A NOTE ON REGIONAL ECONOMIC DEVELOPMENT
——A FACTOR ANALYSIS ON THE JAPANESE CASE*——

By KOZO YAMAMURA**

I. *Introduction*

Causes and consequences of regional differentials of degree of economic development—however measured—have received much attention by economists who are interested in economic development and regional economics. Numerous articles written on the subject, however, are mostly either theoretical or in the form of debates concerning the respective authors' views on goals and means of regional economic development.[1]

In addition to difficulties in obtaining data, one of the major problems of regional economic research has been to establish a generally-agreed-upon measure of "degree of economic development."[2] If the level of per capita income in regions is highly correlated with the level of economic development, it is useful to have such a measure which we could somehow attempt to increase. Such an indicator, over and beyond the general Clarkian type of description, could be of significant use in measuring the results of a policy, the impact of economic growth on respective regions, or in isolating key variables or parameters which cause the indicator to change either relatively or absolutely. If such an indicator shows no relationship, or only a weak one, with the level of regional per capita income, then policy makers need to consider other avenues of increasing the income of low income regions.

This paper attempts to suggest a possible indicator which is derived from a technique suggested by Pal[3] and by use of factor analysis. The resulting indicator, then, will be used to examine some aspects of the Japanese economy between 1955 and 1960. Due to difficulties

* Mr. M.N. Pal of the Indian Statistical Institute, New Delhi, India, was very helpful in offering his comments and guidance. Professor Hisao Nishioka of Aoyama Gakuin University, Japan, commented on the initial draft of this paper. My thanks are also due to the San Diego State College Foundation which provided a small grant for calculation of the data.

** Associate Professor of Economics, San Diego State College, U.S.A.

[1] Important contributions such as Hanna's and numerous others in the U.S., and the recent volume edited by Shinohara in Japan comprise extensive empirical research. But what I am suggesting here is the relative neglect of empirical research which could be more directly useful in policy formulation. For further comments and citations on the Shinohara volume, see footnote 2.

[2] In addition to the Clarkian description of primary, secondary and tertiary sector breakdowns, the share of the secondary sector, value added in manufacturing per worker, income originating in the non-agricultural sector, and so on, have been used. M. Umemura, "Structure and Variation in Regional Employment" in M. Shinohara, ed., *Econometric Analysis of Regional Economic Structure* (Hitotsubashi University Research Series, 1965) uses these and other methods.

[3] M.N. Pal, "A Method of Regional Analysis of Economic Development with Special Reference to South India," *Journal of Regional Science*, Vol. 5, No. 1, 1963, pp. 41–58. For a detailed derivation of the equation (1), see: Pal, "Zur Berechnung eines Kombinierten Konzentrationsindexes: Ein Beitrag zur Methode der Regionalanalyse," *Raumforschung und Raumordnung*, Vol. 20 (1963) pp. 87–93.

in obtaining data at the regional level for a longer period, what we can examine is limited and only useful in suggesting possible uses of such an indicator. Though the time span observed is extremely brief for any long range observations, this was the period in which the Japanese economy grew at an annual average rate of 10.2 per cent in real terms and in which several major laws to "develop less-developed regions" were enacted and put into effect.[4]

This paper proceeds as follows: (1) an "index of concentration" suggested by Pal was calculated for two points of time (1955 and 1960); (2) these indices then were used as raw data in two sets of factor analyses to obtain Z scores which could be considered indicators of "degree of economic development;" (3) these Z scores were regressed on various regionally significant variables; and (4) some tentative observations and possible courses of further research are ventured.

