totals are made up of very large over- and under-estimation of investment in individual years; thus almost all the individual χ^2 values in column 3 as well as the χ^2 for All Sectors are well beyond their permissible values for 7 and 55 degrees of freedom respectively. May one conclude that while the capital-output ratios are not reliable for forecasting investment in an individual year such as 1970 or 1971 they are likely to give satisfactory estimates for investments in the 1969-78 period as a whole? Unfortunately, even this much is by no means certain: cumulative investments are close to actual investments only over the sample period taken as a whole; when, for instance, forecasted investments are totaled up for a part of the sample period (say, 1965-68) they are once again vastly

different from the totals of actual investments. Thus we cannot be certain that even the cumulated values of investment will be reasonable estimates of the true values for any period other than the sample period taken as a whole. Reliance on the very high \bar{R}^2 of equations estimating capital-output ratios would, therefore, have been completely unwarranted. A part of the reason, of course, is that the \bar{R}^2 of equation (1) above relate to the magnitude $[(\text{INVEST})_i / (\Delta \text{GDPFAC})_t]$ when, in fact, we are interested in - and are implicitly using them to tell us something about - the variable $(\text{INVEST})_i$.

Besides the coefficient of determination the two tests most frequently used - or at least displayed - are the t-ratio and the Durbin-Watson statistic. Some comments on the t-ratio were included in the discussion of collinear regressors. The Durbin-Watson statistic does not require extended comment.

The importance of the error term is greatly emphasized in the theory of econometrics, since perhaps the most critical assumption of regression and econometric models is that the error-term is random. Auto-correlation in the residuals may indicate, for instance, that a systematic element has been left out in the explanation; in addition, least squares estimates are no longer efficient (though they are still unbiased and consistent) and the usual significance tests are no longer directly applicable. The question becomes particularly important in autoregressive schemes that utilize lagged variables for in such cases the estimates of 'b' are neither unbiased nor consistent. For all these reasons econometric theory places considerable emphasis on a careful examination of the error term. In this spirit the Durbin-Watson statistic - one of the tests available for examining residuals - is faithfully listed by most models. But in many cases no further attention is paid

to it. Thus, in a particular case, it is stated that "the (Durbin-Watson) statistic shows that the residuals of the German functions are positively auto-correlated at the five percent level. . . the subsequent discussion, however, will disregard all complications arising from auto-correlated errors."^{1/} Nor is this unrepresentative:^{2/} for many of the equations contained in econometric models of developing countries the Durbin-Watson statistic shows significant auto-correlation in the residuals, yet in no case is any attempt made to transform the data or the specification of the equation in any suitable way to avoid the auto-correlation. Of course, it may be that these models do not pay much attention to the statistic because at least 13-15 observations are needed to be able to use its tabulated values and more often than not the models are based on a smaller number of observations. One suspects, however, that the real reason is that having listed the statistic no one bothers about it simply because no one else bothers about it either: everyone is aware of sinning in good company.

Before concluding this review of the relevance of standard statistical tests, it will perhaps be worthwhile to stress one general point. The theory of the sampling distribution of \bar{R}^2 and other indices like the t-ratios assumes that the hypothesis to be tested has been formulated independently of the data that is used to test its validity. When, as in practice, the hypothesis is framed after a scrupulous and in some cases systematic examination of the data (with, for instance, the aid of step-wise regression procedures) the indices lose a good part of their discriminatory power. To validate the hypothesis or the equation that embodies it one needs data from

^{1/} G. Menghes and H. Diehl, "Time Stability of Structural Parameters," Econometric Analysis for National Economic Planning, ed. by P.E. Hart, G. Mills and J.K. Whitaker, University of Bristol, 1964, pp. 299-317; the sentence is from p. 307.

^{2/} For similar remarks see UNCTAD, 1968, pp. 108, 126, 191, 234, et. al.

a different sample - for instance, from a time span different from the sample period used in estimating the equation. As the developing countries seldom have consistent series of more than 5 to 15 years one seldom finds a model-builder verifying his equations or hypotheses by reference to data other than the data used in their estimation and formulation.

VI. Point Forecasts and Policy Prescriptions for Years Far into the Future

Even if it were the case that the data from which the equations were estimated was reliable and ample; that the problems arising from collinearities had been successfully taken care of; that one could be sure that the equation had been correctly specified; and if care had been taken to avoid the kinds of traps sketched in Section V - even then the question as to whether the coefficients and equations could be regarded as reliable guides to the future is a separate one, which has to be answered with reference to whatever information one can muster about that future.

Models currently in use make point forecasts for years well beyond the end of their sample of observations. One can be almost certain that the precise value for Y we predict from an assumed relationship, say, $(Y=a+bX)$, will never be realized. The individual value of Y is in a sense only an average of likely outcomes. If the forecasts from an equation are to be viewed in the context of some significance or probability levels they must be in terms of a Y-distribution and not just a single value of Y.

Now, the interval estimate of Y at $(1 - \alpha)$ level of confidence that one must work with is given by

$$\hat{Y} \pm t_{(n-p-1), (1-\alpha)} s \sqrt{X_0' (X' X)^{-1} X_0}$$

where \hat{Y} is the point forecast of Y , n the number of observations in the sample period, p the number of explanatory variables and X_0 that point in X -space for which the forecasts are being made. The upper and lower boundaries of these interval forecasts curl away from the line of point forecasts the farther away X_0 is from the means of X in the sample period. Typically, individual equations are being used to make forecasts for periods when the independent variables are two to five and sometimes an even larger number of times their mean values in the sample period and $1\frac{1}{2}$ to three and sometimes an even larger number of times their values at the end of the sample period.^{1/} It is not surprising then that when we are conscientious and do make interval projections for the future - even with just one functional form - the upper and lower bounds turn out to be a long way apart from each other.

The point is not merely that one now has a band instead of a single point forecast from an individual (structural or reduced form) equation; these bands have to be combined in some way to yield an interval forecast of, say, the import-export gap to which one may attach a precise probability significance. Unless one has estimated the reduced form equation for the gap directly - in which case there is no need to combine any equations - the attempt to combine conditional forecasts from individual

^{1/} To cite but one instance, in Jere Behrman and L.R. Klein, Econometric Growth Models for the Developing Economy, Wharton School, University of Pennsylvania, (mimeographed), 1968, by the end of the projection period some of the independent variables increase to 38 times their values at the end of the sample period.

equations soon leads one into problems. For in combining the confidence bands of one structural equation with confidence bands of another one soon loses control of probability levels.

We know that if $R_1 \dots R_n$ are random variables with finite variances $\sigma^2 + \dots + \sigma_n^2$ and $S_n = R_1 + \dots + R_n$ then

$$\text{Var}(S_n) = \sum_{k=1}^n \sigma_k^2 + 2 \sum_{j,k} \text{Cov}(R_j, R_k)$$

where the last term contains the sum of each of the $\binom{n}{2}$ pairs R_j, R_k with $j < k$ once and only once. In the models considered here the authors - to avoid problems of simultaneity - assume that the residuals are independent of one another. Therefore, the variance of residuals that one needs in order to obtain a band forecast for, say, the import-export gap, is:

$$\text{Var}(S_n) = \sum_{k=1}^n \sigma_k^2$$

where k refers to the residuals of each of the equations that goes into the reduced form of the gap. For the author's Kenya model this pooled $\text{Var}(S_n)$ comes to £756.9 million - the standard deviation, that is, is around £27.5 million.^{1/} As the model projects the gross external gap to be around £20 million to £25 million a year throughout the entire projection period, an interval forecast based on this wide band is as likely to be correct as it is unhelpful.

In the specific context of the conventional two-gap models that are so widely used now the moral of the foregoing discussion might seem to be that if one is primarily interested in forecasting magnitudes like the

^{1/} This estimate makes no provision for the many variables in the balance of payments and the debt routine which are exogenously specified in the model.

import-export and the investment-savings gaps one should directly estimate the reduced forms for the two gaps from the sample period data. If one attempts to do so then, clearly, one would have to be even less ambitious while choosing the explanatory variables than the simple models being proposed at present. The number of variables that enter the reduced form for any one of the gaps in these models is often quite large and the number of observations is often so small that one would have few degrees of freedom left to estimate the coefficients. Moreover, in attempting to estimate the reduced forms directly one is immediately faced with the difficulty that in the (ex post) sample period data the savings-investment and the import-export gaps will be necessarily identical. Therefore, the only way of obtaining estimates of equations that would yield independent (i.e., not identical) projections of the two gaps for the future will be to specify the model in a way that is very different from the manner in which two-gap models are being specified at present. One may, for instance, fit the reduced forms for the two gaps using different sets of explanatory variables. Alternatively, instead of having different lists of explanatory variables for the two gaps, one could specify the individual equations in the same way as they are specified at present but in addition specify some adjustment procedures that would furnish independent estimates of the two gaps. The difficulty here would be, as was noted towards the end of Section III, that there are many ways of specifying the equations and reconciling the gaps and one's forecasts may well turn out to be quite sensitive to the particular specification or reconciliation procedure one adopts.

It has been stressed above that these simple models are being used for making projections for periods far beyond their sample periods. Many

statisticians who have worked with regressions and seen their forecasts go astray will express serious doubts as to whether regression equations can be used at all for values of variables that are two to five times (to say nothing of variables that are 38 times) beyond the range of values covered by the sample observations. Using these equations implies the heroic assumption that the inter-relationships of the system in the future will be as they have been in the past or that structural change in the future will follow the pattern of whatever change occurred in the sample period when, in fact, the whole effort of the governments of these countries and others connected with their development is to bring about a structural transformation in the economies. Nor is it the case that when information about likely structural changes becomes available it is customarily incorporated in the model. The fact is that in most of the models equations are being used in a fairly mechanical manner for projections 5 or 10 years into the future. But even if the analyst using a very aggregative econometric model sought out and obtained specific information about a likely structural change in, say, the manufacturing sector, how would he incorporate it into his aggregated model? The models in use are at such a level of aggregation that if one were told, for instance, that in Kenya canning of fruit and vegetables and leather and fur processing are going to grow at twice, sugar, spinning and weaving and pulp and paper at about two and a half times and rubber at four times the rate of growth of the manufacturing sector as a whole; or that Kenya is going in for a large fertilizer plant using by-products of its refinery as feedstock, if one were given this kind of specific information one would not be able to make full use of it. All one could do would be to lump it all together to see what effect it has on the overall growth rate of the manufacturing sector or of the economy and work

with that. If by contrast the model had concentrated on incorporating a greater amount of production detail, it would have been able to effectively utilize this kind of specific information. And one need hardly stress that information about the future when it becomes available almost invariably comes in this specific form; seldom does one get a revelation about the overall marginal elasticity of consumer imports or the overall capital-output ratio in a sector.

The position thus is as follows: equations are being used to make forecasts for periods well beyond the range of observation points covered in the sample period; the only information one can usually get about those distant periods relates to specific and particular developments; and the models and equations are at such a level of aggregation that they cannot use this information. Strictly speaking, interval forecasts, though far preferable to point forecasts, do not help partly because of the practical difficulties outlined above and partly because the structural change that makes point forecasts undependable also makes the information on which interval estimates are based obsolete. For both sets of forecasts are based on the same information.

VII. The Moral of the Story

The argument of this paper has not been against quantification or sophistication in economic analysis nor against economic models in general. The paper has focused on a particular type of model that is being used for analysis in developing countries: the highly aggregated macro-economic regression model. Over the last few years models of this kind have become increasingly common and are being used for medium- and long-term forecasts and for policy prescription. This apparent faith in these models is unwarranted. The reasons which have been sketched in this paper can be summarized briefly as follows.

As the sorts of models we have been considering are being constructed for one country at a time they have necessarily to rely on time series data. This data is generally available for only a few years, and when available is of uncertain or very low reliability. As the time series are very short and as almost all variables seem to grow or decay over the short periods for which data about them can be assembled, every attempt to include in the equations all the independent variables that appear to be relevant on a priori grounds soon exhausts the degrees of freedom and is itself frustrated by the prevalence of collinearities. For this and other reasons there is every likelihood that the models and their equations are grossly underspecified or mis-specified. Moreover, in a large number of cases one can expect little help from economic theory, from mechanically applied statistical tests or from appeals to one's intuition.

Any one of these objections taken individually is sufficient to dissuade one from taking these macro-models at all seriously. In combination their effect can fairly be described as devastating.

One is often told that one is safe in using these grossly underspecified equations and models provided one is aware of their limitations. ^{1/} This is no more than a ritualistic incantation. For in practice one either uses a model, an equation or a coefficient or one does not. How does one use one's awareness of their limitations in a concrete way when a model

^{1/} "The drawbacks to the use of capital-output ratios are well known. . . . Despite these drawbacks, the capital-output ratio is often heavily relied on as a tool of analysis. If enough care is exercised in its application, and particularly if separate estimates can be made for individual sectors, it should be possible to indicate the broad order of magnitude of investment requirements for particular rates of growth," UNCTAD, 1968, p. 13.

often involves a number of equations, when little is known about the errors in the data and when little has been done to experiment with alternative specifications?

Before constructing an econometric model and perhaps again after having constructed one, one should regain one's perspective by re-reading some of the iconoclastic literature on regression and econometric analysis (for example, writings of Yule^{1/} Box,^{2/} Geary,^{3/} and others) and on problems presented by the inaccuracy of basic data. National accounts and the manner in which they are put together in a country deserve more attention than they usually receive at present. A plea should also be made for paying greater attention to textbooks that so often warn us against doing so many of the things we do. A classroom is not their only proper place.

^{1/} Op. cit.

^{2/} G.E.P. Box, "Use and Abuse of Regression," Technometrics, Vol. 8, No. 4, November, 1966, pp. 625-629.

^{3/} R.C. Geary, "Some Remarks About Relations Between Stochastic Variables: A Discussion Document," Review of the International Statistical Institute, Vol. 31, No. 2, 1963, pp. 163-181.

WORLD BANK REPRINT SERIES: Number Three

Arun Shourie

**The Use of Macro-economic
Regression Models of
Developing Countries for
Forecasts and
Policy Prescription**

Reprinted from *Oxford Economic Papers*
Volume 24, No. 1, March 1972

THE USE OF MACRO-ECONOMIC REGRESSION MODELS OF DEVELOPING COUNTRIES FOR FORECASTS AND POLICY PRESCRIPTION: SOME REFLECTIONS ON CURRENT PRACTICE¹

By ARUN SHOURIE

I. Introduction

THE need for long-term forecasting of various magnitudes in under-developed economies is increasingly recognized. Lending agencies that make loans with maturities ranging from ten to fifty years naturally want to know all they can about likely developments in these economies for as many years into the future as possible. An international agency like UNCTAD which has learned that it takes a good many years to wrest action from a hundred-odd countries must plan now for steps it would want taken in the mid or late 1970s, and to do so it needs projections relating to those distant periods. National planning agencies often want to examine the problems of their economies in a long-term context and, just as often, they must put together plan documents that give a long-term perspective to current developments, their programmes, and their requests for external assistance.

