This paper is prepared for the
Bank's internal use and is not
for publication. The views are
those of the author and not
necessarily those of the Bank.

INTERNATIONAL BANK FOR RECONSTRUCTION AND DEVELOPMENT

INTERNATIONAL DEVELOPMENT ASSOCIATION

Economics Department Working Paper No. 75

THE RELEVANCE OF ECONOMETRIC MODELS FOR MEDIUM- AND LONGER-TERM FORECASTS AND POLICY PRESCRIPTION

May 6, 1970

Arun Shourie
Quantitative Techniques and
Analysis Division

THE RELEVANCE OF ECONOMETRIC MODELS FOR MEDIUM- AND LONGER-TERM FORECASTS AND POLICY PRESCRIPTION

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 natural-ly 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 programmes 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

has 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, and also of some additional studies using standard econometric techniques. 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 relevance of econometric models.

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] This study follows the admirable practice of presenting the conclusions, the equations and the data all in one volume. Raul Prebisch in his preface to the volume 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] The author's model for Kenya is designed to assess the consequences of differing sectoral growth rates for the investment-savings gap and the import-export gap separately. The model is typical of its kind in that the length of the sample period (9 years) corresponds to the mode of the lengths of sample periods for the UNCTAD, 1968 group; the length of the projection period (10 years) is the same as that for the UNCTAD models; and the individual equations are also fairly representative. The model uses 22 equations and an iterative debt routine to project the two gaps independently.[3]

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

---

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/ The model itself and the data on which it is based is not presented here so as to conserve space. A writeup of it can be obtained from the author on request.

the models are estimated. But the fact is that reliable and ample data is critical to an econometric model. The objective of econometric techniques is to obtain unbiased, consistent and optimal estimates of individual coefficients and equations. If the underlying data is seriously deficient then all 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 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.

1/ "Certainly it would be a nice world if all data had no error; but lacking such a world, shall we turn back the clock to the 19th century or do we do the best we can with what we have and go ahead from there? We opt for the latter!" 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. Similarly 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.

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.[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 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."[2] Trade data, in fact, constitute the only set of figures for these countries 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 to the returns of the country's trading partners - reveals

---

[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)	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).

[2] 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.

very substantial discrepancies.[1], [2]

The underlying data are such that many refinements in estimation procedures, such as multiple stage least squares, are 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.[3] 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 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.[4] 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/ These deficiencies in the data base make nonsense of the current vogue that classifies and ranks countries for purposes of lending policy or analysis according to "performance criteria" such as the savings and investment rates, historical capital-output ratios, rates of growth of GDP and other magnitudes, or proportion of the Government's budget set aside for development expenditure. For examples of this approach see, A.M. Strout and P.G. Clark, Aid, Performance, Self-Help and Need, U.S. AID Discussion Paper No. 20, Washington, D.C., July, 1969; and UNCTAD, Objectives for the Mobilization of Domestic Resources, Mobilization of Resources for Development, 3 TD/B/C•3/75/ADD•1, New York, February, 1970.

3/ Fourteen of the eighteen models in UNCTAD, 1968 use constant price series.

4/ 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: in fact, in the Kenyan case, as evidence from the Industrial Census of 1967 is being compiled and analyzed, it is becoming clear that figures of capital formation in the manufacturing sector will have to be substantially revised in the near future.

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 percent 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 percent in contrast to the 11.3 percent that emerged from the accounts Faaland and Dahl used. Mechanical projections from the author's Kenya model - a model that in its specification is 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.

---

[1] UNCTAD, 1968, pp. 126, 137-138.

The equations contained in the kind of models we are now consider-
ing 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.
Sometimes one obtains such insights as by-products of the model building
exercise. But clearly they cannot by themselves be a sufficient justifica-
tion 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. The statistical theory of models with errors
in variables is far from reassuring in this context. First, the least
squares model itself is scarcely applicable when the independent variables
have errors in them since they are not then independent of the error term.
Second, the theory offers us few rules of thumb that we can use in each of
the many diverse circumstances in which we are liable to encounter the
"errors in variables" problem. Consider, for instance, the following
question. In what direction are the errors in independent variables liable
to bias our coefficients? Elementary theory and a few numerical experiments

will show that we <u>can</u> make a definite statement about the direction of bias

in the case of a simple regression equation - an equation, that is, which

has only one independent variable: if the errors are random they will bias

the coefficient towards zero. But even this general rule does not hold if

we have two or more independent variables and certainly not if the errors

are not random. Many of the equations we fit - and <u>almost all</u> the equations

we would <u>want</u> to fit if multicollinearities did not paralyze us so complete-

ly - have more than two independent variables, and if the pattern in which

national accounts are revised periodically is any guide the errors in our

independent variables are certainly not random. Moreover, the errors are

as liable to be in the dependent as in the independent variables. Third,

all the methods which have been devised to allow for errors in variables

require a good deal of information about the nature of the errors. Since

more often than not one does not know the possible extent or even the

direction of errors in estimation, it is virtually impossible to use any of

these methods to assess the type of bias that errors in some or all of the

variables might be communicating to the final projections or policy recom-

mendations. Introducing large (or still larger) error terms around the

forecasts or the coefficients only confronts us with the problems that are

sketched in Section VI below.

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 coef-

ficients of a fitted regression equation are treated not merely as coeffi-

cients 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 equations:

$$\text{Private consumption expenditure} = -106.04 + \underset{(0.12)}{1.006} \text{ Wage Income} \\ (3.574) \\ + \underset{(6.625)}{0.655} \text{ Non-Wage Income}$$

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 marked-
ly 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 "reason-
able" 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] - a situation in
which the individual coefficients of a multiple regression equation have
little significance. 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 report does
    not furnish data on incomes at constant prices.

To find out whether or not consumption behavior in the two groups is different in a situation where the two types of income are highly collinear we must fit two separate equations relating consumption expenditures of each group to its income and then compare the coefficients.

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] This is precisely the situation in which the bias in the estimation of a coefficient is very considerable.

A further illustration may be taken from a recent study of cost-benefit relationships of education in Kenya.[2] In order to determine the extent to which different factors influence student performance the study fits equations relating examination scores to factors like number of pupils in the school, number of examination candidates in the school, percent repeaters among candidates, average age of candidates, the school's pupil/teacher ratio, the school's expenditure on salaries per pupil, an occasional dummy variable and so on. Having fitted these equations (which account for anywhere from 4 to 18 percent of the variance of examination scores) the report proceeds to draw various inferences from the equations of which the following is representative: "The linear relationship also indicates that expenditure per student would have to be increased by more than K£6 per pupil to affect exam scores by one percentage point. Since the average score among all schools is 1.82, or slightly above average Division II pass, it would be possible to drop expenditure to about K£36 per pupil, or approximately 40 percent of the average in all CSC schools (K£87/pupil) and still achieve an average Division II pass (1.75)."[3]

In another part of the report where regression equations are fitted linking wage, employment and GDP data the report informs us: "The coefficients of sectoral GDP are almost exactly equal for Africans in agriculture and private industry and commerce. For a one percent increase in sectoral GDP, African employment increases by 0.67 percent, salaries being held constant. This implies that real sectoral GDP would have to grow 4.5 percent per annum for these two sectors to absorb a three percent annual increase in the labor force in the two sectors. If sectoral GDP does not grow at 4.5 percent per annum, but the population growth rate remains at the

---

[1] K. Marwah, op. cit., in particular, the discussion of the consumption equations, pp. 234-235.