II. *"Index of Concentration"*

To obtain raw data to be used in the following factor analysis, indices of economic development (I), as defined below, were calculated using the equation (1) adopted by Pal:[5]

$$I_i=\frac{\sigma_{D_i}}{\sigma_{D_i}\cdot\bar{P}_i+\sigma_{P_i}\cdot\bar{D}_i}P+\frac{\sigma_{P_i}}{\sigma_{D_i}\cdot\bar{P}_i+\sigma_{P_i}\cdot\bar{D}_i}D_i\,,\qquad (i=1,\ldots,7)$$

where P_i=a percentage variable, $D_i=\log_e d_i$ where d_i denotes a density variable as shown in Table 1, $\bar{P}_i$ and $\bar{D}_i$=the mean values, and σ_{P_i} and σ_{D_i} are the standard deviation of P_i and D_i respectively. (see Table 1). These coefficients of P_i and D_i in equation (1) have been chosen by Pal as this form satisfies $r_{I_iP_i}=r_{I_iD_i}$, which are the coefficients of correlation of I_i with P_i and D_i respectively. The total variation of P_i and D_i as explained by I_i is:

$$(r^2_{I_iP_i}+r^2_{I_iD_i})=\frac{1}{2}(1+r_{P_iD_i}) \tag{2}$$

where $r_{P_iD_i}$ is the coefficient of correlation (positive value) between P_i and D_i. Pal has also shown that this is the maximum variation that the linearly combined variable I_i of P_i and D_i can explain. The value of $I_i>1$ represents a higher level and $I_i<1$ represents a lower level of "concentration" than that of the mean "concentration" in Japan in 1955 and 1960. Tables 2 and 3 give the results of the calculation of equation (1) for all seven indices and the related coefficients for 1955 and 1960.

---

[4] In Professor Ito's words: "Beginning 1956, Acts to Promote Regional Development have come into being one by one in Tohoku, Kyushu, Shikoku, Chugoku, and Hokuriku regions. When these acts are considered with the Capital (Tokyo) Reorganization Act, it is clear that the development of less developed regions and reexamination of developed regions have become our concern... After completing the postwar reconstruction, the questions concerning the reevaluation of conditions of industrial location, and readjustment of regional income differentials have now come to the forefront replacing the problems of dam building and food supply." Z. Ito, *Kokudo Kaihatsu no Keizai Gaku* (Economics of National Development) (Shunjyusha Publishing Co., 1965) 3rd Edition, p. 64.

[5] The nature of equation (1) can perhaps be better grasped intuitively when it is realized that it can be re-written as follows:

$$I_i=\frac{\frac{P_i}{\sigma_{P_i}}+\frac{D_i}{\sigma_{D_i}}}{\frac{\bar{P}_i}{\sigma_{P_i}}+\frac{\bar{D}_i}{\sigma_{D_i}}}$$

TABLE 1. DEFINITION OF VARIABLES RELATED TO INITIAL INDICES OF CONCENTRATION

Index	P	$d=e^D$
I_1=Index of urbanization	P_1=Those living in cities of over 30.000 as a percentage of total population	d_1=density of urban population per square kilometer of total area of a prefecture
I_2=Index of concentration of secondary industries	P_2=Employees in secondary industry as a percentage of total labor force	d_2=density of employees in secondary industries per square kilometer of total area of a prefecture
I_3=Index of concentration of manufacturing industries	P_3=Employees in manufacturing industry as a percentage of total labor force	d_3=density of employees in manufacturing industries per square kilometer of total area of a prefecture
I_4=Index of concentration of those engaged in financial insurance and real estate (FI & R) activities	P_4=Employees in FI & R as a percentage of total labor force	d_4=density of employees in FI & R per square kilometer of total area of a prefecture
I_5=Index of concentration of those engaged in wholesale and retail industries	P_5=Employees in wholesale and retail as a percentage of total labor force	d_5=density of employees in wholesale and retail per square kilometer of total area of a prefecture
I_6=Index of concentration of those engaged in transportation, communication and public utilities (TC & PU)	P_6=Employees in TC & PU as a percentage of total labor force	d_6=density of employees in TC & PU per square kilometer of total area of a fecture
I_7=Index of concentration of those engaged in service industries	P_7=Employees in service industries as a percentage of total labor force	d_7=density of employees in service per square kilometer of total area of a prefecture

Source: Research Division, Economic Planning Agency of the Japanese Government, *Chiiki Keizai Tokei Yoran* (Essential Statistics of Regional Economies), 1960, and Toyo Keizai Publishing Co.'s 1958, 1959 and 1962 volumes of *Keizai Tokei Nenkan* (Annual Economic Statistics) were used for area of prefectures and the urbanization data.