It is a common practice in making these forecasts, whether medium- or long-term, to use econometric models that were originally developed for short-term forecasting of relationships in developed economies. Thus UNCTAD has a large project under way for constructing some thirty models of various economies; ECAFE recently published econometric models for ten countries; other international organizations, such as the IMF and the World Bank, and national donor agencies, such as the U.S. AID, are making increasing use of them; and consultants and research organizations working on problems of the third world often base their projections and advice on econometric models.

This paper is based on a review of about forty-five such models that have been proposed for various developing countries. All of these models use statistically estimated parameters to provide a quantitative description of certain macro-relationships in developing economies. They might be described as aggregative macro-economic regression models.

Sections II to VI review the models from five points of view: the reliability

¹ I am grateful to P. D. Henderson, B. S. Minhas, E. P. Holland, N. G. Carter, and R. Cheetham for many useful comments and to S. Malik and E. Ramiscal for their help in the computations. None of them is responsible for any errors that remain. The author is a member of the Economics Department of the World Bank. The views expressed in this paper are his personal views and not those of the World Bank group.

of the data from which they are estimated; the significance of the coefficients of their equations; the functional form of their equations; their propensity to use certain statistical tests in a mechanical way; and their use for making forecasts and prescribing policies for years far into the future. Section VII summarizes the principal conclusions and contains some general remarks about the manner in which econometric models are being used at present.

The discussion in Sections II to VI is illustrated by cases from various models that have been proposed in the last five years and from a conventional two-gap model of the Kenya economy put together by the author. Many of the illustrations are taken from a recent UNCTAD study that contains models for eighteen countries.¹ Raul Prebisch in his preface states that 'the study provides a solid basis for the elaboration of new national and international targets for the United Nations Development Decade of the 1970's'.²

II. The data base

It may appear trite or even superfluous to begin yet another article on econometric models by drawing attention to the data from which the models are estimated. But the fact is that reliable and ample data is critical to an econometric model: if the underlying data is seriously deficient then our care in adopting optimal methods of estimation is of little avail. Every economist constructing a model is conscious of this fact and acknowledges the problems connected with the basic data by some ritualistic disclaimers: the data are of uncertain quality but this is the best that is available; the alternative to using it is to abandon all hopes of quantifying policy recommendations; in any case the exercise is worth while for it highlights the particular areas in which the data-collection efforts need to be strengthened.³ Having protected his flanks with these disarming

¹ UNCTAD, *Trade Prospects and Capital Needs of Developing Countries*, United Nations, New York, 1968 (henceforth, *UNCTAD, 1968*). A similar volume has been published by the Economic Commission for Asia and the Far East: ECAFE, *Feasible Growth and Trade Gap Projections in the ECAFE Region*, Bangkok, 1968 (henceforth, *ECAFE, 1968*). This volume contains models for nine countries. The model for Thailand, however, is the same as the one contained in *UNCTAD, 1968*. All subsequent references to *ECAFE, 1968* relate to the eight models other than the model for Thailand.

² *UNCTAD, 1968*, op. cit., p (iii).

³ J. T. Scott and Earl O. Heady, 'Econometricians and the data gap; reply', *American Journal of Agricultural Economics*, vol. 51, No. 1, Feb. 1969, p. 188. *UNCTAD, 1968* notes 'the paucity, inadequacy and unreliability of the relevant statistical data (for, in this case, Nigeria). . . Suffice it to say that in many cases the estimates are rather crude and impressionistic with very little firm basis . . .', p. 150. It goes on to note that the national accounts in Nigeria are intimately tied up with population estimates as for many items the aggregate magnitude is obtained by multiplying an estimated *per capita* figure by the population estimate and that the 1963 census has shown population to be much higher than the figures assumed in the national accounts. 'Nevertheless, use of available data is probably better than not using any data at all and provides a useful check on more qualitative approaches', *ibid.*

assertions the model-builder often proceeds to act in precisely the way he would have acted had the data been abundant and reliable.

The data of most developing countries present three problems to the econometrician.

First, the series for most macro-variables cover only a few years. Consequently, from an econometric point of view the models operate on the subsistence margin in that their relationships are estimated from no more than five to sixteen observations.¹

Second, the figures are of varying and often very low reliability. Estimates of many macro-variables are based on guesses or on rules of thumb derived from scanty empirical evidence. Investigators who have had the patience to examine the data carefully have often concluded that 'we may have to admit that up to now we really are not in a position to tell how the main economic variables are changing in large areas of the third world'.² For these countries trade statistics, in fact, constitute the only data in which most economists are willing to place a good deal of confidence. But even this information when it is examined systematically—for example, when it is compared with the returns of the country's trading partners—reveals very substantial discrepancies.³

The underlying data are such that many refinements in estimation procedures, such as multiple stage least squares, seem to be somewhat pointless. In fact, one does not quite know how to assess even the more basic refinements. Consider the fact that most models in use base their estimates on constant price data.⁴ Relationships are estimated at constant prices so as to separate the income and price effects. More often than not the model builder adopts the national accounts data for a country without examining the deflators used to arrive at the constant price series. To cite one instance: *UNCTAD, 1968* gives projections for Ceylon at constant 1960 prices.⁵ Production in the agricultural sector consists of a few

¹ The frequency distribution of the models in *UNCTAD, 1968* and *ECAFE, 1968* by the number of observations on which they are based is as follows:

Number of observations:	5-7	8-10	11-13	14-16
Number of countries: (i) <i>UNCTAD, 1968</i>	2	7	3	2
(ii) <i>ECAFE, 1968</i>	4	1	1	2

Equations in the four models of *UNCTAD, 1968* that have not been included above are based on varying numbers of observations: Argentina (12 to 16), Ghana (10 to 16), Nigeria (10 to 13), and Philippines (9 to 14).

² L. Whitehead, 'Basic data in poor countries: the Bolivian case', *Oxford University Institute of Economics and Statistics, Bulletin*, vol. 31, No. 3, Aug. 1969, p. 226.

³ As a recent and relevant illustration, cf. S. Naya and T. Morgan, *The Accuracy of International Trade Data: The Case of Southeast Asian Countries*, University of Wisconsin, mimeographed, Aug. 1967.

⁴ Fourteen of the eighteen models in *UNCTAD, 1968* use constant price series.

⁵ *UNCTAD, 1968*, p. 75.

identifiable crops: as the product is homogeneous over space and time output indices are used and we do, in fact, obtain a constant price series that is meaningful. For the other sectors value added at current prices is adjusted by the following odd assortment:

Sector	Deflator
Mining and quarrying	Wage index for workers in industry and commerce.
Manufacturing	
(a) Tea, rubber, coconut processing	Indices of physical output.
(b) Factory and cottage industries excluding (a)	Domestic group of the Colombo cost of living index.
Construction	Domestic group of the Colombo cost of living index.
Electricity, water, and sanitary services	Wage index of Government technical and clerical employees.
Transport, storage, and communication	A mixture of indices of wages of Central Government employees, workers in industry and commerce, and the Colombo cost of living index.
Wholesale and retail trade	Domestic group of the Colombo cost of living index.
Banking, insurance, and real estate	Wage index for employees in industry and commerce.
Ownership of dwellings	Domestic group of the Colombo cost of living index.
Services	Colombo cost of living index.

Weights for the more widely used indices are based on surveys conducted many years ago when the structure of the economy was very different from what it is now. The most widely used deflator—the Colombo cost of living index—for instance, is based on a survey conducted in 1952.¹ The problems are compounded for economies that in addition to having deflators of somewhat limited reliability have had multiple or oft-changing exchange rates.²

For a similar reason one must be wary of reading too much into high coefficients of determination and *t*-ratios before examining the manner in which the relevant data were estimated in the first place. If, for instance, we fit an equation relating imports of construction materials to value added

¹ Some studies explicitly acknowledge that for their countries current price series are more reliable than available constant price series: cf. for Thailand, *UNCTAD, 1968*, pp. 388, 390; for Tanzania, *ibid.*, p. 190; for Kenya, *ibid.*, p. 107; for Uganda, *ibid.*, pp. 233-4.

² For an instructive illustration of the gyrations one must go through to take account of this problem see B. A. de Vries and J. C. Liu, 'An econometric analysis of inflation and growth in Brazil', Washington, D.C., 1969 (mimeographed), presented at the New York meetings of the Econometric Society, Dec. 1969.

by construction in Ceylon we obtain a very satisfactory fit and a highly significant *t*-ratio. The reason for this is simply that value added by the bulk of the construction sector is itself estimated by multiplying the imports of construction materials by (4.25)—after allowing for a three-month lag. The \bar{R}^2 and the *t*-ratios in this case would tell us no more than the national accounts themselves.¹

The third source of difficulties is that the national accounts of many of the countries for which econometric models are being constructed at present are in their infancy and are still being revised periodically. Therefore, each of the models is 'dated' in the sense of being tied to the national income series that were in use at the time when it was estimated. A rough check indicates that of the twenty-two countries for which *UNCTAD, 1968* provides detailed models the accounts for as many as eleven have been revised since 1965 (the last year for the sample period of *UNCTAD, 1968*) in ways that would materially affect the estimates of some important coefficients. The national accounts of Kenya, for instance, have been revised often and very considerably over the last few years. One has only to compare the data used by Paul Clark in his models for the East African economies,² that used by Faaland and Dahl in their model of Kenya,³ and the data in use now, to see that series like gross domestic product at factor cost and gross investment have been revised upward by 15 to 20 per cent.

Clearly, national accounts have to be revised as and when new data or information become available. However, these revisions necessarily reduce confidence in models estimated at an earlier stage of the country's experience of preparing national accounts. Consider the consequences of revising capital formation figures upwards by 15 or 20 per cent: the savings figures in Kenya, as in many other countries, are derived as a residual using the capital formation and balance of payments accounts. Now, the capital formation figures have been revised upwards by one-eighth to one-fifth but the balance of payments account has not undergone a revision of any comparable magnitude. Thus today one gets very different savings coefficients and a very different picture of the accelerations or decelerations in the country's savings efforts in a given period than one would get using

¹ These remarks are suggested by some equations in a recently published model of the Colombian economy: K. Marwah, 'An econometric model of Colombia: a prototype devaluation view', *Econometrica*, Apr. 1969, pp. 228-51. We are told that 'there seems to be a strong complementary relation between the imports of capital goods and investment. A marginal increment of one billion pesos worth of imports of capital goods seems to be accompanied by an addition of 1.49 billion pesos of total investment', *ibid.*, p. 236. The *t*-ratio is more than 16 and the R^2 is 0.981. In Colombia fixed investment in sectors other than construction is estimated by applying multipliers to imported capital goods assigned to the sectors. These multipliers have been changed only once in the last twelve years.

² Paul G. Clark, *Development Planning in East Africa*, East African Publishing House, 1965, especially pp. 34-5.

³ *UNCTAD, 1968*, pp. 104-48.

Similarly, a recently published model of the Colombian economy designed specifically to analyse the consequences of changes in exchange-rate policies treats each of the coefficients as economic propensities and elasticities—even when some important explanatory variables that are left out are highly collinear with the ones that have been included.¹

B. A. de Vries and J. C. Liu in their recent paper on Brazil² provide a very explicit illustration of economists attributing causal significance to each of their coefficients. At each step this instructive paper carefully stresses the number of significant explanatory variables that have not been included in each equation (many of these are probably collinear with the ones that have been included), the poor quality of data, the partial specification of the model itself, and many other difficulties; and yet it interprets the resulting coefficients as if they were the conventional elasticities and propensities to which the economist is accustomed. The paper proceeds to derive trade-offs between inflation and growth.

For a large number of reasons we are not justified in using regression equations and their coefficients for control purposes. We shall discuss two of these reasons. First, there is the fact that in a very large number of cases the explanatory variables that have been included in an equation are highly collinear and, second, the likelihood that the equations are misspecified and that some of the excluded variables may well be collinear with the variables that have been included in the equation.

(a) *Collinearity among regressors*³

We know that if the regressors are orthogonal, if the equation is specified correctly, and if the data is reliable, an unambiguous meaning can be attached to the estimated coefficients. At the other extreme, the coefficients cannot be estimated if the regressors are perfectly collinear. In practice the situation encountered is an intermediate one: over short periods of five to fifteen years a surprisingly large number of aggregative time series have strong trend components and are, therefore, highly though not perfectly collinear.⁴ The use of such variables as independent regressors in multiple regression equations raises difficult problems of interpretation.

¹ K. Marwah, op. cit., in particular, the discussion of the consumption equations, pp. 234-5.

² B. A. de Vries and J. C. Liu, op. cit.

³ The following is by no means a comprehensive treatment of the problems that arise from multicollinearity in regression analysis and their solutions. A brief account is given in J. Johnston, *Econometric Methods*, McGraw Hill, New York, 1963, pp. 201-7. For two recent statements see D. C. Farrar and R. R. Glauber, 'Multicollinearity in regression analysis: the problem revisited', *Review of Economics and Statistics*, vol. xlix, Feb. 1967, and Y. Haitovsky, 'Multicollinearity in regression analysis: comment', *ibid.*, li, Nov. 1969.

⁴ Of the 630-odd estimated equations that are listed in *UNCTAD, 1968* about 168 have more than one regressor. The data given in *UNCTAD, 1968* enabled us to calculate the

Individual coefficients have little significance if the regressors in a multiple regression equation are collinear. In particular if the regressors X_1 and X_2 are not causally independent it is misleading to assume that a coefficient, say b_2 , in the equation ($Y = b_1 X_1 + b_2 X_2 + u$) indicates the extent to which Y will change as a consequence of a unit change in X_2 'other things remaining the same'. To begin with, the high correlation between X_1 and X_2 over the sample period may imply that there is some organic dependence between the two regressors so that when X_2 is altered 'other things'—in this case, X_1 —may not 'remain the same'. The coefficient b_2 in the single equation model ($Y = b_1 X_1 + b_2 X_2 + u$) does not take account of these indirect effects of X_2 on Y via the changes it induces in X_1 . Even if there is reason to believe that there is no organic dependence between X_1 and X_2 and that they 'just happened to have been collinear in the sample period' there are various reasons to be wary of using the coefficients as if they were the partial derivatives of calculus.

