Title
The Social Thresholds: Patterns Associated with Economic Growth

Author
Clemhout, Simone

Publication Date
1967-04-01

Peer reviewed

COMMITTEE ON ECONOMETRICS AND MATHEMATICAL ECONOMICS

Working Paper No. 106

THE SOCIAL THRESHOLDS: PATTERNS ASSOCIATED WITH ECONOMIC GROWTH

by

Simone Clemhout

April, 1967

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

University of California

Berkeley, California

ACKNOWLEDGMENTS

We are extremely grateful to Dale Jorgenson, Muthu Subramanian, N. V. Sovani, and Henry Y. Wan, Jr. for comments and suggestions. All errors of omission and commission are, unfortunately, fully ours. Generally, the staff of UNRISD provided a highly stimulating environment.

Thanks are due to the Computer Centers of CERN, The University of Geneva, and The University of California, Berkeley, for providing the necessary facilities for the processing of the data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES AND CHARTS

PART I

1. Introduction

In the planning of development for growth, the approaches of economists, sociologists, psychologists, and anthropologists are usually conditioned by the orientations of their respective disciplines. Cross-fertilization of ideas has begun in a number of areas but more work is needed to further this fruitful tendency.

Present-day economic theory is advanced in analyzing developed economies and rather rudimentary in treating the problems of developing nations. There are many differences between the behavior of developed and developing economies. Perhaps one of the most important causes for such a distinction is that certain sociological variables play a much more significant role in a developing economy than in a developed one.

There are two alternatives open to economists specializing in economic development: either relegating sociological variables to play the role of exogenous magnitudes or attempting to construct an integrated model encompassing economic as well as social variables. It is along the latter line that the present study is conducted.

The main purpose of this study is:

1. To investigate relationships between certain sociological forces and economic growth.
2. To examine whether there exist levels below which the impact of certain social variables is critically significant. These levels we call "thresholds." This term will be fully explained later.

Prior to dealing with the main topic, let us briefly survey theories of economic development related to the idea of a threshold.

According to the well-known Rostovian hypothesis,[1] the so-called economic "take-off" constitutes a necessary although not sufficient condition for sustained economic growth. But even before the take-off stage, certain "pre-conditions" have to be met and certain infra-structure (e.g. roads, railroads, etc.) has to be built up. A similar theory by Bicanic,[2] studying investment behavior, also emphasized the existence of a threshold for growth.

The precise mechanisms of such models are accepted by some experts, and rejected by others.[3] Generally, economists admit that certain social variables like health and education play roles similar to communication systems, etc. Using Rostovian terms, they can be considered as social infra-structures and <u>unless certain minimum standards in such areas are met, economic development will be very difficult if not impossible</u>. While such views are generally held or conceded by economists and sociologists alike, no precise test has been conducted to verify or to refute such beliefs from observed data. The present study is devoted to checking the underlined statement against the available empirical evidence.

---

[1] W. W. Rostow, "The Stages of Economic Growth," <u>Economic History Review</u> (August, 1959); "The Take-Off Into Self-Sustained Growth," <u>Economic Journal</u> (March, 1956); and <u>The Process of Economic Growth</u> (Oxford, 1953).

[2] R. Bicanic, "The Threshold of Economic Growth," <u>Kyklos</u>, Vol. XV (1962) - Fasc. 1.

[3] W. W. Rostow (ed.), <u>The Economics of Take-Off Into Sustained Growth</u>, Proceedings of a Conference held by the International Economic Association (New York: St. Martin's Press, 1963).

However, before proceeding further, certain methodological disclaimers have to be put forward.

(i) As in any scientific inquiry, complicating the model can usually add realism albeit the simpler the model, the higher its testability. The latter consideration has to be regarded as an important desideratum in the present context. In short, a compromise between realism and testability is inevitable. This explains the assumptions made in the core of this study.

(ii) Sociological variables usually present challenging problems to the empirical researcher for the following reasons:

(a) Many extremely important variables are not quantifiable. The entrepreneurship of the managerial group, and the geographical and vocational mobility of the working class, belong to this category. No extensive magnitude can be found as a surrogate for the former while the percentage of urban population may provide an imperfect indicator for the latter. In the latter case, the indicator could be highly imperfect. Yet it may be the best indicator which can be gathered among the available cross-country data.

(b) The various indices are causally interdependent. From the statistical point of view, using the technique similar to that of factor analysis, one may be able to construct certain "synthetic variables," i.e. factors or indices for the social components. Hence, we may treat the problem of threshold as a study between economic growth and the synthetic social variable. Such an approach, though interesting, involves both methodological and interpretative complications.[1] The resolution of such problems lies beyond the scope of the present study.

From the causality point of view, the interrelationship between economic and social variables are very intricate indeed. For example, high-protein foodstuff may be necessary to build up physical fitness for productive endeavor, yet high disposable income can also be the principal cause for mass consumption of high-protein meals. This dual problem might be approached after the following hypotheses:

Hypothesis 1: Below a certain threshold value, a social variable is a "necessity" in producing national income.

In other words, within this range, that social variable is a tight constraint for growth.

Hypothesis 2: Above the threshold, the social variable is no longer an income determinent. Instead, it becomes a "luxury" for the private consumers or a policy objective.

This implies that, within this range, two cases may emerge.

2.1. The level of a certain social variable becomes a function of the income level. For example, high-calorie diets are the manifestations of an affluent community.

2.2. The level of a certain social variable is not correlated with income. This is because, above the threshold, expenditures influencing that social variable are determined by private and public preference patterns for each economy. The latter may bear no systematic relationship to income levels.

---

[1] J. Drewnowski and W. Scott, The Level of Living Index, Report No. 4 (Geneva: UNRISD, 1966).

Within this broad conceptual framework, alternative approaches can be adopted to analyze the problem. Each of these entails a different characterization of the threshold as well as an estimation procedure for its determination. These will be considered hereafter.

All in all we have considered four main different approaches:

(a) Simultaneous recursive equations system

(b) The convex hull boundary

(c) Factor analysis

(d) Binary analysis

(d)1 Polynomial analysis

(d)2 Phase model

of which (c) and (d)1 have been pursued to some degree. The main method adopted was (d)2--the phase model.

For each approach we shall consider: (1) the appropriate interpretation of the threshold, and (2) the merits and limitations of that approach. For approaches (c) and (d)1 we also have to discuss the empirical results in order to appraise their limitations. The reader will kindly refer to Part II for a discussion of the data. The interpretation of the findings of approach (d)2 is presented in Part III.

1.1. Theoretical and Statistical Preliminaries: Alternative Approaches

(a) Simultaneous recursive equations system. Under this approach, the basic model can be written as a simultaneous recursive equations system:

$$Z(t) = A(t)\, Z(t-1)$$

where $Z(t) = \begin{bmatrix} Z_0(t) \\ \cdots \\ Z_1(t) \\ \vdots \\ Z_n(t) \end{bmatrix}$; $Z_0(t)$ is the economic variable (GNP); $Z_i(t)$ is the i^{th} social variable. $1 \leq i \leq n$ One could, of course, have more than one economic variable.

A(t) is the coefficient matrix for observation t.

A(t) can take alternative values:

1. Before the threshold is reached, the social variables and the economic variable jointly determine each other, e.g. higher income implies better diet; better diet means fewer sick leaves and hence also higher income.

In this case:

$$A(t) = \left[\begin{array}{c|c} a_{oo} & a_{o\cdot} \\ \hline a_{\cdot o} & a_{\cdot\cdot} \end{array}\right]$$

2. After the threshold is reached, the social variables have little effect upon the economic variable; e.g. a diet with more than 3,000 calories intake per day has little effect on GNP.

In this case:

$$A(t) = \left[\begin{array}{c|c} a_{oo} & o \\ \hline a_{\cdot o} & a_{\cdot\cdot} \end{array}\right]$$

The threshold can either be defined as a critical level for Z_o or as a certain critical rectangular region in the social variable space (if the truncated subvector $\begin{matrix} Z_1 \\ \vdots \\ Z_n \end{matrix}$ belongs to this region, then the threshold is passed).

This approach may appear sophisticated. However, its requirement for long time series rules out its feasibility at this stage.

(b) The convex hull boundary. In the simultaneous equations approach it is assumed that before the threshold is reached, each social variable contributes individually to economic achievements and the observed economic variable reflects the aggregate of these influences. Alternatively, we may consider each social variable as a "limitational" factor. The present approach takes such a view. For instance, unless both health standards and the education level reach certain minima, economic well being cannot be raised beyond a given niveau.

Let x_i be the i^{th} social factor. t_i is the "factoral" threshold. y is the economic variable, say, GNP. Then:

$$y \leq f_i(x_i) \qquad \text{for} \qquad x_i \leq t_i \qquad f_i' > 0$$

showing that for $x_i \leq t_i$, there exists an increasing function $f_i(x_i)$ which constitutes a limitation for y.

However, since there are potentially n limitations, only the tightest ones hold as equalities; hence, one can write:

$$y = \underset{1 \leq i \leq n}{\text{Min}} \quad f_i(x_i) \qquad \text{for } x_i \leq t_i$$

The estimation of f_i can be illustrated from the diagram in Figure 1. The crosses indicate the observed points. The dotted line is the upper boundary of the convex hull spanned by the observed points. The points below the boundary, but to the left of x_i reflect (presumably) the limiting effects of other social factors x_j, $j \neq i$. For instance, if x_i is a health index and x_j is an education index, points A and C may show how much

GNP can be attained at levels a and c. When the health index reaches b, presumably the GNP can attain the level β', yet due to an education gap, only level β is attained.

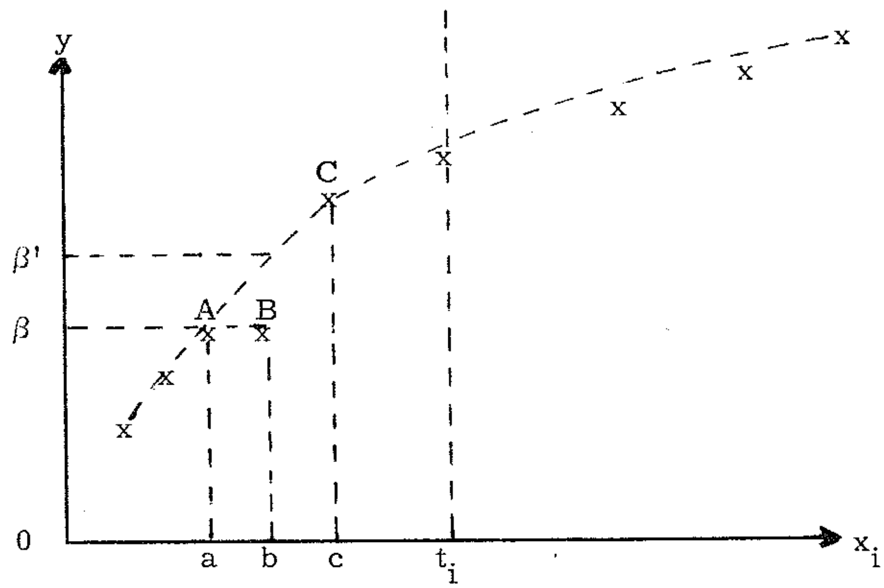


FIGURE 1

On the other hand, once x_i rises beyond t_i, y is very weakly correlated with x_i, then presumably the threshold for the i^{th} factor has been passed.

In view of the complicated estimation problems involved, this approach has not been experimented with.

We may note that the use of polygonal convex hull boundaries here is akin to what Farrell has used in his studies[1] dealing with measurement of productivity in agriculture.

---

[1] M. J. Farrell, "The Measurement of Productive Efficiency," J. Royal Statistical Society (series A), Vol. 120 (1957); and M. J Farrell and M. Fieldhouse, "Estimating Efficient Production Functions under Increasing Returns to Scale," J. Royal Statistical Society (series A), Vol. 125 (1962).

(c) Factor analysis.[1] The above two approaches both handle all the social and economic variables at one time. However, due to the fact that many of the social variables may be highly correlated with each other, direct study of these original variables need not be the most appropriate method. In this connection one may employ factor analysis as a tool to generate "synthetic" variables. Thus, the above said approaches may be conducted on synthesized rather than original variables. An appendix is provided to discuss a summary of the methodological aspects of factor analysis. As a preliminary step, a factor analysis has been conducted upon the data of 32 countries for 1950 and upon the data of 45 countries for 1960. The study of two different years was designed to disclose any structural change over time. The binary correlation coefficients for both years are listed in Tables 1 and 2, and the results of the factor analysis are summarized in Tables 3 and 4. It turns out that most social variables as well as the GNP "crowded" into one common factor. Whether or not this is the result of the insufficiency of our data, it certainly indicates that our search for the threshold cannot be carried out with the help of such an approach on the basis of the limited data available. Although some aspects of the factor analysis are of interest from other points of view, we shall not go into further details here.

---

[1] This factor analysis was performed on the Geneva University School of Physics' IBM 1620 (with disc) according to IBM programme 6.0.091 which uses the method of principal factor analysis based on Hotelling Iterative Procedure. For reasons why this method is preferable, see Raymond B. Cattell: "Factors Analysis: An Introduction to Essentials I. The Purpose and Underlying Models," Biometrics, March, 1965 Vol. 21, No. 1, pp. 190-215. Further explanations on this method are given in the Appendix.