Notes on I's:

I_1: In the 1950 data, two cities which were very close to 30,000 (*i.e.*, over 29,000) were included.

I_2: This consists of manufacturing, mining and construction industries.

I_7: This consists of all who are engaged in the tertial sector less those engaged in FI&R, TC&PU, wholesale and retail sales, and government employees.

It should be made clear that these seven indices were chosen *subjectively* and were those available in usable form by prefectures (a political unit which will be used in this paper in lieu of the term "region"). In no way is it claimed that the present mix of indices chosen yields a definite and more reliable Z score than possible alternative mixes.[6] The mix of indices could be improved when more pertinent data are published for each prefecture and further discussions on index selection yield a more objectively agreed upon mix of indices.[7]

[6] The mix of I's of course needs to be chosen based on some theoretical framework. Those indices selected for this paper are those which are often used. As will be made clear, in the course of our discussion, various mixes of I's are to be chosen depending the nature and purpose of investigation. See footnote 19.

[7] The source cited in Table 1 and the Shinohara volume, *op. cit.*, contain many other data which could be used had they been available for 1955 and 1960. However, they are available for selected years only and those available for a longer time period often appear in unusable form. The question of what are "pertinent" data will be discussed shortly.

The values of correlation coefficients given in column 3 of Tables 2 and 3 are so high that each of the seven I's can be considered as a very good representative of P_i as well as D_i. This comment also applies to the amount of total variation as given in column 4 of the same two tables. Given these two sets of equations in the form of $I_i=aP_i+bD_i$, where a and b are coefficients as shown in Tables 2 and 3, we have computed 2 sets of I_{ij}'s, one for 1955 ahd the other for 1960. ($i=1, \ldots, 7$, and $j=1, \ldots, 46$, or 322 I_{ij}'s for each year.)

TABLE 2. EQUATIONS OF INITIAL INDICES AND THE VALUES OF RELATED COEFFICIENTS, 1955

	Equation of I_i according to equation (1)	Correlation of I_i with each of P_i and D_i $r_{I_iP_i}=r_{I_iD_i}$	Amount of total variation explained equation (2)
1	2	3	4
1	$I_1=0.7940\,P_1+0.1235\,D_1$	.9580	.9177
2	$I_2=2.250\,P_2+0.1699\,D_2$	.9668	.9347
3	$I_3=2.9789\,P_3+0.1994\,D_3$	.9745	.9496
4	$I_4=63.4820\,P_4+0.3394\,D_4$	.9711	.9431
5	$I_5=4.6643\,P_5+0.1537\,D_5$	.9722	.9453
6	$I_6=15.6684\,P_6+0.2020\,D_6$	.9111	.8301
7	$I_7=5.7314\,P_7+0.1613\,D_7$	.9544	.9110

TABLE 3. EQUATIONS OF INITIAL INDICES AND THE VALUES OF RELATED COEFFICIENTS, 1960

	Equation of I_i according to equation (1)	Correlation of I_i with each of P_i and D_i $r_{I_iP_i}=r_{I_iD_i}$	Amount of total variation explained equation (2)
1	2	3	4
1	$I_1=0.8021\,P_1+0.1090\,D_1$	.9640	.9293
2	$I_2=1.8773\,P_2+0.1556\,D_2$	.9685	.9381
3	$I_3=2.4856\,P_3+0.1847\,D_3$	.9762	.9530
4	$I_4=54.9488\,P_4+0.2978\,D_4$	.9758	.9522
5	$I_5=4.4247\,P_5+0.1277\,D_5$	.9633	.9279
6	$I_6=14.6151\,P_6+0.1775\,D_6$	.8962	.8032
7	$I_7=6.0767\,P_7+0.1227\,D_7$	.9197	.8458

Using, then, these calculated I_{ij}'s as raw data, we can now construct a composite index of "economic development" (Z), provided that, as Pal has written, (i) the inter-correlations among the variables are high; (ii) the explained variation in quite high (over 50 per cent) and (iii) there is no difficulty in interpretation of the final composite index. The conditions (i) and (ii) are clearly met as a glance at Tables 2 and 3 shows, and condition (iii) is also met, as will be made clear shortly, in that the use of the final Z scores as a measure of "economic development," the writer believes, is easy to interpret.