First, if the correlations between regressors are indeed as high as (0.95) (as they most often are in the models under review—(0.9985) in Tanzanian equation cited below) the values obtained for the coefficients are very sensitive to sample coverage, data errors, specification of the equation, and even rounding-off errors. The argument can be stated compactly as follows. Let R be the matrix of coefficients of correlation (r_{ij}) between the regressor variables; let b^* denote the vector of standard partial regression coefficients so that the regression coefficient b_j is equal to ($b_j^* S_y / S_{x_j}$) and let g be the vector of coefficients of correlation between the dependent variable and the regressors. Then,

$$Rb^* = g,$$

therefore,

$$b^* = R^{-1}g.$$

Now, R^{-1} is $[(R_{ij})' / |R|]$ where $[R_{ij}]$ is the adjoint matrix of R consisting of cofactors and $|R|$ is the determinant obtained from R . The closer a regressor is to being the linear function of other regressors the closer the value of the determinant $|R|$ will be to zero. With a divisor close to zero, the values of elements in the inverse matrix R^{-1} (from which the vector b^* and ultimately b and the standard error of the individual coefficients are estimated) become extremely volatile: even small changes in the r_{ij} arising

intercorrelations among the regressors in 118 of these equations. A total of 168 intercorrelations among different regressors were involved. The distribution of these intercorrelations is as follows:

Number	Value of the correlation coefficient								
	0.0 to 0.6	> 0.6 ≤ 0.65	> 0.65 ≤ 0.7	> 0.7 ≤ 0.75	> 0.75 ≤ 0.8	> 0.8 ≤ 0.85	> 0.85 ≤ 0.9	> 0.9 ≤ 0.95	> 0.95 ≤ 1.0
	12	6	6	12	6	4	6	13	103

In other words, about 60 per cent of the intercorrelations among supposedly 'independent' variables are higher than 0.95 and about 70 per cent are higher than 0.90.

from sample coverage, or data errors or from the effects of rounding-off can alter the value of $|R|$ greatly. In order to gauge the extent of this volatility it is sufficient to examine the simple correlation between the regressors when the equation has only two regressors. But if there are more than two regressors the value of R can be close to zero even if none of the correlation coefficients is individually close to 1.0. The reason for this is that an individual regressor X_i though not very highly correlated with any one of the other regressors may yet be a linear function of all or some (i.e. more than one) of the other regressors. In this case an inspection of the simple correlation coefficients will not suffice; one must look at the value of the determinant $|R|$ itself.¹

Second, the standard error of the estimate of a coefficient increases as the degree of collinearity increases. When the correlations among regressors are as high as 0.9985 or 0.95 the standard errors are often so high as to make the point estimate of the coefficient virtually useless for policy purposes. In this context two conventions regarding standard errors and t -ratios deserve comment. First, the customary rule of thumb that is used in evaluating the t -ratios is to regard them as acceptable if they are greater than two. This rule is a carry-over from the days when the primary interest was in verifying if the b -coefficient was significantly different from zero; that is, whether the independent and dependent variables were related in a qualitative sense. But if, as in the models we are considering, the equations and their coefficients are being used to make forecasts and to prescribe policy then one would expect that the model builders would look for t -ratios

¹ The diagonal elements of the symmetric matrix (R) are 1.0; the off-diagonal elements can vary between -1.0 and 1.0. Thus the value of the determinant $|R|$ can vary between 1.0 and zero. In the two-variable case if the two regressors are completely independent the matrix (R) is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ so that the determinant is 1.0. If they are perfectly correlated, (R) is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and so the determinant is zero. As the value of the determinant is a function solely of the coefficient of correlation between the two regressors it is sufficient to examine the value of the coefficient. But when more than two regressors are involved the individual correlation coefficients are not always adequate guides to the value of the determinant. Consider the following array of four regressors:

	1	2	3	4	5	6	7	8	9
X_1	2.0	0.0	0.0	3.5	0.0	0.0	4.0	0.0	0.0
X_2	0.0	2.5	0.0	0.0	3.0	0.0	0.0	3.5	0.0
X_3	0.0	0.0	3.0	0.0	0.0	2.5	0.0	0.0	4.0
X_4	2.0	2.5	3.0	3.5	3.0	2.5	4.0	3.5	4.0

As X_4 is a linear combination (in this case just the sum) of the other three the value of the determinant to be used in the regression computations is zero. Yet the individual correlation coefficients are as follows: $r_{12} = 0.0$; $r_{13} = 0.0$; $r_{14} = 0.59531$; $r_{23} = 0.0$; $r_{24} = 0.54972$; and $r_{34} = 0.58600$. Compared with the correlation coefficients one encounters in the models under consideration none of these is alarmingly high. In fact, as there are only nine observations one could even maintain that none of the coefficients is significantly different from zero at, say, the 2.5 per cent significance level.

that are much higher than two. A t -ratio of two, after all, means that the standard error of the coefficient is one-half as large as the coefficient itself. Consequently, the forecasts and the policy conclusions one can draw from it will indeed cover a very wide band. Second, the fact that the standard errors increase with increasing collinearity has been taken to imply that one need not worry about multicollinearity as long as the standard errors do not increase so much as to render the coefficients insignificantly different from zero; as long, say, as the t -ratios are not below the magic figure of two. This is a dangerous deduction. The fact that the standard errors are not unacceptably high may just be a fortuitous result of the computation algorithm one has used or of some special inter-relationships among the variables.

Consider an equation with two regressors: $Y = b_1 X_1 + b_2 X_2 + u$. The variance for b_1 is $S_{b_1}^2 = \sum y^2(1 - R_{y,12}^2)/(n - k - 1) \sum x_1^2(1 - r_{12}^2)$, where $y^2 = (Y - \bar{Y})^2$, $x_i = (X_i - \bar{X}_i)$, $R_{y,12}^2$ is the coefficient of multiple determination and r_{12} is the coefficient of correlation between X_1 and X_2 .¹ The greater the linear dependence between X_1 and X_2 —i.e. the closer r_{12} is to 1.0—the closer the entire denominator is to zero and, hence, given other things, higher is the standard error. But in a very large proportion of multiple regression equations reported, for instance, in *UNCTAD, 1968* while the intercorrelation among regressors is 0.95 or 0.9985 the R^2 term in the numerator term is also close to 1.0. Thus the term $(1 - R_{y,12}^2)/(1 - r_{12}^2)$ may not be very large, the effects of extremely high collinearity being 'absorbed', 'submerged', or 'neutralized' by the equally high $R_{y,12}^2$.²

Technically the estimated coefficients are unbiased and among the class of linear estimators they are the best. Moreover, as the relatively low standard errors testify, they are also relatively 'precise'. A coefficient may

¹ The expression used in the text can be derived as follows in the two variable case. We know that

$$\begin{bmatrix} \sum x_1^2 & \sum x_1 x_2 \\ \sum x_1 x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum x_1 y \\ \sum x_2 y \end{bmatrix},$$

where $y = (Y_i - \bar{Y})$, $x_1 = (X_1 - \bar{X}_1)$ and $x_2 = (X_2 - \bar{X}_2)$.

Thus

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum x_1^2 & \sum x_1 x_2 \\ \sum x_1 x_2 & \sum x_2^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum x_1 y \\ \sum x_2 y \end{bmatrix}.$$

The inverse matrix is:

$$1/(\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2) \begin{bmatrix} \sum x_2^2 & -\sum x_1 x_2 \\ -\sum x_1 x_2 & \sum x_1^2 \end{bmatrix}.$$

The variance for b_1 is

$$S_{b_1}^2 = \text{MSE}[\sum x_2^2 / (\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2)],$$

where MSE is the Mean Square Error.

Thus

$$\begin{aligned} S_{b_1}^2 &= \text{MSE}[1/\sum x_1^2(1 - (\sum x_1 x_2)^2/(\sum x_1^2 \sum x_2^2))] \\ &= \text{MSE}[1/\sum x_1^2(1 - r_{12}^2)] \\ &= \sum y^2(1 - R_{y,12}^2)/[(n - k - 1) \sum x_1^2(1 - r_{12}^2)]. \end{aligned}$$

² Compare Johnston, op. cit., pp. 204-6.

have all these properties and yet may be unreliable for policy purposes. A sample of observations, even when it yields the best linear unbiased estimates of coefficients that have high *t*-ratios, may yet not be rich enough to be able to disentangle the separate effects of the regressors in an economically meaningful sense.

A set of equations reported in *UNCTAD, 1968* provide an instructive illustration. Faaland and Dahl fit the following regression to account for the value of agricultural output in Tanzania:¹

$$P_A = 34.322 + 0.402 (\text{GDP}) \quad \bar{R}^2 = 0.94, \quad DW = 1.59, \\ (0.036)$$

where P_A is the value of agriculture and the expression in parentheses is the standard error of the coefficient. To take account of the fact that demand patterns may be changing because of the growth of *per capita* income and the diminishing importance of the subsistence sector they fit another regression:

$$P_A = -8.001 + 5.698 (\text{GDP}/N) \quad \bar{R}^2 = 0.96, \quad DW = 1.57. \\ (0.397)$$

Then, to see how the relationship between agricultural output and GDP is being modified by changing income levels, they fit an equation containing both GDP and GDP per head as explanatory variables:

$$P_A = -131.716 - 1.226 (\text{GDP}) + 22.816 (\text{GDP}/N) \\ (0.239) \quad (3.345) \quad \bar{R}^2 = 0.99, \quad DW = 2.13.$$

The coefficient of correlation between (GDP) and (GDP/*N*) is 0.9985. The authors point out that the 'negative partial correlation (of P_A) with GDP . . . on *a priori* grounds appears improbable'. Yet the *t*-ratios are high and, if we assume that the third equation is correctly specified, the estimated coefficients are the best linear unbiased estimates.

When collinearities among regressors and the numerical algorithms used for computation lead to patently implausible results one is at once made aware of their distorting influence. But this influence may be present even when the resulting coefficients do not appear to be obviously or dramatically unreasonable. After all the negative and significant coefficient for (GDP) in the third equation cited above is only an extreme illustration. A tightly reasoned and precisely formulated hypothesis can be a great help in such situations. In practice, however, too little attention is paid to formulating precise hypotheses *before* sending the data down to the computers. As the initial hypotheses are not precisely formulated often

¹ *UNCTAD, 1968*, p. 204.

there is little reason to scrutinize or doubt the results obtained. Consequently, the text accompanying the equations just sets out either to rationalize the values that emerge from the calculations or to reject the results in a perfunctory manner.

From current practice it would seem that for many relationships it is possible to rationalize a wide range of values of coefficients including positive as well as negative signs. Examples can be cited about the direction and extent of the relationship between inflation and consumption, the determinants of Government consumption expenditures, changes in consumption expenditures and their impact on investments via their effect on expectations and so on. A single example may suffice. Consider the relationship of the existing capital stock to current or future gross investment. A negative sign or a very low value of the coefficient could reflect the depressing effects of a larger existing stock on net investment—a hypothesis that is not altogether unreasonable in the kind of capital-stock-adjustment or accelerator models in which demand considerations dominate. On the other hand, a positive and numerically significant coefficient could reflect, first, the stimulus that a more adequate infrastructure provides for investment and, second, the purely arithmetical fact that a larger capital stock requires larger replacement outlays. As there seldom are reliable estimates for replacement outlays and, very often, not even sectoral investment detail, one has to make do with treatments such as the following:

The role of the capital stock in an investment function has been of concern to many researchers. Various alternative specifications of the function have been tried and the coefficient invariably bears a positive sign. Others working with developed economies have found it to be negative. Of course, in a developing economy the positive sign may be valid. However, the capital stock variable has not been included in the relevant equation in the model.¹

The scope for rationalizations is even greater when what is at issue is not the sign but the numerical magnitude of a coefficient. Ought we to expect constant returns to scale when fitting aggregate production functions for developing economies? What is a reasonable income elasticity of demand for imported consumer durables? 'Common sense' or one's 'feel for figures' is a poor guide on these issues: its only criterion is that the result under consideration should not look too different from past results. And in many cases the past results in their turn are only based on a more or less unsystematic examination (for instance, with the aid of some earlier inadequately specified regression model) of more or less unsatisfactory data.

We see, therefore, that a very large proportion of the regressors that are used are highly collinear, that the resulting coefficients may have little

¹ M. Dutta and V. Su, 'An econometric model of Puerto Rico', *Review of Economic Studies*, vol. xxxvi (3); No. 107, July 1969, p. 326.

economic meaning and that we may not always be able to detect this either by using statistical tests like the *t*-test or by appealing to our intuition. Given these facts, how do the model builders tackle the problems posed by collinear regressors? The method that is most often mentioned in econometrics texts as a way of dealing with collinearities is to use some external or *a priori* information to supply estimates for the coefficients of all but one of the collinear variables. In practice this 'solution' is seldom adopted—reflecting no doubt the difficulties of obtaining *a priori* estimates for the coefficients. Nor do the models extract the principal components of the regressors and use these instead of the regressors in the equations. Perhaps the reason is that the model builders are interested in the regressors themselves rather than in entities called their 'principal components' which they cannot very readily manipulate for policy or projection purposes. The 'solution' that is most frequently adopted in practice is to drop one or more of the collinear regressors from the estimating equation. Once again a number of examples could be cited, of which two may suffice.

So as to find a function determining Government consumption expenditures, de Vries and Liu try a number of explanatory variables and report three equations:¹

$$\begin{aligned} C_G &= -337.8 + 8.96N & \bar{R}^2 &= 0.95, \quad d = 1.34 \\ & \quad (14.89) \\ &= 31.6 + 0.62T & \bar{R}^2 &= 0.94, \quad d = 1.42 \\ & \quad (14.4) \\ &= -2290 - 0.30T + 48.0N - 61.7t - 5.87P_3 & \bar{R}^2 &= 0.97, \quad d = 1.34 \\ & \quad (2.08) \quad (4.82) \quad (3.79) \quad (3.54) \end{aligned}$$

where *N* is population, *T* is taxes, *t* is time, and *P*₃ is a price index. The entries in parentheses are *t*-ratios. The authors choose the first equation 'because it seems more important in expressing the economic reason behind an increase in government consumption'. The third function is rejected because 'though showing significant parameters for all four variables, the sign of *T* is reverse to the expectation and others are due to multicollinearity'.¹ Notice the extent to which the coefficient for *N* in the equation that is selected (8.96) differs from that in the third (48.0).

On six occasions authors in *UNCTAD, 1968*² postulate that output in the agricultural sector and in 'other sectors, mainly manufactures' is a function of total GDP and of GDP *per capita*. When the output of either of the sectors is regressed on GDP alone or on GDP *per capita* alone it yields highly 'significant' coefficients. When the two regressors are used together the values of the coefficients and their standard errors alter dramatically.

¹ B. A. de Vries and J. C. Liu, *op. cit.*, p. 11.

² *UNCTAD, 1968*, pp. 123-4, 204-5, 248-9.

Some of the coefficients are no longer 'significantly different from zero', though some are—as in the equation for Tanzania's agricultural output cited above. But in each of the six cases after a perfunctory discussion of the standard errors being 'relatively high' or of a sign being 'improbable' on *a priori* grounds the second regressor—*per capita* GDP—is dropped.

The authors seldom pause to notice that by dropping a regressor or a group of regressors to avoid problems of collinearity they are altering the reduced form of their models very considerably and hence the policy conclusions that can be drawn from them. Moreover, and this leads us to the next sub-section, by dropping these variables they introduce serious specification errors into their equations.