[2] Hans H. Thias and Martin Carnoy, Cost Benefit Analysis in Education, A Case Study of Kenya, Economics Department, International Bank for Reconstruction and Development, Washington, D.C., July 1969; especially pp. 175-194.

[3] Ibid., p. 188.

three percent level, the two sectors will not be absorbing a proportional share of the labor force - other sectors, such as the government or subsistence agriculture would have to take an increasing proportion of the labor force if unemployment rates were to remain at the same level . . . . As we indicate above, the low GDP elasticities for Africans in our regression results indicate that GDP annual growth must exceed a certain 'break-even' rate (4.5 percent for Africans in agriculture, private industry and commerce) if the expected increase in the labor force is to be absorbed."[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 independent variables 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) The Prevalence of Multicollinearity

The illustration from the UNCTAD model for Argentina indicates that the independent variables of an equation may be highly collinear. As collinearities between independent variables increase, the individual coefficients become increasingly arbitrary; they also become more sensitive to model specification and to the coverage of the sample. In a situation of this kind the individual coefficients of an equation cannot be interpreted

---

1/ Ibid., pp. 224-227.

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

as the partial derivatives of elementary calculus. Thus, the coefficient
(b_2) in the model $(Y = a + b_1X_1 + b_2X_2 + b_3X_3)$ when, say, X_2 and X_3 are
highly correlated, does not indicate "the extent to which Y will change when
X_2 increases by one unit - other things remaining the same." As X_3 varies
with X_2 "other things" are not liable to "remain the same". If by (b_2) we
are to understand "the extent to which Y will change when X_2 changes by one
unit and X_3 takes on the values appropriate to the altered value of X_2" then
the appropriate coefficient is the $(\hat{b}_2)$ in the simple regression equation
$(Y = a + \hat{b}_2X_2)$ and not the (b_2) in a multiple regression equation that
includes X_3.[1] In fact, apart from the case when the explanatory variables
are truly independent of one another the only instance in which a unique
interpretation can be placed on (b_2) and (b_3) is when X_2 and X_3 are almost
perfectly collinear and then also it is the sum of the two that has meaning
rather than the individual coefficients themselves.

Statisticians have sometimes stressed this consequence of multi-
collinearity and yet, to their despair, models containing highly suspect
individual coefficients continue to be used for control purposes. From this
one can only conclude that many model builders feel that in practice multi-
collinearity is not a serious problem and that its consequences can be
handled. It can be argued, for instance, that one can always (i) make sure
that one's coefficients have reasonable values and the correct signs and
that (ii) in practice one can often introduce a priori estimates for some
of the coefficients. The latter expectation can be dismissed summarily:
in spite of the ingenious work that has been done on "the great ratios" in

---

[1] Cf. R.C. Geary, "Some remarks about relations between Stochastic
Variables: a discussion document," Review of International Statistical
Institute, Vol. 31, No. 2, 1963, pp. 163-181; especially pp. 168-171.

economic development and on quantitative patterns in international develop-
ment, there are hardly any coefficients linking highly aggregated variables
that can be trasferred across countries without the utmost care.

The question of coefficients having a reasonable value and the
correct sign deserves somewhat more extended comment. Most economists doing
empirical work have "hunches" of what the sign of a coefficient or its
approximate value should be: "the income elasticity of demand for imports
of consumer durables just cannot be that low" or "the foregoing equation is
obviously unreliable as the sign of the income coefficient is not what
economic theory leads us to expect." There are two difficulties in taking
these hunches seriously.

First is the obvious point that different individuals have dif-
ferent hunches that are held with varying degrees of conviction. Second,
these expectations or opinions are, in any case, based on a more or less un-
systematic examination (possibly with the aid of some earlier inadequately
specified econometric model) of more or less unsatisfactory data. In such
a situation at least as strong a case can be made for not taking someone's
hunch or some rough and ready economic generalization seriously as for
saying that one should reject a coefficient or an equation on the grounds
that the value of the coefficient or even its sign "is not what economic
theory leads us to expect."

This is especially important when our norms for signs or values
are derived from an economic theory developed in the context of mature and
institutionally very different economies. A well-known economic hypothesis,
for instance, links the bond rate inversely to the ratio of liquid assets
to income in the economy; this hypothesis finds wide confirmation in the data

of developed economies like the U.K. When equations utilizing not just one but a large variety of concepts of the liquidity ratio are fitted to the data for India the coefficients turn out to be woefully insignificant and when they are significant the signs are positive.[1] Does one "refuse to believe the equations" on the ground that they do not accord with economic theory? Can one not make an equally strong case for discarding this hypothesis about the bond rate and other similarly simplistic hypotheses regarding other rates? Should one not - as K.N. Raj suggests - look for alternative explanations of interest rates that are more relevant to an underdeveloped economy - couched, for instance, in terms of the returns from inventory holdings or assets like land and gold?

On many of the issues that crop up in the course of devising or examining a model it is very difficult to formulate an unambiguous expectation about the value of the coefficient or the sign. Does inflation encourage or discourage present consumption? 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? What role should be assigned to the capital stock in an investment function? What variables are the appropriate ones for explaining Government consumption expenditures? The econometrician's logic and rationalizations are often quite obscure when he finally decides to opt for one hypothesis rather than another. 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

---

1/ I am grateful to Muslehuddin Shahi for the relevant regressions.

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 invest-

ment 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]

Similarly, so as to find a function determining Government con-
sumption expenditures, de Vries and Liu try a number of explanatory
variables and report three equations:[2]

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

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).[3]

---

[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.

[2] B.A. de Vries and J.C. Liu, op. cit., p. 11.

[3] Ibid.

The authors do not even consider the fact that by choosing the first equation instead of the second or third on these - somewhat ambiguous - grounds they alter the reduced form of their model and the conclusions one may draw from it quite dramatically.