TABLE 1

ORIGINAL MATRIX OF BINARY CORRELATION COEFFICIENTS, 1950
(32 COUNTRIES)

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}	r_{11}	r_{12}
r_1	1.0000	.7884	.7638	.3620	.4638	-.4920	.6598	.6620	.3377	.6592	.9140	.8195
r_2		1.0000	.9557	.0220	.1858	-.4999	.6557	.5596	.1916	.6166	.7223	.7801
r_3			1.0000	.0360	.1796	-.5161	.6517	.5604	.2396	.6137	.6983	.7690
r_4				1.0000	.9266	.0386	.1848	.3649	-.0825	.2559	.5095	.1871
r_5					1.0000	-.0782	.3079	.4066	-.0694	.3321	.5904	.3284
r_6						1.0000	-.6305	-.4777	-.4368	-.6336	-.4784	-.5203
r_7							1.0000	.6592	.5694	.9254	.6387	.6654
r_8								1.0000	.4792	.7935	.7019	.8033
r_9									1.0000	.6368	.2544	.3194
r_{10}										1.0000	.6679	.7253
r_{11}											1.0000	.8052
r_{12}												1.0000

r_1 Per capita GNP
r_2 Per capita calories
r_3 Per capita protein
r_4 No. of physicians
r_5 No. of hospital beds
r_6 Crude death rate
r_7 Primary enrollment
r_8 Secondary enrollment
r_9 Higher enrollment
r_{10} Adjusted enrollment
r_{11} No. of radios
r_{12} Daily newspaper

TABLE 2

ORIGINAL MATRIX OF BINARY CORRELATION COEFFICIENTS, 1960
(45 COUNTRIES)

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	r_{10}	r_{11}	r_{12}
r_1	1.0000	.7401	.7108	.4290	.5058	-.2716	.6475	.7329	.7428	.7480	.8437	.8405
r_2		1.0000	.9342	.1798	.2459	-.2757	.5769	.6384	.6068	.6826	.5786	.7152
r_3			1.0000	.2105	.2882	-.3184	.5784	.6194	.6323	.6571	.5593	.6759
r_4				1.0000	.9569	-.0634	.2772	.3647	.5339	.3031	.6402	.2988
r_5					1.0000	-.0797	.3698	.4624	.5727	.4075	.6961	.4095
r_6						1.0000	-.5150	-.3221	-.3617	-.4661	-.2461	-.2752
r_7							1.0000	.6615	.7293	.8774	.5410	.6432
r_8								1.0000	.5652	.8282	.5978	.8760
r_9									1.0000	.7496	.8003	.5713
r_{10}										1.0000	.6405	.7861
r_{11}											1.0000	.6491
r_{12}												1.0000

r_1 Per capita GNP
r_2 Per capita calories
r_3 Per capita protein
r_4 No. of physicians
r_5 No. of hospital beds
r_6 Crude death rate
r_7 Primary enrollment
r_8 Secondary enrollment
r_9 Higher enrollment
r_{10} Adjusted enrollment
r_{11} No. of radios
r_{12} Daily newspaper

TABLE 3

FACTOR ANALYSIS
1950

	F_1	F_2	F_3	F_4
ECONOMIC INDICATOR				
Per capita GNP	.8977	.1087	-.1966	.0687
NUTRITION				
Per capita calories	.8253	-.2034	-.4693	.0121
Per capita proteins	.8219	-.2163	-.4316	.0079
HEALTH				
Number of physicians	.3451	.8981	.1956	-.0516
Number of hospital beds	.4660	.8396	.1035	-.1258
Crude death rate	-.6538	.3572	-.1527	.5904
EDUCATION				
Primary enrollment ratio	.8544	-.1757	.2453	-.1162
Secondary enrollment ratio	.8359	.0610	.2112	.2555
Adjusted enrollment ratio	.8831	-.1329	.3407	-.0056
Higher enrollment ratio	.4769	-.4212	.6521	.2209
WORLD VIEW				
Number of radios per 1,000 inhabitants	.8925	.2680	-.1487	.0123
Daily newspaper circulation per 1,000 inhabitants	.8928	-.0536	-.1820	.1776

TABLE 4

FACTOR ANALYSIS
1960

	F_1	F_2	F_3	F_4
ECONOMIC INDICATOR				
Per capita GNP	.9058	-.0071	.1892	-.0031
NUTRITION				
Per capita calories	.7972	-.3403	.3241	-.2805
Per capita proteins	.7928	-.3092	.2657	-.3429
HEALTH				
Number of physicians	.5385	.8006	-.0898	-.0369
Number of hospital beds	.6236	.7414	-.0627	.0333
Crude death rate	-.4223	.3348	.7413	.1549
EDUCATION				
Primary enrollment ratio	.8072	-.2149	-.3318	.1022
Secondary enrollment ratio	.8430	-.1179	.0729	.4254
Adjusted enrollment ratio	.8937	-.2176	-.1565	.1976
Higher enrollment ratio	.8526	.1288	-.1454	-.2524
WORLD VIEW				
Number of radios per 1,000 inhabitants	.8446	.3167	.0563	-.1410
Daily newspaper circulation per 1,000 inhabitants	.8585	-.1725	.2106	.3253

(d) Binary analysis. In this case we consider one social variable and one economic variable at a time. Under this caption two alternative approaches have been attempted:

(d)1. Polynomial curve fitting. We estimate the equation:[1]

$$y = \sum_{j=0}^{n_i} (x_i)^j$$

where x_i is the i^{th} social variable, y is GNP, and n_i is the degree of the polynomial.

The polynomial form is chosen because of its flexibility and convenience of estimation. Presumably, if the curve exhibits the form shown in the following diagram:

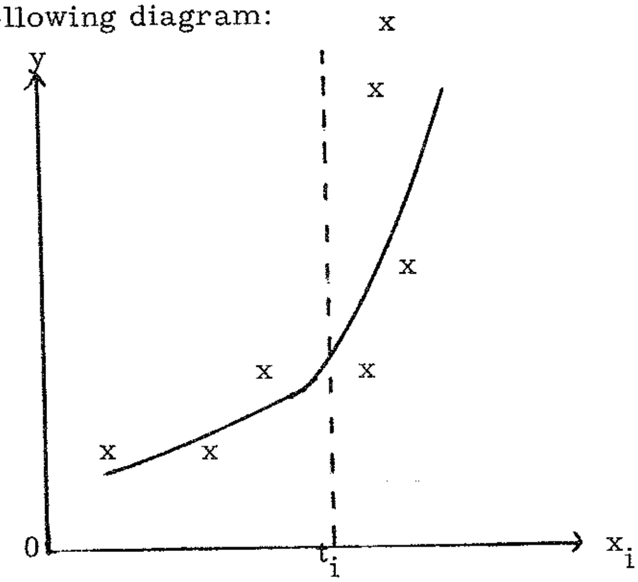


FIGURE 2

[1] G. W. Snedecor, Statistical Methods, (Aures, Iowa: Iowa State University Press, 1956), Ch. 15; E. J. Williams, Regression Analysis (New York: John Wiley and Sons, Inc., 1959); A. S. Householder, A Glossary for Numerical Analysis (Oak Ridge, Tenn.: Oak Ridge National Laboratory), March 19, 1959. ORNL - 2704.

We can make conjectures about the position of the threshold t_i, for instance, as an abrupt change of the slope coefficient. Presumably there would not be any "point threshold" (a polynomial is a smooth function) but an "interval threshold" can be defined if we decide upon some criterion to decide how abrupt a change of slope qualifies as a threshold.

However, the extreme flexibility of the polynomial suffers from the following drawback: within the range of relevant values of x_i, the polynomial can have multiple upturns and downturns. The unsatisfactory feature of such a phenomenon is shown in the next diagram.

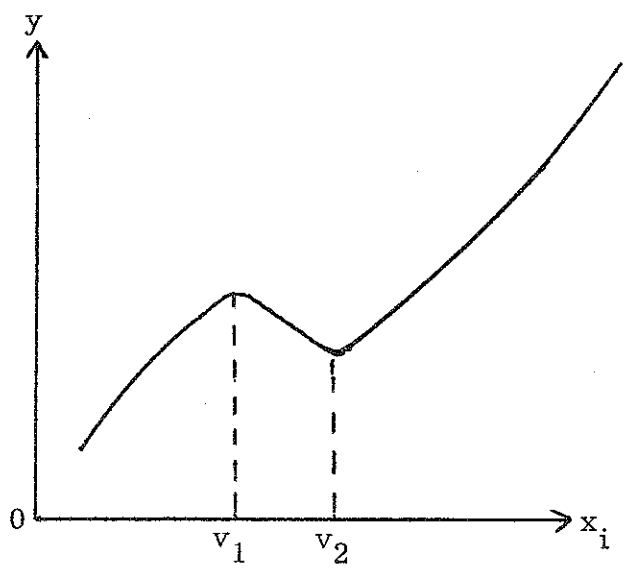


FIGURE 3

A literal interpretation would be:

x_i is helpful for economic development when $x_i < v_1$

x_i is not helpful for economic development when $v_1 < x_i < v_2$

x_i is helpful for economic development when $v_2 < x_i$

This may be contrary to our a priori notions. For instance, health and education are helpful and mortality rate is harmful to growth over all ranges. Although there could be levels of income for which the effect of a particular social variable is not the major influence.

Of course, a n-degree polynomial usually has (n-1) troughs and peaks. The crucial question is: Will such upturns and downturns occur in the relevant range?

As a tentative step, polynomial curves have been fitted[1] for the cross-country data of 1950, 1955, and 1960. In each case the GNP per capita is regressed against various social variables (from 12 to 17 depending on the availability of data for a given year). The sample size varies between 24 and 70. Tables 5 to 8 recapitulate the statistical results and Figures 4 to 7 show some representative graphs of such polynomials.

A brief appraisal of the results can be summarized as follows:

1. Although the fits are statistically significant, the standard errors are somewhat large.

2. The degrees of the polynomials are high and usually quite a number of "turning points" are located in the relevant range. When there is a succession of such turning points within a narrow range, a straightforward interpretation of such points may not be forthcoming without further assumptions. For instance, one may wish to "smooth" the

---

[1] The curve fitting was processed according to the Summer 1965 version of programme G 200 on the electronic computer Control Data 6.600 at the European Center for Nuclear Research (CERN), Geneva.

TABLE 5

POLYNOMIAL CURVE FITTING: STATISTICAL TESTS
1950

Types of Indicators	Degree Polynomial	Number of Countries	Degree of Freedom	Root Mean Square
NUTRITION				
Per capita calories per day	4	33	28	284.24
Total proteins grams per day	6	33	26	341.05
HEALTH				
Inhabitants per physician	3	42	38	374.40
Number of hospital beds per 1,000 inhabitants	3	38	34	288.04
Infant mortality death per 1,000 live births	2	41	38	287.65
EDUCATION: FORMAL TRAINING				
Primary enrollment ratio	10	61	50	260.43
Secondary enrollment ratio	4	61	56	236.13
Higher enrollment ratio	5	44	38	373.53
Adjusted enrollment ratio	10	63	52	254.21
Enrollment ratio of all types	...	...	...	...
Lagged total education enrollment ratio	...	...	...	...
INFORMAL TRAINING: WORLD VIEW				
Number of radios per 1,000 inhabitants	10	56	45	134.63
Number of T.V. per 1,000 inhabitants	...	...	...	...
Movie attendance per capita	8	24	15	370.87
Illiteracy rate	3	54	50	214.53
Daily newspaper circulation per 1,000 inhabitants	7	51	43	195.68
Urbanization percentage of population in cities above 20,000 inhabitants	6	46	39	322.04

TABLE 6

POLYNOMIAL CURVE FITTING: STATISTICAL TESTS
1960

Types of Indicators	Degree Polynomial	Number of Countries	Degree of Freedom	Root Mean Square
NUTRITION				
Per capital calories per day	4	45	40	381.34
Total proteins grams per day	5	54	48	362.26
HEALTH				
Inhabitants per physician	2	39	36	422.93
Number of hospital beds per 1,000 inhabitants	3	38	34	302.84
Infant mortality death per 1,000 live births	2	41	38	334.76
EDUCATION: FORMAL TRAINING				
Primary enrollment ratio	4	68	63	369.97
Secondary enrollment ratio	10	68	57	
Higher enrollment ratio	7	53	45	360.85
Adjusted enrollment ratio	8	70	61	296.55
Enrollment ratio of all types	3	24	20	361.24
Lagged total education enrollment ratio	8	52	43	372.94
INFORMAL TRAINING: WORLD VIEW				
Number of radios per 1,000 inhabitants	10	59	48	181.95
Number of T.V. per 1,000 inhabitants	10	41	30	390.54
Movie attendance per capita	5	49	43	443.12
Illiteracy rate	2	27	24	351.08
Daily newspaper circulation per 1,000 inhabitants	9	70	60	249.69
Urbanization percentage of population in cities above 20,000 inhabitants	3	28	24	402.90