(b) *Specification errors: the choice of explanatory variables*

The second and in practice an equally important group of reasons for which it may be misleading to treat individual coefficients as the partial derivatives of calculus has to do with the set of explanatory variables that have *not* been included in an equation. The standard argument for least squares proceeds as follows: let *X* and *Z* be matrices and *Y*, *b*, β , and θ be vectors; assume that the true model is

$$Y = X\beta + u. \quad (i)$$

Least squares procedures estimate β by *b* where

$$b = (X'X)^{-1}X'Y. \quad (ii)$$

Substituting for *Y* from (i) and taking expectations we get

$$E(b) = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'E(u). \quad (iii)$$

But as $(X'X)^{-1}X'X = I$ and $E(u) = 0$

$$E(b) = \beta. \quad (iv)$$

That is, *b* is an unbiased estimator of β . Suppose, however, that the correct model is

$$Y = X\beta + Z\theta + w \quad (v)$$

while the model used for estimation is

$$Y = X\beta + v.$$

That is, let us assume that the analyst has left out some of the explanatory variables because, as we saw in the preceding section, he wants to avoid problems arising from collinearities among regressors, because he has misconstrued the causal relationships, because he cannot obtain satisfactory data about them, or for some other reason; then we obtain biased estimates of β . As before,

$$b = (X'X)^{-1}X'Y.$$

Substituting for Y —but this time from (v)—and taking expectations we get

$$E(b) = (X'X)^{-1}X'X\beta + (X'X)^{-1}X'Z\theta + (X'X)^{-1}X'E(w). \quad (\text{vi})$$

As $(X'X)^{-1}X'X = I$ and $E(w) = 0$ we have

$$E(b) = \beta + A\theta, \quad (\text{vii})$$

where A is $(X'X)^{-1}X'Z$ and is known as the alias matrix. Thus the omission of a variable that has some causal links with Y means not only that in the end we do not have an estimate of the coefficient linking the excluded variable with Y but also that the estimates of coefficients linking all the other X_i to Y may be vitiated. Incidentally, if the Z are included when in fact Y is independent of them, or if they are left out when (though influencing Y) they are linearly independent of X , then the estimates of β are not vitiated. If $\theta = 0$ or if $A = 0$, then $E(b) = \beta$ in (vii).

If we are interested only in forecasting and have reason to believe that the future relationship between (X) and (Z) will be the same as in the past, there is nothing wrong in using an abbreviated model [$Y = f(X)$] and having the (X) do the work for themselves as well as for the (Z) . But if we are going to use the coefficients for policy prescription then this is precisely the situation when leaving out (Z) as in the abbreviated model will give us highly misleading estimates of the coefficients linking (Y) and (X) . Instead of leaving (Z) out, the model-builder should search out the parent variables that explain (Z) and (X) themselves—if such variables exist—and substitute those variables into the equation. If the search does not yield the ultimate explanation for the collinearity of (X) and (Z) one can indeed argue that both sets of variables, (X) and (Z) , should be included even if the t -ratios of one or both are not significant so that mechanical regression procedures do not attribute to one of them the variance of the dependent variable that the two together help explain. From this negative point of view a variable may be very important even if it is 'statistically insignificant'.

We have already stressed that many econometric models *are* being used for control purposes also. The question that arises then is: how often is it that in practice explanatory variables that have a quantitatively significant influence on the dependent variables get left out from one's estimating equations? One has only to scan the text (and perhaps the footnotes) accompanying the equations of a model to realize that—for one reason or another—every author has had to leave out quite a few variables from most equations that he would have liked to include.¹ Apart from the fact that

¹ One repeatedly comes across passages like the following: 'The suitable explanatory variable for raw material imports seems to be the level of output in secondary industry (manufacturing, construction, and electricity). However, in the absence of necessary data, the following two alternative forms (linking imports of raw materials to fixed investment and to GDP) were obtained', *UNCTAD, 1968*, p. 360. 'The model fails to include all the relevant variables affecting inflation and growth either because of the lack of statistical

for many of the important explanatory variables the relevant data is not available, and that some of them are very difficult to quantify anyhow, there is the pervasive difficulty that economic theory offers insufficient guidance as to what constitutes an adequate list of explanatory variables. Some models assume very simple functional relationships; others are more venturesome and imaginative, making the birth-rate a function of *per capita* consumer expenditure and time; rates of exchange a function of capital transfers into the country in the preceding period, the ratio of the official exchange rate to the GDP deflator and the 'capacity to import' which in turn is to be a function of the price index of the country's imports; and investment a function of earnings of the country from tourism and net credit extended to the private sector.

Now, there is no logical reason why any of these or more esoteric variables should not have been included; but the list does suggest that when one opts for one or a very small number of explanatory variables one has possibly not exhausted the field for potentially relevant candidates. Thus the risk is always present that the estimates of the coefficients are misleading. In a sense, the fact that in devising a model the model-builder almost invariably has a specific purpose in mind (for instance, to project the two gaps, to analyse the consequences of different exchange-rate policies, to estimate the 'trade-offs' between various rates of growth and inflation) itself greatly influences his choice of explanatory variables. This becomes apparent when three or four models for the same economy are compared. The sets of equations in them differ much more in their list of explanatory variables than they do in their lists of dependent variables: a model aimed at analysing the choices between inflation and growth attempts to include price variables at as many points as possible; another shows a marked preference for the exchange rate; a third for ratios depicting structural changes in the economy. As the purpose influences the selection of explanatory variables, it may indirectly influence the values and signs of the coefficients derived and, thus, the numerical conclusions that are drawn from the model.

Each of the structural equations in a model can be specified in a variety of ways. When we use the model as a whole—for example, focusing on the import-export gap (a magnitude that depends on a large number of equations)—the uncertain specifications of each structural equation are communicated to the final result; we cannot be certain that the biases will cancel and not reinforce each other. This is especially important because

data or because of the inability to identify them properly', B. A. de Vries and J. C. Liu, *op. cit.*, p. 1; 'Since the data about the availability of tractors and consumption of agricultural products are not available, the former is excluded and the latter is substituted by a proxy variable, aggregate consumption', *ibid.*, p. 7. And so on.

often, as in the case of the import-export gap, we are interested in what are relatively small differences between fairly large variables so that even small errors in the latter are liable to affect the former in a major way; in fact, there often are other variables in the system—a country's debt-servicing obligations are a case in point—that depend on the *cumulative* values of these small differences. Moreover, just as there is no unique way of specifying a single equation so is there no unique way of specifying the model as a whole. For example, the conventional two-gap models are designed to project the investment-savings gap and the import-export gap independently and are by their very nature overdetermined. There are many ways of reconciling the two gaps and the outcome in terms of the final projections and policy prescriptions will differ depending on which particular reconciliation procedure is adopted. Considering the numerous possibilities open to him in specifying a model and its individual relationships the model-builder may find it very difficult to be sure that his projections and policy prescriptions are anything more than a mechanical result of the particular way in which he has specified his model.

IV. Functional form of the equations

One way in which a model-builder may mis-specify the relationship between a dependent variable and a set of independent variables is to assume that the relationship is, say, linear when in fact it is, say, quadratic. This is a particular case of the more general class of specification errors considered above; but a few additional comments are needed.

The choice between one functional form and another is not a matter of indifference. First, each form implies a very specific view of underlying economic relationships. There is nothing in economic theory or in the empirical research conducted thus far which tells us that the most reasonable assumption about the future is that the marginal rate will remain unchanged (the assumption one makes when one uses a linear equation of the $Y = a + bX$ variety). It may be just as reasonable to assume that both the average and marginal rates will remain constant and use the relationship ($Y = bX$); or to assume that both will vary but that the elasticity will remain unchanged and use ($\log Y = \log a + b \log X$); or to assume that the elasticity will fall as Y rises and use a semi-log relationship ($Y = a + b \log X$); or that the elasticity will fall as X increases using, then, the log-inverse function ($\log Y = a - b/X$); or that the average and marginal rates as well as the elasticity will vary without any particular restriction as, say, in the quadratic equation ($Y = a + bX + cX^2$). In dealing with a few problems one can invoke some consistency conditions (for instance the homogeneity conditions in demand analysis) to limit the choices open to the model-builder or to constrain the coefficients in some particular way. But, for

most of the relationships that highly aggregated models of the kind considered here try to quantify, our knowledge of economic relationships is just not specific enough to suggest unambiguously that we should use one form and not another.

Second, one cannot just abdicate responsibility and choose that particular form which gives 'the best fit' to the sample period data. For in many cases various alternative forms yield similar \bar{R}^2 s. But, it may be asked, if different forms fit the sample period data equally well, does it make much practical difference which is used for making projections or drawing policy conclusions? Indeed it does. Consider the following equations which have been computed from Kenyan data for 1960-8.

		<i>F</i> -ratio	\bar{R}^2
(DIRTAX) _t	= 0.04847 (GDPFAC) _t (0.00191)	36.997	0.822
(DIRTAX) _t	= - 9.03007 + 0.07468 (GDPFAC) _t (0.39871) (0.00704)	112.502	0.933
log(DIRTAX) _t	= - 2.54335 + 1.48174 log (GDPFAC) _t (0.00969) (0.13391)	122.441	0.938
(DIRTAX) _t	= 34.15114 - 0.18211 (GDPFAC) _t + (0.19761) (0.05425) + 0.00037 (GDPFAC) _t ² (0.00008)	240.242	0.984
(MCAPTL) _t	= 0.27497 (INVEST) _t (0.01945)	34.920	0.614
(MCAPTL) _t	= - 5.28821 + 0.36031 (INVEST) _t (1.02273) (0.05305)	46.125	0.849
log(MCAPTL) _t	= - 1.37567 + 1.44553 log (INVEST) _t (0.03465) (0.25389)	32.417	0.797
(MCAPTL) _t	= - 27.87714 + 1.14340 (INVEST) _t - (0.96160) (0.56760) - 0.00604 (INVEST) _t ² (0.00436)	27.047	0.867
log(MCAPTL) _t	= 1.74490 - 28.71198 [1/(INVEST) _t] (0.05169) (2.49690)	132.228	0.943

In the first set of four equations receipts from direct taxes are made a function of GDP at factor cost and in the second set of five equations imports of capital goods are made a function of gross fixed investment. Figures in parentheses are standard errors. Each of the equations yields reasonably good fits for the sample period and the *t*-ratios are satisfactory for all of them. In Figs. 1 and 2 the different equations are used to project the dependent variables into the future. In each figure identical values of the independent variable are used for all the equations. The diagrams show very clearly that the different forms yield widely divergent projections for the future.

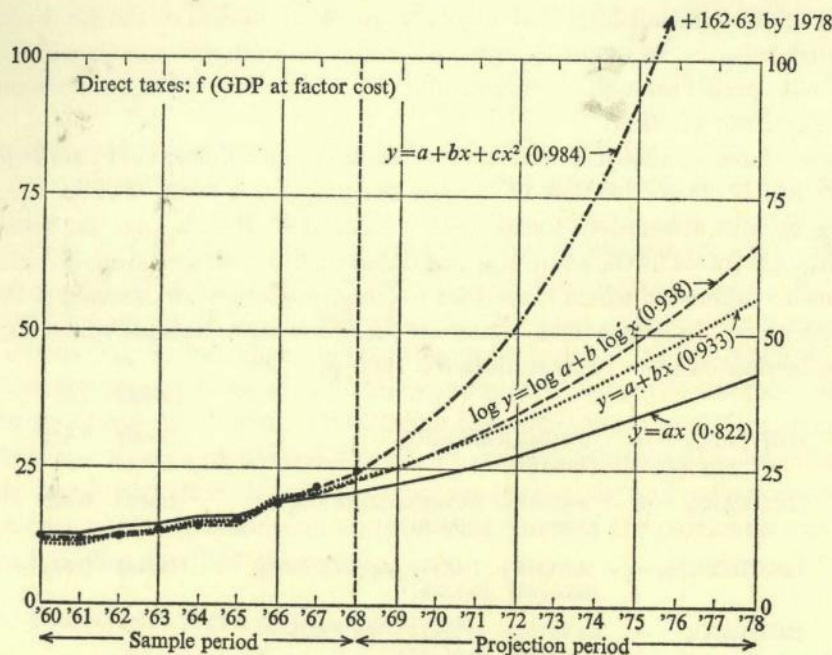


FIG. 1

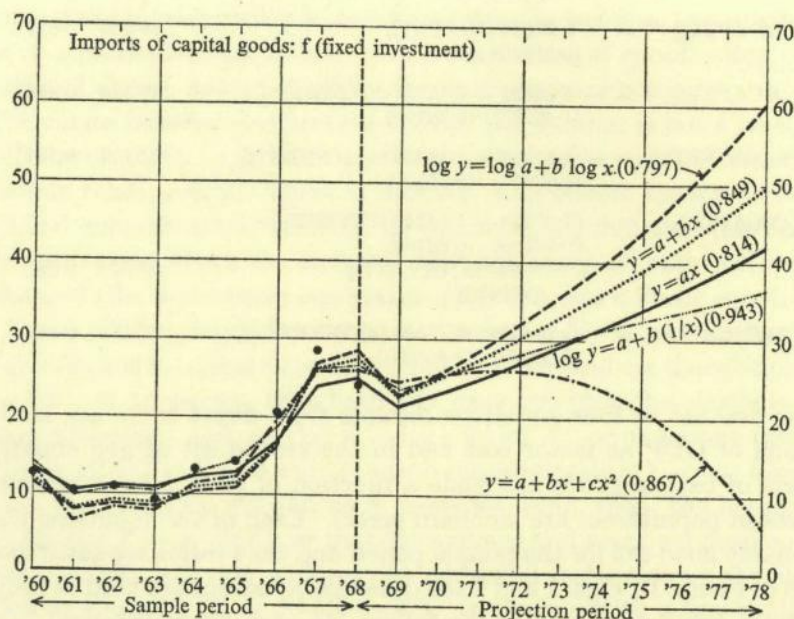


FIG. 2

FIG. 1 and FIG. 2. Effects of functional form on projections. Actual values in the sample period are represented by (●). Values of R^2 are given in parentheses.

The same sort of differences arise when one uses the first derivatives of the reduced form equations to assess the consequences of a change in some of the independent variables. While the first derivatives of equations like $(Y = aX)$ or $(Y = a + bX)$ are constant and unambiguous, those of equations like $(Y = aX^b)$ or $(Y = a + bX + cX^2)$ vary with the value of X . Therefore, their absolute values and, in many cases, the rankings of, say, individual industries according to their effect on the import-export gap, may be very different depending on whether one takes X to be equal to the mean of the sample period or assigns to it the value it is projected to have at the end of the projection period. As the values of the derivatives vary, so must one's policy prescription.¹

We do not have to seek far for the reasons that account for these differences. Least squares estimates of coefficients are fairly sensitive to extreme values. As the sample consists of only five to fifteen observations even a few aberrant values can affect the estimates very considerably. Moreover, as the equations are used to project variables many time periods beyond the sample period or to gauge the impact of very substantial changes in the 'policy variables', even small differences in the marginal rates of change implicit in the equations get magnified in a very substantial manner.

