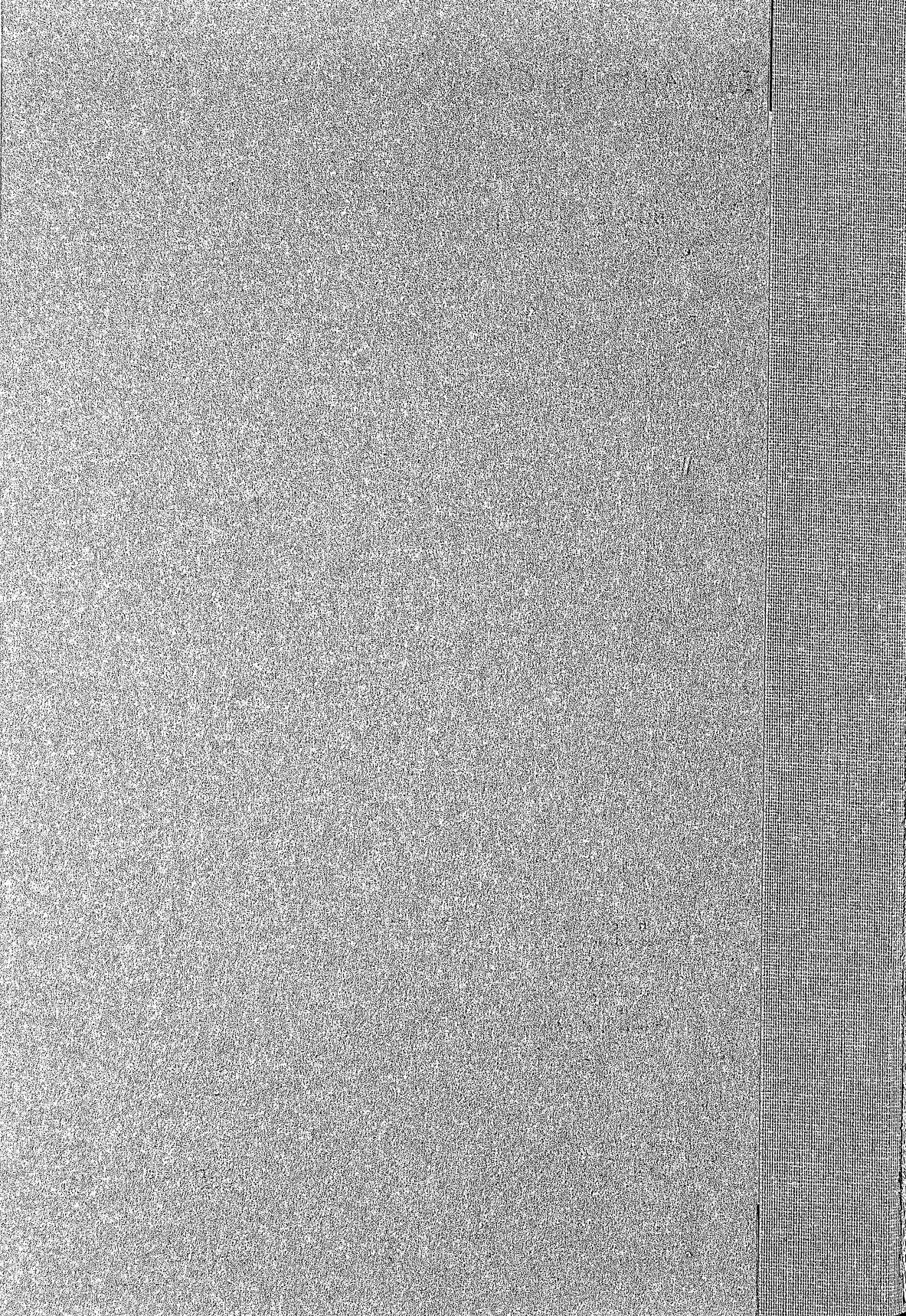


MichU  
CenRED  
R  
#12



# THE AMERICAN ECONOMIC REVIEW

LIBRARY

FOR RESEARCH ON ECONOMIC DEVELOPMENT  
March 1970

THE UNIVERSITY OF MICHIGAN

VOLUME LX, NUMBER 1

GEORGE H. BORTS

Managing Editor

WILMA ST. JOHN

Assistant Editor

Board of Editors

GARY S. BECKER

BARBARA R. BERGMANN

ACDISH N. BHAGWATI

GREGORY C. CHOW

ALAN F. CHRIST

C. FERGUSON

IRVING FRIEND

HARRY G. JOHNSON

ALVIN L. MARTY

MARC NERLOVE

G. WARREN NUTTER

WERNON SMITH

WARREN L. SMITH

LEONARD ULMAN

• Manuscripts and editorial correspondence relating to the regular quarterly issue of this Review should be addressed to George H. Borts, Managing Editor of THE AMERICAN ECONOMIC REVIEW, Brown University, Providence, R.I. 02912. Manuscripts should be submitted in duplicate and in acceptable form. See *Instructions* for guidance in preparing manuscripts will be provided upon request to the editor.

• No responsibility for the views expressed by authors in this Review is assumed by the editors or the publishers, The American Economic Association.

• Copyright American Economic Association  
1970

## Articles

- Trends in the Activities Generating Technological Progress  
*William Fellner* 1
- About the Growth Path of Firms  
*Romesh K. Diwan* 30
- Portfolio Choice, Investment and Growth  
*Duncan K. Foley and Miguel Sidrauski* 44
- Toward a Theory of Nonprofit Institutions: An Economic Model of a Hospital  
*Joseph P. Newhouse* 64
- Cyclical Variation in the Interindustry Wage Structure  
*Michael L. Wachter* 75
- Monetary Growth Theory in Perspective  
*Jerome L. Stein* 85
- Differential Tariffs, Negative Value-Added and the Theory of Effective Protection  
*Augustine H. H. Tan* 107
- A Geometric Treatment of Averch-Johnson's Behavior of the Firm Model  
*E. E. Zajac* 117
- Migration, Unemployment & Development: A Two-Sector Analysis  
*John R. Harris and Michael P. Todaro* 126
- On the Economic Welfare of Victims of Automobile Accidents  
*R. A. Holmes* 143
- Production Scheduling, Intermediate Goods, and Labor Productivity  
*W. H. Locke Anderson* 153
- Multiperiod Consumption-Investment Decisions  
*Eugene F. Fama* 163

# An Econometric Model of Development: Comment

By SARA S. BERRY\*

Sixty years ago, Joseph Schumpeter began his *Theory of Economic Development* with a discussion of the partial character of all economic analysis. "Economic development is . . . simply the object of economic history which in turn is merely a part of universal history . . . Because of this fundamental dependence of the economic aspect of things on everything else, it is not possible to explain *economic* change by previous *economic* conditions alone" (p. 58). Schumpeter was writing primarily about the historical experience of European societies who shared many common cultural and institutional features; his point is, if anything, even more relevant for contemporary students of economic development who try to explain patterns of economic change for a much greater variety of societies. Yet, although many economists have recognized that differences in the economic performance of different societies cannot be explained by economic factors alone, few have attempted systematic analyses of the respective roles of economic and non-economic factors in generating economic change. Consequently, the literature on economic development offers few general propositions about the quantitative importance of political and cultural factors for development, or about the way in which such factors interact with economic forces to alter the structure or volume of an economy's productive capacity.

In their article "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," Irma Adelman and Cynthia Taft Morris have undertaken an important step toward filling this crucial gap in the literature. Specifically, they have attempted, using stepwise regression analysis, to estimate the importance of a large number of variables for explaining differences in the economic performance of sev-

enty-three non-Communist countries during the 1950's and early 1960's.<sup>1</sup> They then used the regression coefficients to calculate multipliers measuring the impact of changes (of one standard deviation) in eighteen of the independent variables on a country's "development potential."

They included a number of social and institutional variables which are not readily quantifiable and explored relationships about which there are few widely accepted theoretical propositions. The significance of their conclusions about how and why countries develop economically depends not only on the form of their statistical analysis, but also on their identification and quantification of relevant variables and on their interpretation of the statistical associations discovered. In all these respects the authors faced formidable problems, which they have not always solved very successfully.

## I. *The Independent Variables*

In constructing quantitative indicators for various institutional and social factors which may affect countries' economic development, Adelman and Morris follow the general procedure of classifying countries into several categories according to the factor in question and then assigning values to the categories at arbitrary intervals on a linear scale.<sup>2</sup> They justify this procedure on the grounds that all their variables are continuous and that "the only difference between our case and the more usual one is that deficiencies in information led us to scale our variables by means of a yardstick which has demarcations at coarse intervals only (e.g.,

<sup>1</sup> In their article, Adelman and Morris label their original dependent variable "development potential" but it consists of a classification of countries according to several indicators of past economic performance. (p. 1188, fn. 6)

<sup>2</sup> Detailed explanations of how Adelman and Morris defined and measured the variables in their model are given in their book, *Society, Politics, and Development*.

\* The author is an assistant professor of economics at Indiana University.

feet or yards instead of inches)" (p. 1204). This argument is misleading since, for a number of variables, the "intervals" to which Adelman and Morris have assigned equal numerical values do not represent equal quantitative changes in the underlying variables. For example, in constructing their index of "social mobility" ( $X_{26}$ ), the authors assigned scores of 90 to countries with more than 40 percent of the population aged 5 to 19 in school and more than 10 percent of the active males in "middle class" occupations; 50 to countries with 25 to 40 percent of the 5-19 age group in school and 5-10 percent of the active males in middle-class occupations; and 10 to countries not classed in the above groups (pp. 34-35). Thus, the 40 point difference between the first two groups represents a 1 to 15 percent difference in the proportion of school-aged children in school, while the 40 point difference between the latter two groups represents a 1 to 25 percent difference in this indicator. In addition, some countries received lower scores than these criteria warranted because they were judged to have "prohibitive cultural and ethnic barriers to upward social mobility." This criterion further obscures the meaning of an interval on the social mobility scale because it is not operationally defined. In short, what the intervals seem to measure is as much the authors' opinion (or the opinions of anonymous "country experts") about the extent of social mobility in different countries as any objective evidence on the subject.<sup>3</sup>

Some of the independent variables included in Adelman and Morris's model are not only arbitrarily measured but also ambiguously defined. A case in point is their treatment of dualism, an overworked term

in development literature which has long ceased to have any precise meaning. The original proponent of the term, J. H. Boeke, defined it as "the clashing of an imported social system with an indigenous social system of another style" (p. 4). Others have used the term to characterize economies with various kinds of sectoral differences, e.g., in technology and factor proportions, in the use of money and exchange, in the institutions available for capital accumulation and technical change, or in people's attitudes toward work, profit and efficiency. In order to use the concept of dualism in a quantitative model, one is faced with the task of sorting out various possible definitions into observable variables whose relationship to development potential can be interpreted meaningfully.