TABLE 7

SOCIAL THRESHOLDS: THE PER CAPITA INCOME TRENDS OF SOCIAL GROWTH - 1950

Types of Indicators	Degree of Poly-nomial	Number of Coun-tries	Thresholds per capita GNP: Ranges	Thresholds per capita GNP: Average of Ranges	Thresholds, Social Variables: Ranges	Thresholds, Social Variables: Average of Ranges	Maximum of Threshold Span: Per Capita GNP	Maximum of Threshold Span: Social Vari-ables
NUTRITION								
1. Per capita calories/day	4	33	140 1,100 250 1,220	195 1,160	2,300 3,250 1,900 3,175	2,100 3,212	250	2,550
2. Total proteins grams/day	6	33	100 750 270 1,170	185 960	66 97 53 88	60 93	280	73
HEALTH								
3. Inhabitants/physician	3	42	100	100	45 105	75	200	36
4. Number of hospital beds /1,000 inhabitants	3	38	140 200	170	2 4	3	200	3
5. Crude death rate	---	56	---	---	---	---	---	---
6. Infant mortality	2	41	100 150	125	100 115	108	140	86
EDUCATION: FORMAL TRAINING								
7. Primary enrollment ratio	10	61	98 210	154	47 74 35 66	41 70	210	53
8. Secondary enrollment ratio	4	61	100 200	150	3 14	9	210	15
9. Higher enrollment ratio	5	44	100 620 740	100 680	.16 1.19 .75	.16 .97	100	.16
10. Adjusted enrollment ratio	10	63	110 220	165	35 50	43	220	62
11. Enrollment ratio of all types	---	---	---	---	---	---	---	---
12. Lagged total enrollment ratio	---	---	---	---	---	---	---	---
INFORMAL TRAINING: WORLD VIEW								
13. Number of radios/1,000 inhabitants	10	56	80 930 240 450 1,250	160 450 1,090	0 282 50 124 237	25 124 260	500	164
14. Number of T.V./1,000 inh.	---	---	---	---	---	---	---	---
15. Movie attendance per capita	8	24	100 535 750	100 535 750	.9 7.5 14.5 2.0 5.3 11.0	1.45 6.4 12.75	200	3.2
16. Illiteracy rate	3	54	120 170	145	42 67	54	170	28
17. Daily newspaper circulation	7	51	140 1,170 300 1,270	220	174 400 82 351	128 376	300	213
18. Urbanization	6	46	50 540 200 730	125 685	15 44 8 33	11 34	200	22

TABLE 8

SOCIAL THRESHOLDS: THE PER CAPITA INCOME TRENDS OF SOCIAL GROWTH - 1960

Types of Indicators	Degree of Poly-nomial	Number of Coun-tries	Thresholds per capita GNP: Ranges	Thresholds per capita GNP: Average of Ranges	Thresholds, Social Variables: Ranges	Thresholds, Social Variables: Average of Ranges	Maximum of Threshold Span: Per Capita GNP	Maximum of Threshold Span: Social Vari-ables
NUTRITION								
1. Per capita calories/day	4	45	180 200	190	2,250 2,050	2,150	200	2,350
2. Total proteins grams/day	5	54	80 790 210 1,040	145 915	59 101 48 87	54 94	220	67
HEALTH								
3. Inhabitants/physician	2	39	0 200	100	32 75	54	200	32
4. Number of hospital beds /1,000 inhabitants	3	38	140 200	170	.9 2.7	1.8	210	2.7
5. Crude death rate	---	---	---	---	---	---	---	---
6. Infant mortality deaths /1,000 live births	2	41	85 150	117	85 115	100	100	73
EDUCATION: FORMAL TRAINING								
7. Primary enrollment ratio	4	68	80 160	120	40 22	31	210	50
8. Secondary enrollment ratio	10	68	50 600 250	150 600	20 45 10	15 45	250	28
9. Higher enrollment ratio	7	53	140 770 190 900	165 835	.43 1.82 .18 1.37	.30 1.59	190	.54
10. Adjusted enrollment ratio	8	70	30 210 850 160 240 1,200	95 225 1,025	37 64 94 26 55 87	32 60 91	240	66
11. Enrollment ratio of all types	3	24	230 280	255	43 29	36	280	49
12. Lagged total education enrollment ratio	8	52	90 840 200 1,010	145 925	1.0 3.6 .5 5.6	.75 4.6	200	1.6
INFORMAL TRAINING: WORLD VIEW								
13. Number of radios/1,000 inhabitants	10	59	90 230 940 130 290 1,350	110 260 1,145	31 125 288 13 87 237	22 106 263	290	144
14. Number of T.V./1,000 inh.	10	41	260 650 360 930	310 790	0 60 10 37	5 49	350	9
15. Movie attendance per capita	5	49	70 900 120	95 900	1.3 10.5 2.1	1.5 6.4 12.8	200	3
16. Illiteracy rate	2	27	10 100	55	58 75	66	160	35
17. Daily newspaper circu-lation	9	70	140 280	210 770	25 210 75 138	50 174	300	82
18. Urbanization	3	28	80 110	95	7 12	10	110	12

FIGURE 4

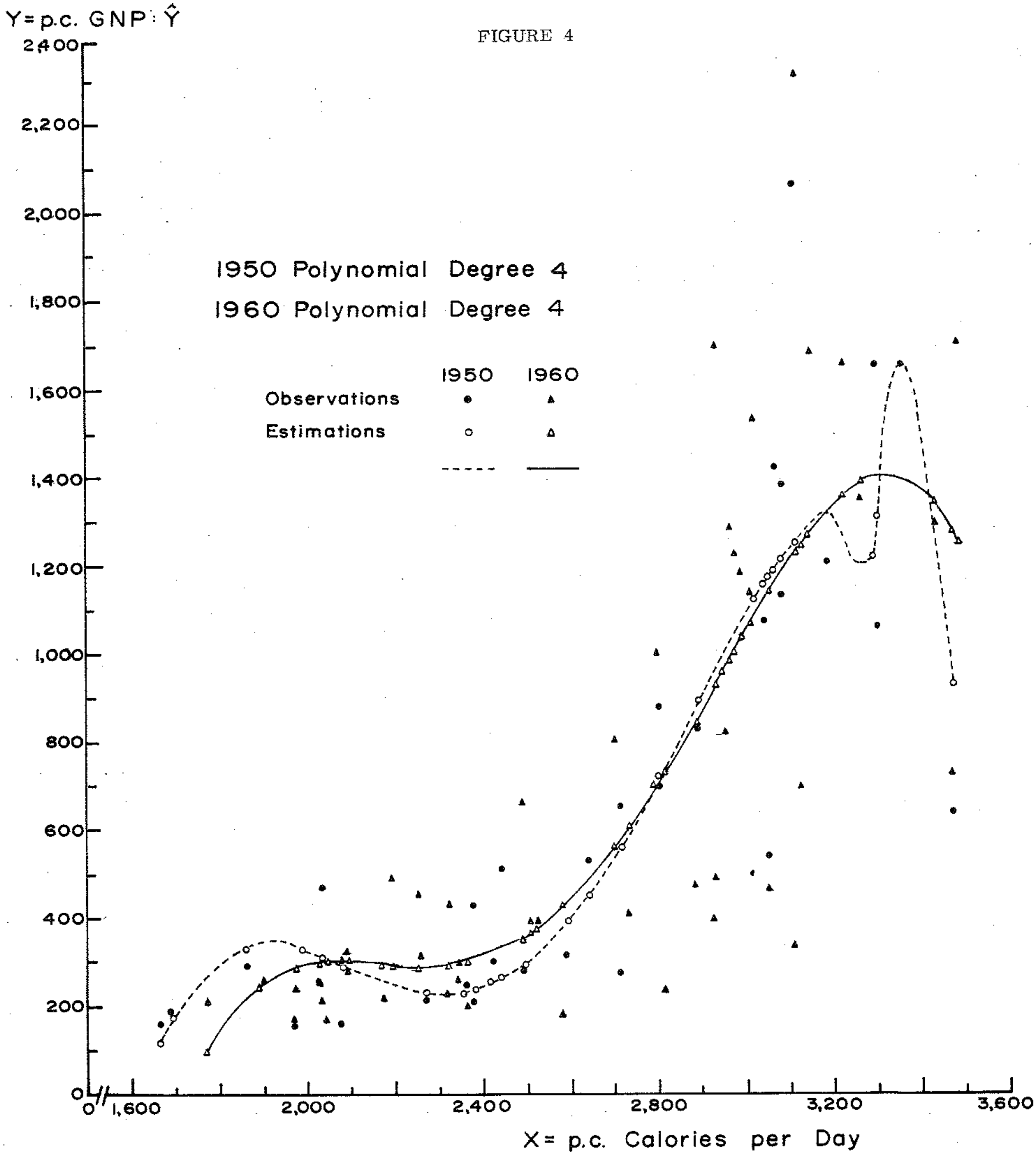
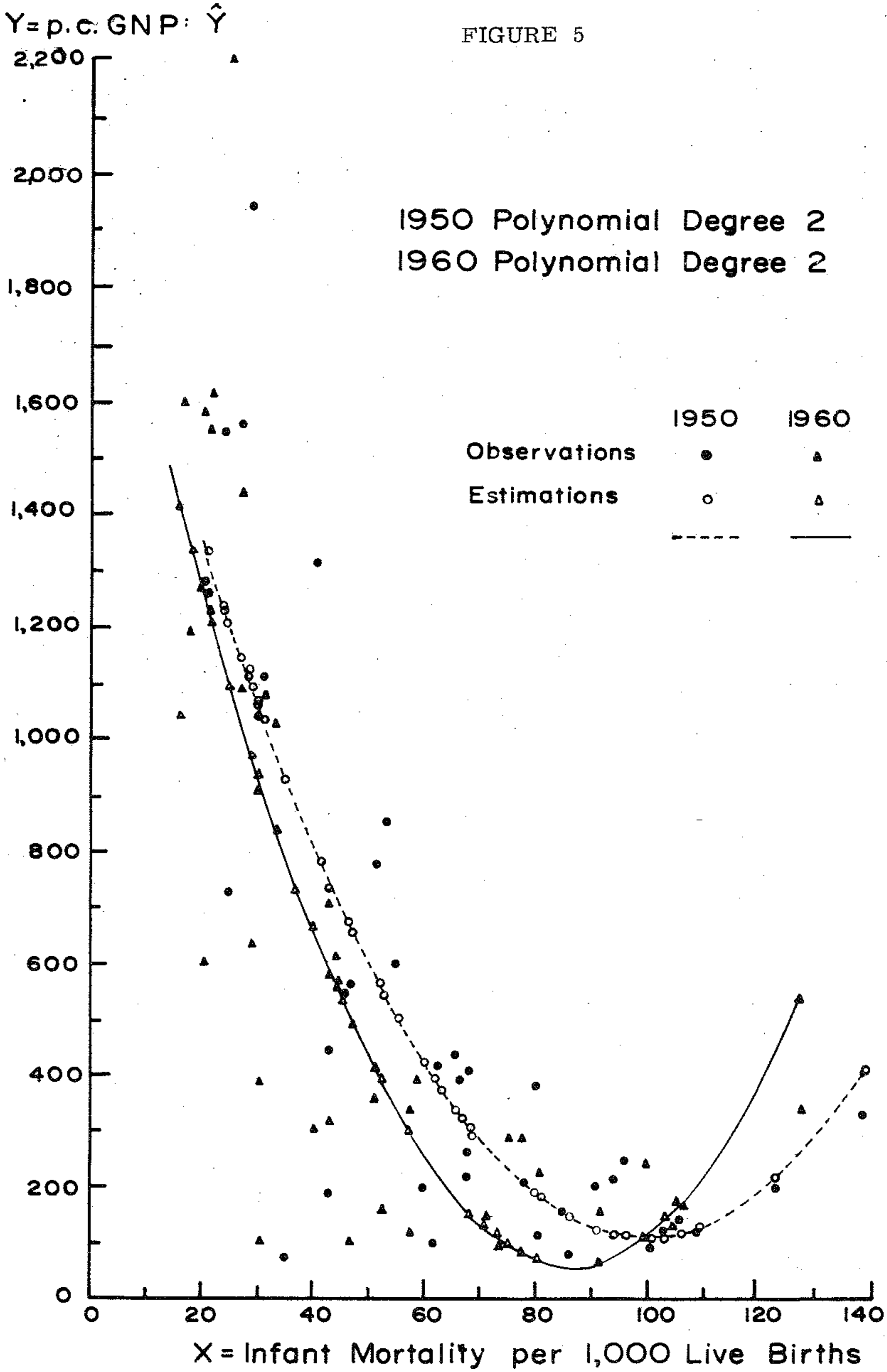

Y = p.c. GNP: Ŷ
2,400
2,200
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
1950 Polynomial Degree 4
1960 Polynomial Degree 4
1950
1960
Observations
Estimations
0
1,600
2,000
2,400
2,800
3,200
3,600
X = p.c. Calories per Day

FIGURE 5



Y = p.c. GNP: Ŷ
2,200
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
1950 Polynomial Degree 2
1960 Polynomial Degree 2
1950
1960
Observations
Estimations
0
20
40
60
80
100
120
140
X = Infant Mortality per 1,000 Live Births