(b) Specification Errors: The Choice of Explanatory Variables

The second and in practice an equally important reason 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 Y, X and Z be matrices and b, β and θ be vectors; assume that the true model is

$$Y = X\beta \qquad\qquad\qquad (i)$$

Least squares procedures estimate β by b where

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

Substituting for Y from (i)

$$b = (X' \ X)^{-1} X' \ (X)\beta \qquad\qquad\qquad (iii)$$

But $(X' \ X)^{-1} (X' \ X) = I$ and so

$$b = \beta \qquad\qquad\qquad (iv)$$

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

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

while the model used for estimation is

$$Y = X\beta$$

that is, if some of the explanatory variables (Z) have been left out because the analyst misconstrued the causal relationships or because he could not obtain satisfactory data for them or for some other reason, then

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

Substituting once again for Y - but this time from (v),

$$b = (X' X)^{-1} X'(X\beta) + (X' X)^{-1} X'(Z\theta)$$

$$. = \beta + (X' X)^{-1} X'Z\theta \text{ or} \tag{vi}$$

$$b = \beta + \delta\theta \tag{vii}$$

where δ gives X as a function of $Z(X = \delta Z)$. Thus, 'b' can be an under- or over-estimate of β. As the variables in Z have not been included and, as, most probably, one does not have or has not tabulated the information about them it is almost impossible to determine <u>a priori</u> the extent or - in some cases - even the direction of the bias in estimating β. For the extent of this bias depends both on θ and on the variability of Z. 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 <u>not</u> 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 $/Y = f(X)/$ and having the (X) do the work for themselves as well as for the (Z).[1]

---

[1] Though the selection of a single most appropriate set of variables for forecasting is not always a straightforward matter: the step-wise forward and the step-wise backward procedures often yield different answers as to which variables should be included and which should be left out. Efficient programmes are, however, being devised which can get around the difficulties to come up with sets of explanatory variables that are optimal from some points of view. Cf., E.M.L. Beale, M.G. Kendall and D.W. Mann, "The Discarding of Variables in Multivariate Analysis," <u>Biometrika</u>, 1967, Vol. 54, Nos. 3-4, pp. 357-366.

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 any-

how, there is the pervasive difficulty that economic theory offers insuf-

ficient 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 electric-
ity). 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.

Similarly, the cost benefit study of Kenyan education cited earlier
emphasizes that there are many causal factors that could not be included
in the equations - a fact that is also self-evident from the extremely
low proportion of variance that is accounted for by the independent
variables that have in fact been included: "The combined explanatory
power of these variables is only moderate. . . which is hardly surprising
given the large number of educational variables (e.g., turnover of teachers,
distances to school, standards of equipment, etc.) and the still larger
number of socio-economic factors which could not be taken into consider-
ation because of lack of data." Op. cit., p. 176. The report also
stresses that in all probability there are biases in the measurement of
some of the explanatory variables. "On the whole the data appears to be
fairly reliable, with the obvious exception of the share of repeaters and,
possibly, average age of candidates. In the case of Meru Country, in
particular, the understatements must be of major proportions (only 11
out of 41 schools reported any repeaters at all, against 46 out of 48
in Muranga). No allowance can be made for this bias, which might explain
why the percentage of repeaters is not a significant explanator of exam
performance. Average age of candidates, which is probably less biased
downward due to misreporting, was the variable which ranked third in
terms of significance." Op. cit., p. 175.

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.

Most models in use utilize linear equations only. When they do experiment with alternative functional forms they almost always adopt the linear or the log-linear form whichever has the higher coefficient of determination. This preference for linear relationships is rationalized on the grounds that "everything is linear in the small" and that "the only simple mathematics is linear mathematics." As for things being linear in the small the fact is that the kind of models reviewed here are <u>not</u> being used to study the effects of small changes only; typically the explanatory and dependent variables increase manifold during the sample and projection periods. The second reason - about the simplicity of linear relationships - also sounds a bit dubious as most of us now have access to a large variety of computing devices and as only some special types of non-linearities are difficult to handle.

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. And the
reason is that one is using the equation to project magnitudes five or ten
years beyond the end of the sample period when the values of the variables
are likely to be - or are assumed to be - very different from what they were
in the sample period. Consider the equations displayed on the following
page. They have been computed from the Kenyan data for 1960-1968. Each of
the equations yields reasonably good fits for the sample period and the t-
ratios are satisfactory for all of them. But, as the diagrams in <u>Figure 1</u>
show very clearly, they yield widely divergent projections for the future.

The same sort of differences arise when one uses the first deriv-
atives 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.

	D-W	F-Ratio	$\overline{R}^2$

$$(XGOODS)_t = \underset{(.00593)}{.19540} (GDPFAC)_t$$

0.906 30.379 0.792

$$(XGOODS)_t = \underset{(2.15356)}{-0.23133} + \underset{(0.03803)}{0.19607} (GDPFAC)_t$$

0.907 26.584 0.762

$$\log(XGOODS)_t = \underset{(0.01362)}{-0.92168} + \underset{(0.18832)}{1.08349} \log(GDPFAC)_t$$

0.874 33.102 0.801

$$(XGOODS)_t = \underset{(1.35559)}{-213.55198} + \underset{(0.37217)}{1.46463} (GDPFAC)_t - \underset{(.00054)}{0.00183} (GDPFAC)_t^2$$

2.396 39.379 0.906

$$\log(XGOODS)_t = \underset{(0.07526)}{2.29960} - \underset{(24.21042)}{160.32999} \underline{/1/(GDPFAC)_t \underline{/}}$$

1.042 43.856 0.843

$$(MINTER)_t = \underset{(0.01393)}{0.46053} (\textstyle\sum_{2}^{6} GDPFAC)_t$$

1.586 39.486 0.832

$$(MINTER)_t = \underset{(1.28848)}{3.08999} + \underset{(0.07103)}{0.42618} (\textstyle\sum_{2}^{6} GDPFAC)_t$$

1.566 36.000 0.814

$$\log(MINTER)_t = \underset{(0.01364)}{-0.29636} + \underset{(0.15543)}{0.97935} \log(\textstyle\sum_{2}^{6} GDPFAC)_t$$

1.702 39.701 0.829

$$(MINTER)_t = \underset{(0.92077)}{-87.30748} + \underset{(0.74255)}{2.4828} (\textstyle\sum_{2}^{6} GDPFAC)_t - \underset{(.00403)}{0.01120} (\textstyle\sum_{2}^{6} GDPFAC)_t^2$$

3.360 39.101 0.905

$$(MINTER)_t = \underset{(25.11969)}{-134.01597} + \underset{(13.02679)}{90.22789} \log\left[\textstyle\sum_{2}^{6}(GDPFAC)_t\right]$$

1.867 47.974 0.854

$$(DIRTAX)_t = \underset{(.00191)}{0.04847} (GDPFAC)_t$$

0.378 36.997 0.822

$$(DIRTAX)_t = \underset{(0.39871)}{-9.03007} + \underset{(0.00704)}{0.07468} (GDPFAC)_t$$

1.124 112.502 0.933

$$\log(DIRTAX)_t = \underset{(0.00969)}{-2.54335} + \underset{(0.13391)}{1.48174} \log(GDPFAC)_t$$

1.129 122.441 0.938

$$(DIRTAX)_t = \underset{(0.19761)}{34.15114} - \underset{(0.05425)}{0.18211} (GDPFAC)_t + \underset{(0.00008)}{0.00037} (GDPFAC)_t^2$$