Adelman and Morris's definition does not accomplish this task, but tends to reflect the confusion in the literature. For example, they assign the lowest score on their dualism scale to the "largely agrarian societies having extremely small exchange sectors." At the other end of the scale are "countries with continuous interaction between modern and non-modern elements" (p. 1214). Apparently we are to assume that the larger the exchange sector and the less agrarian the society, the greater the degree of interaction between modern and non-modern elements. But it is not clear, either in theory or in practice, that the exchange sector always grows at the expense of the agrarian sector. In many West African countries, for example, commercial activity and the use of money are widespread, yet the majority of the population is still engaged in agriculture. Indeed, the authors appear to recognize this fact in discussing intermediate points along the scale: "countries in which the growth of an indigenous small-scale cash-crop sector using conventional techniques evolves at the expense of a traditional subsistence sector" rank relatively high on the scale. This sug-

<sup>3</sup> Moreover, by assigning arbitrary scores to different cultural or political characteristics of a society, the authors interject their own preferences into the estimated relationships between those characteristics and the society's economic performance. Another observer who chose to "evaluate" the three levels of social mobility defined above at scores of 10, 5 and 1 would obtain a different estimate of the regression coefficient relating differences in social mobility to differences in the authors' index of "development potential." For example, see R. D. Luce and H. Raiffa's discussion of Ar-

<sup>4</sup> Adelman and Morris's terminology is somewhat confusing here—they state that "positive changes in dualism involve reductions in the cleavage between modern and traditional sectors" (p. 1194) or, the less

gests that it is the extent to which goods are sold for money, rather than the composition of output or the occupational structure which best reflects the degree of interaction between modern and non-modern elements. If this is what the authors have in mind, however, it is difficult to see why they argue that "societies in which a foreign-financed and directed modern sector is superimposed upon a predominantly agrarian society" should rank below countries where cash-cropping is widespread among the indigenous population. In countries such as Zambia or Malawi (which rank low on Adelman and Morris's scale), many of the indigenous population regularly exchange goods and services (e.g., labor services) with foreign-financed and directed firms while continuing to engage in traditional agriculture. Thus there is a good deal of monetary exchange between modern and non-modern elements. Evidently, the authors had some other form of interaction in mind, but they do not state it explicitly. They have combined so many criteria (not all of which are closely intercorrelated) in their definition of dualism that is not clear what they are trying to measure or what significance ought to be attached to their statistical estimates of the associations between dualism and other variables in the model.

Moreover, in ranking individual countries on their dualism scale, Adelman and Morris have not applied their own criteria very consistently. For example, South Africa ranks well above Zambia, although in both economies a "foreign-financed and directed modern sector" exists side by side with a low-income agrarian sector; Uganda, where the bulk of the cash crops are produced on small farms by indigenous farmers, is scored below Rhodesia and Kenya, both of which have substantial foreign-financed and directed sectors. Since the authors do not justify their rankings of individual countries in detail, it is impossible to say what other criteria they used. However, available data suggest that South Africa does have a higher per capita income than Rhodesia which, in turn, has a higher per capita income than does Uganda. If, in effect, Adelman and

Morris's index of dualism measures levels of economic performance rather than some independent structural characteristic of the economies in question, then the strong correlation they find between dualism and development potential is neither surprising nor very instructive.

I have discussed the ambiguities in Adelman and Morris's definition and measurement of dualism at some length to illustrate what seems to be a fairly extensive problem in their analysis. There is not room in the present discussion to examine each of the nineteen variables in their model, but dualism is not the only one whose meaning is ambiguous because the authors have either combined several types of information in one index or defined variables in such a way as to assume relationships they are trying to measure.<sup>5</sup>

## II. *The Regression Coefficients*

Even for those variables which are unambiguously defined, it is not always clear why the authors interpret their results as they do. In reading their initial premise, that "there are no firmly validated theories of the process of socio-economic and political change [so that] we consciously avoided a priori specification of the functions we wished to fit" (p. 1184), one anticipates a fairly thorough discussion of alternative interpretations of these functions once they are fitted. However, Adelman and Morris's explanations of their regression coefficients and the multipliers derived from them are very brief. In many cases they consider only one of several possible interpretations of a statistically significant relationship. Take, for example, the first regression equation which the authors present (after the discriminant

<sup>5</sup> In other instances they have simply not taken the trouble to express themselves clearly. In defining "the importance of the indigenous middle class," one presumes that the authors do not really mean that "this classification is based upon the relative size . . . of indigenous people in middle-class occupations" (p. 1214), but the fact that they say so does not increase the reader's confidence in their other stated opinions or conclusions.

function which is used to construct the dependent variable):

$$(1) X_{12} = .60X_4 + .33X_{21} \quad R^2 = .57$$

where  $X_{12}$  is the degree of improvement in financial institutions;  $X_4$  is the degree of industrialization; and  $X_{21}$  is the extent of dualism. According to Adelman and Morris this equation shows that increased demand for credit in both industry and agriculture has a favorable effect on the performance of financial institutions. Since, as we have seen, dualism is not quite the same thing as the agricultural sector's demand for credit, their interpretation of the second regression coefficient seems a bit arbitrary. More important, however, is their failure to explain why the line of causation should run from industrial expansion to improved financial institutions rather than vice versa. In general terms, it seems just as plausible to argue that the growth of the banking system facilitates industrial expansion by helping to mobilize potential savings and to inform investors of their availability. In fact, in their 1967 study, Adelman and Morris argued that changes in structural variables might be considered causes of changes in economic performance because "it is difficult to see how, in the short run at least, causality could run in the opposite direction" (p. 175). Evidently they have changed their minds; one would be interested in their reasons for doing so.

Another difficulty of interpretation arises out of the authors' technique of analysis.<sup>6</sup> Stepwise regression is useful in this kind of study in that it permits one to select from a large number of variables the few which are most closely associated with a particular dependent variable. Moreover, the technique selects first that independent variable which explains the greatest proportion of the variation in the dependent variable; second, the independent variable which explains the largest share of the remaining unexplained variation in the dependent variable, and so on. The disadvantage of stepwise regression

is that in selecting variables for inclusion in the regression equation, if any of the independent variables are intercorrelated so that they affect the dependent variable jointly as well as independently, the technique attributes all of the joint effect to the variables selected for inclusion. Thus, some variables rejected by the stepwise regression analysis as not adding significantly to the  $R^2$  for a particular relationship may, in fact, have an important influence on the dependent variable. In the present example, if a rejected variable (say,  $X_2$ , the gross investment rate) is intercorrelated with  $X_4$ , the regression coefficient for  $X_4$  includes the joint effect of  $X_2$  and  $X_4$  on  $X_{12}$ . Thus, the conclusion that the gross investment rate does not have a significant effect on improving financial institutions in developing countries may be incorrect.

### III. *The Multipliers*

In the last section of their paper, Adelman and Morris provide "a statistical explanation of the economic and noneconomic forces which directly and indirectly determine a country's capacity for economic growth" (p. 1184). Specifically, they use the regression coefficients of the model to calculate multipliers which estimate the quantitative impact of changes in different socio-economic and political variables on development (measured by the discriminant function), and from the multipliers they draw conclusions about the relative importance of different variables for understanding how and why economies develop. Their conclusions seem questionable on several grounds. In the first place, there is no more evidence for the multipliers than there was for the regression coefficients concerning causal sequences. One may argue just as plausibly from their data that development will generate significant changes in outlook, economic structure and financial institutions as vice versa. Second, because their regression coefficients may include joint effects of variables not included in the equations (to say nothing of those not specified to begin with), the meaning of the multipliers is somewhat ambiguous. For example, the multiplier for the improvement

<sup>6</sup> I am indebted for this point to my colleague Jeffery Green.

in financial institutions, 2.309, does not mean that if the rate of improvement in financial institutions rises by one standard deviation then development performance will rise by 2.309 standard deviations *ceteris paribus*, because other things are not necessarily equal. But if the multipliers cannot be interpreted in this fashion, it is not clear how they provide us with a valid ranking of the independent variables in terms of their impact on development.

Finally, the ambiguities in the authors' initial definitions and measurements of the independent variables are in no way mitigated by the statistical analysis. Since we still do not know, for instance, whose outlook weighs most heavily in Adelman and Morris's index of modernization of outlook or what kinds of information they used to measure differences in outlook, their conclusion that greater modernization of outlook will have a significant impact on economic development is not very enlightening. In the case of dualism, insofar as the authors' classification of countries was based on dif-

ferences in economic performance, their conclusion is tautological; otherwise it is merely confusing. In short, although Adelman and Morris have undertaken an important and formidable task, they do not seem to have solved enough of the problems inherent in any attempt to quantify interactions between economic and noneconomic forces to have produced significant or conclusive results.

#### REFERENCES

- I. Adelman and C. T. Morris, "An Economic Model of Socio-Economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58 1184-1218.
- , *Society, Politics and Economic Development*. Baltimore 1967.
- J. H. Boeke, *Economics and Economic Policy of Dual Societies*. New York 1953.
- R. D. Luce and H. Raiffa, *Games and Decisions*. New York 1957.
- J. A. Schumpeter, *The Theory of Economic Development*. Cambridge, Mass. 1936.



# An Econometric Model of Development: Comment

By PETER ECKSTEIN\*

In the December 1968 issue of this *Review*, Irma Adelman and Cynthia Taft Morris present a cross-section analysis of the recent growth experience of underdeveloped countries. Their "econometric model of development" seeks to supplement economic analysis and country case studies as means of isolating those policies most likely to promote rapid development. The argument of this note is that their model not only fails to deal directly with some of the inherent problems of cross-country comparisons but also neglects some of the prime rules of scientific inquiry and offers, therefore, no reliable guide to policy-making. In commenting on their work, I shall also make reference to their book, *Society, Politics and Economic Development*, and to their journal article which appeared in May 1968.

The model has three stages: a discriminant analysis of 73 non-Communist underdeveloped countries grouped according to "development potential"; a series of multiple regression equations for the four explanatory variables that emerge as important from that analysis; and a set of "multipliers," based on these two analyses and designed to measure the impact on "development potential" of some 18 political, social, and economic variables. Each of these stages will be examined in turn.