FIGURE 6

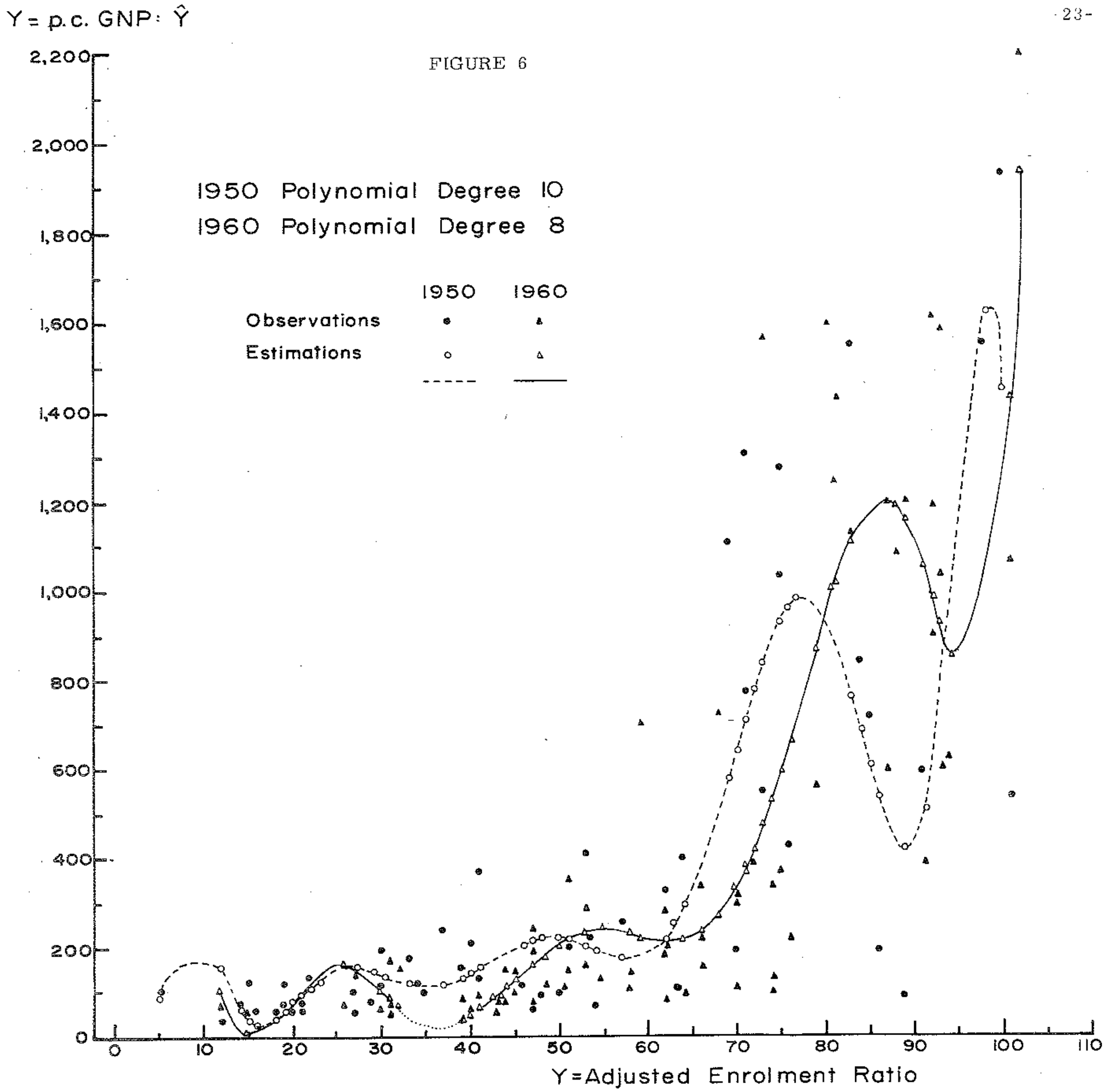
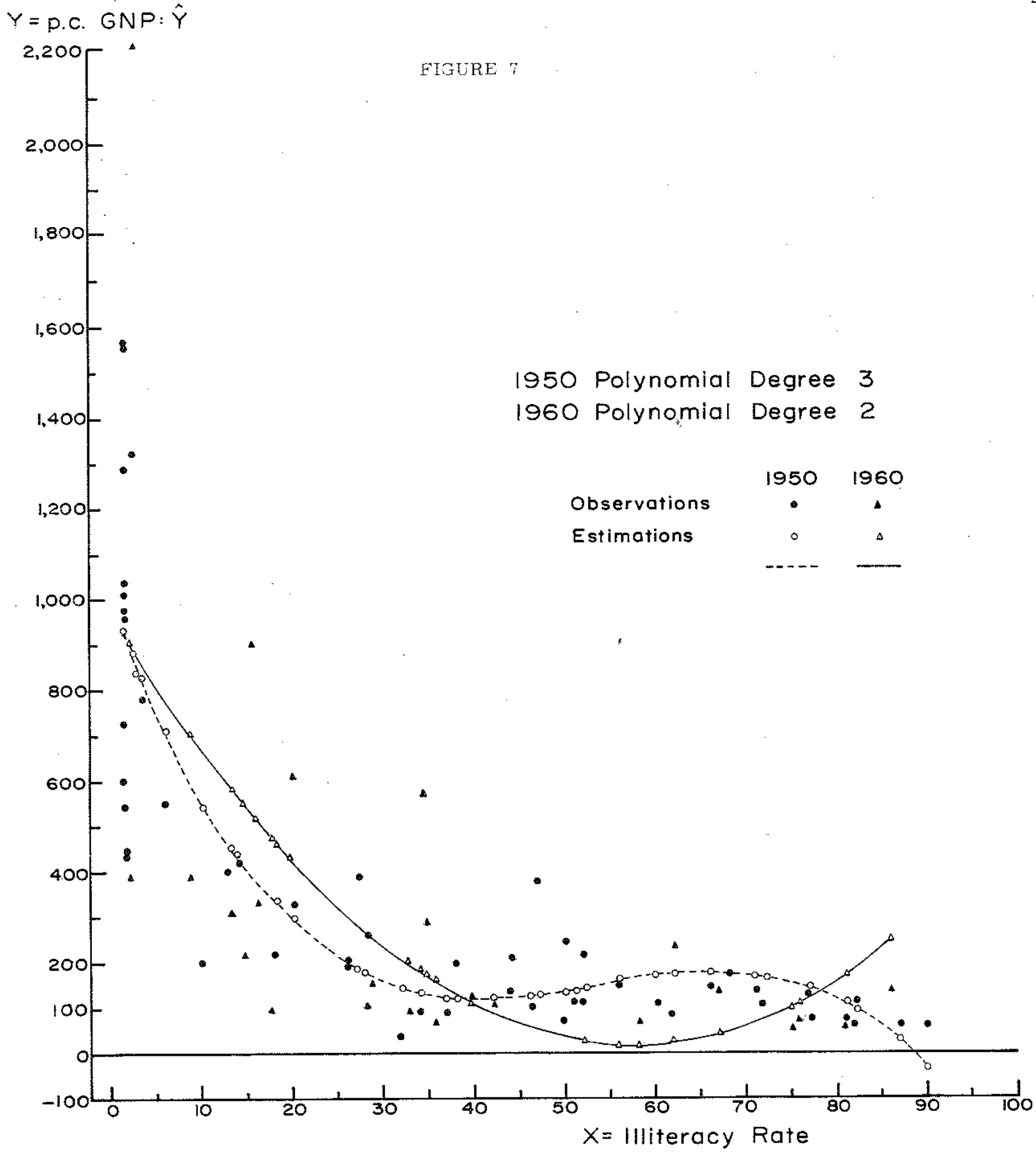

Y = p.c. GNP: Ŷ
2,200
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
0
1950 Polynomial Degree 10
1960 Polynomial Degree 8
1950 1960
Observations
Estimations
0 10 20 30 40 50 60 70 80 90 100 110
Y=Adjusted Enrolment Ratio

FIGURE 7


Y = p.c. GNP: Ŷ
2,200
2,000
1,800
1,600
1,400
1,200
1,000
800
600
400
200
100
0
−100
1950 Polynomial Degree 3
1960 Polynomial Degree 2
1950
1960
Observations
Estimations
0
10
20
30
40
50
60
70
80
90
100
X= Illiteracy Rate

polynomial by "averaging" the span of turning points over a range.[1]

3. Observed points scatter farther away from the fitted curve in higher ranges.

4. There exists some general patterns for each binary relation over the different years. On the other hand, the shift of functional relationships over time is difficult to explain systematically.

Due to reasons discussed before, the existence of numerous turning points rules out the fitting of polynomial as an appropriate approach. Nonetheless, important clues about the structure of the problem are provided by such a study. Because despite the "wiggles and waggles" of the polynomial curve, the scatter diagrams indicate that the deviations from the "trend" (i.e. the polynomial) are generally larger at the upper ranges of x_i. This suggests either one of the following two possibilities, or both:

a. There exists more than one statistical regimes underlying the samples.

b. There may be heteroscedasticity[2] in the samples.

Using the residuals' distribution obtained from the polynomial analysis, one could eliminate the heteroscedasticity from the original data. This procedure being time consuming has not been attempted here.

---

[1] For instance, taking the average of the values for the social variable (or GNP) as they occur at the lowest and highest turning points in a given span. Such "critical levels" for the polynomial curve fitting are given in Tables 7 and 8.

[2] References on this problem are: S. Golberger, Econometric Theory (New York: John Wiley and Sons, Inc., 1964); J. Johnston, Econometric Methods (New York: McGraw-Hill Book Co., Inc., 1963); and E. Malinvaud, Methodes Statistiques de l'econometrie (Paris: Dunod, 1964).

1.2 The Main Approach: Binary Analysis--The Phase Model

(d)2. The phase model. Our consideration and experimentation of alternative approaches convince us that the best practical approach is what we shall call the "phase model." We shall first discuss the way in which the model is applied and then briefly sketch the statistical basis for this method.

A. A statistical procedure devised by Quandt[1] to partition the observed data into two subsamples shows that countries whose social variable is below a certain switching point belong to one subsample and the rest of the observations belong to another. Linear regressions are then run for each subsample. Both a test for the existence of a switching point and the determination of its location are conducted through a likelihood analysis. These will be further discussed in Item B.

The results of such a procedure give rise to four possible cases:

	Low Income Sample	High Income Sample
Case 1.	Strong correlation (-)	Strong correlation (-)
Case 2.	Strong correlation (-)	Weak correlation (.)
Case 3.	Weak correlation (.)	Strong correlation (-)
Case 4.	Weak correlation (.)	Weak correlation (.)

---

[1] A detailed account of the procedure can be found in Richard E. Quandt, "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," Journal of the Am. Stat. Assoc., 53 (December, 1958), 873-880; and "Tests of the Hypothesis That A Linear Regression System Obeys Two Separate Regimes," ibid., 55 (June, 1960), 324-330.

Cases 1 and 2 are consistent with both hypotheses listed above. In each case the switching point is the presumed threshold. Case 3 does not support our a priori notion of a threshold; the latter concept implies that below a certain threshold level, a social variable is indispensable to economic achievement. Case 4 represents a situation where little influence can be drawn about the threshold. Hence, the determination of a switching point does not necessarily imply the existence of a threshold.

We can now turn to the technical aspects of the problem.

B. Statistical Test of the Phase Model

In estimating the switching point from one regime to another, we postulate the existence of two relationships:

$$Y = a_1 S_I + b_1 + u_1 \qquad (1)$$

$$Y = a_2 S_I + b_2 + u_2 \qquad (2)$$

where Y is per capita income, S_I is the social indicator, and u_1 and u_2 are normally and independently distributed error terms with mean zero and standard deviation σ_1 and σ_2. There are a total of T observations.

Assume that the first t observations are generated by (1) and the last $T-t$ by (2).

The densities of u_1 at point i and u_2 at point j are:

$$(1 / \sqrt{2\pi}\,\sigma_1) \quad \exp\left[- (1 / 2\sigma_1^2)\,(y_i - a_1 X_i - b_1)^2 \right]$$

and

$$(1 / \sqrt{2\pi}\,\sigma_2) \quad \exp\left[- (1 / 2\sigma_2^2)\,(y_i - a_2 Y_j - b_2)^2 \right]$$

The likelihood of a sample of t observations from (1) and $T-t$ observations from (2) are therefore:

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_1}\right)^t \exp\left(-\frac{1}{2\sigma_1^2}\sum_{i=1}^{t}(y_i - a_1 X_i - b_1)^2\right)$$

and

$$\left(\frac{1}{\sqrt{2\pi}\,\sigma_2}\right)^{T-t} \exp\left(-\frac{1}{2\sigma_2^2}\sum_{j=t+i}^{T}(y_j - a_2 X_j - b_2)^2\right)$$

The logarithm of the likelihood function of the entire sample is:

$$L = -T\log\sqrt{2\pi} - t\log\sigma_1 - (T-t)\log\sigma_2$$

$$-(1/2\sigma_1^2)\sum_{i=1}^{t}(y_i - a_1 X_i - b_1)^2 - (1/2\sigma_2^2)\sum_{j=t+1}^{T}(y_j - a_2 X_j - b_2)^2$$

Substituting for values of parameters and coefficients we obtain:

$$L(t) = -T\log\sqrt{2\pi} - t\log\hat{\sigma}_1 - (T-t)\log\hat{\sigma}_2 - \frac{T}{2}$$

which gives the logarithm of the maximum likelihood for a given value of T and is a function of t alone.