3.069 240.242 0.984

$$(MCAPTL)_t = \underset{(.01945)}{.27497} (INVEST)_t$$

0.612 34.920 0.814

$$(MCAPTL)_t = \underset{(1.02273)}{-5.28821} + \underset{(0.05305)}{0.36031} (INVEST)_t$$

0.942 46.125 0.849

$$\log(MCAPTL)_t = \underset{(0.03465)}{-1.37567} + \underset{(0.25389)}{1.44553} \log(INVEST)_t$$

0.866 32.417 0.797

$$(MCAPTL)_t = \underset{(0.96160)}{-27.87714} + \underset{(0.56760)}{1.14340} (INVEST)_t - \underset{(0.00436)}{0.00604} (INVEST)_t^2$$

1.446 27.047 0.867

$$\log(MCAPTL)_t = \underset{(0.05169)}{1.74490} - \underset{(2.49690)}{28.71198} \underline{/1/(INVEST)_t \underline{/}}$$

1.864 132.228 0.943

EFFECTS OF FUNCTIONAL FORM ON PROJECTIONS

EXPORTS OF GOODS: f (VALUE ADDED BY SECTORS 2 TO 6)

$y = a + bx$ (0.762)

Log y = Log $a + b$ Log x (0.801)

$y = ax$ (0.792)

Log $y = a + b$ (1/x) (0.843)

$y = a + bx + cx^2$ (0.906)

-350.10 by 1978

+162.63 by 1978

DIRECT TAXES: f (GDP AT FACTOR COST)

$y = a + bx + cx^2$ (0.984)

Log y = Log $a + b$ Log x (0.938)

$y = a + bx$ (0.933)

$y = ax$ (0.822)

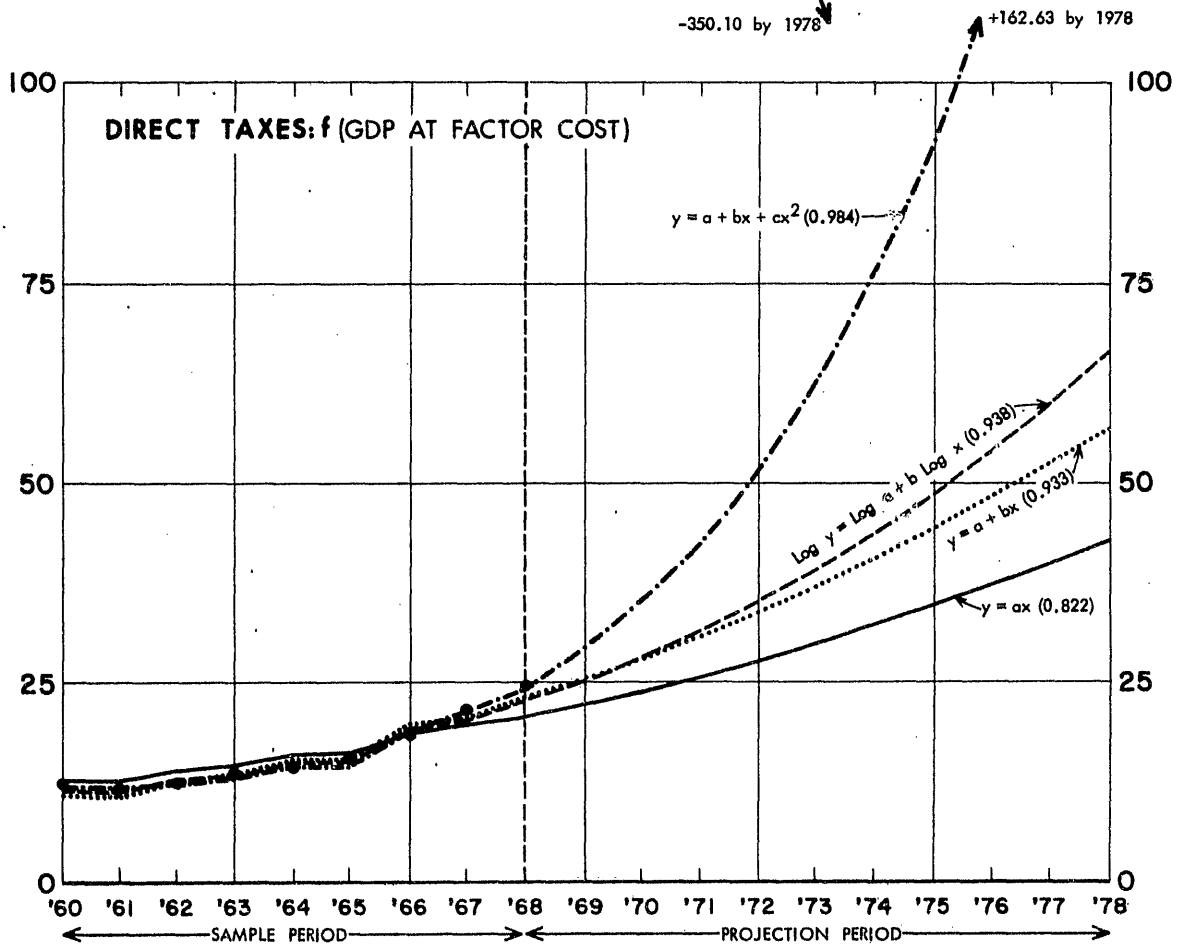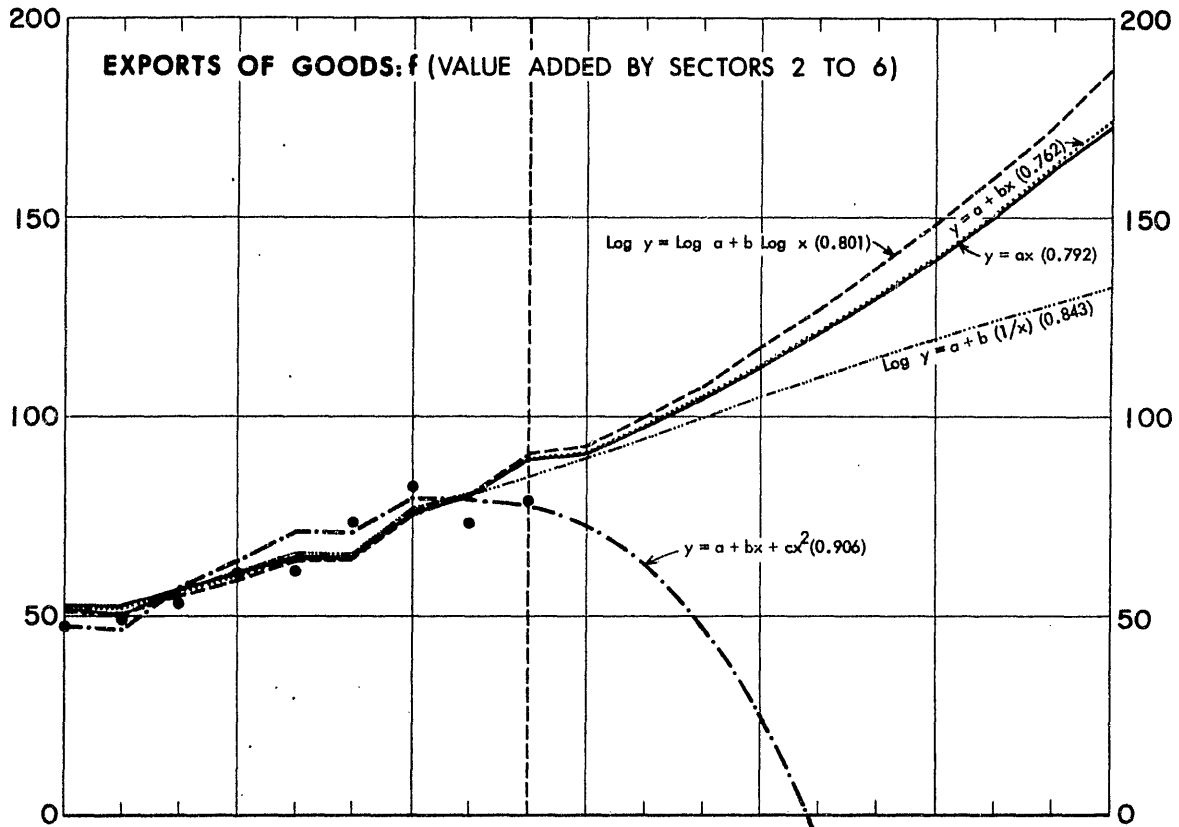SAMPLE PERIOD · PROJECTION PERIOD