The composite index (Z) then was calculated for 1955 and 1960 using a principal-components-factor analysis which is frequently used in econometric studies.[8] Under the above assumptions, and presuming that regional differences account for the majority of the indices' variation,

the first principal component, or factor, will logically be a regional dimension.[8] The formula for such a score is:

$$Z_j=\sum_{i=1}^{7}\left[\frac{I_{ij}-\bar{I}_i}{\sigma_{I_i}}\right]\cdot F_i$$

TABLE 4. COMPOSITE INDEX Z FOR 46 PREFECTURES 1955 AND 1960

1	2	3 Z_{1955}	4 Z_{1960}	1	2	3 Z_{1955}	4 Z_{1960}
1	Hokkaido	−5.218	−4.326	24	Mie	− .119	− .400
2	Aomori	−6.158	−6.320	25	Shiga	−1.828	−2.754
3	Iwate	−7.947	−7.624	26	Kyoto	8.407	7.995
4	Miyagi	−2.345	−2.061	27	Osaka	18.513	18.089
5	Akita	−7.221	−7.411	28	Hyogo	7.725	8.161
6	Yamagata	−5.052	−4.951	29	Nara	− .204	2.017
7	Fukushima	−5.625	−6.344	30	Wakayama	.176	.551
8	Ibaraki	−4.239	−3.978	31	Tottori	−3.403	−2.777
9	Tochigi	−2.292	−2.406	32	Shimane	−5.406	−5.490
10	Gunma	−1.574	−1.577	33	Okayama	−1.195	− .948
11	Saitama	3.688	4.314	34	Hiroshima	1.681	1.528
12	Chiba	.129	.791	35	Yamaguchi	1.549	1.506
13	Tokyo	22.029	21.371	36	Tokushima	−3.785	−2.909
14	Kanagawa	14.406	13.921	37	Kagawa	1.934	2.036
15	Niigata	−3.290	−3.156	38	Ehime	− .690	− .732
16	Toyama	.113	−1.068	39	Kochi	−5.635	−4.777
17	Ishikawa	.507	.999	40	Fukuoka	7.033	7.813
18	Fukui	−1.062	−2.014	41	Saga	.072	− .272
19	Yamanashi	−2.873	−3.505	42	Nagasaki	.453	.556
20	Nagano	−4.593	−4.809	43	Kumamoto	−1.212	−3.330
21	Gifu	−1.691	−1.758	44	Ohita	−3.285	−3.151
22	Shizuoka	2.653	2.912	45	Miyazaki	−5.607	−5.920
23	Aichi	8.906	8.699	46	Kagoshima	−6.426	−6.494

[8] The principal components analysis is a technique, which has been widely used in psychometric research. The basic idea of the technique is to construct new indices that are linear functions of the original variables, are independent of one another, and account for as much of the variance of the original variables as is possible with a minimum number of new or constructed variables. The original formulation of this method is found in H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psycholopy* (September and October, 1933). This metnod is also described fully in H.H. Harman, *Modern Factor Analysis* (University of Chicago Press, 1960), pp. 160–191.

The subject of factor analysis covers a broad range of models and techniques and has immense literature concerning it. The following four articles listed are all econometric studies which use the principal components analysis: Meyer, J.R., and Kraft, G., "The Evaluation of Statistical Costing Techniques as Applied to the Transportation Industry." *American Economic Review*, LI (1961); Rutherford, R.S.G., "The Principal Factors Approach to Index Number Theory," *Economic Recood* (November 1954); Stone, R., "On the Interdependence of Blocks of Transactions," *Journal of Royal Statistical Society*, Supplement, IX (1947); and Tintner, G., "Some Applications of Multivariant Analysis to Economic Data," *Journal of the American Statistical Association*, XI (1946).

where: Z_j is the composite (Z) score for the j^{th} prefecture, or region;
I_{ij} is the value of index i for the j^{th} prefecture;
I_i is the mean value of index i;
σ_{I_i} is the standard deviation of index i;
and F_i is the factor loading for the i^{th} index on the regional factor.
The values of these scores for 1955 and 1960 are shown in Table 4.