## I. The Discriminant Function

Discriminant analysis is a method of finding those characteristics which most clearly set off the members of one group from the members of other groups. Like regression analysis, it is a means of explaining or pre-

dicting a dependent variable (here called a criterion variable) in terms of several independent (or prediction) variables. Unlike regression analysis, only the prediction variables of discriminant analysis are defined continuously. The criterion variable is the group membership of the observation where this membership is known beforehand for all the cases in the sample.<sup>1</sup>

Adelman and Morris begin their analysis by classifying the countries by their average annual rates of growth of per capita Gross National Product (*GNP*) between 1950-51 and 1963-64: "high" for rates above 2 percent, "intermediate" between 1 and 2 percent, and "low" below 1 percent. They explicitly reject, however, growth of *GNP* per capita as a measure of the "prospects for sustained economic growth" on the grounds that there are many cases of "growth without development." They therefore modify the membership of these groupings in several ways (discussed below) in order to arrive at a classification according to high, intermediate, and low "development potential." Although they begin with 29 potential prediction variables, their final discriminant function requires only four to assign each of the seventy-three countries a score closely approximating the mean for its own group and to account for "over 97 percent of the overall variance in membership." The four prediction variables are "the degree of improvement in financial institutions," "the degree of modernization of outlook," "the extent of leadership commitment to economic development," and "the degree of improvement in agricultural productivity."

Unfortunately, the impressive statistical power of the function seems largely the result of the ways in which the authors have defined and measured the criterion variable and the prediction variables.

\* Assistant professor of economics, University of Michigan, and research associate, Center for Research on Economic Development. I am indebted to Elliot Berg, Lester Taylor, Wolfgang Stolper, and Shanti Tangri for careful and critical readings of earlier drafts. This note represents a condensation of my larger discussion paper, "Quantitative Measurement of Development Performance."

<sup>1</sup> See, for example, Kendall's *A Course in Multivariate Analysis*.

*Defining "Development Potential"*

There are two general reasons for regretting the decision to move from individual estimates of country growth rates to groupings of countries by development potential. By arbitrarily dividing the sample into only three groups, they have lost information about the substantial differences in growth performance within each of the groups.<sup>2</sup> Moreover, by attempting to transform groupings by growth rates into groupings by development potential, the authors have rested the entire analysis on a concept that could easily elude the most careful efforts at definition and measurement.

In practice, however, the authors neither define development potential prior to the analysis nor measure it independently of the prediction variables. They move from estimates of growth of per capita *GNP* to their concept of development potential in three distinct steps, each of which reduces the extent to which the variable being explained is observed independently of the variables being employed to explain it.

First, six countries with high growth rates are declared "unclassified" as to development potential and initially excluded from the analysis. Two—Panama and Burma—are completely unexplained, while four are African countries "in which, as of about 1960, over 90 percent of the population was in the traditional subsistence sector where per capita *GNP* had not changed significantly since 1950" (Adelman and Morris, 1967, pp. 89-90). For these four, then, poor performance on one of the prediction variables—improvement in agricultural productivity—was the explicit basis of the declassification, but five of the six countries are predominantly low on all four prediction variables.

Next, because they are interested in "development" rather than "growth," the authors seek to eliminate from the high-potential category those countries whose high

<sup>2</sup> For example, Turkey (with a per capita *GNP* growth rate of 2.1 percent) is grouped (and therefore assumed identical in "potential") with Greece (5.9 percent) but not with Costa Rica (1.9 percent).

growth rates "had only a limited overall impact within the economy." They therefore relegate to the intermediate category eleven fast-growing countries that do not rank "at least moderately high" with respect to five out of seven "economic performance characteristics."<sup>3</sup> Unfortunately, all seven of these characteristics are included among the twenty-nine variables which the computer later scans in order to build the discriminant function. Not surprisingly, two of the seven—improvement in financial institutions and improvement in agricultural productivity—that are here used to categorize countries as to development potential are also among the four that are eventually "found" to be the best predictors of the categories to which countries belong.<sup>4</sup>

The third step by which growth-rate groupings are transformed into development potential groupings is the most disturbing. On the basis of the rankings already described, the authors develop an initial discriminant function which contains four prediction variables—three of the four mentioned above, plus "degree of improvement in physical overhead capital" instead of the agricultural productivity variable. When scores on this discriminant function are calculated for the sixty-seven countries in the analysis, twelve countries are found to have ratings on the prediction variables more characteristic of countries in different categories of development potential, while three are marginal.

Adelman and Morris react to these discrepancies by observing that their confidence in the initial classifications for potential was "not especially high" in the first place (Adelman and Morris, May 1968, p. 277). They do not explain how their discriminant function could possess more validity than the criterion variable from which it is directly derived, or how their prediction variables—

<sup>3</sup> The seven are: change in degree of industrialization; improvement in agricultural productivity; improvement in physical overhead capital; improvement in financial institutions; improvement in the tax system; and improvement in human resources.

<sup>4</sup> Examination of the ratings for the high-growth countries confirms that those rated low on either of these two characteristics were far more frequently demoted to the "intermediate potential" category.

apparently measured no more carefully or objectively than the criterion variable—could greatly exceed it in validity. Nevertheless, they proceed to employ both the discriminant function and the initial classifications of the prediction variables not only to classify the six countries for which development potential had been treated as unknown, but also to declassify the fifteen countries which the function did not fit well. They treat these new classifications and declassifications not as *hypotheses* but as *the equivalent of data*—that is, as the input to (or exclusions from) another discriminant analysis. The discriminant function resulting from this analysis is then used to reclassify the unknown development potential of the fifteen countries, and these classifications are treated as part of the data input to a third discriminant analysis.

When this procedure is complete, all seventy-three of the countries have been classified—although for twenty-one countries observed growth performance has been treated as irrelevant, and classification has been based *entirely* on ratings on the four prediction variables included in the first two discriminant functions. When the third and final function is calculated from those classifications, it accounts for “over 97 percent of the overall variance,” and it renders “the separation between groups . . . considerably better . . . and the dispersion within the groups . . . substantially reduced” (Adelman and Morris, May 1968, p. 278). This should not be surprising since the new function employs three of the four prediction variables used in classifying so many of the observations.<sup>5</sup>

Thus, there are three operations by which a classification of countries by growth rates

<sup>5</sup> In all there are 32 reclassifications in which 26 of the 73 countries end up in new categories. In my discussion paper I have shown that the final discriminant function produces very similar scores for countries with high and intermediate rates of *GNP* growth, though the low-growth countries generally score lower. The ability to distinguish between countries ranked high and intermediate on development potential is largely achieved when the first discriminant function is used to reclassify the 6 originally unclassified countries and to declassify the 15 which fit it poorly.

evolves into a classification by development potential, and each operation brings the implicit definition of high development potential closer to being nothing more than high performance on the four final prediction variables. There can, of course, be no scientific interest in even the strongest of statistical relationships between variables that have been defined to be similar.

#### *Defining and Measuring the Prediction Variables*

At least some of the statistical power of the discriminant function is based on underlying associations between the four prediction variables and the rate of growth of *GNP*.<sup>6</sup> Unfortunately, we cannot conclude from these associations that any of the four is an independent cause of faster economic growth. Two of the variables could easily have been influenced by the same empirical phenomena that lay behind the growth-rate categorizations; the other two may merely serve as proxies for more basic explanatory variables.

“Leadership commitment to economic development” is one of several qualitative or “judgmental” variables; these were initially quantified by the authors on the basis of recent country studies and subsequently checked by country experts (largely officials of the U.S. Agency for International Development). Countries were ranked highest on leadership commitment if the government and central economic leadership had undertaken “concerted efforts” to promote growth, “serious attempts” to alter “institutional arrangements unfavorable to growth” and “some reasonably effective development planning” (Adelman and Morris, Dec. 1968 p. 1216; 1967, p. 80). The difficulty is that a judge must rely on indirect evidence to decide whether all that commotion over planning and development really represents “concerted” and “serious” effort, and the best indirect evidence may be whether or not

<sup>6</sup> The simple correlation coefficients with rate of growth are: .35 for leadership commitment, .61 for agricultural productivity, .54 for modernization, and .69 for financial institutions. The statistics are taken from Adelman and Morris (1967).

the effort is succeeding. Can a leadership really be judged "committed," can planning really be judged "effective," can institutional reform really be judged a contribution to growth, if the country is in fact not growing? Until more independent measures can be developed, we must suspect that unavoidable biases of judgment, not underlying causal relationships, account for some of the association between leadership commitment and growth.

A second prediction variable, "improvement in agricultural productivity," was intended to be a measure of total factor productivity. For economies dominated by agriculture, even accurate and complete data would be expected to show a high correlation between output per worker (or per factor) in agriculture and output per worker (or per capita) in the economy as a whole. We cannot find a great deal of theoretical interest in an empirical finding that across countries in which agriculture accounts for most of *GNP*, agricultural productivity and *GNP* growth rates are highly correlated.

In practice, the authors concede that even "reasonably reliable" data were often not available for either variable—and particularly not for the many low-income African countries—and they were forced to rely on purely "qualitative indications" or "estimates" (Adelman and Morris 1967, pp. 87–89, 104–108). There must have been available very little concrete evidence, then, on which to move from estimates of growth of output per worker in agriculture to quantitatively different estimates of the growth of output per factor in agriculture or of output per capita in the larger economy.

"Modernization of outlook," another judgmental variable, could easily be a proxy for such basic economic factors as level of *GNP* per capita or the extent of subsistence agriculture. In particular, a high negative correlation of "modernization" with the "size of the traditional agricultural sector" ( $-.82$ ) implies very few differences in the rankings on the two variables. Those differences may have been because of actual differences in popular "outlook," but they also may have

been because the concomitants of high or low growth colored the interpretations of the experts. Indeed, the fact of growth itself could easily have been taken as evidence of popular support for "programs of political and economic modernization"—one of the major components of the definition of modernization of outlook. Even small differences of this sort would be sufficient to account for the greater explanatory power of this variable when it enters the discriminant function.

The variable "improvement in financial institutions" could easily be a surrogate for the "widespread pattern of dynamic improvements" that the authors find "well correlated" with it—"changes in the degree of industrialization, improvements in agricultural productivity, improvements in tax institutions, and to a lesser extent, improvements in human resources" (Adelman and Morris, May 1968, p. 269). Its statistical superiority over these other variables (several of which were also involved in the definition of the criterion variable) may be partly explained by the fact that one of its two components is "the approximate increase in the real value of private domestic liabilities to the banking system" (Adelman and Morris 1967, p. 121), apparently measured absolutely, not in relation to a growing *GNP*. Thus, even if all countries maintained a constant ratio of financial liabilities to *GNP*, this measure would be larger for the faster-growing economies.