FIGURE 1

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

IBRD-4728

EFFECTS OF FUNCTIONAL FORM ON PROJECTIONS

IMPORTS OF INTERMEDIATE GOODS: f (GDP AT FACTOR COST)

$y = ax$ (0.832)

Log y = Log a + b Log x (0.829)

$y = a + bx$ (0.814)

$y = a + b$ Log x (0.854)

$y = a + bx + cx^2$ (0.905)

-368.46 by 1978

IMPORTS OF CAPITAL GOODS: f (FIXED INVESTMENT)

Log y = Log a + b Log x (0.797)

$y = a + bx$ (0.849)

$y = ax$ (0.814)

Log $y = a + b$ (1/x) (0.943)

$y = a + bx + cx^2$ (0.867)
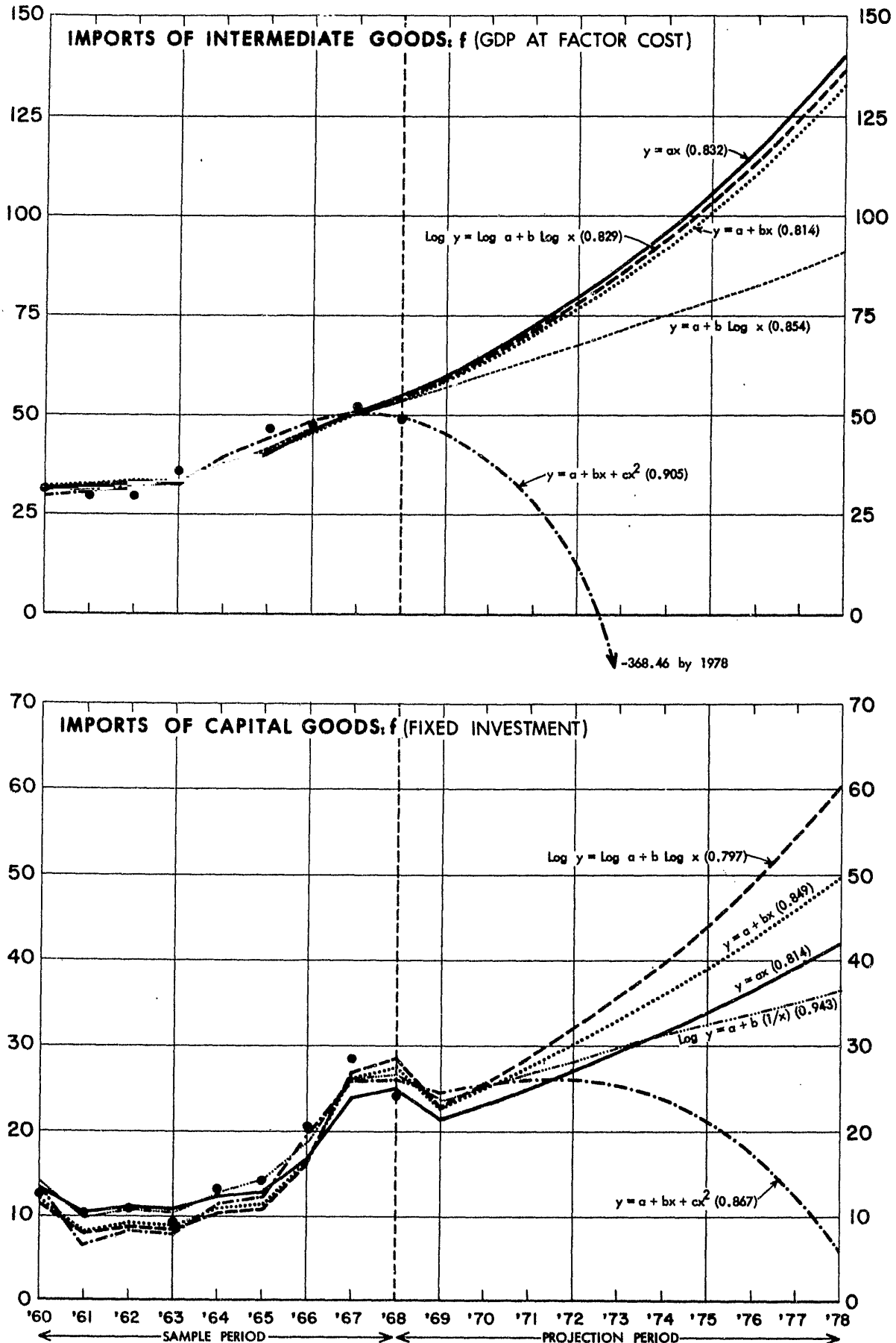
SAMPLE PERIOD PROJECTION PERIOD

FIGURE 1 contd.

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

IBRD–4729

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 <u>even close to the center of gravity of the observed point set</u> 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 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

---

1/ Cf., H. Theil, "Specification Errors and the Estimation of Economic Relationships," <u>Revue de l'Instut International de Statistique</u>, Vol. 25, No. 1/3, 1957, pp. 41-51.