Finally, the choice of a particular functional form is not without its consequences for the efficacy of initial estimates of the coefficients. It can be shown—and the case is just a particular variety of specification errors considered above—that in general a linear equation will not yield an unbiased estimate of the true marginal rate even close to the centre of gravity of the observed point set when the true underlying relationship is, say, quadratic or log-linear. Biases arise both because of non-linearity in the true relationship and because of an erroneous assessment of disturbances.²

V. Mechanical use of certain statistical tests

The question now comes to mind: do not standard statistical tests alert us to the sorts of deficiencies that have been illustrated above? For example, if some important explanatory variables have been left out of an equation can it not be assumed that the value of the coefficient of determination will be extremely low or that the residuals will be highly auto-correlated

¹ If the data has some non-linearities in it non-linear equations like the quadratic one are bound to explode when extended far enough into the future. The point being made in the text is not that quadratic equations should always be used for long-term projections even when they yield patently absurd results; rather that the data often contain significant accelerations or decelerations, and if so these should be examined instead of being mechanically suppressed in a linear equation.

² Cf. H. Theil, 'Specification errors and the estimation of economic relationships', *Revue de l'Institut International de Statistique*, vol. 25, No. 1/3, 1957, pp. 41-51, deals with this and other issues raised in Section IV.

and that the standard tests will draw our attention to the excluded variables? Similarly, if the independent variables are highly collinear will the low values of the t -ratios not warn us about their consequences? These questions need to be considered at some length; the answers unfortunately are far from reassuring.

Most models display statistics relevant to three statistical tests: the coefficient of determination, the t -test for individual coefficients, and the Durbin-Watson test for auto-correlation in the residuals. Among these the coefficient of determination receives the greatest attention. In fact, many models display an almost obsessive concern with the \bar{R}^2 .

Consider some extracts from a paper projecting employment patterns over twenty years:

'Group 0. For professional workers, the semi-logarithmic form gave the best overall fit and appeared to cope with elements of non-linearity in the relationship. The equation chosen was:

$$S_0 = -12.05 + 2.94 \ln Y + e, \quad (0.21)$$

with R^2 equal to 0.843. Only per capita income shows a regression coefficient significant at the five percent level, although the linear form (with smaller R^2 of 0.811) shows population size significant as well. . . .

'Group 2. A semi-logarithmic form was shown to fit the data on clerical workers quite well ($R^2 = 0.784$), with no evidence against the linearity of this relationship. Although growth rate as well as income was significant in the simple linear fit, R^2 was lower (0.766). Hence, the equation chosen was:

$$S_2 = -13.12 + 3.22 \ln Y + e. \quad (0.28)$$

'Group 4. For farmers and related occupations, the best fit is found with the semi-logarithmic relationship:

$$S_4 = 153.33 - 19.34 \ln Y + e, \quad (1.40)$$

with R^2 of 0.838. The scatter diagram revealed no problems with the linearity of this fit. The simple linear form shows population and growth rate significant (at the 10 percent confidence level) as well as income, but R^2 was much lower at 0.677. . . .

'Group 9. As with sales workers, neither equation form is very successful in explaining the variance. The semi-logarithmic form, however, had the higher R^2 (0.278, significant by F -test), and will be used despite the appearance of non-linear elements in the corresponding scatter diagram.

The equation used is:

$$S_9 = 5.55 + 1.63 \ln Y - 0.78 \ln P + e. \dots^1 \quad (0.53) \quad (0.31)$$

Similarly, the Colombia model referred to above cites two equations explaining investment in construction:

$$\frac{I_{hH}}{P_{cnt}} = -0.9595 + 0.1312X + 1.208 \left(\frac{I_{mct}}{P_{icnt}} \right)_{-1} \quad \bar{R}^2 = 0.977, \quad d = 2.327$$

and

$$\begin{aligned} \frac{I_{hH}}{P_{cnt}} = & -0.311945 + 0.39590 [X - X_{-1}] + 0.5920 \left(\frac{I_{mct}}{P_{icnt}} \right)_{-1} + \\ & + 0.94580 \left(\frac{I_{hH}}{P_{cnt}} \right)_{-1} \quad \bar{R}^2 = 0.912, \quad d = 1.590. \\ & (0.11380) \quad (0.3550) \quad (0.09470) \end{aligned}$$

The coefficients for (I_{mct}/P_{icnt}) are very different in the two equations and this should have provided an important clue to the 'reliability' of the coefficients for control. However, the text merely notes that the former equation 'was accepted in the final analysis on the basis of its predictive value'.²

In a superficial sense this preoccupation with the \bar{R}^2 is well rewarded. The models are fairly simple and highly aggregated representations of the economies of developing countries.³ Few of them incorporate any institutional detail or any information about production relationships in the economy. Yet if one considers their structural equations individually one is struck by their very impressive coefficients of determination. It is often and rightly said that the process of growth and change in developing countries is a complex and in many respects an incompletely understood phenomenon. The equations of these models, on the other hand, seem to suggest that the broad features of growth and change in a host of dissimilar countries over the past five to fifteen years can be represented by extremely

¹ J. S. Scoville, 'The occupational structure of employment, 1960-1980' in *Sectoral Aspects of Projections for the World Economy, First International Seminar on Long-Term Economic Projections*, United Nations, New York, 1969, vol. 3, pp. 77-109, especially pp. 86-90.

² K. Marwah, op. cit., pp. 233 and 236-7. (X) is GNP at 1958 prices; (I_{mct}) is value of imports of construction materials, and (P_{cnt}) and (P_{icnt}) are indices of prices of construction materials in pesos and dollars.

³ The eighteen models in *UNCTAD, 1968* utilize about 370 estimated equations (though, as noted elsewhere, about 630 estimated equations are reported in the text). Their distribution by the number of regressors used in the equation is as follows.

Total number of equations	Equations with 1 regressor	Equations with 2 regressors	Equations with 3 regressors	Equations with 4 or more regressors
370	281	67	21	1

simple equations with no more than one or two, or at the most four or five, explanatory variables.¹

Even though the point has been made that complex social phenomena can often be represented by fairly simple mathematical expressions,² the high degree of success of the models and their simple equations, in that they seem to explain a wide range of dependent variables, comes as a considerable surprise. However, this success of the models in explaining developments in particular economies is almost entirely deceptive. And the models pay a substantial price for this apparent success. The foregoing extracts illustrate one consequence of relying so heavily on the coefficient of determination. By mechanically adopting the equation with the higher \bar{R}^2 the authors in each case exclude a number of explanatory variables from the particular segment of their model. The reduced form and, perhaps, also the dynamics of the model and hence the policy conclusions that can be derived from it are all altered drastically. Now, it so happens that the coefficients of determination can often be affected by so many extraneous factors, which may have nothing to do with functional or causal relationships between variables, that their use can be very misleading—especially when the models or equations are to be used as a basis for policy prescription. Three sets of considerations will be used to illustrate some common situations in which reliance on the \bar{R}^2 can mislead the model-builder.

By far the most frequently encountered situation is the one that Yule documented forty-five years ago:³ two series may display 'nonsense correlations' because, even though unrelated in a causal sense, they have pronounced trend or cyclical components. Consider, for instance, Kenya's exports and imports of services other than tourism. These consist mainly of freight, transportation, insurance, and similar items. Let us postulate that most of these arise from the movement of goods in foreign trade and, therefore, let us regress them against the single regressor: the total value of goods imported into and exported from Kenya. For different equation forms the \bar{R}^2 for the 1960–8 data for the exports of services are between

¹ On a rough count the eighteen models in *UNCTAD, 1968* between them use about 370 estimated equations and the eight models of *ECAFE, 1968* list about 120 computed equations. The distribution of these equations by their coefficients of determination is as follows:

Coefficient of determination:	< 0.3	> 0.3 & < 0.4	> 0.4 & < 0.5	> 0.5 & < 0.6	> 0.6 & < 0.7	> 0.7 & < 0.8	> 0.8 & < 0.9	> 0.9	
% of total equations									
(i) <i>UNCTAD, 1968</i>		2.2	1.3	6.1	6.1	11.2	11.7	23.9	37.5
(ii) <i>ECAFE, 1968</i>		2.5	5.0	4.2	3.4	4.2	11.8	18.5	50.4

² M. G. Kendall, 'Natural law in the social sciences', *Journal of the Royal Statistical Society, Series A*, 1961, Vol. 124, Part 1, pp. 1–16, *en passant*.

³ G. Udny Yule, 'Why do we sometimes get nonsense correlations between time series?', *Journal of the Royal Statistical Society*, Jan. 1926, vol. lxxxix, Part 1, pp. 1–64.

0.912 and 0.950 and for the imports of services between 0.715 and 0.779. The F -ratios are highly significant as are the t -ratios. The fact, however, is that Kenya's foreign exchange earnings from 'exports of services' and expenditures on 'imports of services' are related only in an incidental way to its imports and exports of goods: first, Kenya has no large shipping lines in which these goods may be ferried and, second, imports of goods as used in computing the equations are recorded c.i.f. so that whatever impact they have on freight and transportation costs is already taken account of elsewhere. Kenya's earnings from freight and transport arise more from the imports and exports of *Uganda* and *Tanzania* rather than its own imports and exports. The incidental way in which its own imports and exports of goods enter the picture is that a small part of Kenya's earnings from exports of services consist of port-charges and part of these are collected from the ships that ferry goods specifically meant for or specifically originating in Kenya. In this case, therefore, the \bar{R}^2 would certainly have misled us if we abandoned our search for explanatory variables once we had hit upon the total value of goods imported into and exported from Kenya. But, it may be argued, what is wrong in using the equations for projections or policy prescription? After all, the variables *have* moved together in the past. The difficulty is that the less certain one is about the existence of a causal connection between the dependent and the explanatory variables, the less confident one is about asserting that 'as they have moved together in the past it is reasonable to assume that they will move together in the future'. In fact, taking the 'exports of services' as an example, one would be on firmer ground in asserting that these are likely to grow at a rate *lower* than the growth of the total value of goods imported into and exported from Kenya: Uganda's exports and imports together are expected to grow at a rate about one-half of the rate being currently projected for Kenya's imports and exports of goods and Tanzania has definite plans of developing Dar-es-Salaam which will mean that it will be relying somewhat less on Mombasa in the future. To avoid pitfalls of this kind models of developed economies almost invariably estimate their relationships from data that has been transformed—for instance, by taking its first differences—to avoid the effects of auto-correlation in the series. But hardly any of the newer models for developing countries is estimated in terms of first differences or similarly transformed data.¹

¹ Not one among the eighteen models of *UNCTAD, 1968* estimates its relationships using data that has been transformed to remove the effects of auto-correlation in the original series. The argument here is not that models should always be calculated from first differences: in fact, one can think of situations in which taking first differences may itself introduce spurious correlations into the data. The argument simply is that in many of the equations of the aggregative models considered here the high \bar{R}^2 may well be due to factors like trend components in the series. The difficulties outlined in the text are not

The second illustration of how a spuriously high \bar{R}^2 may be obtained relates not to a failure to transform data when it should have been transformed but to transforming it in such a way as to generate once again the kind of situation that Yule warned us against¹—in which one gets 'conjunct series with conjunct differences'. This illustration is provided by the output–investment functions used in *ECAFE, 1968* and *UNCTAD, 1968*.² Fourteen of the eighteen models of *UNCTAD, 1968* and seven of the eight models of *ECAFE, 1968* use the following output function:

$$Y_t = f\left(\sum_{i=0}^{t-1} I_i\right),$$

where Y is GDP at factor cost and I is gross fixed investment. The function yields extremely impressive fits: the \bar{R}^2 being almost invariably higher than 0.95. The two series Y and I have pronounced trend components in almost all cases; by cumulating I we are generating a new series which not only has high auto-correlation but the first differences of which are also auto-correlated. The nature of the effect on the coefficients of determination is perhaps best demonstrated by a simple example. Consider three series: Y_t , X_t , and Z_t where Z_t is $\sum_{i=0}^{t-1} X_i$. Y_t is growing at 10 per cent and X_t fluctuates around a mean of about 11.5.

Y_t	X_t	$Z_t = \sum_{i=0}^{t-1} X_i$	Y_t	X_t	$Z_t = \sum_{i=0}^{t-1} X_i$
1.10000	12	0	2.35794	10	89
1.21000	11	12	2.59374	12	99
1.33100	10	23	2.85311	11	111
1.46410	14	33	3.13842	13	122
1.61051	9	47	3.45226	9	135
1.77156	10	56	3.79749	14	144
1.94871	12	66	4.17724	8	158
2.14358	11	78			

confined to time series models. Just as trends often result in spurious correlations in time series models so may factors like scale mislead us in models based on cross-section data. Liberty is not the only thing that demands eternal vigilance.

¹ G. Udry Yule, op. cit.

² The same function is used in Kanta Marwah, op. cit. The function is used for fourteen of the eighteen countries for which *UNCTAD, 1968* contains models. For Argentina and India imports of capital and intermediate goods have been included in addition to cumulative investment. Among the remaining four countries the function was fitted for Kenya, Uganda, and Tanzania also but the regressions revealed a negative correlation between investment and GDP for Kenya and no significant correlation for Uganda and Tanzania and so the authors assumed a capital–output ratio of 3:1 for Kenya and 2.5:1 for the latter two (*UNCTAD, 1968*, pp. 124–5, 205, and 249). For Thailand there is no separate function, the relevant magnitudes being derived from an identity.

In fitting Y_t to X_t we get:

$$Y_t = 3.00572 - 0.06106X_t \quad R^2 \quad F\text{-ratio} \\ (0.407) \quad 0.0125 \quad 0.165$$

The figure in parentheses is the t -ratio. But when we cumulate X_t and fit Y_t to Z_t as in the *UNCTAD, 1968* and *ECAFE, 1968* output functions we get:

$$Y_t = 0.81828 + 0.01933Z_t \quad 0.970 \quad 430.172 \\ (20.740)$$

The very impressive coefficient of determination and the significant t -ratio are due solely to the fact that X_t has been cumulated and do not in any way imply a reliable, stable, or significant relationship between X_t and Y_t . The reader can indeed experiment with any set of positive random numbers and obtain similar results.

The final illustration of how \bar{R}^2 may be misleading involves regressing one ratio on another. We know that gross fixed investment in a sector $(\text{INVEST})_i$ consists of net investment $(\text{INVNET})_i$ and depreciation $(\text{INVDEP})_i$:

$$(\text{INVEST})_i = (\text{INVNET})_i + (\text{INVDEP})_i.$$

Assuming that a given proportion (d) of the capital stock is replaced every year and that the capital stock bears a fixed relation to the value added in a sector

$$(\text{INVDEP})_i = d(\text{KSTOCK}) = d.k(\text{GDPFAC})_i = b(\text{GDPFAC})_i.$$

Therefore $(\text{INVEST})_i = (\text{INVNET})_i + b(\text{GDPFAC})_i$.