Our examination of the four prediction variables in the discriminant function, then, provides no basis for confidence that they do represent important causal forces explaining differences in rates of economic growth or the potential for growth.

## II. Regression Analysis

The regression stage of the model is an effort to explain the four prediction variables used in the final form of the discriminant function. These are introduced as the dependent variables in four multiple regression equations. Five additional regression equations are developed to explain some of the independent variables in the first set, and

four more are then developed to explain some of the independent variables in the second set.

In building an equation for each dependent variable, the authors apparently began with all twenty-eight other variables as candidates for inclusion and let the computer choose those most strongly associated with the dependent variable. In a few cases "variables were omitted from equations in which their inclusion would lead to a clear-cut violation of the direction-of-causality constraint," but the criterion was applied only "sparingly." The unfortunate result of this restraint is that very few of the actual regression equations appear to describe genuine causal relationships between independent and dependent variables. There are three general sources of doubt.

#### *Definitional Overlap*

Several of the dependent variables have definitions which seriously overlap those of the independent variables being used to explain them. For example, regression equation (6) finds "the character of agricultural organization" most strongly explained by the "size of the traditional sector." Yet "agricultural organization" runs from "communally owned agricultural lands in which the marketing of crops is only of incidental importance," up to large and viable "commercial owner-operated farms." The "traditional sector" is defined as "traditional subsistence agriculture in which . . . marketing of surpluses [is] of incidental importance"; it specifically excludes "modern commercial agriculture." Equation (7) in turn finds "the size of the traditional sector" to be negatively dependent on "the extent of dualism." Yet the low point of the "dualism" scale is "the largely agrarian society having an extremely small exchange sector." All we really know from these relationships is that there are many countries with a large share of the population in subsistence agriculture—whether this be classified as high on a "traditional sector" dimension or low on "dualism" or "character of agriculture" dimensions. If a high degree of correlation among

these dimensions holds any interest at all, it is only as a reliability check on the coding techniques.<sup>7</sup>

#### *Implausible Causation*

Many of the regression equations associate effects with very implausible "causes." One general difficulty is that variables representing short-run rates of growth are sometimes used to explain variables which express long-run levels. In equation (11) "the extent of literacy" (level for population aged 15 and over in 1958) is explained in part by "the rate of improvement in human resources" (measured by attendance ratio of school-age population in 1961). The two variables should be related—some countries have stronger traditions of mass education than others—but hardly causally, given that almost no one aged 15 or over in 1958 was of school age in 1961. In equation (12) the size in 1961 of the "indigenous middle class" is explained by the extent of social mobility in 1961.

A second difficulty is that some of the equations seek to explain very "hard" economic realities in terms of rather "soft" social or political rankings. For example, equation (8) explains the degree of "dualism" in the economy as partly dependent on "modernization of outlook." Equation (6) explains the "character of agricultural organization" partly in terms of "leadership commitment to development." Any implication that such transient attitudinal variables as "commitment" and "outlook" are significant causes of the size of the subsistence sector is better postponed to the day when the scientific bases of Couéism are more firmly established.

#### *Inferences of Mutual Causation*

Strong associations are far too readily interpreted as evidence of mutual causation. When two variables are strongly associated, it is not unexpected that each will figure prominently in a multiple regression equation for the other. In five cases where this occurs the authors do not hesitate to attrib-

<sup>7</sup> For other examples, see Eckstein, pp. 17-18.

ute this result to "reciprocal causation" or, more grandly, to a "mutually reinforcing feed-back relationship." For example, the variable "rate of development of human resources" is found in the equation explaining the "change in the degree of industrialization," while the industrialization variable is found in the equation explaining the human resources variable. In the authors' interpretation, not only does human investment remove constraints on industrialization, but industrialization creates a necessary demand for human investment. The authors offer no justification for automatically discarding three alternative interpretations of the observed association: the unidirectional impact of industrialization on human investment, the unidirectional impact of human investment on industrialization, or the impact of a third variable (e.g., governmental activism) on both phenomena. Nor do they consider whether the regression coefficients should be adjusted downward from their observed levels so that each does not reflect individually a full association that they are assumed to be causing together.

### III. *The Multipliers*

The culmination of the econometric model is the calculation of multipliers estimating "the relative impact of the various economic and noneconomic forces represented in the model upon the potential for economic development." Only the four prediction variables are assumed to have a direct impact on development potential; but all eighteen independent variables in the regression equations are assumed to have an indirect impact through the network of relationships to the prediction variables which those equations imply. There are two important problems with this multiplier analysis.

#### *Utilization of Cross-Section Relationships*

As econometricians have long recognized, there are substantial limitations on our ability to use cross-section analysis either to infer intertemporal relationships from the past or to project such relationships into the future. The inference of past time-series relationships requires that they have been

neither exaggerated nor obscured by such factors as: differences among countries in their economic institutions; differences in the environment—external and internal—within which different economies function; causal interactions among variables; disequilibria in relationships at the time of observation; and the impact of the relative performance of other nations. Even if the cross-section estimates actually do reflect intertemporal relationships, their use as a basis for projections into the future requires that no significant lags will exist in a country's response to a new level of performance; and that dynamic considerations—technological progress, changing consumer tastes, and a shifting structure of world demand at higher world income levels—can safely be ignored.<sup>8</sup> The authors do not address themselves to any of these thorny issues. It seems unlikely, however, that these requirements are met by very many of the relationships described in their model.

#### *Lack of Independent Basis*

The authors never attempt to establish a logical relationship between the discriminant analysis and the multipliers. In the earlier analysis all twenty-nine variables were examined for their relationship to development potential, and four of these were found sufficient to account for more than 97 percent of the variance between groups. Are we to take that result as meaningful, or must we go beyond it and see as significantly related to development a whole slew of variables that were already tried and found unable to account for more than 3 percent of the remaining variance? If we must go beyond the results of the discriminant function, would we not be more interested in the *direct* association of each of the variables with development potential itself, rather than the indirect association through the four variables in the discriminant function?

These questions could have practical implications, because there are wide discrepancies between the results of two analyses.

<sup>8</sup> See Eckstein, pp. 21-22, for possible applications to the model.

Two prediction variables—the degree of improvement of financial institutions and the degree of modernization of outlook—emerge from the multiplier analysis as still the most important influences on development potential. However, the other two prediction variables—leadership commitment to economic development and improvement in agricultural productivity—fall to eighth and tenth place, respectively. What advice, then, are we to give to the development planner who might wish to act on the basis of the values in the Adelman-Morris analysis? Should he seek to inculcate leadership commitment and to increase agricultural productivity, as suggested by the discriminant analysis? Or should he attempt to reduce dualism, to make physical overhead capital more “adequate,” to accelerate industrialization, and to increase investment in human resources—all of which have greater importance in the multiplier analysis?

Anyone tempted to employ the Adelman-Morris multipliers should examine the highly indirect relationships on which they are constructed. For example, the seventh largest multiplier is that for “rate of improvement in human resources.” Fully 61 percent of its value comes from making possible more rapid industrialization, and in turn fully 85 percent of the multiplier for industrialization comes from its impact on “improvements in financial institutions.” The latter variable carries the largest multiplier of all, and both the earlier ones are largely based on it. The implication then, is that one should attempt to influence a variable that has no directly-demonstrated relationship to development primarily because it may influence another variable that also has no directly-demonstrated relationship to development but that may influence a third variable which may have a direct relationship. It may be that “For want of a nail . . . a kingdom was lost,” but more than mere statistical association would be required to establish so tenuous a chain of causation.

#### IV. *Structured vs. Unstructured Models*

Throughout their joint work—including the factor analysis not covered in this dis-

cussion—Professors Adelman and Morris have relied on completely unstructured statistical techniques. In both the discriminant and the regression analysis, they “consciously avoided a priori specification of the functions” and instead presented the computer with a long list of variables and “let the data specify the model.” Thus, they are forced to judge variables entirely by their contribution to the explanatory power of the equation. The high multicollinearity in their data, however, renders this doubly dangerous. One typical result of collinearity is that coefficients of the variables have very large standard errors of estimate. Even when a coefficient is significantly different from zero (as are most of those in the Adelman-Morris regression equations), it is very unlikely to be significantly different from the coefficients of other closely-related variables. Thus, discriminant functions and regression equations are built on the bases of very narrow statistical differences, and their compositions are easily determined by random influences or small biases. Yet the authors take their results at face value and assume, for example, that the twenty-five variables not included in the discriminant function are important only insofar as they help explain the four variables that are included. When the data are highly collinear, however, the case is stronger than ever for using as much logic, theory, and prior knowledge as the investigators can muster in order to preclude a capricious selection among available explanatory variables.

The authors argue that their unstructured statistical techniques are necessary at this stage in the study of development, “since there are no firmly validated theories of the process of socio-economic and political change.”<sup>9</sup> Yet even if we have very few answers—established quantitative measures of the empirical relationships among variables—certainly we know enough about the logical relationships among variables and

<sup>9</sup> Note that the authors themselves must have had such a theory in mind when they required that a fast-growing country score high on five out of seven particular variables in order to qualify as having “high development potential.”

levels of analysis to be able to pose questions in an ordered manner. For example, the rate of growth of *GNP* per capita might be expressed as the rate of growth of *GNP* minus the rate of growth of population. The growth of total *GNP* might then be disaggregated into a weighted sum of either the growth of sectors or the growth of individual factors and their overall productivity. The growth of population might be expressed as the difference between crude birth rates and the rates of mortality and net emigration. Given these disaggregated measures of performance, development might be conceived as a three-stage process in which the country's "resources"—its natural wealth, physical capital, and human attributes—become translated through "economic action," governmental policies and private behavior, into "economic performance."<sup>10</sup> The point, however, is not the operational or theoretical merit of a particular framework, but merely that the absence of "firmly validated theories" does not prevent us from conceptualizing a process in a way that may lend some order—and greater meaning—to empirical analysis.

### V. Conclusions

There are many problems inherent in any attempt to separate out the strands of causation in the complex web of differences in the performance and potential of underdeveloped countries. Unfortunately, Professors Adelman and Morris have not adequately addressed themselves to these problems. There may be methods of direct observation that would permit the objective measurement of such subjective phenomena as "commitment" and "outlook"; but it is impossible to be confident that purely "judgmental" measures can ever be purged of the judge's preconceptions as to their causes and concomitants. Measures of the rate of growth of

per capita *GNP* in underdeveloped countries are themselves fraught with great uncertainty; but even the best of empirical measures would prove an inadequate foundation for the elusive concept of "development potential." Comparisons across countries can offer clues as to intertemporal relationships; but careful justification would be required before policy implications could safely be inferred from even the best-established of cross-section findings. No statistical technique can by itself separate out causes, effects, and joint effects or even the most important associations among closely correlated variables; but completely unstructured techniques, in which theory and prior knowledge are neither tested nor used to order the relationships are the least likely to produce meaningful results.