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 un-

fortunately 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:

$$S_9 = 5.55 + 1.63 \ln Y - 0.78 \ln P + e.$$
$$(.53) \qquad (.31). \ldots "[1]$$

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

$$\frac{I_{hH}}{P_{cnt}} = -0.9595 + 0.1312 \ X + 1.208 \left(\frac{I_{mcnt}}{P_{icnt}}\right)_{-1}$$
$$(0.0068) \qquad (0.187)$$

$$\overline{R}^2 = 0.977, \ d = 2.327$$

and

$$\frac{I_{hH}}{P_{cnt}} = -0.311945 + 0.39590 \left[X - X_{-1}\right] + 0.5920 \left(\frac{I_{mcnt}}{P_{icnt}}\right)_{-1}$$
$$(0.11380) \qquad\qquad (0.3550)$$

$$+ 0.94580 \left(\frac{I_{hH}}{P_{cnt}}\right)_{-1}$$
$$(0.09470)$$

$$\overline{R}^2 = 0.912, \ d = 1.590$$

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]

In a superficial sense this preoccupation with the $\overline{R}^2$ is well rewarded. The models are fairly simple and highly aggregated representations of the economies of developing countries. Few of them incorporate any institutional detail or any information about production relationships in

---

[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.

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]

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 con-
siderable 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

---

[1] On a rough count the 18 models in UNCTAD, 1968 between them use 376
estimated equations and the 8 models of ECAFE, 1968 list 119 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
No. of equations								
(i) UNCTAD, 1968	8	5	23	23	42	44	90	141
(ii) ECAFE, 1968	3	6	5	4	5	14	22	60

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

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
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 equations will be used to illustrate some common situations
in which reliance on the $\bar{R}^2$ is definitely misleading.

The equations of most models in use have been estimated from
data that have not been transformed in any way to remove the effects of
auto-correlation in the series. Many of the series have strong trend
components so that when they are regressed one on another the coefficients
of determination turn out to be very impressive for a reason that may or may
not have anything to do with any functional relationship between the
variables. 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 auto-
correlation in the original series.

Philosophically the position is an ambiguous one. For instance, in the case of the author's Kenya model imports and exports of services (excluding tourism) consist mainly of freight, transportation, insurance and similar items. Let us assume that most of these arise from the movement of goods in foreign trade and, therefore, let us postulate that the imports and exports of services are a function of the imports and exports of goods. The 1960-68 data yields the following equations:

	D-W	$\overline{R}^2$	F-Ratio

$(XSERVC)_t = 0.15755(MGOODS + XGOODS)_t$ 0.903   .924   60.603
      (0.00295)

$= -5.16306 + 0.18747(MGOODS + XGOODS)_t$ 1.721   .935   73.388
    (0.465)   (0.022)

$= 9.721 + 0.00431(MGOODS + XGOODS)_t + .00055(MGOODS + XGOODS)_t^2$
  (0.527) (0.53106)   (.00160)   2.105   .912   28.674

$\log(XSERVC)_t = -1.236 + 1.193 \log(MGOODS + XGOODS)_t$ 1.754   .950   95.548
          (0.007) (0.122)

$(MSERVC)_t = .06159(MGOODS + XGOODS)_t$ 2.192   .779   17.614
     (.0017)

$= 0.389 + .059(MGOODS + XGOODS)_t$ 2.178   .725   14.184
    (0.334) (.0157)

$\log(MSERVC)_t = -1.026 + 0.917 \log(MGOODS + XGOODS)_t$ 2.083   .715   13.551
         (.0141)(0.249)

The linear equations fit the raw, untransformed data reasonably well. When first differences are used we get the estimates displayed on the following page. In this case the quadratic equations alone seem to fit the data satisfactorily. One may take either of two positions: it may be that the (bulk) steady component of $(MSERVC)_t$ and $(XSERVC)_t$ _is_ related to the (bulk) steady component of $(MGOODS + XGOODS)_t$ as reflected by the good fits obtained from untransformed data, and that this is what we are essentially interested

	D-W	$\overline{R}^2$	F-Ratio

$\Delta(\text{MSERVC})_t = \underset{(.02443)}{0.06864} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         1.823 0.570 5.308

$= \underset{(.39739)}{-0.47489} + \underset{(.039969)}{0.09205} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         1.503 0.523 5.380

$\log(\Delta\text{MSERVC}) = \underset{(.01780)}{-0.01525} + \underset{(.69360)}{1.30120} \log \underline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         1.554 0.386 3.519

$\Delta(\text{MSERVC}) = \underset{(0.21915)}{-0.87119} + \underset{(0.04164)}{-0.00728} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7} + \underset{(0.00234)}{0.00656} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}^2$

         1.144 0.855 12.777

$\Delta(\text{XSERVC}) = \underset{(0.03750)}{0.14416} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         1.719 0.000 -1.433

$= \underset{(.35354)}{1.51363} + \underset{(0.03531)}{0.06953} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         2.620 0.418 3.878

$\log\Delta(\text{XSERVC}) = \underset{(0.00638)}{0.02209} + \underset{(0.24857)}{0.54963} \log \underline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7}$

         2.497 0.493 4.889

$\Delta(\text{XSERVC}) = \underset{(0.25077)}{1.19174} - \underset{(0.04765)}{0.01116} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})_{t}\underline{7} + \underset{(0.00268)}{0.00533} /\overline{\Delta}(\text{MGOODS} + \text{XGOODS})\underline{7}^2$

         2.161 0.707 5.834

in verifying; or that if $(MSERVC)_t$ and $(XSERVC)_t$ had been truly related to $(MGOODS + XGOODS)_t$ the year-to-year changes in them would have shown a higher correlation.[1] Even though the basic position may be ambiguous the likely effects on one's research strategy are clear: if one uses untransformed data there is every likelihood that high coefficients of determination will make one complacent a bit too early in the game; working with first differences or detrended data one would at the very least be forced to push farther in the search for explanatory variables.

Of course, the fact that the quadratic equation fits the first differences of the data reasonably well does not imply that it reflects any causal factors at work. Consider the foregoing equations once again. 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 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

---

[1] Marschak expresses this ambiguity explicitly: "The point about 'spurious correlation' is not the _existence_ of some common cause but its _unsuspected_ existence. If. . . the existence of the common cause is the very hypothesis we are testing (or a part of it) then there is nothing spurious about the correlation. . . .(Similarly) serial interdependence may, again, be the result of the very relations the economist studies as his hypothesis. For example, if the supply of a given year determines the current price (demand function) but is itself determined by the price of the preceding year (supply function) then there is a relationship holding for any two successive prices. (But) relations of this kind may constitute the very hypothesis the economist wants to test. The parameters of such a theoretical or 'true' relation are the parameters he has to estimate. The attention of economic statisticians has been too often absorbed by the question of 'removing' these relations." J. Marschak, "Economic Interdependence and Statistical Analysis," in _Studies in Mathematical Economics and Econometrics_, ed. by O. Lange, F. McIntyre and T.O. Yntema, University of Chicago, Illinois, 1942, pp. 135-150; the extracts are from page 148.

freight and transport arise more from the imports and exports of <u>Uganda</u> and
<u>Tanzania</u> 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 $(MGOODS + XGOODS)_t$. But, one may
argue, what is wrong in using the equation $(MSERVC)_t = f(MGOODS + XGOODS)_t$?
After all, the variables <u>have</u> 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
$(XSERVC)_t$ as an example, one would, in fact, be on firmer ground if one
asserted that these are likely to grow at a rate <u>lower</u> than the growth of
$(MGOODS + XGOODS)_t$ of Kenya: Uganda's exports and imports together are
expected to grow at a rate about one-half of the rate being currently pro-
jected 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.