Dividing both sides by $\Delta(\text{GDPFAC})_i$ we get

$$\frac{(\text{INVEST})_i}{\Delta(\text{GDPFAC})_i} = \frac{(\text{INVNET})_i}{\Delta(\text{GDPFAC})_i} + \frac{b(\text{GDPFAC})_i}{\Delta(\text{GDPFAC})_i} \quad (1)$$

$$k_{gi} = k_{ni} + b \frac{1}{r_i},$$

where k_{gi} is the gross capital–output ratio in sector (i), k_{ni} the net capital–output ratio, and r_i the sectoral rate of growth. On fitting equation (1) to the data we obtain the relevant gross capital–output ratios.¹ Given the sectoral growth rates (r_i) we obtain numerical values of the capital–output ratios that can be used for determining, say, the amounts of investment required to assure a certain specified growth of value added in the sectors. For the sample period we may estimate the (r_i) by fitting the trend equation [$Y_t = Y_0(1+r)^t$ or $\log Y_t = \log Y_0 + t \log(1+r)$] to the sample period data.

¹ The preceding formulation of arriving at capital–output ratios is taken from A. M. Strout, 'Savings, imports and capital productivity in developing countries', First World Congress of the Econometric Society, Rome, Sept. 1965. For an alternative way of handling changing growth rates see J. Vanek, *Estimating Foreign Resource Needs for Economic Development*, McGraw Hill, New York, 1967. Strout uses the function for cross-section data. Vanek's estimates of coefficients are not derived from regression techniques.

The sectoral growth rates and the gross capital-output ratios estimated from the Kenyan data for 1960-8 along with the coefficients of determination are given below.

(1) Sector	(2) Rates of growth in the sample period	(3) \bar{R}^2 for col. 2	(4) Capital- output ratios for the sample period	(5) \bar{R}^2 for equations estimating capital- output ratios
Agriculture, forestry, and fishing	5.39	0.801	1.58	0.979
Mining and quarrying	1.84	0.000	5.20	0.624
Manufacturing and repairing	9.27	0.966	2.87	0.987
Building and construction	3.80	0.134	3.63	0.961
Electricity and water	8.79	0.966	4.22	0.986
Transport, storage, and communications	8.84	0.981	4.85	0.751
Other services	6.86	0.969	2.61	0.970
TOTAL: When estimated directly	6.56	0.964	2.69	0.999

The coefficients of determination for the trend equations and for the capital-output ratios are satisfactory for all the sectors other than mining and quarrying. In addition, the capital-output ratios do not seem unreasonable. Would we be justified in presuming that the satisfactory \bar{R}^2 and the reasonable values of the ratios indicate that we now have a satisfactory basis for forecasting sectoral gross investment?

Notice that the denominator $(\Delta \text{GDPFAC})_t$ is the same on both sides of equation (1). In fact, it turns out that the variance of $(\Delta \text{GDPFAC})_t$ is much greater than the variance of either $(\text{INVEST})_t$ or $(\text{GDPFAC})_t$. So that the reason we get the very impressive \bar{R}^2 for the capital-output ratios is largely that the variance of $(\Delta \text{GDPFAC})_t$ is just explaining itself. The literature contains formal tests relevant to this situation¹ but they are seldom employed in practice. Whether one is going to regard the very high values of \bar{R}^2 as 'spurious' or not depends to a large extent on one's predilections; for philosophically the situation is an ambiguous one.²

¹ Cf. E. Kuh and J. Meyer, 'Correlation and regression estimates when the data are ratios', *Econometrica*, vol. 23, 1955, pp. 400-16; and A. Madansky, 'Spurious correlation due to deflating variables', *Econometrica*, vol. 32, 1964, pp. 652-5.

² 'The question of spurious correlation quite obviously does not arise when the hypothesis to be tested has initially been formulated in terms of ratios, for instance in problems involving relative prices. Similarly, when a series such as money value of output is divided by a price index to obtain a 'constant dollar' estimate of output, no question of spurious correlation need arise. Thus, spurious correlation can only exist when a hypothesis pertains to un-deflated variables and the data have been divided through by another series for reasons extraneous to but not in conflict with the hypothesis framed as an exact, i.e. non-stochastic, relation', E. Kuh and J. Meyer, op. cit., pp. 401-2.

Perhaps the following practical consideration will be more persuasive than lengthy arguments about whether or not the correlations are properly regarded as spurious. The sectoral capital-output ratios reported above were used along with actual year-to-year changes in sectoral (GDPFAC) to estimate fixed investment in the sample period. The following χ^2 result from a comparison of forecasted and actual investment in 1961-8.

*Comparisons of actual and forecasted gross fixed investment
in Kenya, 1961-8*

Sector	χ^2 : Total 1961-8	χ^2 : Years 1961-8 taken individually
(1)	(2)	(3)
Agriculture, livestock, forestry, and fishing	0.002	347.635
Mining and quarrying	0.037	141.052
Manufacturing and repairing	0.255	18.071
Building and construction	0.098	268.240
Electricity and water	0.026	8.395
Transport, storage, and communications	0.505	14.011
Other sectors	0.186	40.502
All sectors	1.109	837.906

Column 2 indicates that the forecasts of cumulated investment for the period as a whole are indeed satisfactory; the χ^2 for 'all sectors' in column 2 is less than the permissible limit with six degrees of freedom. But the cumulated totals are made up of very large over- and underestimation of investment in individual years; thus almost all the individual χ^2 values in column 3 as well as the χ^2 for 'all sectors' are well beyond their permissible values for seven and fifty-five degrees of freedom respectively. May one conclude that while the capital-output ratios are not reliable for forecasting investment in an individual year such as 1970 or 1971 they are likely to give satisfactory estimates for investments in the 1969-78 period as a whole? Unfortunately, even this much is by no means certain: cumulative investments are close to actual investments only *over the sample period taken as a whole*; when, for instance, forecasted investments are totalled up for a *part* of the sample period (say, 1965-8) they are once again vastly different from the totals of actual investments. Thus we cannot be certain that even the cumulated values of investment will be reasonable estimates of the true values for any period other than the sample period taken as a whole. Reliance on the very high \bar{R}^2 of equations estimating capital-output ratios would, therefore, have been completely unwarranted. A part of the reason, of course, is that the \bar{R}^2 of equation (1) above relate to the magnitude

$[(\text{INVEST})_i/(\Delta\text{GDPFAC})_i]$ when, in fact, we are interested in—and are implicitly using them to tell us something about—the variable $(\text{INVEST})_i$.

Besides the coefficient of determination the two tests most frequently used—or at least displayed—are the t -ratio and the Durbin-Watson statistic. Some comments on the t -ratio were included in the discussion of collinear regressors. The Durbin-Watson statistic does not require extended comment.

The importance of the error-term is greatly emphasized in the theory of econometrics, since a critical assumption of regression and econometric models is that the error-term is random. Auto-correlation in the residuals may indicate, for instance, that a systematic element has been left out in the explanation; in addition, least squares estimates are no longer efficient (though they are still unbiased and consistent) and the usual significance tests are no longer directly applicable. The question becomes particularly important in auto-regressive schemes that utilize lagged variables for in such cases the estimates of 'b' are neither unbiased nor consistent. For all these reasons econometric theory places considerable emphasis on a careful examination of the error-term. In this spirit the Durbin-Watson statistic—one of the tests available for examining residuals—is faithfully presented for most equations. But in many cases no further attention is paid to it. Thus, in a particular case, it is stated that 'the (Durbin-Watson) statistic shows that the residuals of the German function are positively auto-correlated at the five percent level . . . the subsequent discussion, however, will disregard all complications arising from auto-correlated errors'.¹ Nor is this unrepresentative:² for many of the equations contained in econometric models of developing countries the Durbin-Watson statistic shows significant auto-correlation in the residuals, yet in no case is any attempt made to transform the data or the specification of the equation in any suitable way to avoid the auto-correlation. Of course, it may be that these models do not pay much attention to the statistic because at least thirteen to fifteen observations are needed to be able to use its tabulated values and more often than not the models are based on a smaller number of observations. One suspects, however, that having listed the statistic no one bothers about it simply because no one else bothers about it either: everyone is aware of sinning in good company.

Before concluding this review of the relevance of standard statistical tests, it will perhaps be worth while to stress one general point. The theory of the sampling distribution of \bar{R}^2 and other indices like the t -ratios assumes

¹ G. Menghes and H. Diehl, 'Time stability of structural parameters', *Econometric Analysis for National Economic Planning*, ed. by P. E. Hart, G. Mills, and J. K. Whitaker, University of Bristol, 1964, pp. 299-317; the sentence is from p. 307.

² For similar remarks see *UNCTAD, 1968*, pp. 108, 126, 191, 234.

that the hypothesis to be tested has been formulated independently of the data that is used to test its validity. When, as in practice, the hypothesis is framed after a scrupulous and in some cases systematic examination of the data (with, for instance, the aid of step-wise regression procedures) the indices lose a good part of their discriminatory power. To validate the hypothesis or the equation that embodies it one needs data from a *different* sample—for instance, from a time span different from the sample period used in estimating the equation. As the developing countries seldom have consistent series of more than five to fifteen years one seldom finds a model-builder verifying his equations or hypotheses by reference to data other than the data used in their estimation and formulation.

VI. Point forecasts and policy prescriptions for years far into the future

Even if it were the case that the data from which the equations were estimated was reliable and ample; that the problems arising from collinearities had been successfully taken care of; that we could be sure that the equation had been correctly specified; and if care had been taken to avoid the kinds of traps sketched in Section V—even then the question as to whether the coefficients and equations could be regarded as reliable guides to the future is a separate one, which has to be answered with reference to whatever information one can muster about *that* future.

Models currently in use make point forecasts for years well beyond the end of their sample of observations. We can be almost certain that the precise value for Y we predict from an assumed relationship, say,

$$(Y = a + bX),$$

will never be realized. The individual value of Y is in a sense only an average of likely outcomes. If the forecasts from an equation are to be viewed in the context of some significance or probability levels they must be in terms of a Y -distribution and not just a single value of Y .

Now, the interval estimate of Y at $(1-\alpha)$ level of confidence that one must work with is given by

$$\hat{Y} \pm t_{(n-p-1), (1-\alpha)/2} s \sqrt{(X'_0(X'X)^{-1}X_0)},$$

where \hat{Y} is the point forecast of Y , n the number of observations in the sample period, p the number of explanatory variables, and X_0 that point in X -space for which the forecasts are being made. The upper and lower boundaries of these interval forecasts curl away from the line of point forecasts the further away X_0 is from the means of X in the sample period. Typically, individual equations are being used to make forecasts for periods

when the independent variables are two to five and sometimes an even larger number of times their mean values in the sample period and one and a half to three and sometimes an even larger number of times their values at the end of the sample period.¹ It is not surprising then that when we are conscientious and do make interval projections for the future—even with just one functional form—the upper and lower bounds turn out to be a long way apart from each other.

The point is not merely that one now has a band instead of a single point forecast from an individual (structural or reduced form) equation; these bands have to be *combined* in some way to yield an interval forecast of, say, the import-export gap to which one may attach a precise probability significance. Unless one has estimated the reduced form equation for the gap directly—in which case there is no need to *combine* any equations—the attempt to combine conditional forecasts from individual equations soon leads one into problems. For in combining the confidence bands of one structural equation with confidence bands of another one soon loses control of probability levels.

We know that if $R_1 \dots R_n$ are random variables with finite variances $\sigma_1^2, \dots, \sigma_n^2$ and $S_n = R_1 + \dots + R_n$ then

$$\text{Var}(S_n) = \sum_{k=1}^n \sigma_k^2 + 2 \sum_{j,k} \text{Cov}(R_j, R_k)$$

where the last term contains the sum of each of the $\binom{n}{2}$ pairs R_j, R_k with $j < k$ once and only once. In the models considered here the authors—to avoid problems of simultaneity—assume that the residuals of different equations are independent of one another. Therefore, the variance of residuals that one needs in order to obtain a band forecast for, say, the import-export gap, is:

$$\text{Var}(S_n) = \sum_{k=1}^n \sigma_k^2,$$

where k refers to the residuals of each of the equations that goes into the reduced form of the gap. For the author's Kenya model this pooled $\text{Var}(S_n)$ comes to £756.9 million—the standard deviation, that is, is around £27.5 million.² As the model projects the gross external gap to be around £20 million to £25 million a year throughout the entire projection period, an interval forecast based on this wide band is as likely to be correct as it is unhelpful.

In the specific context of the conventional two-gap models that are so

¹ To cite but one instance, in Jere Behrman and L. R. Klein, *Econometric Growth Models for the Developing Economy*, Wharton School, University of Pennsylvania (mimeographed), 1968, by the end of the projection period some of the independent variables increase to thirty-eight times their values at the end of the sample period.

² This estimate makes no provision for the many variables in the balance of payments and the debt routine which are exogenously specified in the model.

widely used now the moral of the foregoing discussion might seem to be that if one is primarily interested in forecasting magnitudes like the import-export and the investment-savings gaps one should directly estimate the reduced forms for the two gaps from the sample period data. If one attempts to do so then, clearly, one would have to be even less ambitious while choosing the explanatory variables than the simple models being proposed at present. The number of variables that enter the reduced form for any one of the gaps in these models is often quite large and the number of observations is often so small that there are few degrees of freedom left to estimate the coefficients. Moreover, in attempting to estimate the reduced forms directly there is the immediate difficulty that in the (*ex post*) sample period data the savings-investment and the import-export gaps will be necessarily identical. Therefore, the only way of obtaining estimates of equations that would yield independent (i.e. not identical) projections of the two gaps for the future will be to specify the model in a way that is very different from the manner in which two-gap models are being specified at present. One may, for instance, fit the reduced forms for the two gaps using different sets of explanatory variables. Alternatively, instead of having different lists of explanatory variables for the two gaps, one could specify the individual equations in the same way as they are specified at present but in addition specify some adjustment procedures that would furnish independent estimates of the two gaps. The difficulty here would be, as was noted towards the end of Section III, that as there are many ways of specifying the equations and reconciling the gaps the forecasts may well turn out to be quite sensitive to the particular specification or reconciliation procedure adopted.