Note that I_1 to I_7 are mutually correlated significantly as shown in Tables 5 and 6, so all of these variables can be treated together to yield a single characteristic of "degree of economic development." Also, the proportion of the total variance of the seven I's explained by the one factor Z is high enough, viz., .8894 for 1955 and .8975 for 1960, to be taken as the general character of these seven variables. In addition, with little difficulty we can accept Z as the composite index of the "degree of economic development" when we note that both Z_{1955} and Z_{1960} are positively correlated with each of the seven I_i's. The coefficients of correlation (factor loadings) between Z_{1955} and each I are: .926, .958, .939, .965, .980, .910, and .950; and between Z_{1960} and each I are: .968, .952, .922, .971, .967, .894, and .923.

TABLE 5. CORRELATION MATRIX, 1955

I	I_1	I_2	I_3	I_4	I_5	I_6	I_7
I_1	1.000	.852	.843	.866	.911	.789	.879
I_2		1.000	.979	.896	.924	.847	.854
I_3			1.000	.885	.903	.797	.815
I_4				1.000	.962	.866	.918
I_5					1.000	.848	.943
I_6						1.000	.892
I_7							1.000

TABLE 6. CORRELATION MATRIX, 1960

I	I_1	I_2	I_3	I_4	I_5	I_6	I_7
I_1	1.000	.921	.899	.941	.935	.815	.868
I_2		1.000	.984	.886	.883	.812	.797
I_3			1.000	.858	.839	.759	.742
I_4				1.000	.966	.842	.906
I_5					1.000	.829	.922
I_6						1.000	.856
I_7							1.000

The Z scores seen in Table 4 should strike those who are familiar with the Japanese economy as extremely "reasonable" sets of figures.[9] Tokyo and Osaka—the two largest industrial centers of Japan enjoy the two highest Z scores, followed closely by Kanagawa (which includes Yokohama City). These super-industrial prefectures are then followed by Aichi

[9] The values of Z_{1955} and Z_{1960} of Table 4 are comparable by prefecture only in the specific year indicated. If the raw data of 1960 are fed into the equation system with coefficients as obtained in 1955, we can compute another index, say Z'_{1960} with base 1955. This will permit us to analyze the differential pattern of growth by prefectures.

(Nagoya), Kyoto, Hyogo (Kobe), and Fukuoka—all of which are either highly "developed" and or contiguous to the super-industrialized three prefectures cited above. Also, a list of prefectures having the largest negative Z scores includes such prefectures as Kagoshima, Miyagi, Aomori, Iwate, Yamagata, Fukushima, Shimane and others which the Japanese customarily refer to as "*Kōshin Chiiki*" or "under-developed regions."

III. *Correlation between Z Scores and Selected Data*

Given the nature of Z scores, one expects some economic data to be highly correlated with them. Table 7 below shows the results of regressing four such data on Z scores.

TABLE 7. COEFFICIENT OF DETERMINATION BETWEEN Z SCORES AND SELECTED DATA

	R^2 with Z_{1955}	R^2 with Z_{1960}
1. Manufacturing Investments[(a)]	.7626	.6945
2. Manufacturing Value Added Per Person[(b)]	.6918	.6775
3. Migration[(c)]	.6043	.7191
4. Per Capita Total Tax	.8356	.8366

Sources and Notes.

(*a*) Computed from *Industrial Statistics* of 1955 and 1960 volumes, compiled by the Analytical Statistics Section, Ministry of International Trade & Industry.

(*b*) Value added=Value of Output-Input Costs-Taxes-Depreciation taken from the source cited in (*a*).

(*c*) This could be negative (emigration) or positive (immigration). The data was taken from the source cited for Table 1.