The greatest problems with the study, however, are not the inherent ones. In order to persuade us of the validity of their model as a whole, the authors would have to demonstrate first the soundness of both their discriminant and their regression analyses and then the legitimacy of their "multipliers" as a separate level of inquiry. Yet in neither statistical analysis do they establish either the independent measurement of the variables they are relating or the structural logic of the relationships they are describing. They also fail to make clear why the "multipliers," describing as they do indirect associations with development potential, should be taken more seriously than any direct associations that might be observed.

The general conclusion the authors offer is that "the important impediments to increasing capacity to develop" that emerge from the model "are social and political as well as economic." This is intrinsically a rather plausible statement. Unfortunately, neither its plausibility nor that of the specific relationships described has been augmented by any of the statistical results which comprise the "econometric model of development." On the contrary, our analysis of the model suggests that formidable obstacles remain to be confronted before the measurement of cross-sectional statistical associa-

<sup>10</sup> This would imply that very unlike processes are being mixed in the authors' attempt to compare in one function the relative importance of two "resources" (modernization of outlook and leadership commitment), one set of "actions" (improvement in financial institutions), and one element of "performance" (improved agricultural productivity).



tions can begin to challenge economic analysis—the creative interaction of general theory and the knowledge of particular institutions—as the fundamental tool of policy-making for economic development.

## REFERENCES

I. Adelman and C. T. Morris, *Society, Politics and Economic Development: A Quantitative Approach*, Baltimore 1967.

———, "Performance Criteria for Evaluating Economic Development Potential: An

Operational Approach," *Quart. J. Econ.*, May 1968, 82, 260-80.

———, "An Econometric Model of Socio-economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58, 1184-1218.

P. Eckstein, "Quantitative Measurement of Development Performance: A Critique of the Adelman-Morris Model," Discussion Paper No. 7, Center for Research on Economic Development, Ann Arbor, April 1969.

M. G. Kendall, *A Course in Multivariate Analysis*, London 1957.

# An Econometric Model of Development: Reply

By IRMA ADELMAN AND CYNTHIA TAFT MORRIS\*

The comments of Peter Eckstein and Sarah Berry highlight the major methodological issues in our cross-section analyses of socio-economic and political change in underdeveloped countries. Our purpose in this reply is to clarify those characteristics of our research design and data which distinguish them from more conventional approaches; and to note our agreement or disagreement with the major comments. Our reply touches on both our earlier and current work since "An Econometric Model of Development" represents only one segment of our research and treats only sketchily the underlying methodological issues.

Before proceeding further, we would like to emphasize the limitations we have imposed on our work. We have not claimed to offer a "reliable guide to policy-making." We have not attempted to explain differences in rates of change of per capita *GNP*. We certainly have not claimed at this early stage of our inquiries to have "separated out causes, effects, and joint effects." We have used an empiricist approach to construct an exploratory model and have described cause and effect relationships *implied by the model*. We can make no better correction of Eckstein's view of our purposes than to repeat the final statement of our paper:

... it is to be hoped that the various noneconomic and economic features of low-income countries that appear *in the present model* to have particularly strong effects upon development potential may give some indication to social scientists of the most profitable directions for research into the determinants of intercountry variations in overall capacity to perform well economically in the long run.<sup>1</sup>

\* The authors are, respectively, professor of economics at Northwestern University and professor of economics at The American University.

<sup>1</sup> Adelman and Morris (December 1968, p. 1212). The italics have been added.

## I. *Statistical Methodology: General Comments*

In designing our recent research on the interrelationship between economic and non-economic forces in economic development, we have used statistics as a tool for indicating the structure of the underlying phenomena rather than as a device for testing hypotheses. We have chosen this approach as most efficient in view of the extremely small quantity of validated knowledge concerning the "laws" governing the complex interactions between economic, social, and political forces in the process of economic development and modernization. Since the statistical theory for testing a given hypothesis against more than one alternative is at an early stage of development,<sup>2</sup> the hypothesis-testing approach is only efficient when an operationally testable hypothesis can be formulated, and the equally plausible alternative hypotheses capable of formulation are few in number and clearly distinguishable. Neither of these prerequisites appears to be satisfied in the field of research under discussion.

The literature of economics, history, sociology, and anthropology have produced a large number of reasonable, yet not well validated theories which would justify the choice of a multiplicity of alternative variables for the explanation of intercountry variations in economic development. Therefore we see no grounds for Eckstein's view that an a priori choice among competing hypotheses is either less arbitrary or more meaningful than a selection made by a reasonable and fairly robust empirical procedure. We need hardly add that we reject his recommendation that the problems posed by the

<sup>2</sup> For a discussion of this point, see E. L. Lehman who develops the theory of hypothesis testing against multiple alternatives for the case in which the alternative hypotheses can be parameterized. The latter condition is seldom met for the kind of alternative hypotheses relating to our present research.

presence of interrelated qualitative influences be solved by a priori specifications limited to directly measurable and (we suppose) weakly intercorrelated influences. It so vitiates the formulation of the problem that it in effect would require that we change the subject of our inquiry.

The only specific suggestion for an alternative model which Eckstein makes relates only marginally to the subject of our research. Furthermore, it by no means escapes the methodological problems posed by the existence of a multiplicity of interrelated influences in economic development. His three-stage model would face a serious problem of direction of causality since the private and government actions by which his resources are translated into economic development are, in fact, major causes of changes in the quantities of inputs. In addition, there are significant intercorrelations among the various actions as well as among the various elements of economic performance. Eckstein's failure to escape these difficulties reinforces our view that the problems he cites are inherent in our subject of inquiry and are not quite so amenable to solution as he supposes, even with a priori specified models.

It is our belief that there are many different ways to approach subjects of scientific inquiry, each with its own special defects and advantages as well as defects and advantages held in common with other approaches. We are not recommending the rejection of structured techniques nor objecting if others prefer to apply them to our subject. We simply consider them less suitable than unstructured techniques for our particular research given the general state of knowledge and our particular resources at this point in time.

### *The Robustness of our Research*

Mr. Eckstein and Mrs. Berry have accused us of arbitrariness particularly with respect to decisions we have made in preparing our various definitional schemes. Yet they are both certainly aware that arbitrary assumptions do not, in and of themselves, invalidate a model. The first

point we would like to stress in discussing the arbitrariness of our assumptions is that the choices we faced were not, as our critics imply, between arbitrary decisions we have made and nonarbitrary decisions we might have made. On the contrary, because of the unsatisfactory state of the general theory of social change and modernization, the fact that the concepts in which most theoretical discussions are couched are not operationally defined, the inadequacies of factual information on developing countries, and the prohibitive cost of direct measurement, our choices of assumptions for data construction invariably involved a selection among arbitrary alternatives. In making the necessary decisions, we relied upon the same considerations which influence the decisions of most other investigators; that is, a priori reasoning, judgments regarding data availability, our common sense, and an evaluation of the sensitivity of the results to our choices. However, whereas the various simplifying assumptions on which conventional economic composites are based are familiar, there are few precedents for the choices which we had to make in order to prepare our definitional schemes.

In discussing the arbitrariness of our assumptions, Eckstein in particular has employed some fine rhetoric to indicate that a number of decisions made in the course of our research *could* affect the results we obtained. Yet nowhere does he even hint that the evaluation of the extent of arbitrariness of any assumption in any model requires, among other things, a study of the *actual* sensitivity of the results of the model to reasonable alternative a priori specifications. We would not, ourselves, claim to have avoided completely assumptions to which our results will prove sensitive upon further testing. Nevertheless, since neither Eckstein nor Mrs. Berry has performed any sensitivity tests, despite the fact that the raw material for doing so was available to them, their confidence that the results would vary greatly with specific alternative definitions, measurement of the variables, and choices of procedures is rather surprising.

For example, Eckstein contends that we

TABLE 1—COMPARISON OF FINAL MULTIPLIERS  
RESULTING FROM THE USE OF DISCRIMINANT  
FUNCTIONS  $D_1'$  AND  $D_2'$ <sup>a</sup>

Variable	Multipliers		Ranks based on Multipliers	
	$D_1'$	$D_2'$	$D_1'$	$D_2'$
$X_{15}'$	2.212	2.309	1	1
$X_{31}'$	1.705	2.169	3	2
$X_{21}'$	1.832	2.13	2	3
$X_8'^b$	1.545	1.98	4½	4
$X_4'$	1.545	1.624	4½	5
$X_{24}'$	.913	1.062	6	6
$X_{33}'$	.873	.955	7	7
$X_{49}'$	.727	.940	9	8
$X_{28}'$	.754	.884	8	9 <sup>c</sup>
$X_7'^c$	.095	.823	13	10
$X_{25}'$	.685	.797	10	11
$X_{50}'$	.414	.536	11	12
$X_{27}'$	.347	.407	12	13
$X_{20}'^c$	.033	.288	15	14
$X_{47}'$	.052	.018	14	15
$X_{25}'$	-.232	-.300	17	16
$X_{46}'$	-.253	-.310	18	17
$X_6'$	-.168	-.320	16	18

<sup>a</sup> See text for explanation.

<sup>b</sup> One needs to add to this the change in physical overhead capital

<sup>c</sup> Agriculturally related.

ought to have used the first discriminant function ( $D_1'$ ) obtained in our original discriminant analysis rather than the second ( $D_2'$ ) as a starting point for our econometric model. Table 1 presents the multipliers obtained using both  $D_1'$  and  $D_2'$ . It indicates that the ranking and order of magnitude of the multipliers of the first nine variables are virtually unaffected. The primary effect is to reduce the already low multipliers of the agricultural variables while leaving the order of magnitude and ranking of the other multipliers essentially unchanged. This was to be expected, since the replacement of the index of improvements in physical overhead capital by the indicator of improvements in agricultural productivity is the principal way in which  $D_2'$  differs from  $D_1'$ . The finding that the final multipliers are fairly insensitive to alternative specifications of the structure of the present econometric model is further reinforced by the virtual identity of the multipliers obtained through the stepwise procedure with those which result from

the application of the techniques of canonical analysis to the same data. (See Adelman, Geier, and Morris, May 1969.) As is well known, canonical analysis is equivalent to a limited information simultaneous equation procedure. Thus, the ranking in the model of the important forces affecting development potential appears rather insensitive to the choices made in constructing both the discriminant function and the model.