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 trans-
formed but to transforming it in such a way as to generate precisely 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:

$$(GDPFAC)_t = f\left[\sum_{i=0}^{t-1} (INVEST)_i\right]$$

where (GDPFAC) is GDP at factor cost and (INVEST) is gross fixed investment.

The function yields extremely impressive fits: the $\bar{R}^2$ being almost invariably

higher than 0.95. The two series (GDPFAC) and (INVEST) have pronounced trend

components in almost all cases; by cumulating (INVEST) 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 percent and X_t fluctuates between 10 and 12.

---

[1] 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.

[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 magnitudes being derived from an identity.

Y_t	X_t	$Z_t \equiv \sum_{i=o}^{t-1} X_i$
10.000	10.0	0.0
11.000	11.0	10.0
12.100	10.0	21.0
13.310	12.0	31.0
14.641	10.0	43.0
16.105	11.0	53.0
17.716	10.0	64.0

In fitting Y_t to X_t we get:

$$Y_t = 14.94085 - 0.13127(X_t)$$
$$(16.72678) \quad (1.57852)$$

$$\overline{R}^2 = 0.000 \qquad DW = 0.222$$

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:

$$Y_t = 9.69366 + 0.11954(Z_t)$$
$$(0.19086) \quad (0.00493)$$

$$\overline{R}^2 = 0.990 \qquad DW = 0.954$$

The very impressive coefficient of determination and the significant t-ratios are due solely to the fact that we 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.[1] We know that gross fixed investment in a sector $(INVEST)_i$ consists of net investment $(INVNET)_i$ and depreciation $(INVDEP)_i$:

$$(INVEST)_i = (INVNET)_i + (INVDEP)_i$$

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

$$(INVDEP)_i = d(kSTOCK) = d \cdot k(GDPFAC)_i = b(GDPFAC)_i$$

Therefore

$$(INVEST)_i = (INVNET)_i + b(GDPFAC)_i$$

Dividing both sides by $\wedge(GDPFAC)_i$ we get

$$\frac{(INVEST)_i}{\wedge(GDPFAC)_i} = \frac{(INVNET)_i}{\wedge(GDPFAC)_i} + \frac{b(GDPFAC)_i}{\wedge(GDPFAC)_i} \qquad (1)$$

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

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. 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 $\underline{/}Y_t = Y_o(1+r)^t$ or $\log Y_t = \log Y_o + t \log(1+r)\underline{/}$ to the sample

---

[1] The following 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. Strout uses the function for cross section data.

period data. The sectoral growth rates and the gross capital-output ratios estimated from the Kenyan data 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, $(\Delta GDPFAC)_{it}$, is the same on both sides of equation (1). In fact, it turns out that the variance of $(\Delta GDPFAC)_{it}$ is

much greater than the variance of either $(INVEST)_{it}$ or $(GDPFAC)_{it}$. So that

the reason we get the very impressive $\bar{R}^2$ for the capital-output ratios is

largely that the variance of $(\Delta GDPFAC)_{it}$ 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

predilictions; for here too in a sense we have the Marschak-type situation.[2]

Perhaps the following practical consideration will be more

persuasive than lengthy arguments about whether or not the correlations are

properly regarded as spurious. In Table 1 below the sectoral capital-

output ratios reported above are used to estimate fixed investment in the

sample period. The cumulative investment for 1961-68 is very close to the

actual figure - the χ^2 in Column 12 are within the permissible limits - but

this cumulative total is made up of very large over- and under-estimation

of investment in individual years; thus almost all of the χ^2 in Column 10

are well beyond their permissible values for eight observations. May one

---

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. Simi-
larly, when a series such as money value of output is divided by a
price index to obtain a 'constant dollar' estimate of output, no ques-
tion 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 extra-
neous 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.

conclude that while the capital-output ratios are not reliable for fore-
casting 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, estimated investments of Table 1 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 $\overline{R}^2$ of equations estimating capital-output ratios
would, therefore, have been completely unwarranted. A part of the reason,
of course, is that the $\overline{R}^2$ of equation (1) above relate to the magnitude
$\underline{/}(\text{INVEST})_i/(\Delta\text{GDPFAC})_t\underline{}$ 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. Neither of them requires any extended comment.

Two remarks will suffice for the t-ratio.

First, we may recall the facts already noted that multicollinearity
among independent variables makes the b-coefficients and their standard-
errors highly unstable, and that it is very difficult to avoid multicolli-
nearity without ending up with an incompletely specified model. It is
indeed true that as collinearities between independent variables increase
the standard errors of the coefficients will also increase and ultimately
approach infinity and that the increasing standard errors, in turn, will

Table 1: FORECASTS OF GROSS FIXED INVESTMENT AND ACTUAL GROSS FIXED INVESTMENT IN KENYA OVER THE SAMPLE PERIOD, 1961-1968

(£ million)

(1)	1961 (2)	1962 (3)	1963 (4)	1964 (5)	1965 (6)	1966 (7)	1967 (8)	1968 (9)	χ^2: Years 1961-68 taken individually (10)	Total 1961-1968 (11)	χ^2: Total 1961-196 (12)
Agriculture, Livestock, Forestry and Fishing	-6.27 (6.70)	27.62 (6.04)	20.27 (7.98)	3.98 (8.88)	-23.10 (8.98)	45.52 (11.72)	3.35 (12.81)	5.23 (13.10)	347.635	76.60 (76.21)	0.002
Mining and Quarrying	-2.29 (0.09)	-0.21 (0.10)	0.57 (0.07)	-0.88 (0.04)	-0.57 (0.05)	1.35 (0.09)	2.24 (0.16)	0.78 (0.60)	141.052	0.99 (1.20)	0.0366
Manufacturing and Repairing	3.76 (6.10)	1.03 (6.70)	4.90 (6.00)	13.99 (6.95)	10.63 (7.55)	12.76 (9.51)	9.43 (14.90)	11.06 (14.13)	18.071	67.56 (71.84)	0.255
Building and Construction	-1.01 (1.88)	-0.94 (1.47)	-14.66 (1.53)	3.92 (1.37)	2.79 (2.09)	12.12 (1.66)	13.06 (4.33)	5.22 (4.80)	268.240	20.50 (19.13)	0.098
Electricity and Water	0.00 (1.72)	4.66 (2.23)	0.29 (2.38)	2.87 (1.32)	1.31 (2.17)	2.40 (2.46)	5.91 (6.17)	3.63 (3.38)	8.395	21.07 (21.83)	0.026
Transport, Storage and Communications	3.68 (7.05)	5.09 (7.25)	8.20 (7.00)	11.93 (10.85)	12.03 (8.21)	20.18 (13.35)	8.05 (16.66)	11.74 (17.18)	14.011	80.90 (87.55)	0.505
Other Sectors	2.98 (14.13)	9.09 (16.54)	13.77 (14.60)	33.10 (15.72)	17.24 (17.48)	33.84 (22.41)	23.93 (32.27)	31.35 (37.79)	40.502	165.30 (170.94)	0.186
Total Fixed Investment	0.85 (37.67)	46.34 (40.33)	33.34 (39.56)	68.91 (45.13)	20.33 (46.53)	128.17 (61.20)	65.97 (87.30)	69.01 (90.98)	148.574	432.92 (448.70)	0.555