The results—all of which are signitificant at .001 level—confirm the usual thesis that: (1) economic development is highly correlated with investments in manufacturing industries, or more accurately, those prefectures enjoying a higher "degree of economic development" experience a larger amount of manufacturing investment; (2) more industrialized prefectures enjoy a higher productivity per person both because of a higher capital-labor ratio and more efficient labor; (3) population flows from low Z to high Z prefectures, and (4) the higher the Z score, the higher the per capita tax paid.[10]

Besides these obvious findings, a few interesting facts are brought out by relationships between Z scores and some other data. When the ratio of national tax to total tax was regressed on Z scores, the results were .3788 for 1955 and .5703 for 1960. (Both are significant at .001 level.) This appreciable increase in R^2 undoubtedly reflects the conscious taxation policy of the government to aid less developed prefectures.[11]

[10] Points (3) and (4) are, in a sense, spurious correlations in that both migration and tax paid are highly correlated with income level. For example, when the 1955 level of income was regressed in migration in 1955, .7308 was obtained. But this need not be so as it is quite conceivable that at some point of time or in a different country, Z-Tax and Z-Migration correlations could be statistically insignificant while Z-Income is significant, or *vice versa*.

[11] An excellent and extensive statistical survey and policy description on this point is found in K. Emi, "Regional Economy and Fiscal Structure," in the volume edited by Shinohara, *op. cit.*, p. 89–121.

The relationship between Z scores and the amount of social investments made by all levels of government—hereafter referred to as *SOC*—was found to be significant. The coefficient of determination between Z_{1955} and SOC_{1956} was .2628, and between Z_{1960} and SOC_{1961} was .3373, and both are significant at .001 level. This relationship between "degree of industrialization" and *SOC* has been discussed often. One's policy position could range from insisting on seeing a negative R to a high positive R, i.e., one could recommend a new super highway in Shimane prefecture ($Z=-5.406$) or an additional port facility in Osaka ($Z=18.513$).

Between these two positions an economy needs to find an optional way to distribute *SOC* so that the Z scores of all or some of the prefectures could be changed relative to one another. In deciding the course of policy, such factors as resource allocation, economies of agglomeration, and specific policy goals must be evaluated. With further examinations covering a longer period and involving a finer breakdown of the nature of *SOC*, one may be able to use the possible relationship between Z and *SOC* in policy making.[12]

In the Japanese case, regional economists seem to feel the allocation of *SOC* has not been successful. For example, Professor Ito, discussing the government policy of *SOC* between 1955 and 1959, stated that:

> It (the policy) was focused on increasing output—such as electricity—rather than economic benefits of development. As is clear from the rapid expansion of the (government) designated areas of development..., the policy resulted in a diffused "develop all regions" plan rather than the selective development plan we wanted to see adopted.[13]

Or, in the words of a government report, *SOC* "merely followed on the heels of economic development"[14] in its efforts to ease the problem of bottlenecks rather than following a preconceived and thought-out plan. As in the areas of tax policy, regional economists can make significant contributions in making firm recommendations on the use and allocation of *SOC*, based on detailed and well-conceived empirical research. Realizing the voluminous amount of literature on *SOC* and its relationship to economic development, the writer hastens to add that the above observations are a few passing remarks only to reemphasize the importance of empirical research on *SOC*.

Relationships between the level of economic development and the "equality" of income distribution have also been a subject of frequent discussion. At the international level many authors have supported the view that less developed nations tend to have more unequal distributions of income.[15] What is the relationship between Z scores and the pattern of income distribution?

To examine this question, we have correlated Z with L-coefficients suggested by Morgan,[16] and have also compared variance (to be defined below) and Z scores. Introducing a makeshift

---

[12] Professor Emi in the reference cited above (p. 127) uses *SOC* data which were separated into 8 subcategories. As these data were classified by 9 regions rather than each prefecture, I was unable to use them in this paper.

[13] Ito, *op. cit.*, p. 63–64.

[14] The Japanese government's *White Paper on Construction*, 1960, p. 4, quoted in Ito, *op. cit.*, p. 39.