In order to obtain the switching point, we have calculated the value of the likelihood function $L(t)$ for all possible values of t and have selected as the maximum likelihood estimate that value of t which corresponds to the maximum maximorum.[1]

A likelihood ratio test proves useful in testing the hypothesis that no switch occurred against the single alternative that one switch took place. The likelihood ratio λ is defined as:

$$\lambda = \frac{L(\hat{W})}{L(\hat{\Omega})} = \frac{\hat{\sigma}_1^{\,t}\,\hat{\sigma}_2^{\,T-t}}{\hat{\sigma}^T}$$

---

[1] *Ibid*.

from which we obtain:

$$- 2 \log \lambda = -2 \frac{(t \log \hat{\sigma}_1) [(T-t) \log \hat{\sigma}_2]}{T \log \hat{\sigma}}$$

Under certain conditions,[1] the Chi square distribution with n-m degrees of freedom is an acceptable approximation to the distribution of $-2 \log \lambda$ for large T where n is the dimension of Ω and m the dimension of w. Ω is the entire parameter space, while w is the subspace. If the Chi square distribution is used as an approximation to $-2 \log \lambda$, the relevant degrees of freedom are n-m.[2]

So far we have discussed the analytic methods considered. In Part II we shall review the data selected for our empirical testing. The interpretation of the findings are presented in Part III.

---

[1] A. M. Mood, Introduction to the Theory of Statistics (New York: McGraw-Hill Book Company, Inc., 1950), pp. 257-259.

[2] The author has produced a program for this procedure. The computations were performed on the IBM 7090 at the Computer Center of the University of California, Berkeley.

PART II

Selection and Adjustment of Data: Brief Comments

To study the kaleidoscopic interplay between social and economic factors only a sagacious selection of strategic and precisely defined variables can pinpoint the essence of the problem. Constrained by the limited scope of the present project, our choice of variables is determined by the following four criteria:

1. Substantiality and relevancy -- Variables of minor import or unrelated to the socioeconomic interactions are excluded
2. Quantifiability -- Nonmeasurable factors are ruled out by the statistical nature of the investigation
3. Availability -- Collection of first-hand information is out of the question under the circumstances
4. Reliability -- The use of dubious data inputs threatens to produce misleading findings.

It is deemed advisable to examine the nature of the data for a twofold reason:

1. To indicate what data would be most suitable under ideal circumstances even though they are unavailable at present.
2. To outline our procedures in data selection and adjustment. We have no intention of rationalizing the expediencies we were forced to adopt, but it is imperative to indicate the reservations we have about our ultimate findings.

We have considered eighteen social variables but only one economic variable. This is partly because economists usually use income as a measure of over-all well being. And partly because the multidimensionality of the sociology of development requires us to consider a set of social facets since no single indicator is broadly accepted as a social indicator.[1]

Since our interest in the concept of threshold is relevant mostly with respect to developing countries, in principle one can restrict the sample to that group. However, the distinction between developed and developing countries is an arbitrary one. Results of the analysis of cumulative frequency distribution for a study of patterns of influences of variables undertaken by N. Ginsburg in Atlas for Economic Development lead to the conclusion that "natural" groups of "developed" or "underdeveloped" countries with similar sets of ranks do not exist. Rather the evidence suggests a spread of countries along a continuum.[2] It is for these reasons that the observations are taken for all countries developing as well as developed ones.

An Economic Variable: Income

Due to lack of better alternatives, we are forced to use as the economic index the per capita gross national product (p.c. GNP) in

---

[1] J. Drewnowski and W. Scott, op. cit.

[2] Norton Ginsburg, Atlas of Economic Development (Chicago: The University of Chicago Press, 1961), p. 113. Perhaps if one insists on such a classification of developing vs. developed economies, it is interesting to note that the frequency distribution of per capita income is bimodal (log normal for developing countries and rectangular for developed ones) and that the separation between the two segments of the curve occurs at 425 U.S. dollars (GDP at 1960 prices). This figure is obtained under certain restrictive assumptions. See M. Subramanian, "Developed vs. Developing Countries," mimeographed, UNRISD, 1965 (available on request).

terms of U. S. dollars in 1958 prices. This is far from a happy choice for the following three reasons:

1. International comparison of income is hazardous. Exchange rates reflect poorly relative purchasing power. The existing studies offer no satisfactory means of adjustment.

2. Per capita income in a cross-country comparison conceals differences in both (a) the income distribution and (b) the age distribution.

3. Developing countries usually contain (a) a substantial subsistence sector which is not yet monetized and (b) a plethora of extremely microscopic retail services (e. g. peddling one cigarette at a time). Both defy efforts at adequate national income accounting at present.

However, a preliminary factor analysis showed that income and eleven social variables share a common factor. This suggests that there exists a single basic structure, namely, growth. Hence, the use of income as a dependent variable in subsequent binary analysis may not be unduly misleading.

Social Variables

Based upon the above four criteria for variable selection, we have included eighteen social variables which can be classified into four groups as follows:

Nutrition

per capita calories per day
per capita total proteins grams per day

Health

Number of inhabitants per physician

Number of hospital beds per 1,000 inhabitants

Crude death rate[1]

Infant mortality rate per 1,000 live births

Education: Formal Training

Primary enrollment ratio

Secondary enrollment ratio

Higher enrollment ratio

Total of all types enrollment ratio

Adjusted enrollment ratio

Lagged total enrollment ratio

Informal Training: World View

Illi Illiteracy rate

Daily newspaper circulation per 1,000 inhabitants

Number of radios per 1,000 inhabitants

Number of T. V. per 1,000 inhabitants

Movie attendance per capita

Urbanization, percentage of population in cities above 20,000 inhabitants

Important factors such as political and jurisprudence institutions have to be left out due to nonquantifiability. Future progress for the quantification of such data is most desirable.

The demographic framework plays a pervasive role throughout all four groups, either directly (e.g. crude death rate) or indirectly (e.g. when income or illiteracy is measured on a per capita basis). Yet due

---

[1] The death rate is totally uncorrelated with income per capita as exemplified by scatter diagrams between per capita income and the death rate for 56 countries in 1950 and 58 countries in 1960. Consequently, this data is used in our factor analysis but not in the binary analysis.

to the unavailability of reliable data for many countries, even birth rate data are excluded here, to say nothing of the age distribution, etc.

In the group-by-group discussion of the data below, Group 1 and Group 2 require less elaboration; their importance to sheer survival is self evident. More space will be devoted to the other two groups.

Group 1: Nutrition

We adopt the twin conventional indices for an adequate diet: calories and protein intakes.[1] However, should such data exist, a more meaningful measurement will be the ratio of actual intake to body requirement.[2] These data depend upon weather, age, sex, body weight, and activities. In fact, estimated protein requirements among different countries vary as much as 27 percent. It is hoped that more adequate data will be available in the future.

Group 2: Health

The interaction of nutrition and medicare, on one hand, and birth and death rates,[3] population age distribution, and labor supply, on the

---

[1] It is suspected that fat intake may not be crucial. See C. Clark and M. Haswell, The Economics of Subsistence Agriculture (New York: St. Martin's Press, 1964).

[2] Margaret McArthur, "Some Factors Involved in Estimating Calorie Requirements with Special Reference to Persons Engaged in Agricultural Labour in Asian Countries," Journal of the Royal Statistical Society, Series A, 127, Part 3 (1964).

[3] The death rate is the number of deaths per 1,000 population in an age and sex interval per unit of time.

other hand, is well known.[1] So we shall not elaborate further on this subject here.

Group 3: Education: Formal Training

We shall summarize the important issues involved in measuring education because this is a crucial area whose significance is seldom matched by the adequacy of the treatment devoted to it.

1. With respect to its form, education can be divided into:

   (a) Formal education. Only for this type of education are statistical data available.

   (b) Informal education. Pre-primary,[2] adult education, on-the-job training, training in the armed forces, etc.

2. With respect to its measurements, it presumably can be divided into:

   (a) The flow of students.[3] On this data are available.

   (b) The stock of educated personnel.[4] Presumably, at any

---

[1] A. J. Coale and E. M. Hoover, Population Growth and Economic Development in Low-Income Countries (Princeton, N. J.: Princeton University Press, 1958).

[2] Various studies on the causes of school dropouts have indicated that the potential or pre-school education in minimizing the school drop-out rate is considerable. A sample study undertaken by the India International Center Council for Social Development would seem to indicate that the school dropout rate during the first years of primary school might be reduced by as much as 70 percent for children having attended pre-primary school.

[3] The Conference of European Statisticians Meetings of the ECE in Geneva (September, 1965) considered a number of studies dealing with the setting up of an adequate and meaningful flow system.

[4] An "ideal" data for the stock of knowledge (skill) would be a composite of the following requirements. The amount of school years would

point of time, the effect on economic performance of this data is more important that that of the flow data. A cross-country comparison of the stock data usually will be biased against the developing countries because newly-expanded programs have not yet cumulated enough fruits. Since technical knowledge becomes obsolete over time, a vintage approach may be desirable. At any rate, data are not readily available.

3. With respect to its quality, education may be judged either from:

(a) An Opportunity-Initiative analysis. How many people are eligible for training and how many have the motivation to participate?

(b) A Cultural-Vocational dichotomy. What portion of the curriculum is devoted to the acquisition of practical skill, what portion to general acculturation?

---

have to be adjusted for length, attendance, etc. The curriculum would have to be weighted for quality to reflect consistency through time as well as for international consistency, i.e. vintages would be introduced. In addition, such vintages would need to be depreciated for obsolescence. An approximation to this "ideal" data has been achieved by Denison's distribution of knowledge over generations. Of course, for developing countries such distributions might be difficult to establish due to lack or scarcity of censuses. One method adopted by Schultz was to use an analogy with additions to the stock of capital in measuring the amount of education available in the population at large. Taking into account income foregone by students and the resources used in providing formal education, he reaches a total of annual investment in education from 1900 to 1956 in the United States. E. F. Denison, "Measuring the Contribution of Education and the Residual to Economic Growth," in The Residual Factor and Economic Growth (Paris: OECD, Study Group in the Economics of Education, 1964); T. W. Schultz, "Capital Formation By Education," Journal of Political Economy, 67 (December, 1960), 571-587; and J. Vaisey, The Economics of Education (London: Faber and Faber, 1962).

The importance of such investigation into the qualitative aspects is obvious, yet the existing source data denies us any opportunity to pursue such approaches.

In our study six ratios are adopted:[1]

(a) Primary education enrollment ratio. This measures the percentage of the population between 5 and 14 enrolled in primary schools. No adjustment is made to correct the effect of differential durations of primary education in various countries. The same drawback applies to (b) below.

(b) Secondary education enrollment ratio. This measures the percentage of the population between 15 and 19 enrolled in secondary schools.

(c) Higher education enrollment ratio. This measures the percentage of the population between 20 and 24 enrolled in colleges.

(d) Adjusted enrollment ratio. This combines, respectively, the numerators and denominators of (a) and (b) in an attempt to eliminate biases due to different education systems.

(e) Enrollment ratio of all types. This is similarly computed for (a), (b), and (c).

(f) Lagged total education enrollment ratio. In cognizance of the lagged effect of education on growth, an index[2] is derived

---

[1]For supplementary details, see F. Harbison and C. A. Myers, Education, Manpower and Economic Growth: Strategies of Human Resource Development (New York: McGraw-Hill Book Company, 1964), pp. 29-30.

[2]Some degree of arbitrariness in the choice of weights needed for the aggregation of this index is unavoidable due to imperfect information. Statistically note that since the dispersion at any point is the square of

based upon the primary enrollment ratio of 1950. The secondary enrollment ratio of 1955 and the higher education enrollment of 1960.

Group 4: Informal Training: World View

Economic development sometimes is affected as much by informal communications with the outside world as by formal education. Some sociologists maintain that mass media play a key role in transitional economies. The transformation of systems of value, the revolutionization of aspiration levels, and the conceptualization of world environment are all of basic importance in preconditioning a developing society. The quickening of the speed of adaptation and the dissemination of daily information are all fruits of a modern communication network which enables an economy to function at high efficiency. In such aspects mass media outrange classroom education both in the spectrum of its audience and in its power of persuasion.

While there is general agreement on the importance of "informal training," its measurement is still a challenging task. Newspaper circulation, T.V. and radio facilities, and movie attendance are the more direct

---

the value of the observation at that point, it seems desirable to use the reciprocal of the value as weights, i.e.

$$\frac{(\frac{1}{E_p} \cdot E_p) + (\frac{1}{E_s} \cdot E_s) + (\frac{1}{E_H} \cdot E_H)}{\frac{1}{E_p} + \frac{1}{E_s} + \frac{1}{E_H}} = \text{aggregate enrollment ratio}$$

Where E_p, E_s, and E_h are the enrollment ratios in primary, secondary, and higher education respectively.

indicators for this crucial yet nebulous element. We include two more indices, the literacy (or its converse, the illiteracy) rate and the stage of urbanization, since both determine, to a high degree, the effectiveness of various communication channels.[1]

---

[1] D. Lerner, The Passing of Traditional Society (Glencoe, Ill.: The Free Press, 1958). He points out that his survey of 54 countries indicates that empathy or the process of psycho-sociological identification is strongly influenced by mass media. He also shows that about ten percent of the population must be urbanized in cities of at least 50,000 inhabitants before cultural growth (change) occurs. For a development of this theory, see my paper "The Psycho-Sociological Nature of Nativistic Movements and the Emergence of Cultural Growth," Anthropos, 61 (1966).