Note: The forecasts are based on the capital-output ratios listed in the text. Figures in parenthesis indicate actual gross fixed investment.

ultimately make the t-ratios for one or more of the coefficients insignificant. But this fact cannot be taken to mean that the t-test is sufficient in the sense that it will adequately alert us to the collinearities. It will not suffice. The t-ratio is just the ratio of the estimated coefficient to its standard error. Therefore, as the degree of collinearity increases it becomes just as arbitrary as the b-coefficients and the effort to partition the variance of the dependent variable between highly collinear independent variables becomes arbitrary even if at a particular stage in the analysis t-ratios for one or all the independent variables still happen to be significant.

A set of equations reported in UNCTAD, 1968 provide an instructive illustration. Faaland and Dahl fit the following regression to account for the share of agriculture in the total GDP of Tanzania:1/

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

where P_A is value added in 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:

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

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:

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

They conclude that as the negative coefficient for GDP "on a priori grounds appears improbable" their data "apparently do not give us any reliable basis for assessing how the functional relationship between agricultural output and GDP is modified by changes in income levels."2/ But a glance at the matrix

---

1/ UNCTAD, 1968, p. 204.

2/ Ibid.

of correlation coefficients would have shown that the two independent variables were so highly correlated that any attempt to separate their effects on the basis on the sample period data alone was bound to fail.1/ Note that in the third equation the t-ratios for both the explanatory variables are well above the conventional cut-off point of 2.0.

Second, 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.

1/ The same experiment with identical explanations and similar results is repeated for Kenya (ibid., pp. 123-124) and Uganda (ibid., pp. 248-249). The matrix of simple correlation coefficients for the Tanzanian data given in UNCTAD, 1968, page 222, is as follows:

	P_A	GDP	GDP/N
P_A	1	0.9984	0.9997
GDP	0.9984	1	0.9985
GDP/N	0.9997	0.9985	1

The value of the determinant of the correlation coefficients between independent variables taken by themselves can vary between one (when the variables are completely independent of one another) and zero (when they are perfectly correlated). With two independent variables the matrix of correlation coefficients in the former case is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and in the latter case it is $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. For the Tanzanian data it is $\begin{bmatrix} 1 & 0.9985 \\ 0.9985 & 1 \end{bmatrix}$. The value of the determinant thus is (0.003).

It is now customary to display the Durbin-Watson statistic with most equations. 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 a fairly large proportion 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

---

[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.

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 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. This is
particularly important since even a casual glance at the different models
will show that the individual equations used sometimes explain no more than
25 percent of the variance in the dependent variable.

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 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 $R_1 \ldots R_n$ are random variables with finite variances $\sigma^2 + \ldots + \sigma_n^2$ and $S_n = R_1 + \ldots + R_n$ 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

---

[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.

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 re-
duced 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.

Lest one is tempted to dodge this issue by trying out all combina-
tions of high, medium and low values of each equation so as to "present
management with a range of answers," we note that the number of combinations
one would have to try out is truly astronomical. In the very small Kenya
model, 22 equations go to determine the import-export gap - excluding the
debt routine which also involves discrete assumptions about the terms of
lending and about the proportions that are likely to be financed on different
terms. Taking the high, medium and low values of each we obtain 3^{22} combina-
tions. That is, one would have to try out about 31.3 billion combinations
if one were serious about conducting a "sensitivity analysis by complete
enumeration." Even this enormous number of combinations does not provide for
high, medium and low variants of the seven sectoral growth rates that are the
exogenous variables in the model or of the many exogenously specified items
in the balance of payments. Nor does it provide for alternative functional
forms or specifications of each equation; if, for instance, one had a linear,
a log-linear and a quadratic form for each equation the number of combinations
would increase to 9^{22} or about 983 billion billions. The task, needless to
say, would not only be to do the computations but also to analyze them in
some way - by studying indices like the debt service ratio, the ratio of net
to gross flows and so on. In practice, therefore, one would have to adopt
some mechanical rules for ruling out some combinations altogether. Thus one
may impose some "consistency" requirements: if, for instance, one believes
that an acute shortfall in exports will almost certainly result in a regime
of tight import controls one can reject all the combinations involving the
low-exports high-imports alternatives. The difficulty, of course, is that
for some the low-exports high-imports combinations will indeed be the
interesting ones and they will argue that in this age of foreign aid,

---

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.

compensatory financing and special drawing rights these combinations are not
altogether impossible. Alternatively, one may impose "feasibility" require-
ments: if the quadratic equation for imports of capital or intermediate goods
results in Kenya importing negative quantities of these goods one could rule
out all combinations involving these equations on the grounds that we know
enough about Kenya's production plans to be confident that Kenya is not going
to become an exporter of these goods in the next five years. But these
consistency or feasibility conditions will not really help us out for there
is no reason to suppose that the range of answers that results from the
manipulations will be any less than the one that emerges from the procedure
indicated above.

In the specific context of the conventional two-gap models 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 invest-

ment-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 imme-

diately faced with the difficulty that in the (ex post) sample period data

the savings-investment and the import-export gaps will be necessarily iden-

tical. 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 to-
wards 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 proce-
dure 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 sam-
ple 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 effor 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 aggre-
gation 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 to go 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 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 fore-casts are based on the same information.

VII. <u>The Moral of the Story</u>

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 <u>a priori</u> 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 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 does one use one's "awareness of their limitations" 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 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,[2] Ehrenberg,[3] Box,[4] Geary,[5] and others) and on problems presented by the inaccuracy of basic data (as for instance the writings of Morgenstern). 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] "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] A.S.C. Ehrenberg, "Bivariate Regression Analysis is Useless," Applied Statistics, Vol. 8, No. 3, November, 1963, pp. 161-179.

[4] G.E.P. Box, "Use and Abuse of Regression," Technometrics, Vol. 8, No. 4, November, 1966, pp. 625-629.

[5] R.C. Geary, "Some Remarks About Relations Between Stochastic Variables: A Discussion Document," op. cit.