Mr. Eckstein discusses certain variables in our discriminant function which he implies entered merely by chance. We have tested the robustness of the choice of variables in our final discriminant function to alternative statistical criteria for adding and deleting variables. An alternative stepwise discriminant procedure which uses the  $F$ -ratio test instead of the generalized distance criterion yields results identical to  $D_2'$ .

Mrs. Berry criticises the arbitrariness involved in our choice of scoring scale. In our reply to O. T. Brookins' criticism of our scoring scheme we presented factor analytic results based upon three alternative transformations of our numerical data inputs (logarithmic, squaring and rank), (see Adelman and Morris, forthcoming). Not only does the mean absolute difference in factor loadings never exceed .08, but in only one instance out of 267 for the three test analyses is the loading sufficiently different to lead to the assignment of a variable to a factor different from that to which it originally was assigned.

How many conventional econometric models perform as well? In different estimates of consumption functions with similar data but over different (though overlapping) time periods, the marginal propensity to consume out of quarterly disposable income was found to vary between .1 and .6, the long-run marginal propensity to consume to range from .60 to .98, and the estimated coefficient of past peak income to be anything from a significant value of  $-.625$  to an insignificant value of  $-.241$ . (See Griliches, Maddala, et al.; Zellner; Duesenberry et al.) In alternative estimates of production functions for the United States the elasticity of output with respect to labor was estimated

to be as different as .18 and 0.95 despite the use of the same functional form. (See Klein and Goldberger, Valavanis-Vail.) In the study of investment behavior, estimates of the elasticity of capital stock with respect to relative price range anywhere from a value of .8 to a value of .09, and with respect to output anywhere from a value of .6 to a value of 1.0. (See Eisner.) Yet, all these widely divergent estimates were obtained by well recognized econometricians sharing a common theoretical framework and applying similar econometric techniques to analogous bodies of empirical information. Thus, a comparison of the sensitivity of the results of our analyses with the sensitivity of results of other econometric work indicates that our research is quite robust, not only in absolute terms but also by the standards of more conventional econometric studies.

Since we have used unstructured techniques of statistical analysis for the purposes of hypothesis-making, there is yet another test of validity applicable to our research, namely, whether the hypotheses suggested by the various statistical manipulations of our data are mutually consistent. After all, the use of inductive techniques is likely to provide a good foundation for theory construction only if the hypotheses suggested by different analyses are both sensible when viewed in isolation and mutually compatible.

While it is somewhat too early in our work to apply this criterion of robustness with great stringency, we can cite several general propositions whose validity is reinforced by the several different pieces of empirical analyses. First, our work shows that in the analysis of economic development, the division between endogenous and exogenous variables along traditional disciplinary lines can produce misleading results. This conclusion emerges from our socio-political factor analyses published in 1967 in which social and political forces explain 79 percent of intercountry differences in economic growth rates for low-level underdeveloped countries, 49 percent for the intermediate-level, and 61 percent for the high-level (pp. 187, 213, 241). It is reinforced by our discriminant analysis of May 1968 in which only 2 out of 4 vari-

ables explaining variations in development potential are economic (p. 277). It is further strengthened by our paper in this *Review* in which we found that 67 percent of the relationships of the economic variables are with noneconomic variables and that, of the multipliers which exceed .80, only 4 are purely economic (p. 1201 ff).

Second, in studying economic performance, one should distinguish among a minimum of three different typologies of underdevelopment. In our factor analysis of 1967 we found systematic differences between the patterns of interactions among economic and noneconomic influences upon rates of economic growth for the three subsamples representing successive levels of development (pp. 6-7, and Chaps. V-VII). This evidence is reinforced by our canonical analysis of May 1969 in which we found systematic variations in relationships between policy instruments and policy goals for countries at different levels of development (pp. 416-23).

Third, for an average underdeveloped country, the connections between economic and sociological forces are more intimate than the connections between economic and political forces. In our 1967 full-sample factor analysis of per capita *GNP*, 40 percent of intercountry differences in income per head were explained by social and socio-economic variables and only 26 percent by political forces (p. 151). In our econometric model in this *Review*, 90 percent of the interconnections between economic and noneconomic variables were between variables in the economic block and variables in the social and socio-economic blocks and only 10 percent were with political forces (pp. 1201-3).

Fourth, the most important political variable for economic performance is the nature of leadership attitudes towards economic development. This conclusion emerges from our 1967 socio-political factor analysis of rates of economic growth for high-level underdeveloped countries in which extent of leadership commitment by itself accounted for approximately 59 percent of intercountry differences in rates of economic growth. It is reinforced in our May 1968 discriminant analysis in which leadership commitment in

economic development was the only political variable in the discriminant function. This conclusion is also emphasized by the multiplier results of our econometric model in which the extent of leadership commitment to economic development was the only political variable to enter the "top ten" list.

Thus there would appear to be a fair extent of congruence among hypotheses derived from various different statistical analyses performed by us with our data.

We have been quite aware that the appropriate scientific procedures for evaluating a model include fairly extensive testing of the sensitivity of the results to reasonable changes in its specifications. Consequently, an important part of our past and present research has been devoted to such tests.<sup>3</sup> Preliminary tests suggest that our results are not very sensitive to reasonable variations in the specification of our

<sup>3</sup> In addition to the sensitivity tests reported upon in the text, while preparing our data we tested the sensitivity of our factor analyses to changes in the definition of selected variables for which we had a choice of feasible possibilities. In defining the degree of centralization of political power in two different manners, for example, we found our results substantially unchanged. Compare the factor analysis results based on the first definition in our 1965 paper (p. 562) with those based on the second definition in our 1967 study (p. 58 and p. 151). We defined improvements in human resources in several ways: one stressing primary education; one emphasizing primary education for low-level countries and secondary education for more developed countries; and one based on a combined index of secondary and higher education. Since our results for the different samples did not change significantly with these variants, we chose the third alternative as the simplest to prepare and as reasonable a priori as the others.

We have also tested the sensitivity of our factor analyses to changes in the loss of information due to our grouping procedures by varying the number of categories for several indicators (e.g. changes in per capita GNP, gross investment rates, crude fertility rates) for which we had point estimates for the majority of countries in the sample. Since the results proved relatively invariant to changes in the number of class intervals, we chose that number of brackets which gave us both reasonable discrimination among countries and reasonable reliability of the judgmental information used to make interval estimates for countries without point data.

data,<sup>4</sup> to our scoring techniques, or to the detailed structure of our econometric model. In addition, different statistical approaches appear to lead to mutually reinforcing conclusions.

## II. *Data: Theory and Taxonomy*

Our research efforts of the past six years have been directed to the interaction of non-economic and economic influences in the economic development of underdeveloped countries. We have chosen to approach this subject through quantitative analyses of those social, political, and economic characteristics of underdeveloped countries which, according to a priori reasoning, descriptive historical studies, and the relevant social science literature seemed likely to affect the economic performance of developing countries. Our subject of inquiry and choice of approach required us, therefore, to devise measures of qualitative non-economic and economic institutional characteristics for which the full range of underdeveloped countries could be ranked.

Direct measurement of the relevant influences was rarely possible because of overwhelming data deficiencies. Choices of indirect measurements were constrained by the small number of feasible alternatives for the measurement of each influence. Finally, in choosing to include countries for which all data were poor (usually those at the lowest development level) and thereby to avoid a major bias present in most cross-section studies of developing countries, we were restricted to measurement by intervals sufficiently wide so that descriptive and judgmental information could be used to make reasonably reliable interval estimates.

Our critics have failed to distinguish clearly

<sup>4</sup> As would be expected, our results have proved sensitive to changes in definition which involved major conceptual revisions. Our original measure of "intensity of nationalism and sense of national unity" associated with leadership characteristics. We replaced it in response to criticisms by Everett Hagen with a measure of the "degree of national integration and sense of national unity" which associated with indicators of social development. Compare the results cited in fn. 3. (See also p. 55 of our 1967 study.)

between difficulties peculiar to our data and those which our data share with conventional economic and psychological data. We shall therefore discuss what are the essential attributes of our data. In particular, we shall point out their multidimensionality, their qualitative and ordinal character, and our translation of ordinal ranks into cardinal scores. We will then consider in some detail the conceptualization and definition of qualitative indicators.

#### *Multidimensionality of Data*

Almost all our variables are composites of several component elements. Consequently, as for all composites including such indices as *GNP*, cost of living, and intelligence, either a formal theory or an a priori conceptualization of the phenomenon is required in order to construct an aggregate index. What distinguishes our data from conventional aggregates is not the use of theory in measurement, but the fact that the a priori judgments we have made in aggregating the components of each index are less securely based in explicit theory than is the case for conventional composites.

#### *Qualitative Character of Data*

Most of our indicators are either purely qualitative or partly qualitative variables in which descriptive information and expert opinion were used to classify individual countries. The ranking of individuals by expert opinion is a completely accepted procedure in psychology; nevertheless, there are several pitfalls in its application. Experts may interpret concepts and definitions differently; their opinions about the facts may differ; and they may be biased. We attempted to minimize the effects on our data of these sources of variation in expert opinion; we used interviews to probe differences in expert reactions to both our definitions and the individual country classifications. Moreover, we attempted to define our concepts operationally prior to the consultation of experts. By using relatively objective criteria for ranking individual countries, we sought to eliminate systematic errors in our

data arising from expert biases of various sorts.<sup>5</sup> With respect to the extent of leadership commitment for economic development, for example, these criteria included: the presence or absence of cooperation between the government ministries, planning agencies, and national banks, engaged in central guidance of the economy in actions to promote growth; the presence or absence of leadership measures for institutional changes to promote economic growth (e.g., land reform); the presence or absence of full-time planning groups engaged in both planning for and execution of national plans. The application of these criteria permitted us to rank the majority of countries with the use of descriptive information. Experts were consulted only in order to resolve doubtful cases. Our critics would do well to study carefully the criteria for classifying countries specified in our actual definitional schemes. Mr. Eckstein appears to have based his criticism on his intuitive view of the concepts suggested by the titles of the indicators, while Mrs. Berry has based her criticisms on her reactions to our general statements preceding the definitions.

While it is obvious that we have not succeeded in eliminating all expert biases of various sorts, we reject Eckstein's contention that there is a *systematic* bias in our data due to the tendency of experts to assume that good growth performance is achieved by committed leadership. The evidence supporting our view<sup>6</sup> is that the simple correlations between rates of growth of per capita *GNP* and most of our socio-economic and political variables are significantly different for the subsamples representing different levels of development; the correlations between economic growth rates and leadership commitment to development, for example, are .59, .16, and -.02, for the high, intermediate, and low groups of countries, respectively. To maintain that the

<sup>5</sup> For a useful discussion of the nature and varying importance of different kinds of errors in measurement, see our 1967 study (pp. 199ff.)