PART III

Interpretation of Results

The statistical results of the phase model are summarized in Table 9. Table 10 lists the values of income and of the respective social indicators corresponding to the switch points. Figures 8 to 10 give a pictorial representation of the results shown in Table 9. The interpretation of these results will be presented in two parts:

(1) General, technical remarks. For all the 17 variables considered, there exist statistically significant switching points which partition each respective sample into two subsamples. In 16 cases the regression coefficient for the whole sample is higher than that coefficient for either subsample. However, the second subsample usually has much wider dispersion (variances) for both the social and economic variables. This tends to apply a large (weighted) average variance to both subsamples (especially the first) and hence overstate the overall regression coefficient. On the other hand, a comparison of the correlation coefficients for the two subsamples yield the following:

R is higher in:

First subsample	7 cases
Second subsample	10 cases

The result of the t-tests for the regression coefficient can be read from Table 9.

(2) Detailed, interpretive survey. Relating the criteria for the threshold

TABLE 9

PHASE MODEL: STATISTICAL TEST - 1960

Types of Indicators	Number of Observations	Switch Point t	-2 log λ	χ square distribution[a]		Y = a.S$_I$ + b + u: Constant	t test[a,b]	Slope	t test[a]	S.E.R.	R
NUTRITION											
1. Per capita calories/day	58	28	59.8658	.001	1-28	- 106.2699	.65	.1259	.30	90.3909	.2263
					29-58	-2,637.4255	.001	1.1755	.001	453.6847	.5722
					1-58	-1,580.7184	.001	.8206	.001	348.8595	.7474
2. Total proteins grams/day	58	35	75.5006	.001	1-35	- 118.7607	.30	4.9560	.02	102.2526	.4068
					36-58	832.5740	.55	1.9730	.90	506.4814	.0287
					1-58	-1,015.9615	.001	21.3987	.001	369.7430	.7123
HEALTH											
3. Inhabitants/physician (inverse)	39	21	41.8553	.01	1-21	141.3752	.05	305,876.0586	.05	139.8875	.4209
					22-39	1,409.1784	.01	-219,168.5723	.55	482.7484	-.1550
					1-39	111.6015	.40	659,278.2266	.001	423.7441	.6385
4. Number of hospital beds/ 1,000 inhabitants	38	20	41.5801	.01	1-20	105.0669	.20	46.8572	.05	118.1353	.4996
					21-38	1,050.8186	.02	8.8590	.80	446.3033	.0656
					1-38	83.1091	.45	85.2198	.001	383.1932	.7139
5. Infant mortality deaths/ 1,000 live births(increase)	41	17	33.8434	.01	1-17	212.8836	.05	402.5656	.95	93.1424	.0165
					18-41	- 16.8093	.95	26,481.2546	.01	429.5339	.5716
					1-41	- 157.9956	.20	29,456.4526	.001	344.3380	.7823
EDUCATION: FORMAL TRAINING											
6. Primary enrollment ratio	68	24	82.7149	.001	1-24	49.1503	.30	1.6334	.20	49.8505	.3137
					25-68	- 966.3370	.05	23.2668	.001	453.4827	.5224
7. Secondary enrollment ratio	68	46	105.2016	.001	1-46	87.8613	.01	4.3227	.01	89.2134	.4528
					47-68	833.4911	.05	3.4660	.50	460.7512	.1618
					1-68	- 12.0274	-0.1881	13.9499	9.4536	328.8978	.7584
8. Higher enrollment ratio	53	21	52.8325	.001	1-21	112.9185	.01	112.0466	.20	77.4048	.3237
					22-53	266.0264	.20	357.8047	.001	459.1701	.5567
					1-53	97.2731	.25	432.5133	.001	373.3088	.7199
9. Adjusted enrollment ratio	70	31	92.5659	.001	1-31	12.0013	.78	2.7908	.02	64.2687	.4455
					32-70	-1,700.6196	.001	30.7154	.001	405.7169	.6705
					1-70	- 579.6504	.001	16.5775	.001	347.5248	.7188
10. Enrollment ratio of all types	24	7	26.0862	.001	1-7	51.4113	.55	2.8690	.40	42.8824	.3896
					8-24	- 682.0327	.30	24.4123	.05	468.5417	.5005
					1-24	- 441.1670	.20	20.0837	.01	401.6813	.6127
11. Lagged total education	52	20	50.5607	.001	1-20	110.9444	.01	41.6170	.20	80.0543	.3189
					21-52	287.1562	.20	119.5298	.01	477.8945	.4997
					1-52	101.6248	.30	149.6963	.001	390.8651	.6897
INFORMAL TRAINING: WORLD VIEW											
12. Number of radios/1,000 inhabitants	61	39	94.0817	.001	1-39	72.5581	.001	1.8459	.001	65.7491	.7397
					40-61	473.8915	.02	1.9342	.01	381.9607	.6349
					1-61	84.1494	.10	2.8733	.001	268.1552	.8581
13. Number of T.V./1,000 inhabitants	40	18	41.0123	.01	1-18	185.5106	.001	6.0498	.50	93.7656	.1675
					19-40	602.0957	.001	4.2126	.01	405.5081	.6327
					1-40	338.6643	.001	5.8613	.001	350.3162	.7654
14. Movie attendance per capita	49	15	52.8286	.001	1-15	71.0527	.04	31.5569	.08	60.3994	.4535
					16-49	502.3732	.03	28.2632	.15	534.7483	.2493
					1-49	204.0527	.07	52.1085	.001	470.2484	.5178
15. Illiteracy rate (inverse)	26	11	25.1773	.01	1-11	116.9511	.10	- 160.6630	.95	53.7927	-.0152
					12-26	328.1462	.001	173.8214	.75	209.7799	.0935
					1-26	213.2525	.001	588.5352	.20	191.4182	.2728
16. Daily newspaper circulation/ 1,000 inhabitants	70	34	106.6053	.001	1-34	72.1054	.001	1.7348	.001	39.4906	.6001
					35-70	117.5911	.40	2.9366	.001	366.1267	.7236
					1-70	57.3744	.20	3.1264	.001	265.8134	.8469
17. Urbanization, percentage of population in cities above 20,000 inhabitants	27	7	24.7016	.001	1-7	- 10.9839	.80	16.3316	.01	47.4397	.9227
					8-27	217.7243	.55	13.2384	.15	473.4671	.3597
					1-27	47.1196	.80	16.7431	.001	412.0149	.5954

[a] Significant on the percent level (or better).

[b] The t test establishes whether the coefficient is significantly different from zero. t test degrees of freedom, n-2.

TABLE 10

PHASE MODEL: SOCIAL THRESHOLDS - 1960

Indicators	Number of Obser-vations	Switch point	Threshold value of: Income[a]	Social indicator[a]	
NUTRITION					
1. Per capita calories/day	58	28	162	2,353	(4)
2. Total proteins grams/day	58	35	232	75.3	
HEALTH					
3. Inhabitants/physician (inverse)	39	21	318	.0008	
4. Number of hospital beds/1,000 inhabitants	38	20	429	5.6	
5. Infant mortality death/1,000 live birth (inverse)	41	17	237	.0194	(4)
EDUCATION: FORMAL TRAINING					
6. Primary enrollment ratio	68	24	127	48	(5)
7. Secondary enrollment ratio	68	46	270	38	(2)
8. Higher enrollment ratio	53	21	224	.65	
9. Adjusted enrollment ratio	70	31	121	57	
10. Enrollment ratio of all types	24	7	148	32	
11. Lagged total education enrollment ratio	52	20	224	1.8175	
INFORMAL TRAINING: WORLD VIEW					
12. Number of radios/1,000 inhabitants	61	39	260	132	
13. Number of T.V./1,000 inhabitants	40	18	252	7.0	
14. Movie attendance/capita	49	15	170	3.10	
15. Illiteracy rate (inverse)	26	11	105	.0243	
16. Daily newspaper circulation/ 1,000 inhabitants	70	34	164	56	
17. Urbanization, percentage of population in cities above 20,000 inhabitants	27	7	357	23.5	(1)

[a]These values are approximations taken on the basis of the last three years preceding and including the switching point (except as indicated in parentheses). The exact values would be as estimated by regression.

FIGURE 8

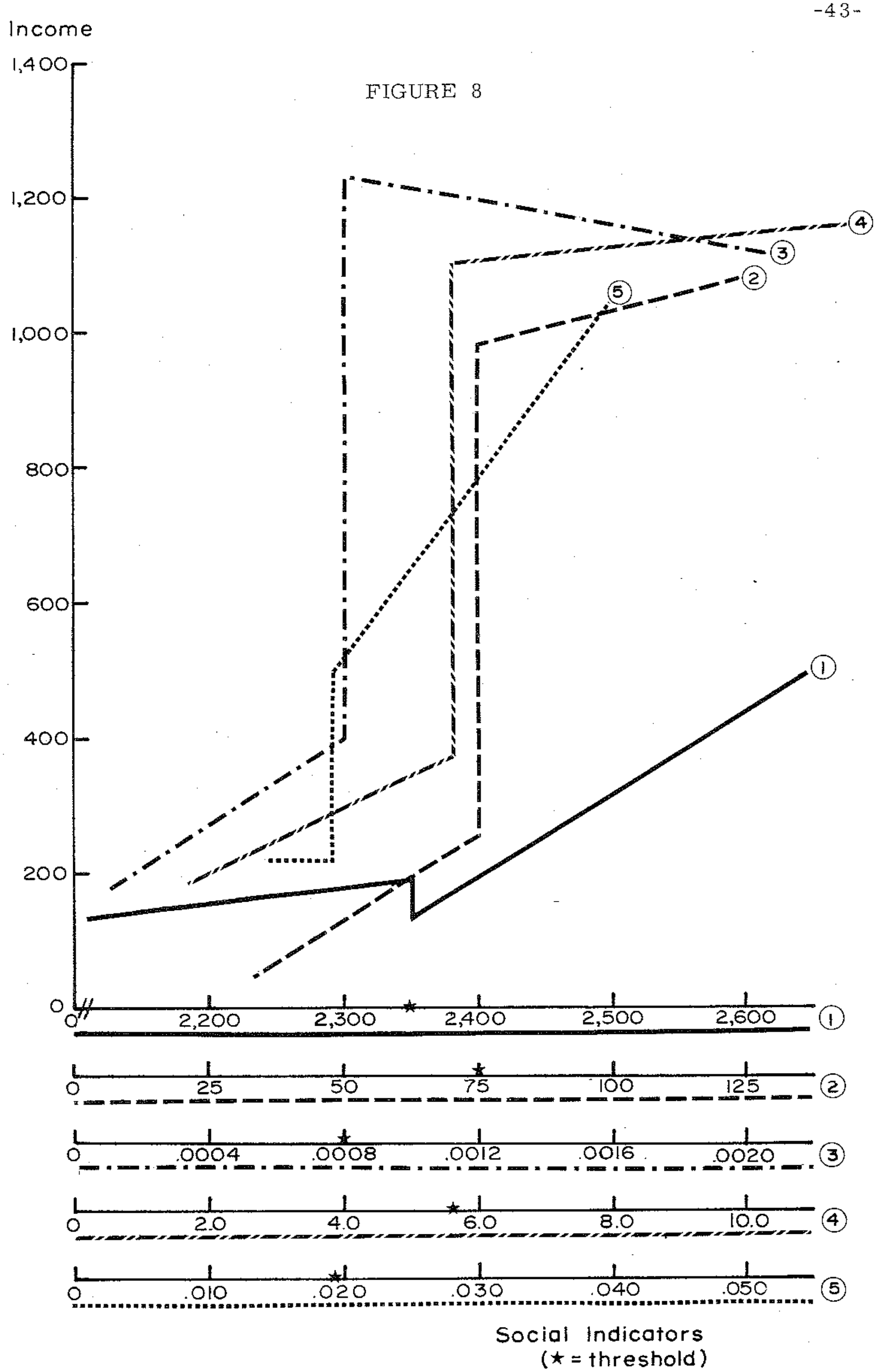

Income
1,400
1,200
1,000
800
600
400
200
0
0
2,200
2,300
2,400
2,500
2,600
0
25
50
75
100
125
0
.0004
.0008
.0012
.0016
.0020
0
2.0
4.0
6.0
8.0
10.0
0
.010
.020
.030
.040
.050
Social Indicators
(★ = threshold)