It has been stressed above that these simple models are being used for making projections for periods far beyond their sample periods. Many statisticians who have worked with regressions and seen their forecasts go astray will express serious doubts as to whether regression equations can be used at all for values of variables that are two to five times (to say nothing of variables that are thirty-eight times) beyond the range of values covered by the sample observations. Using these equations implies the heroic assumption that the interrelationships of the system in the future will be as they have been in the past or that structural change in the future will follow the pattern of whatever change occurred in the sample period when, in fact, the whole effort of the governments of these countries and others connected with their development is to bring about a structural transformation in the economies. Nor is it the case that when information about likely structural changes becomes available it is customarily incorporated in the model. The fact is that in most of the models equations are being used in a fairly mechanical manner for projections five or ten years

into the future. But even if the analyst using a very aggregative econometric model sought out and obtained specific information about a likely structural change in, say, the manufacturing sector, how would he incorporate it into his aggregated model? The models in use are at such a level of aggregation that if one were told, for instance, that in Kenya canning of fruit and vegetables and leather and fur processing are going to grow at twice, sugar, spinning, and weaving and pulp and paper at about two and a half times, and rubber at four times the rate of growth of the manufacturing sector as a whole or that Kenya is going in for a large fertilizer plant using by-products of its refinery as feedstock one would not be able to make full use of this kind of specific information. All one could do would be to lump it all together to see what effect it has on the overall growth rate of the manufacturing sector or of the economy and work with that. If by contrast the model had concentrated on incorporating a greater amount of production detail it would have been able to utilize this kind of specific information effectively. And it need hardly be stressed that information about the future when it becomes available almost invariably comes in this specific form; seldom is there a revelation about the overall marginal elasticity of consumer imports or the overall capital-output ratio in a sector.

The position is thus as follows: equations are being used to make forecasts for periods well beyond the range of observation points covered in the sample period; the only information usually obtainable about those distant periods relates to specific and particular developments; and the models and equations are at such a level of aggregation that they cannot use this information. Interval forecasts, though far preferable to point forecasts, do not help partly because of the practical difficulties outlined above and partly because the structural change that makes point forecasts undependable also makes the information on which interval estimates are based obsolete. For both sets of forecasts are based on the same information.

VII. The moral of the story

The argument of this paper has not been against quantification or sophistication in economic analysis nor against economic models in general. It has focused on a particular type of model that is being used for analysis in developing countries: the highly aggregated macro-economic regression model. Over the last few years models of this kind have become increasingly common and are being used for medium- and long-term forecasts and for policy prescription. The apparent faith in these models is unwarranted.

As the sorts of models we have been considering are being constructed for one country at a time they necessarily have to rely on time series data. This data is generally available for only a few years, and is of uncertain or very low reliability. As the time series are very short and as almost all

variables seem to grow or decay over the short periods for which data about them can be assembled, every attempt to include in the equations all the independent variables that appear to be relevant on *a priori* grounds soon exhausts the degrees of freedom and is itself frustrated by the prevalence of collinearities. For this and other reasons there is every likelihood that the models and their equations are grossly mis-specified. Moreover, in a large number of cases one can expect little help from economic theory, from mechanically applied statistical tests, or from appeals to intuition.

Any one of these obvious objections taken individually is sufficient to dissuade one from taking these macro-models at all seriously. In combination their effect can fairly be described as devastating.

It is often said that it is safe to use these grossly under-specified equations and models provided one is aware of their limitations.¹ This is no more than a ritualistic incantation. For in practice one either uses a model, an equation, or a coefficient, or one does not. How is awareness of their limitations usable in a concrete way when a model often involves a number of equations, when little is known about the errors in the data, and when little has been done to experiment with alternative specifications?

Before constructing an econometric model and perhaps again after having constructed it, one should regain one's perspective by rereading some of the iconoclastic literature on regression and econometric analysis (for example writings of Yule,² Box,³ Geary,⁴ and others) and on problems presented by the inaccuracy of basic data. National accounts and the manner in which they are put together in a country deserve more attention than they usually receive at present. A plea should also be made for paying greater attention to textbooks that so often warn us against doing so many of the things we do. A classroom is not their only proper place.

International Bank for Reconstruction and Development, Washington, D.C.

¹ 'The drawbacks to the use of capital-output ratios are well known. . . . Despite these drawbacks, the capital-output ratio is often heavily relied on as a tool of analysis. If enough care is exercised in its application, and particularly if separate estimates can be made for individual sectors, it should be possible to indicate the broad order of magnitude of investment requirements for particular rates of growth', *UNCTAD, 1968*, p. 13.

² *Op. cit.*

³ G. E. P. Box, 'Use and abuse of regression', *Technometrics*, vol. 8, No. 4, Nov. 1966, pp. 625-9.

⁴ R. C. Geary, 'Some remarks about relations between stochastic variables: a discussion document', *Review of the International Statistical Institute*, vol. 31, No. 2, 1963, pp. 163-81.

WORLD BANK REPRINT SERIES: Number Five

Arun Shourie

**Growth, Poverty
and Inequalities**

INTERNATIONAL BANK FOR
RECONSTRUCTION AND DEVELOPMENT
1818 H Street, N.W., Washington, D.C. 20433 U.S.A.
Telephone number: EXecutive 3-6360
Cable address: INTBAFRAD

Reprinted from Foreign Affairs
An American Quarterly Review, January 1973

GROWTH, POVERTY AND INEQUALITIES

By Arun Shourie

Reprinted From

FOREIGN
AFFAIRS

AN AMERICAN QUARTERLY REVIEW



JANUARY 1973

FOREIGN AFFAIRS

AN AMERICAN QUARTERLY REVIEW



JANUARY 1973

ON POWER

- The Relevance of Traditional Strategy . . . *Michael Howard* 253
Can Nuclear Deterrence Last Out the Century? . . . *Fred Charles Iklé* 267
The Changing Essence of Power *Seyom Brown* 286
Who After Mao? *Mark Gayn* 300
Japan's Role in a New World Order *Kei Wakaizumi* 310
Untying Thailand *Maynard Parker* 327
Growth, Poverty and Inequalities *Arun Shourie* 340
- ## WESTERN EUROPE
- Stuck Fast *John Newhouse* 353
Why the Malaise? *Edward L. Morse* 367
America's Move *Benjamin S. Rosenthal* 380
One Nigeria *Jean Herskovits* 392
Law and National Security *John Norton Moore* 408
Recent Books on International Relations *Elizabeth H. Bryant* 422
Source Material *Donald Wasson* 440

WILLIAM P. BUNDY
Editor

JAMES CHACE
Managing Editor

ELIZABETH H. BRYANT
Book Editor

JENNIFER SEYMOUR WHITAKER
Assistant Editor

Editorial Advisory Board

A. DOAK BARNETT
JAMES H. BILLINGTON
HARVEY BROOKS

CARL KAYSER
HENRY A. KISSINGER
WILLIAM L. LANGER

JOHN J. McCLOY
HARRY C. McPHERSON, JR.
WILLIAM M. ROTH

Published quarterly by Council on Foreign Relations, Inc. Editorial and Business Offices, 58 East 68 Street, New York, N.Y. 10021. All mail, including correspondence about subscriptions, should be sent to this address. Cable address Foraffairs, New York. Subscriptions, \$10.00 a year, post-free to any address. The Editors will consider manuscripts submitted, but assume no responsibility regarding them. Reprints of articles are available for purchase. Printed at 10 McGovern Avenue, Lancaster, Pennsylvania.

Vol. 51, No. 2. Copyright 1972, Council on Foreign Relations, Inc. Printed in the U.S.A.

GROWTH, POVERTY AND INEQUALITIES

By *Arun Shourie*

GROWTH is a besieged deity. An increasing number of economists and policy-makers are becoming convinced that it is imprudent for a country to devote all its efforts toward maximizing the rates of overall growth—and wait for the benefits to trickle down to all sections of the population. Trickle-downism is thus on the wane. Developing countries are now being warned that rapid growth is liable to take too long to alleviate the miseries of the poor, and that for long periods rapid growth may indeed worsen the lot of large numbers—hence they should launch “direct attacks” on poverty.

This new advice rests on two assumptions: first, that developing countries have indeed been preoccupied with growth to the exclusion of distributive questions and that it is this single-minded pursuit of overall growth which has kept the governments from undertaking programs that might have benefited the poor; second, that “direct attacks” on poverty—consisting of special programs to improve health, sanitation and nutrition, and to provide drinking water, low-cost housing and the like to larger numbers along with large-scale rural public works that provide jobs and convert local labor and materials into productive assets—that these will alter the character of growth and ensure that henceforth its nature is such as to benefit the poor. Neither of these premises is valid. I will argue that the reasons why growth has been such as to have primarily benefited the well-to-do are much more fundamental and that direct forays against poverty—even if they consist of large and expensive programs—will not alter the character of growth unless they are accompanied by a thoroughgoing political transformation involving drastic redistribution of income, assets and the control over institutions. Finally, I shall argue that such a transformation would make growth more meaningful and not decelerate the rates of overall expansion.

II

Leading Pakistani economists who were closely associated with the high rates of growth of the Ayub period and Western advisers who were in Pakistan in the early sixties are now saying

that their determination to achieve high growth rates was misplaced and that their euphoria over the rates which were achieved misled them into believing that the country was well on the way toward solving its economic problems. The dramatic events of 1971 that broke Pakistan asunder have attracted much attention to these laments.

Consider, however, the counter-examples of Ceylon and India. Decision-makers in these countries have never been mesmerized by growth rates. Ceylon has been engaged in a costly, sustained and very "direct" attack on poverty for the last two decades. Successive governments have been providing subsidized health, education, transport services and food, with little regard for the effects of these subsidies on the country's ability to grow, on its ability to adjust to the declining prospects of its principal exports and on its foreign indebtedness. Yet economic discontent—especially among the educated youth—erupted in large-scale violence two years ago, and today the country's economic policies are stalemated. With neither growth nor structural transformation, and with a severe problem of short-term indebtedness, Ceylon is finding it extremely difficult to continue its direct attacks on poverty.

In the last 25 years Indian leaders—Nehru in particular—have displayed a lively concern for the poor. There is scarcely a proposal for channeling a larger proportion of the benefits of growth to the poor that has not been enacted and for which institutions, procedures and controls have not been devised. Special programs have been formulated to help the dispossessed tribes and castes; numerous laws have been enacted to redistribute land and to reform tenurial arrangements; the controls that have been instituted to ensure that resources are used only for setting up industries of high priority have been so comprehensive that—quite literally—not a single industrial unit (other than the purely household enterprise) has been established and operated without government approvals and sanctions at every step; corporate laws have been codified and tightened; inheritance, wealth and gift taxes have been introduced and the rates of income tax have been made steeper and steeper. A legion of commissions and committees have often been in session examining the effects of the country's economic policies on people in different income groups and in different regions. Yet today official documents tell us that the number of people living in abject pov-

erty has not been diminished in spite of 25 years of development effort—that still 200 to 250 million subsist on a diet that is deficient even in calories—and that the controls and programs which were instituted to channel benefits to the poor have in fact ensured them for the rich and the powerful.

These counter-examples document the obduracy of economic forces and suggest that before the debate between growth and distributive justice becomes more strident we must correctly summarize the experience of the last two decades. One danger is that sloganeering about equity will choke off all growth: a nervous self-consciousness arising from the apprehension that they may be accused of being obsessed with growth—that their progressive *bona fides* may be called into question—may paralyze governments and keep them from taking decisions that are needed for higher productivity. Indian industrial policies have been paralyzed in precisely this way for the past two years. The equally grave danger is that some governments will henceforth embellish their plans with a few programs attacking poverty directly—and then sit back with the feeling that indeed they have taken the necessary steps to alter the character of growth. Valuable time can thus be lost in both ways: by decisions that are deferred and by decisions that place their faith in a few ornamental programs.

III

The experience of the last two decades teaches us that the distribution of income, assets and influence over institutions determines the character of subsequent growth. Given the extremely unequal distribution of incomes and assets, industries are set up to produce goods which the rich want, while import policies aim either at importing the materials needed for producing these goods or at importing the goods themselves. This point can be documented by examining the kinds of consumer industries that have been established in India, Pakistan and Ceylon over the last two decades; it is even more dramatically brought out by considering the pattern of growth in countries such as Brazil, Colombia, Peru and Rhodesia where the richest five percent receive 38, 40, 48 and 60 percent respectively of total income. It is hardly surprising that industries do not get established for meeting the needs of the poorest 20 percent who, in the four countries listed above, receive only 3.5, 2.2, 4.0 and 4.0 percent of total income.

Indeed, the extent to which the richest five percent affect the pattern of what is to be produced in the manufacturing sector—the sector that often attracts the most enterprising and which directly or indirectly absorbs a very large share of the government's promotional efforts and resources—is even greater than their high shares in total income. The poorer sections spend a much larger proportion of their incomes on basic necessities like foodstuffs and coarse clothing than do the richer sections. The share of the latter, therefore, in the market for industrial goods is generally 10 to 20 percent higher than their share in total income.¹

Controls over the establishment and operation of industries and over the distribution of resources that are in short supply (such as foreign exchange), or efforts to present projects meeting the needs of the poor in a more attractive light by, for example, shuffling shadow prices—none of these is able to counter the overwhelming pressures of the marketplace. Private entrepreneurs do not set up units to produce goods that the poor need, but instead offer goods which they cannot afford. And if governments undertake to supply these goods on the assumption that they must accord a greater weight to the needs of the poor than the poor are able to exercise through the market, they soon run into resource problems of the kind that have plagued Ceylonese governments.

This lesson has an important corollary. The longer redistributive steps are delayed the larger the proportion of resources—enterprise, skilled manpower, capital equipment, foreign exchange—that will have congealed in industries catering to the demands of the rich, and the more difficult it will be to reorient the pattern of industrial production. Efforts to reorient this pattern forcibly will result in much unutilized equipment and major dislocations of labor. The economy is not akin to a simple and versatile machine like a lathe that can be quickly made to change from boring holes to chiseling rods just by turning a few knobs and replacing one part or two by another. Rather it is like an educational system which has been patterned over a few decades

¹The rich industrialist and the decision-maker in poor countries often ask "Which of the things that we are producing is particularly inessential or luxurious?" As their consumption aspirations are influenced by the consumption standards of the developed countries they find it difficult to identify goods as luxurious or superfluous. It is the wrong question, put to the wrong people. Instead, we should ask the very poor about what *they* need or what *they* would want in the foreseeable future. Once this is done we at once realize how much of the country's total investment effort has gone into goods that are so very distant from the horizons of the poor and so very irrelevant to their needs.

on the system of a former colonial power. Decades will pass before the system is transformed, the teachers are retrained, new syllabi and texts are developed, new schemes of scholarships and stipends are devised, new facilities are acquired and new examination methods are developed.

Apart from the fact that they do not command much purchasing power, the poor are numerous, dispersed and unorganized. Because of these features many conventional measures are of little avail. Minimum-wage legislation is an example: it does not help the unemployed and may, in certain situations, actually swell their ranks; even for those who have a job off-and-on, legislating a minimum for the wage that they receive on the days on which they are employed is little help because the number of days for which they have a job is not guaranteed; nor can such legislation be enforced in the myriad fields and small establishments in which the vast majority of workers are employed. Moreover, as they are dispersed and unorganized and have little to fall back on, the poor can seldom agitate and bargain collectively—as industrial laborers, who are in the richest third of the urban population, can.