<sup>6</sup> It is readily available in the Appendix of our 1967 study.

barely statistically significant correlation of .35 for leadership for the full sample is due to bias, it is therefore necessary to assert that the bias of experts varies significantly with the level of development—a rather implausible contention. To reemphasize our basic point, we do not claim that some or even many experts were not biased in various ways but only that these biases do not appear to have systematically distorted our results.

#### *Cardinal Translation of Scores*

The third special characteristic of our data, the assignment of numerical scores to ordinal data is criticized by Mrs. Berry. We have discussed this in detail in our forthcoming paper, but let us note here again that our procedure is an accepted and conventional one in the field of psychology. It is no more arbitrary than a host of other commonly accepted econometric simplifications and abstractions. The practical point, surely, is the sensitivity of the results to the choice of scale; as already indicated, our preliminary tests suggest that the structure of interrelationships underlying our data are surprisingly insensitive to wide variations in methods of scoring.

It should be noted that the choice of a cardinal scale with equal intervals even for a measurable characteristic such as school enrollment ratios is in no way free of possible bias. It involves a selection among such alternatives as raw scores, squares, logarithms or other transformation. This choice must be based in part on the presumed relationships between school enrollment ratios and whatever underlying characteristic (such as the quantity of skilled human resources) they are presumed to measure as well as on the presumed relationship between that underlying characteristic and the purpose of the investigator's inquiry. In our own data preparation, a priori reasoning together with the other sources of knowledge available to us were used in order to choose breaking points between intervals.

#### *The Conceptualization of Qualitative Indicators*

The major difficulties of our data lie, in

our opinion, neither in the fact that they are qualitative and based on judgmental information nor in our translation of ordinal ranks into cardinal scores. The core difficulties lie rather in the conceptualization and definition of phenomena for which there are no theories as precise and acceptable as those on which qualitative economic measures are based. There are few available theories suited to defining continua along which the full range of underdeveloped countries can be ranked for most of the characteristics in our study. In addition, there are little data to fit such theories as are current. Furthermore, our efforts at definition were constrained by the frequent necessity to use indirect rather than direct evidence. As generally recognized, the use of indirect evidence requires inferences concerning "presumed connections, usually causal, between what is directly observed and what [a] term signifies . . ." (Abraham Kaplan, p. 55).

The process of conceptualization and definition which we followed in constructing our composite indicators is a procedure well tried in the history of scientific inquiry—in the physical as well as the social sciences. Abraham Kaplan in *The Conduct of Inquiry* has the following to say about the derivation of scientific concepts in the early stages of scientific inquiry:

In short, the process of specifying meaning is a part of the process of inquiry itself. In every context of inquiry we begin with terms that are undefined—not indefinables, but terms for which that context does not provide a specification. As we proceed, empirical findings are taken up into our conceptual structure by way of new specifications of meaning, and former indications and references in turn become matters of empirical fact . . .

What I have tried to sketch here is how such a process of "successive definition" can be understood so as to take account of the openness of meaning of scientific terms. *For the closure that strict definition consists in is not a precondition of scientific inquiry but its culmination.* To start with we do not know just what we mean by our terms, much as we do



not know just what to think about our subject-matter. We can, indeed, begin with precise meanings, as we choose; but so long as we are in ignorance, we cannot choose wisely. It is this ignorance that makes the closure premature. [p. 77-78, Italics added]

In the preparation of our qualitative multidimensional indicators, we began with a priori definitions. Next, we studied the descriptive data in order to see how well actual country situations fit our formulation of the concept. The inadequacies of the initial fit were then used to reformulate the concept to fit better the characteristics of the real world. We then consulted expert opinion and again reformulated the definitions. We continued this process of confronting successive reformulations with information on actual country situations until we were able to classify the 74 countries in our sample with reasonable confidence.

#### *Example of the Dualism Indicator*

The construction of our indicator of the extent of socioeconomic dualism illustrates well the interaction of conceptualization and testing against the actual world which took place in the preparation of our data. We choose this indicator for discussion because Mrs. Berry criticizes it in her comment.

In constructing our dualism indicator, our major conceptual difficulty was the ranking of two intermediate categories, for we characterized both ends of the continuum by the absence of marked dualism. Dualism is absent at the lower end by reason of the overwhelming predominance of nonmarket subsistence agriculture combined with the extremely limited growth of a market-oriented sector; while the upper end is not markedly dualistic because the intermingling of modern and traditional elements throughout the economy resulted in the absence of a clearcut geographic cleavage between a market-oriented sector and a distinct predominantly nonmonetized traditional sector. In our initial a priori definition, we conceived of a single intermediate category characterized by sharp social, economic, and technological contrasts between a geo-

graphically distinct and important plantation, extractive, or industrial sector using advanced technology and a major subsistence nonmarket agricultural sector.

When we confronted our original a priori definition with descriptive data from country studies, we immediately discovered that it made no provision for countries with a geographically quite distinct, important, and rapidly growing market sector characterized by the predominance of indigenous cashcropping with conventional techniques. The countries in which peasant cashcropping dominate the market sector differ significantly from those in which extractive or plantation enterprises using advanced technologies dominate the market sector. In particular, resource flows in the latter countries are limited primarily to the intermittent labor flows, while in the former countries, resource flows included movements of land, labor, and capital.

In reformulating the dualism indicator we therefore gave weight to the extent of intersectoral resource flows as well as to socioeconomic and technological contrasts and to the presence or absence of a relatively clearcut cleavage between market and nonmarket sectors. We defined two intermediate categories for countries with both a geographically distinct and important market sector and a relatively large predominantly nonmonetized traditional sector. We gave higher scores to countries in which: (a) the market sector was characterized by the predominance of indigenous cashcropping; (b) less marked contrast existed between technologies and styles of life; (c) more extensive economic interaction occurred between traditional and modern sectors. We did so because of our judgment that these countries were further along the path to the pervasive intermingling of modern and traditional elements throughout the economy characteristic of the upper end of the spectrum summarizing the extent of dualism.

The reformulated four-way classification scheme enabled us to rank the great majority of countries in our sample. An examination of additional sources of country information led to only minor reformulations of

the category descriptions which were necessary to take better account of the diversity of individual country situations.

In criticizing our dualism indicator, Mrs. Berry did not mention the actual definitional scheme which consists of the set of descriptions of the categories just discussed. Consequently, her attack on individual country classifications is most misleading. She does not state anywhere that her criticism is primarily in terms of a criterion, the extent of foreign finance and direction, *which is not even mentioned in the definition*. It is true that in our general comments we point to foreign finance and direction as typical of countries in the lower-ranking categories of dualism. However, because of those country situations Mrs. Berry cites, we explicitly did not include foreign finance and direction as a classificatory principle in the definitional scheme.<sup>7</sup>

On the general subject of the reasonableness of our individual country classifications, it is self-evident that, in making 3,034 classifications, we have probably made some misclassifications which will be pointed out to us in the course of time. It is nevertheless our opinion that the number and size of these errors are significantly less than in conventional quantitative economic data on underdeveloped countries.

<sup>7</sup> The particular countries Mrs. Berry mentions as all having foreign financed and directed sectors can be classified reasonably well in terms of the criteria given in the definition. To illustrate with our ranking of South Africa, Rhodesia and Uganda: according to our information, South Africa was the only one of the three which in 1960 did not have at least a moderately definite cleavage between an important industrial and/or agricultural exchange sector and a single relatively large geographically distinct predominantly nonmonetized sector; it is consequently the only one of the three which met the stated criteria for category a. The minus score to South Africa was given because of the existence of relatively small predominantly nonmonetized agricultural pockets in its economy. Rhodesia and Uganda differed in 1960, according to our information, because the former had a significantly more important industrial, mining, and agricultural exchange sector than the latter and a relatively less important geographically distinct predominantly nonmonetized traditional sector. On both counts, Rhodesia was placed in category b rather than c.

### III. *The Discriminant Function*

Mr. Eckstein's criticisms of our discriminant analysis consist of an attack on the original grouping by development potential; objections to our procedures in obtaining the final discriminant function; criticisms of the variables in the function; and criticisms of our interpretations. We will consider each in turn.

#### *Grouping by Development Potential*

The focus we chose for our discriminant and regression analyses was economic development in the broad sense of institutional transformations which create a capacity for widespread and continuous economic growth. For our purpose, the use of per capita *GNP*, however precise, was inappropriate since, as is well known, the raising of the growth rate of per capita *GNP* can take place, even over periods of a decade or so, without being accompanied by fundamental institutional change in other than a sporadic and limited way. In choosing to devise a less precise but more appropriate classification of countries, we considered that, *for our purpose*, a "gain" of information on the breadth of recent economic change would more than compensate for any "loss" of information on per capita *GNP*.<sup>8</sup>

The classification we devised is akin to a more recent effort by Paul Clark and Alan Strout to construct measures of economic performance of developing countries. To quote their recent paper:

The concept of a developing country's economic or development performance can cover many different phenomena. The growth process consists of a complex and convoluted series of cause and effect relationships ranging from plans and rhetoric to eventual improvements in the level and conditions of living. Performance can apply to the effectiveness of pursuing either intermediate or ultimate goals and can be observed at a wide variety of points in the development process.

<sup>8</sup> It should be noted that we had already devoted an entire book to a factor analytic study of intercountry differences in rates of growth of per capita *GNP*.

Thus performance can include such diverse aspects as (a) choice of effective policies to promote desired growth and development; (b) effective implementation of policies chosen; (c) mobilization of additional resources for growth and development; (d) efficiency of use of both domestic and foreign resources; (e) structural and other changes necessary for longer-run economic, political, and social growth; (f) final effects on level and distribution of income and welfare. [p. 3]

In grouping countries by development potential, we sought, as do Clark and Strout, to distinguish between countries in which registered increases in per capita *GNP* were narrowly based and those in which they were accompanied by widespread economic change. Our three-way classification of countries by development potential would rank countries somewhere between the two ranks (by growth and policy performance) on the Clark and Strout index since our grouping is based on both kinds of performance criteria.<sup>9</sup>

We agree with Eckstein that independent evaluations of country development potential comparable to those used, for example, by Clark and Strout would be preferable to the wide range of indirect evidence on which our classification by development potential was based. However, because of the breadth of information we took into account, we see no particular reason to expect that intensive field research would produce country rankings very different from those we obtained.