FIGURE 9

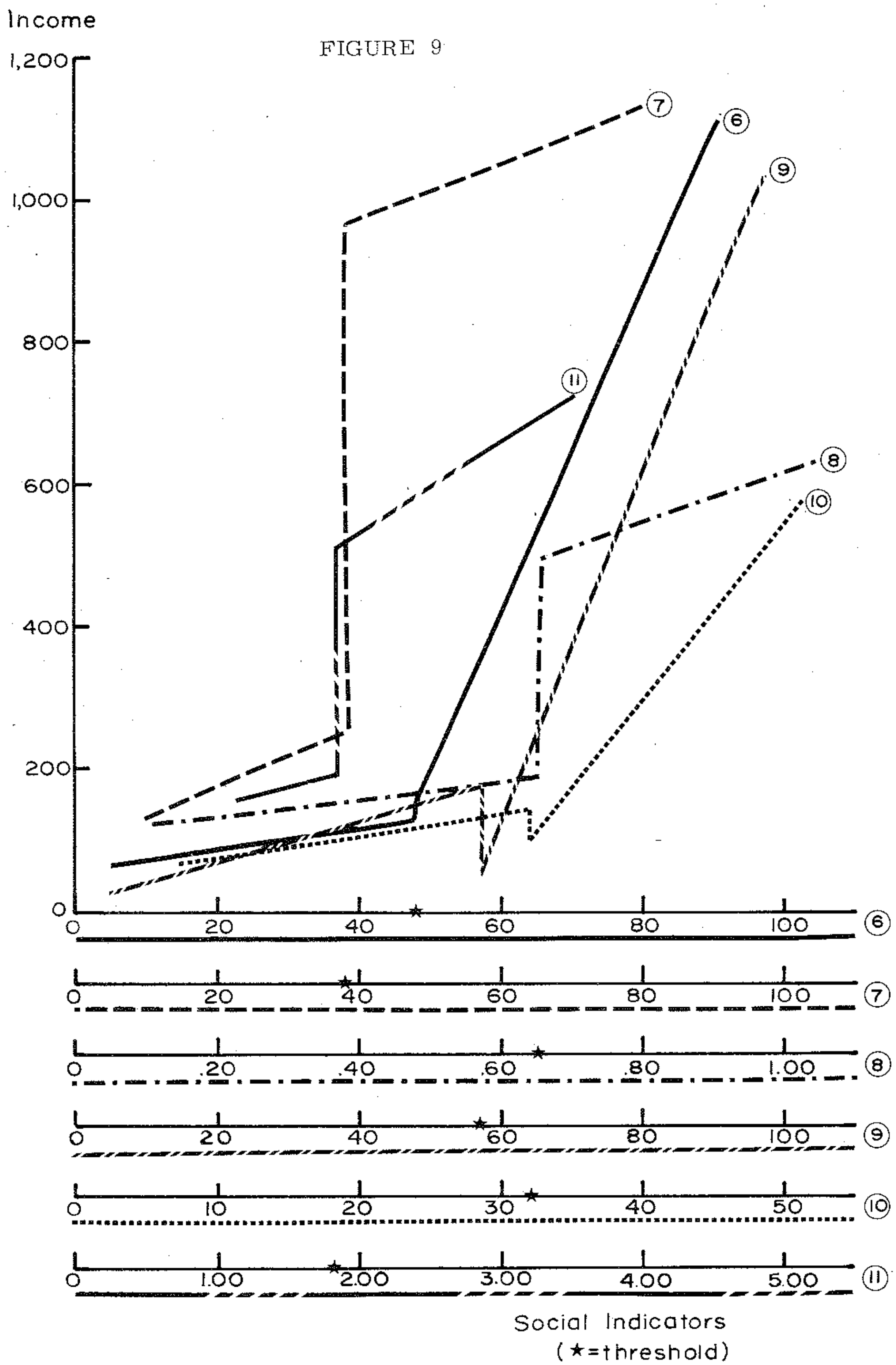
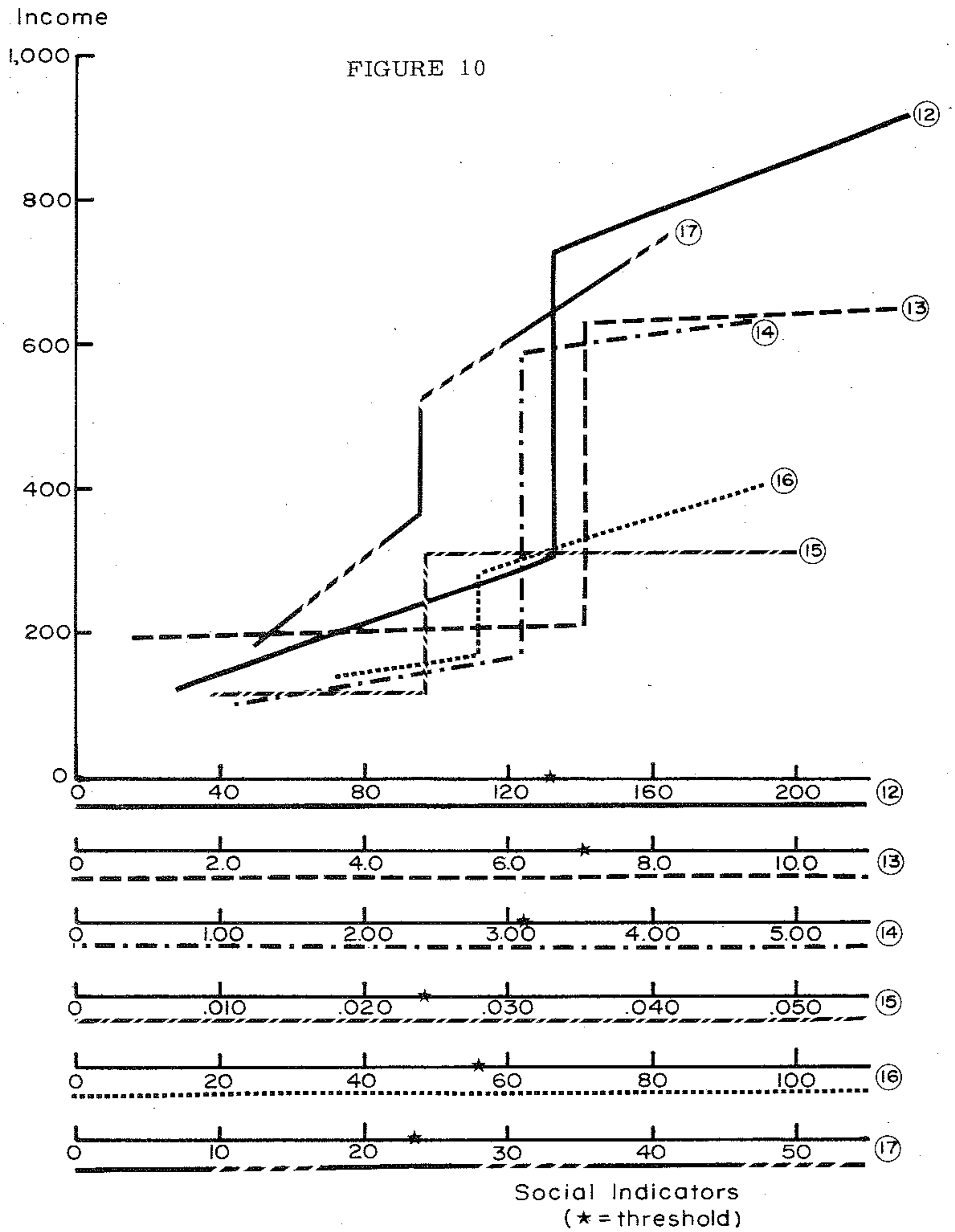

Income
1,200
1,000
800
600
400
200
0
0 20 40 60 80 100 (6)
0 20 40 60 80 100 (7)
0 .20 .40 .60 .80 1.00 (8)
0 20 40 60 80 100 (9)
0 10 20 30 40 50 (10)
0 1.00 2.00 3.00 4.00 5.00 (11)
Social Indicators
(★=threshold)

FIGURE 10


Income
1,000
800
600
400
200
0
(12)
(13)
(14)
(15)
(16)
(17)
0 40 80 120 160 200
0 2.0 4.0 6.0 8.0 10.0
0 1.00 2.00 3.00 4.00 5.00
0 .010 .020 .030 .040 .050
0 20 40 60 80 100
0 10 20 30 40 50
Social Indicators
(★ = threshold)

considered under (d)2 and the results as shown in Table 9, we obtain the following table:

1.	Per capita calories	Case 3 (. -)	Threshold not shown
2.	Protein intake	Case 2 (- .)	Threshold shown
3.	Physicians	Case 2 (- .)	Threshold shown
4.	Hospital beds	Case 2 (- .)	Threshold shown
5.	Infant mortality	Case 3 (. -)	Threshold not shown
6.	Primary schooling	Case 3 (. -)	Weak threshold
7.	Secondary schooling	Case 2 (- .)	Threshold shown
8.	Higher education	Case 3 (. -)	Weak threshold
9.	Adjusted enrollment	Case 1 (--)	Threshold shown
10.	All enrollment	Case 3 (. -)	Weak threshold
11.	Lagged enrollment	Case 3 (. -)	Weak threshold
12.	Radio	Case 1 (--)	Threshold shown
13.	T.V.	Case 3 (. -)	Threshold not shown
14.	Movies	Case 2 (- .)	Threshold shown
15.	Illiteracy	Case 1 (. .)	Threshold not shown
16.	Newspapers	Case 1 (--)	Threshold shown
17.	Urbanization	Case 4 (- .)	Threshold shown

The interpretation of the results by groups is as follows:

<u>Group 1: Food</u>

Within this group there exists a threshold for the protein intake but none for the calorie intake as evidenced by whether there exists a strong correlation in the first subsamples (low-income part). This may mean that in the developing world today it is the poor quality rather than the inadequate quantity that forms the effective bottleneck for economic performance.

Group 2: Health

There exist thresholds for physicians and hospital beds. There is none for the (reciprocal of) the infant mortality rate. The latter was included as an indirect indicator for general health, but it seems that the relation on economic performance is not very strong.

Group 3: Education

Six indices are tested all in all. Among three "original" indices, the secondary enrollment ratio shows a threshold. The primary enrollment and the higher education enrollment show a weak[1] threshold. Presumably, the difference in primary school system (4 vs. 6 years) reduces the statistical significance. But it is also possible that the universality of primary education is, in fact, inconsequential to the early stages of development, since an unskilled worker needs little formal training. On the other hand, higher education enrollment need not be important since students can be trained in developed countries and skilled personnel can be hired from abroad. The other three "derived" indices show weak thresholds with the possible exception of the adjusted enrollment ratio. The explanation can be that either the method of adjustment is still not good enough or that there does not exist a well-marked threshold (except for secondary education).

Group 4: World View

Among the six indices, urbanization, radio, newspaper circulation, and movie attendance show prominent signs of thresholds. T.V. and the

---

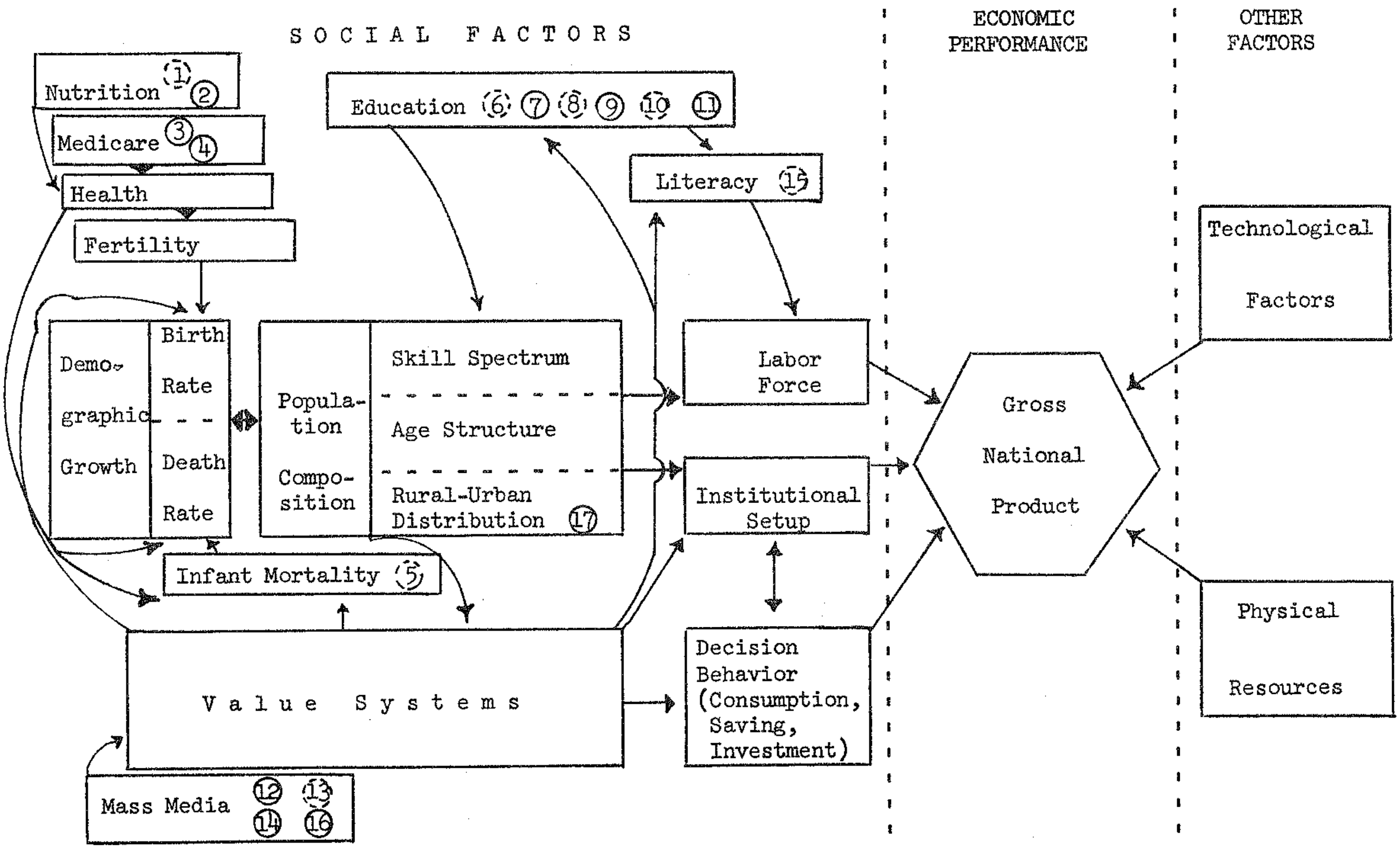
[1] The correlation coefficient is rather low.