More far-reaching measures than minimum-wage legislation—measures such as land reforms—also fail primarily due to the fact that the poor are unorganized and dispersed. I shall note later on that the laws attempting to reform land tenure and ownership have been flawed. The basic reason, however, why land reforms in many countries have not fulfilled their original promise has been their paternalistic character. The expectation has been that the patterns of land ownership and operation can be transformed by passing laws in national and state capitals, by relying on itinerant officers to implement them and by making the reforms justiciable. In brief, the expectation has been that outsiders—the distant legislator, the itinerant officer, the scholarly judge—can transform the most important feature of village life, the ownership and operation of land. But how are such men to monitor every lease, to examine the precise relationship between each sharecropper and his landlord, to ascertain the extent to which each landlord actually tills his own land? And how is a lowly tenant—who must continue to live in the village long after the officer has driven off; whose lease, in all probability, is an oral one; who is, most likely, illiterate, weak and indebted to his landlord—to gather written evidence, muster witnesses and accuse

his landlord in an open court that passes all the Western norms of justiciability?

In a milieu of this kind—in which the poor are desperately poor, in which they are dispersed and unorganized—almost everything a government does ultimately ends up benefiting the rich and powerful more than the poor and the weak. Credit agencies that are set up ostensibly to free the poor tenant and small farmer from the rapacious hold of the money-lender end up channeling the majority of their funds to the large landlords—who themselves often combine the roles of landlord, trader and money-lender, who are often in a position to finance their own investments, and who, therefore, loan the extra (official) funds that they receive to those who do not have equal access to institutional credit and must, for that reason, pay higher interest rates. Canals that are dug to provide water for all end up providing to the large and strong a much more adequate and reliable service than to the small and weak. State trading corporations that are set up to wipe away the profits of intermediaries end up working through agents who, it later turns out, are none other than the traders these agencies were meant to replace. Public sector enterprises that are established with the hoary objective of ensuring the commanding heights of the economy for the state end up working inefficiently and thus ensuring even higher profits for private units in the same industry that manage to work efficiently. And so on.²

Thus, the pattern of growth can only be changed if the distribution of income, of assets and of the control over institutions is drastically altered. And here a blunt truth must be faced: a drastic alteration of this kind can only be brought about by a government that derives its strength from the dispossessed. A leadership that exists on the sufferance of a privileged group such as the army, or of the rural organization which is in the hands of the kulaks or large landowners, or which is dependent for its finances on the very industrialists and urban rich whose privileges it denounces—such a leadership can do little.

Political change will have to be accompanied on the purely

² Each of these assertions can be amply documented from Indian experience. For perceptive accounts of two field trips that illustrate the situation for water and credit, see Daniel Thorner, "The Weak and the Strong on the Sarda Canal" in "Land and Labor in India," by Daniel and Alice Thorner, eds. New York: Asia Publishing House, 1962, p. 14-20; and "Agricultural Cooperatives in India: A Field Report," by Daniel Thorner. New York: Asia Publishing House, 1964.

economic plane by innovations of many kinds, particularly new forms of organization and new types of financial assets. The need for the former is illustrated by the sorry state of government-owned enterprises in countries such as India and Ceylon. Unless these can be made to function well, industrial progress—including the production of goods that are needed by the poor—will be thwarted. State-owned enterprises with enormous amounts of capital locked up in them will remain riddled by poor management, labor-management duels, inter-union rivalry and unproductive relations between them and their supervising ministries, while large private industrialists—who alone in the private sector can now muster the resources required by plants producing fertilizers, steel, aluminum and so on—are hamstrung, not allowed to establish or expand their enterprises in the name of reducing inequalities and preventing the concentration of economic power.

In the interim period during which steps have been initiated to curb the incomes of the rich through, for example, ceilings on urban property and the like, the need for new forms of financial assets is also obvious. Unless new financial assets are created, for example, government paper that provides an adequate hedge against inflation, the rich in countries such as India and Ceylon will begin to transform their wealth into foreign accounts and into assets (such as diamonds) that can be easily concealed and transported. Such developments can only have deleterious effects on the economies of these countries.

IV

If such a political and economic transformation is indeed brought about, if income, assets and the control over institutions are indeed redistributed, will all this retard growth? In the literature and public discussion on growth one comes across three apprehensions about the likely effects of egalitarian policies: (a) that these will divert scarce organizational, administrative and other talent—the talent and time of those who will be trying to implement the policies as well as of those who will be engaged in circumventing them; (b) that the average savings rate of the economy will be lowered, partly because the rich, who account for most of the economy's household savings, will not be able or willing to save as much, and partly because the government's budgetary resources will be frittered away in subsidies as they

have been, for example, in Ceylon; and (c) that incentives to work and invest will be reduced.³

The first apprehension—that egalitarian measures may divert scarce administrative and entrepreneurial talent—seems incontrovertible when one considers the amount of time that hordes of officials seem to spend implementing laws and regulations intended to redress the balance between the rich and the poor. Thus, for instance, many observers in India have been so discouraged by the Indian experience in land reforms—by the apparent time and effort that have gone into them and by what has been accomplished—that they are saying that instead of expending its time and resources on such far-reaching measures as ceilings, the government should focus on issues like regulation of rents and tenancy reform. The fact, however, is that disproportionate administrative and other effort has not been taken up by these measures either because they are egalitarian or because they are far-reaching. On the contrary, such effort as has been put into them has been made necessary (and rendered futile) by the exceptions, exemptions and loopholes that have been incorporated into the laws. The spate of land reform legislation enacted by Indian states and in Pakistan over the years constitutes an excellent illustration of this fact. If tenancy laws are enacted which exclude sharecroppers from the definition of “tenants;” if laws intended to confer occupancy rights on tenants insist that the tenant can acquire these rights only after 12 years of uninterrupted tenancy; if laws limiting rents are passed without any provisions for security of tenure; if landowners are allowed to evict tenants and resume land for “personal cultivation” and the latter is defined so loosely that the owner need not even stay in the village to establish his status as one cultivating his own land; if tenancy laws are passed that explicitly exclude new tenancies from their coverage; if ceilings apply to individuals and not households; if the intention to impose ceilings is announced long before they are enacted and if generous grace periods are allowed in which landlords may dispose of their land in any way; if such laws are passed and if no effort is made to organize or mobilize those who will benefit from the legislation—then obviously much administrative effort will have to be set aside for

³ Mr. McNamara has commented briefly on the second and third apprehensions in his recent address to the Governors of the World Bank. See Robert S. McNamara, “Address to the Board of Governors,” International Bank for Reconstruction and Development, September 25, 1972.

implementing the laws. And those who might lose under the laws will devote at least as much effort to taking advantage of the loopholes that have been so generously provided as in busying themselves with productive tasks. But clearly there is nothing inherent in egalitarian measures that requires such inordinate diversion of administrative and other resources.

As for the savings apprehension, its arithmetic is clear enough: household savings account for 60 percent to nearly 100 percent of the gross domestic savings in different developing countries and the rich are responsible for most of the household sector's saving. Household savings surveys in different countries suggest that the rate of saving rises very rapidly with income: while the poorest households have negative savings, the richest five percent turn out to save 40 to 50 percent of their income. The conclusion seems simple and straightforward—until one examines the true nature of “savings” and then of incentives.

In surveys from which such results are borrowed, the concept of “savings” includes the acquisition of assets such as houses, gold, silver and jewelry, and consumer durables (such as air-conditioners, cars, radios, furniture, etc.). About 30 percent of household “savings” identified in these surveys consists of assets such as these three.⁴ Glorifying this kind of “saving” as being necessary for growth is, as Gunnar Myrdal has so often emphasized, to become a prisoner of words. Indeed much of the “saving” of the rich is just “consumption”—and consumption of goods, like smuggled gold, that are a direct drain on the economy's balance of payments and of goods, like some of the consumer durables, that have a high import content and the production of which requires much capital. On the other side—if the potential effects on output are any criterion—much of the “consumption” that the poor would undertake if some income were to be redistributed in their favor, and which we are told is undesirable because it will diminish the rate of savings in the economy, is more akin to “investment” than to consumption. More food, better health services, modest amounts of protection against the elements, better public transport—all these will improve productivity, reduce absenteeism and improve labor discipline.

⁴ The reader will notice here a case of double counting: elsewhere I have noted that a large part of the consumption of the rich consists of goods that are capital- and import-intensive. Yet here many of these appear as a part of “savings.” The discrepancy is deliberate, for the “savings” concept referred to in the text does in fact characterize many surveys and thus underlies the inferences that are drawn from them.

The definition of "savings" apart, the savings apprehension must be examined in the context of the character of growth in any economy.⁵ Those who fear a fall in the rate of savings reason as follows: growth requires investment; net inflows from the rest of the world apart, investment requires savings; to obtain savings we must retain inequalities and ensure a high and secure rate of return to savings. Moreover, they argue that the rich will not save unless they are able to enjoy the incomes from such savings in the form of luxury consumption; hence we must not stifle the consumption of such goods and services in the economy.

Two points should be noted in this context. First, the goods (air-conditioners, private automobiles, etc.) and services (foreign travel, accounts in foreign banks, etc.) that the rich demand are capital-intensive as well as import-intensive. As capital and foreign exchange are both scarce, this strategy—of inducing the rich to save by offering luxury consumption standards in return—may well be an inefficient way of maximizing the growth of employment and, indeed, of output itself.⁶ The second point is more fundamental: whom does the growth that results from this strategy benefit? The rich provide not only the savings but also command a disproportionate amount of total effective demand in the economy. Consequently, as was emphasized earlier in the article, their tastes and requirements determine the pattern of production; thus the growth that results from their savings is directed primarily at catering to their demands.

⁵ For reasons of space, I shall not deal with the question of the effects egalitarian measures may have on public savings beyond noting the following points: (a) subsidies—such as the ones used in Ceylon—are certainly not necessary for pursuing egalitarian and employment-oriented goals; if the laws are unambiguous, land reforms, ceilings on the ownership of urban property, alteration in the relative prices of labor, capital and foreign exchange, etc., require minimal public expenditure; (b) these (*sic*) measures may well increase the total levels of output and hence, if the fiscal system is at all responsive to aggregate production, the government's revenue receipts.

⁶ This point is made forcefully by Dudley Seers and his associates in "Towards Full Employment, A Programme For Colombia," International Labour Office, Geneva, 1970—especially Chapter 10. Some of the luxuries that the rich consume are labor-intensive: handicrafts of various kinds (jewelry, carpets, hand-embroidered textiles) and services of domestic servants are cases that are often cited. Each of these must be examined in itself for, apart from their apparent labor-intensity, they have little in common. Jewelry and many kinds of textiles are highly import-intensive. The artisan and laborer employed in many of these labor-intensive industries—carpet-weaving is a case in point—are notoriously underpaid; intermediaries and financiers, who are by no means poor, appropriate disproportionate shares of the price of the product. The real incomes of domestic servants and the numerous attendants on estates place them among the richest 30 to 40 percent of the population. Moreover, in considering these cases we must not assume that if these goods and services were not being produced those currently employed in producing them would be unemployed. A different set of goods would be demanded and labor needed to produce these.

As for incentives, the first question to ask is the following one: how many times must the incomes of the rich be of the incomes of the poor for the former to have adequate incentives? Such figures as are available (and which are known to understate by a considerable margin the incomes of the richest in a country) indicate that for countries with per capita incomes of less than \$200, average incomes among the richest five percent are anywhere from 10 to 35 times those of the poorest 20 percent in each country. The corresponding multiples for countries with per capita incomes between \$200 and \$350 range from 13 times to 73 times with an average around 31 times. Yet the corresponding ratio for the United States is said to be around 5:1, while visitors to China report that the highest salaries in an enterprise are seldom more than three to four times the lowest. These data with their related evidence on savings rates, growth rates and the like hardly suggest that high multiples have been conspicuously successful in inducing the rich to exert themselves.

Moreover, may it not be that relatively high incomes and high consumption standards for the richest constitute a disincentive for the masses? May they not convince the masses that the system does not reward their labor and sacrifices proportionately? And may this conviction not ultimately incite the masses to destroy icons like nonviolence, parliamentary methods and constitutionalism?

The second point is equally fundamental. In the end it is hardly satisfactory to justify high incomes and high consumption standards for the few by asserting that they provide the few with incentives to engineer a pattern of growth which ensures high incomes and high consumption standards for the few.

V

The last two decades thus teach us that the poor will not obtain a proportionate share in the benefits of growth until the character of growth is radically altered, that this alteration can only be brought about as part of a major political transformation and that were such a transformation to be brought about there is little reason to suppose that it would affect the rates of overall growth adversely. The conclusion that the character of growth cannot be altered unless the political milieu is thoroughly transformed is not a counsel of despair. It does not imply that the developing world must sit back for years and wait for distant

political changes. In many countries the leadership can initiate the transformation now.

India is an outstanding case in point. Mrs. Gandhi's unique rapport with the Indian masses puts her in a position to transform her party into a party of the masses—a party that mobilizes the dispossessed so that they may acquire the self-confidence to assert their rights and the capacity to police the reforms, a party whose members work with the masses in constructive programs, a party that is no longer dependent on money raked in from the very rich, a party whose dedication to the interests of the masses is proven by the life-styles and the social practice of its members. The last is not to be dismissed lightly; on the contrary, it is the logical place to begin. So many in India are so desperately poor that very large resources must be deployed if the government is to effect any perceptible improvement in their lives. The resources that are needed are so large that even those who are not well-to-do by any absolute standard (for example, factory workers who, as was noted above, are in the richest third of the urban population) must be asked to share the burdens. How will the leadership make groups such as industrial labor accept sacrifices in their modest real incomes when these groups—and indeed the people at large—know that the very leadership that now wails against “the lack of commitment” in the privileged section of the population and which reminds workers that just having a job is a privilege is itself in many ways the most privileged group?

To persuade an alert and well-organized group like industrial labor to accept sacrifices for the very poor, the political leadership in India, Ceylon and other countries must give up its present life-styles—its insistence on living in luxury houses, on traveling in air-conditioned comfort, on receiving imported cars, telephone services, electricity and other amenities at rates far lower than are charged to the people they are meant to represent and serve. It will have to eliminate the impression which now pervades many a country, that while denouncing all sorts of privileges the leaders have feathered their personal nests and those of their relatives by the unabashed use of the privilege *sans* equal, that of dispensing patronage—a privilege that has been greatly enlarged by the spate of controls that has been instituted in most countries over the establishment and operation of industries and the use of scarce resources like foreign exchange.

Unless the leadership in these countries initiates such a transformation little will come of the “direct attacks” on poverty. Governments like Mrs. Gandhi's will have won elections on brave and bold slogans and then locked up the promises of reform in almirahs of resolutions. They will have introduced bills with sweeping and even with revolutionary objectives and then incorporated caveats which nullify the laws completely. They will have proclaimed the necessity of sweeping institutional change and then frittered away valuable time in the theatrics of nationalizing one industry or another. Finally, they will have launched bold programs for helping the poor and then, as they cannot effectively tax the privileged on whom they depend, be unable to raise the resources these programs require.