#### *Procedures for Obtaining Discriminant Function*

The iterative procedure we used to obtain our second discriminant function is attacked by Eckstein as unsound. We disagree. Our aim, as explained in our original article, was to seek the best means for classifying countries into performance groups using a small number of performance characteristics. The procedure we followed is that developed by E. Forgy for deriving the optimal partition among a group of objects. This procedure

involves starting with a partition (in our study, we used the best partition we could devise on the basis of the external evidence on country performance available to us) and then reassigning each object to the group to whose center of gravity it is the closest. In Forgy's procedure, the value of the newly formed partition with higher value no longer results.<sup>10</sup>

#### *The Variables in the Discriminant Function*

We can hardly disagree with several of Eckstein's comments on the limitations of the variables contained in the discriminant function. First, we certainly have *not* obtained "independent causes of faster growth rates." The role of the variables in the discriminant function was merely "to classify countries into performance groups, using a relatively small number of performance characteristics." Second, in both the present article and the original discriminant study, we ourselves have been at pains to emphasize the fact that statistical associations may represent causality in either direction or be the result of common forces affecting both dependent and independent variables. We would, however, be fascinated to find any meaningful statistical analysis in which common underlying forces do not influence both sides of the equation.

Third, as we pointed out in both the present article and in our May 1968 article on the discriminant analysis, each variable in the discriminant function certainly represents closely related influences not directly measured by the variable itself as well as those directly represented. This possibility is surely self-evident for any statistical analysis in which the influences measured are proximate ones only. In our view, the appropriate method for understanding better the influences represented by a particular variable in a set of statistical results is to analyze carefully those variables which are the next-best alternatives to the particular variable chosen at each step in the analysis.

<sup>9</sup> Exact comparisons cannot be made because the time period for their measures is different from that for ours.

<sup>10</sup> For a good general discussion of the problems and procedures involved in deriving optimal classifications of *N* objects into categories, see H. P. Friedman and J. Rubin and the references cited therein.

We have recently rerun the discriminant function with a stepwise program having an  $F$ -ratio criterion for the entry and rejection of variables and giving  $F$ -ratios for all variables at each step.<sup>11</sup> The index of modernization of outlook is consistently the first variable to enter the discriminant function. As would be expected, the variables which are the next-best alternatives (as indicated by high original  $F$ -ratios combined with relatively large decreases in  $F$ -ratios as a result of the entry of the modernization index) represent a complex of closely interrelated economic and social influences, including the level of development of financial, agricultural and industrial structures, the extent of social mobility, the importance of the indigenous middle class, and the extent of secondary and higher education. Thus, the index of the modernization of outlook is undoubtedly closely related to basic economic influences.<sup>12</sup> Nevertheless, since the program was free to select direct economic measures and did not do so, it seems reasonable to interpret the choice of the modernization index as indicative of some independent influence of the noneconomic forces represented by the measure itself. With as many as 74 observations, and a difference of 1.3 in  $F$ -ratio, there is no reason to suppose that the choice of the modernization index rather than the leading economic index was the outcome of purely random influences.

The fourth point about the variables in the discriminant function on which we agree with Eckstein is that a positive relationship between improvements in agricultural productivity and growth performance (actual or potential) was to be expected since agriculture forms a very important sector in less-developed countries.<sup>13</sup> One would simi-

larly expect improvements in industrial technology to be closely related to growth performance in advanced economies. But, what is the point? Is one to eliminate variables for which there is a strong a priori reason to expect a high intercorrelation? Is not an identification of such variables the main function of a priori theorizing?

#### IV. *The Stepwise Regression Model*

We now turn to the specific criticisms by Mr. Eckstein and Mrs. Berry of the nature of the stepwise procedure we employed and of the interpretation of our stepwise results.

##### *Statistical Procedure*

The particular stepwise regression procedure we applied is the one recommended by Draper and Smith for deriving *optimal* empirical relations from a body of data. Specifically, at each step in the procedure, a reexamination is made of the contribution of each variable in the equation as if it had been the most recent variable entered. This is done by computing the  $F$ -ratio for each variable in the regression and comparing it with a preselected percentage point of the appropriate  $F$  distribution. Any variable whose contribution is no longer significant is removed from the equation. Thus, the regressions obtained represent the best, in a least squares sense, summary of the multivariate associations present in the data. Despite Fisher's dictum against "fishing expeditions," this use of regression analysis to indicate and summarize the underlying multivariate associations seems to us a perfectly legitimate and technically more powerful extension of such measures of two-by-two associations as contingency tables and correlation matrices which are often used to investigate the presence or absence of interactions.

##### *Interpretation of the Model*

The major problems of interpretation raised by our critics are (1) the problem of direction of causality, (2) the problem created by the fact that any variable entering a

<sup>11</sup> Unfortunately, no program giving information on all variables is available for the stepwise discriminant procedure which applies the criterion of the greatest contribution to explaining the generalized distance between group means.

<sup>12</sup> See Adelman and Morris (1967, row 4 of Appendix Table A-1) (correlation matrix for social, political and economic indicators: full sample).

<sup>13</sup> The simple correlations between rates of change of per capita *GNP* and improvements in agricultural productivity for the high, intermediate, and low samples,

respectively, are .78, .28, and .44. Thus the correlation is higher for countries with smaller agricultural sectors.

statistical equation may be a surrogate for influences other than those directly measured by the variable, and (3) the problem of interpreting cross-section results as representing historical processes.

Our original decision to apply sparingly the direction-of-causality criterion for omitting variables was based, as we stated in the paper, on the general state of ignorance regarding the potential interactions among the variables included in our study. We nevertheless agree that greater restrictiveness is desirable and have rerun the stepwise analyses with more restricted lists of potential causal indicators established by a priori reasoning based on an extensive examination of the relevant social science literatures. These lists do not, however, exclude cases of genuine mutual interaction since the exclusion of an interacting independent variable produces results which are biased by the omission of relevant influences.

We reject Eckstein's implication that an a priori selection among closely related causal variables yields a meaningful solution to the problem posed by variables which are proxies for influences other than those directly represented. The only effective method for understanding these influences is to study explicitly at each step in the analysis the *F*-ratios of the omitted variables, the net correlations between the dependent and the omitted variables, and the pattern of actual country residuals. These pieces of information and the matrix of simple correlations together can give a reasonable idea of the forces for which each variable in the regression equations may be presumed to stand.

We agree with Eckstein and stressed in our book (pp. 265-66) that, strictly speaking, cross-section analyses cannot be interpreted to have any specific time dimension. Consequently, an important part of our current research consists of cross-section studies for several different historical periods and individual country time-series studies of the process of economic development.

#### V. Conclusion

In conclusion, we firmly believe that the

empirical approach to scientific inquiry forms an extremely useful *complement* to the usual structured techniques of the economist for the study of the process of economic development. We reject the view that, in the investigation of the complex interactions among noneconomic and economic influences in development, a priori models are likely to produce much more meaningful results than models based upon empirically determined regularities. The multiplicity of plausible hypotheses regarding causes of variations in the capacity of low-income countries to perform economically is so great and the body of validated knowledge so small that a priori choices among candidate variables are inevitably quite arbitrary and tend to reflect primarily the disciplinary preferences of the investigator. We do not deny, and indeed are ourselves convinced that models with a priori specified functions are very important for the analysis and understanding of particular subsets of interactions about which a reasonable amount is known.<sup>14</sup> We only propose strongly that alternative empirical procedures may be more fruitful for the initial exploration of those wider interactions involved in economic development which, by crossing disciplinary lines, involve relationships about which very much less is known.

The evaluation of any methodology involves an estimate of its success in achieving its purpose. As we stated clearly in the introduction of the paper under discussion, the purpose of our work has been to provide a good starting point for a process of experimentation in which further empirical testing and theoretical reasoning must interact. This we believe that we have done in view of the demonstrated robustness of our research and in view of the mutual compatibility of the hypotheses which emerge from our different analysis.

#### REFERENCES

- I. Adelman, M. Geier, and C. T. Morris, "Instruments and Goals in Economic Develop-

<sup>14</sup> These subsets are usually encompassed by a single discipline.

- ment," *Amer. Econ. Rev. Proc.*, May 1969, 59, 409-26.
- I. Adelman and C. T. Morris, "An Econometric Model of Socio-Economic and Political Change in Underdeveloped Countries," *Amer. Econ. Rev.*, Dec. 1968, 58, 1184-1218.
- , "A Factor Analysis of the Interrelationship Between Social and Political Variables and Per Capita Gross National Product," *Quart. J. Econ.*, Nov. 1965, 69, 555-78.
- , "Performance Criteria for Evaluating Economic Development: An Operational Approach," *Quart. J. Econ.*, May 1968, 82, 260-80.
- , *Society, Politics, and Economic Development*, Baltimore 1967.
- , "Performance Criteria for Evaluating Economic Development Potential: A Reply," *Quart. J. Econ.*, forthcoming.
- P. Clark and A. Strout, *Aid, Performance, Self-Help, and Need*, Agency for International Development Discussion Paper No. 20, Washington 1969.
- N. R. Draper and H. Smith, *Applied Regression Analysis*, New York 1966.
- J. S. Duesenberry, O. Eckstein, and G. Fromm, "A Simulation of the United States Economy in Recession," *Econometrica*, Oct. 1960, 28, 749-809.
- R. Eisner, "Investment and the Frustrations of Econometricians," *Amer. Econ. Rev. Proc.*, May 1969, 59, 50-79.
- E. Forgy, "Cluster Analysis of Multivariate Data: Efficiency vs. Interpretability," *Biometrics*, No. 1 (abstract)
- H. P. Friedman and J. Rubin, "On Invariant Criteria for Grouping," *Amer. Statist. Assn.*, Dec. 1967, 78.
- Z. Griliches, G. S. Maddala, R. L. Nelson, and N. Wallace, "Notes on Estimate Quarterly Consumption Functions," *Econometrica*, July 1962, 30, 491-500.
- A. Kaplan, *The Conduct of Inquiry: Methodology for Behavioral Science*, San Francisco 1964.
- L. R. Klein and A. S. Goldberger, *An Econometric Model of the United States, 1952*, Amsterdam 1955.
- E. L. Lehman, *Theory of Testing Hypotheses*, Lectures, Mimeographed Notes by J. B. L. Berkeley, no date.
- S. Valavanis-Vail, "An Econometric Growth, U.S.A. 1869-1953," *Amer. Econ. Rev. Proc.*, May 1955, 45, 208-212.
- A. Zellner, "The Short-Run Correlation Function," *Econometrica*, Oct. 1960, 28, 552-67.