(reciprocal of) illiteracy rate do not possess thresholds at all. For illiteracy, the interpretation is similar to that of primary enrollment. For T.V, the initial cost of setting up the network excludes its easy adoption by the low-income countries. The important point is that even in the cases where thresholds for the indicators of world view can be found, the interpretation has to be cautious. None of these may constitute an important bottleneck for growth, with urbanization as a possible exception. But they may influence as well as indicate social attitudes which are crucial to development.

The various binary studies reflect overall interrelationships which can be depicted in Chart 1. From this diagram we note the pervasive role played by the population structure and demographic factors. Together with the value systems, etc., these variables influence the (quality of the) labor force, the institutional setup, and the decision behavior which, in turn, affect the economic performance. The population structure is determined by a three-way scheme: the skill spectrum (or stock of skill), the age distribution and the rural-urban distribution. That education and urbanization affect the quality of the labor force, socio-political institutions, and value systems was to be readily accepted. However, the effect on the age structure requires a little elaboration. One index of the age structure is the dependency ratio (the percentage of people under 15 and over 65), which influences the family size, the number of children to be educated, and the segment of the population within the productive age. All these elements in turn influence the saving versus consumption behavior through the value systems. The age distribution, and its structural modifications, are

CHART I

DETERMINANTS OF ECONOMIC PERFORMANCE



determined by changes in mortality and fertility or the death and birth rates, whose difference is the rate of demographic growth.[1]

The death rate is generally determined by the age structure of the population, morbidity and health conditions. On the other hand, infant mortality which materially affects the general death rate in low-income countries is particularly sensitive both to the improvements and deterioration in environmental circumstances affecting health.[2] This, however, is influenced by such factors as social values, family size, etc., besides the other factors. Birth rate is influenced by such sociological factors as the age of marriage, the motivation for large families (potential support for old age, etc.), as well as biological fertility which is influenced by health.

Health is affected by nutrition and medicare. The systems of values are modified by education and mass media communications. Education, in turn, is reflected in the literacy rate and enrollment ratios.

---

[1]In order to grasp the full impact of the diagram, it may be well to remember certain facts distinguishing developing from developed countries. For instance, the United Nations reports that in developing countries the proportion of children under 15 years of age exceeds 40 percent of the population, and is as high as 45 percent in a few cases, while the corresponding proportion in the developed countries varies generally within the range of 24 to 32 percent. See UN Inquiry Among Governments on Problems Resulting from the Interaction of Economic Development and Population Changes, UN/E/3895/Rev. 1, 23 Nov. 1964. The "dependent" age groups may vary from 49 to 53 percent in Africa to 36 to 38 percent in industrialized countries. See A. A. Igun, "Demographic Approach to the Problems of Social and Economic Development in Africa," UN World Population Conference, W.P. 360 (September, 1965).

[2]G. Z. Johnson, "Public Health Activities as Factors in Levels and Trends of Mortality and Morbidity in Developing Countries," UN/WPC/WP/418.

As we can see from the diagram, there are nonquantifiable factors like systems of value. There are also variables for which adequate statistical series are not available (dependency ratio, etc.). For the phase model we were able to select 17 social indices and these are represented by circled numbers in our chart. For some the data show evidence that a threshold exists in the observed data range in the sense that they seem to present a tight constraint for GNP. For others, no such evidence reveals itself in the study. The former are represented by solid circles; the latter by dotted ones.

Owing to the imperfection of our method (more satisfactory approaches are not readily usable at present) and the available data, perhaps the results can serve only as a broad indication. Better data and a multi-variable (rather than binary) approach may yield better results.

METHODOLOGICAL APPENDIX

Factor Analysis

Factor analysis is a statistical method for separating out common determinants of the behavior of several variates. More precisely, it aims to describe the probable number of common ingredients in the variables, given an array of observed correlations between them. The technique is especially useful when it is suspected that the observed variates are composites of more fundamental elements, but where these basic elements are not directly accessible to the experimenter. In such cases, the physical (manipulative) decomposition of the variate may be replaced by an inferred virtual decomposition.

The problem confronted by factor analysis, and its techniques, may be given a simple geometrical representation. Suppose that the experimenter has measured n variates over a population sample. The observations comprise n-tuples, i.e. points in n-dimensional space. The experimenter is concerned with determining the number of factors that will explain best the pattern of the sample. Essentially, this means that he seeks to approximate the point distribution by some linear structure--a line, place, or the whole space, according as the variates are deemed to share one, two, or more common factors.

There is an element of ambiguity which must be noted at this point; it centers on the fact that any finite number of factors will almost never explain all the variance of the sample. Rather, in general our best hope is to explain a large amount of the variance among the variables by looking at a few common factors, and leave a relatively small amount to be explained by appeal to other factors; the so-called "specific factors." The ambiguity emerges because one may trade dimension of the fitted hyperplane for "communality" (degree of commonness of the factors); it becomes at this point necessary to specify a criterion for the best balance between communality and dimensionality. Some of the criteria proposed for the determination of factor dimension are:

1. The minimum rank attainable by using arbitrary coefficients of communality down the diagonal of the covariance matrix, allowing for small errors of observation in the convariance matrix.
2. By iterative procedures simultaneously determining dimension and communalities "to produce a fit to the original correlation matrix which is satisfactory at a given significance level."[1]

However they might be determined, once the communalities and dimensionalities are known, one may construct a "reduced" correlation matrix, with the communality coefficients down the diagonal and the off-diagonal elements as before. This becomes our "true" correlation matrix: the existence of the specific factors and of error factors shows itself in coefficients of self-correlation of less than unity.

Having reduced the matrix, one may now proceed to the geometric problem--fitting the (dimension now determined) hyperplane to the distribution. The intuitive idea here is that of extracting the direction of highest correlation from the matrix and calling it the first factor, then

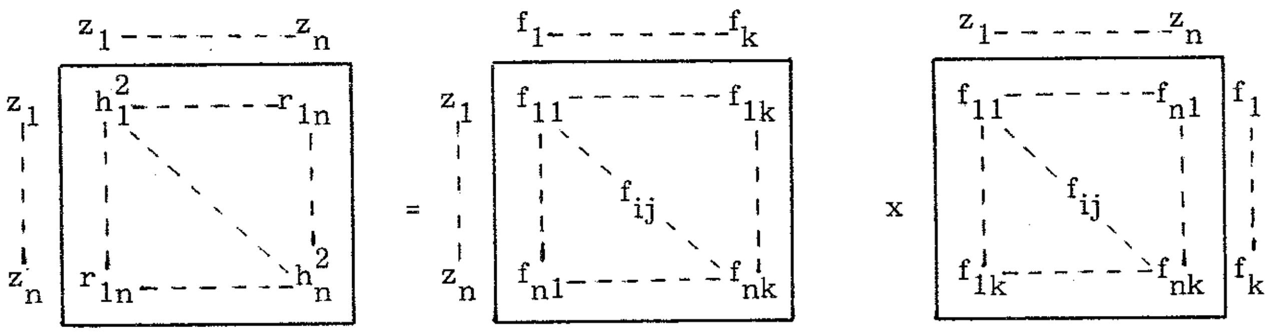
---

[1]For further explanations on Hotelling Iterative Procedure, see Raymond B. Cattell, op. cit.

reducing the matrix once more to eliminate the effect of this factor, and repeating the process until all the factors are extracted. As it happens, this is the same as the mathematical procedure of analyzing the eigen-values of the original reduced matrix, so the discussion from this point on will be approached from this mathematical viewpoint.

The rank of the reduced matrix, k, will be determined by whatever dimensionality criterion is employed. Generally, there will be k eigenvalues $(v_{1i}, \ldots, v_{ni})$ and k eigenvalues λ_i.

Now, the reduced matrix R is to be accounted for completely by the common factors, a situation which can be summed up in the matrix formulation hereafter:



R = the correlation matrix

Factor analysis

Transposed factor matrix

where $z_1, \ldots z_n$ are the original variables, $r_{1n}, \ldots$ are correlation co-efficients, $h^2, \ldots$ are communalities, $f_1, \ldots, f_k$ are factors.

The problem is reduced accordingly to finding the coefficients (f_{ij}) of the factor matrix. It may be stated without proof, that if

$$\begin{bmatrix} f_{11} & \cdots & f_{1k} \\ \vdots & f_{ij} & \vdots \\ f_{n1} & \cdots & f_{nk} \end{bmatrix} = \underset{\text{eigenvectors}}{\begin{bmatrix} v_{11} & \cdots & v_{1k} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nk} \end{bmatrix}} \underset{\substack{\text{square roots of} \\ \text{eigenvalues}}}{\begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_k} \end{bmatrix}}$$

then the desired result will appear.

A final note is necessary here on "rotation." What has really been determined is a hyperplane; this may be obscured since the hyperplane is that spanned by the eigenvectors. There is nothing special about this particular set of vectors, even though we may call them the "common factors." Any other k vectors spanning the same hyperplane would do as well, and in general one would expect that an orthogonal transformation of the vector set might be required for conceptual simplification of the common factor array.

DATA SOURCES

Nutrition

(a) Per capita calories per day: FAO Yearbook and FAO worksheets as communicated to the Institute.

(b) Per capita total proteins per day: idem.

Health

(a) Inhabitants per physician: Years 1950 and 1960: W. Galenson and G. Pyatt, The Quality of Labour and Economic Development (Geneva: ILO, 1964), Table 5, p. 105; Year 1955: UN: Statistical Yearbook, 1957, Table 179, p. 595.

(b) Number of hospital beds per 1,000 inhabitants: idem.

(c) Crude death rate: UN, Demographic Yearbook, 1950, pp. 55, 60.

(d) Infant mortality death per 1,000 live births: Years 1950 and 1960: W. Galenson and G. Pyatt, op. cit.; Year 1955, UN Demographic Yearbook 1957, Table 9, p. 209.

Education: Formal Training

(a) Primary education enrollment ratio: UNESCO, Statistical Yearbook, 1963, Table 9.

(b) Secondary education enrollment ratio: idem.

(c) Higher education enrollment ratio: UNESCO, Statistical Yearbook, 1963, Table 15.

(d) Adjusted enrollment ratio (primary and secondary combined): UNESCO, Statistical Yearbook, 1963, Table 9.

(e) Enrollment ratio of all types (primary, secondary, and higher combined): Demographic Yearbook, 1963 (New York: United Nations, 1963), Table 15.

Informal Training: World View

(a) Number of radios per 1,000 inhabitants: UNESCO, Yearbook, 1963; UN, Statistical Yearbook, various years.

(b) Number of T.V. per 1,000 inhabitants: idem.

(c) Movie attendance per capita: idem.

(d) Illiteracy rate: UN, Compendium of Social Statistics, 1963; B. M. Russett, et al., World Handbook of Political and Social Indicators (New Haven: Yale University Press, 1964); UN Statistical Yearbook, various years.

(e) Daily newspaper circulation per 1,000 inhabitants: idem.

(f) Urbanization, percentage of population in cities above 20,000 inhabitants: UN, Compendium of Social Statistics, B.M. Russett, et al., op. cit.

LIST OF COUNTRIES

Argentina (47)*
Australia (48)
Austria (49)
Belgium (50)
Bolivia (1)
Brazil (2)
Burma (3)
Cameroon (4)
Canada (51)
Ceylon (5)
Chile (52)
China Taiwan (6)
Columbia (7)
Congo (Leo) (8)
Costa Rica (9)
Denmark (53)
Dominican Rep. (10)
Ecuador (11)
El Salvador (12)
Fed. of R. and N. (13)
Finland (54)
France (55)
Germany Fed. Rep. (56)
Ghana (14)
Greece (15)
Guatemala (16)
Honduras (17)
Hong Kong (18)
India (19)
Indonesia (20)
Irak (69)
Ireland (57)
Israel (58)
Italy (59)
Jamaica (21)
Japan (60)
Jordan (22)
Kenya (23)
Korea (South) (24)
Lebanon (25)
Madagascar (26)
Malaya, Fed. of (27)
Mexico (28)
Morocco (29)
Netherlands (61)
New Zealand (62)
Norway (63)
Pakistan (30)
Panama (31)
Paraguay (32)
Peru (33)
Philippines (34)
Portugal (35)
Puerto Rico (64)
Spain (36)
Sudan (37)
Sweden (65)
Switzerland (66)
Syria (38)
Tanganyika (39)
Thailand (40)
Tunisia (41)
Turkey (42)
Uganda (43)
United Kingdom (68)
United States (67)
Uruguay (44)
Venezuela (70)
Viet Nam (South) (45)
Yugoslavia (46)

---

*The numbers refer to these countries as they appear on scatter diagrams.

TABLES OF DATA*

1 to 3	Economic indicators Developing Countries	1950 - 55 - 60
4 to 6	Economic indicators Developed Countries	1950 - 55 - 60
7 to 9	Social indicators Developing Countries	1950 - 55 - 60
10 to 12	Social indicators Developed Countries	1950 - 55 - 60
13 to 15	Subsidiary Social indicators Developing Countries	1950 - 55 - 60
16 to 18	Subsidiary Social indicators Developed Countries	1950 - 55 - 60
19	Subsidiary Social indicators Developing Countries	1960
20	Subsidiary Social indicators Developed Countries	1960

*These tables are kept on file at the Institute. All sources are official sources.