[15] For example, see Irving B. Kravis, "International Differences in the Distribution of Income," *Review of Economics and Statistics*, Vol. 42 (November 1960); Theodore Morgan, "Distribution of Income in Ceylon, Puerto Rico, The United States and the United Kingdom," *Economic Journal*, Vol. 63 (1953); and Harry T. Oshima, "The International Comparison of Size Distribution of Family Incomes with Special Reference to Asia," *Review of Economics and Statistics*, Vol. 44 (November 1962).

device because of data constraints,[17] L-coefficients were calculated as follows:

$$L=1-\sum_{j=1}^{11}[X_{ij}-S_{i(j-1)}+S_{ij}]$$

where S_{ij}'s are shares of income of j^{th} income class in i^{th} prefecture. S_{ij}'s were calculated by taking the ratio of (X_{ij} times midpoint of the j^{th} income class) over $\sum_{j=1}^{11}$ (X_{ij} times midpoint of j^{th} income class) for all i. As the class interval of each income class was 10,000 yen, a midpoint of 5,000 yen for the lowest income class of "below 10,000 yen" and a midpoint of 105,000 yen for the highest income class of "over 100,000 yen," were used. (See Table 8.)

TABLE 8

Prefecture	Variance (V)	L-Index	Prefecture	Variance (V)	L-Index
1	.0229	.309	24	.0005	.364
2	.0026	.373	25	.0007	.352
3	.0109	.403	26	.0098	.364
4	.0004	.376	27	.0276	.338
5	.0014	.344	28	.0140	.363
6	.0033	.353	29	.0032	.344
7	.0015	.369	30	.0013	.352
8	.0016	.372	31	.0028	.369
9	.0008	.351	32	.0224	.392
10	.0030	.361	33	.0038	.381
11	.0054	.343	34	.0004	.368
12	.0043	.343	35	.0011	.362
13	.0443	.326	36	.0115	.381
14	.0372	.321	37	.0018	.354
15	.0024	.350	38	.0087	.394
16	.0098	.327	39	.0137	.404
17	.0048	.344	40	.0123	.339
18	.0017	.349	41	.0008	.347
19	.0061	.352	42	.0007	.371
20	.0067	.356	43	.0031	.390
21	.0023	.355	44	.0087	.369
22	.0013	.357	45	.0236	.399
23	.0024	.389	46	.0655	.460

We must be extremely careful in interpreting the results. The relationship between L and Z_{1960} (income data were not available for 1955) was .1859 (significant at the one per cent

[16] James Morgan, "The Anatomy of Income Distribution," *Review of Economics and Statistics*, Vol. XLIV (August 1962), p. 281. L is the ratio of area between the Lorenz curve and diagonal over the area under diagonal. Morgan suggested tnat "For eight or more groups his approximation should be quite close." Our data covers 11 income groups as indicated in the text.

[17] In the absence of data for respective share of income for each income class, the share of each income class was obtained by multiplying the mid-point of each income by the number of households in that income class. As one cannot assume that distribution of income in each income class is normally distributed around the mid-point, this method introduces undefined degree of bias in our L-coefficients.

level). Note that the figure is positive, contradicting, though weakly, the usual thesis that, at least at the international level, economic development brings about a more equal distribution of income. The nature of this coefficient of determination is perhaps better appreciated in the following analysis *vis-à-vis* "variance."

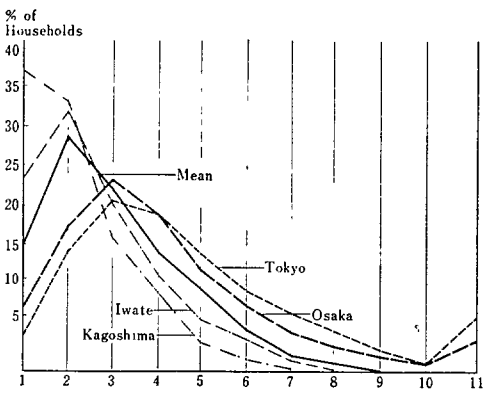
"Variance," referred to above, was computed to check against Z scores in search of a possible relationship. Variance (V) was calculated as follows: First mean of the j^{th} income class M_j ($j=1, \ldots, 11$) was obtained by $\sum_{j=1}^{46} X_{ij}/46$ where X_{ij} is the number of households in the i^{th} prefecture and in the j^{th} income class as a percentage of the total households in the i^{th} prefecture. After M_j were obtained, 46 variances were calculated as follows:

$$\sum_{j=1}^{11} (X_{ij}-M_j)^2 = V_i \qquad (i=1, \ldots, 46)$$

The result yielded V_i's ranging from .0004 (V_3) to .0655 (V_{46}). Since these are variances, one must interpret large V_i's to mean merely an indication of variance from national means. Stating it differently, V_i's are only useful in indicating, when large, that the pattern of income distribution in large V prefectures somehow differ from that of an "average prefecture."

Taking, for example, the 12 largest V_i's, we find that they are prefectures 1, 3, 13, 14, 27, 28, 32, 36, 39, 40, 45, and 46. When each of these prefectures is examined, it immediately becomes clear that V's tend to be larger when absolute values of Z scores are larger (either positive or negative). The reason for this is seen in Diagram 1 below. The solid line indicates the relative income distribution of the national mean of M_j's, and the other four lines indicate two high Z score prefectures, Tokyo and Osaka, and two low Z score prefectures, Iwate and Kagoshima.

DIAGRAM 1. PATTERNS OF INCOME DISTRIBUTION OF TWO HIGH-Z PREFECTURES (TOKYO AND OSAKA) AND TWO LOW-Z PREFECTURES (IWATE AND KAGOSHIMA) *vis-à-vis* NATIONAL "MEAN" PATTERN



These prefectures which have high V's and large negative Z's are the prefectures whose modal class is located at the same class as the national pattern (10,000 to 20,000 yen), but its frequency is higher, while the frequency of higher income classes is significantly below the national mean. These two deviations thus account for the larger V's observed. On the

other hand, in those prefectures which enjoy high positive Z scores and high V's are those which have the modal class in the third income class (20,000 to 30,000 yen) and distributions lose much of their skewness. The latter's V's are large for exactly opposite reasons as those of the negative Z score prefectures. When a rank correlation coefficient was calculated between positive Z scores and V's (17 prefectures) a high .758 (significant at the one per cent level), and between absolute value of negative Z scores and their V's (29 prefectures), .400 (significant at the five per cent, but not at the one per cent, level) were obtained.

The coefficient of determination between L-coefficients and Z scores, when considered with our observations related to V's, makes it clear that Z's of less developed regions must be raised to decrease L-coefficients.[18] However, this must be accomplished without increasing (or hopefully decreasing) observed significant positive relationships between L and Z, and with increasingly smaller values of V's for all prefectures. This is merely to state that low income prefectures' incomes ought to be raised and the level and pattern of income distribution of less developed prefectures should approach those of highly developed prefectures. When relationships among Z, L and V are closely scrutinized over a longer period of time, one could hope that the nature of the relationship between the pattern and the level of income distribution and the degrees of economic development would be made clearer for the purposes of policy making and theoretical analyses.

IV. *Suggestions for Further Research and Policy Discussions*

The contents of this paper, the writer is first to admit, have merely served as a suggestion for further investigation. When additional data—especially for a longer time horizon—can be obtained and further discussions succeed in evolving a more objective (acceptable or agreed upon) Z score,[19] it would aid considerably in gaining "greater knowledge of regional growth processes and related behavioral patterns than is now available."[20]

[18] The positive relationship between L and Z_{1960} perhaps shows the tendency of a concentration of economic power in the hands of a limited number of persons. To interpret the results observed in the text, one must approach the evaluation of the results with full cognizance that inter-regional income distribution is subject to different sets of conditions from those faced by international income distribution.

[19] One direction this type of empirical research ought to enter much more thoroughly is the area of human investment and the causes and effects of reginal differentials Some areas of investigation suggested by Professor Schultz clearly merits our further attention (as one of I indices):

> I shall concentrate on five major categories: (1) health facilities and services, broadly conceived to include all expenditures that affect the life expectancy, strength and stamina, and the vigor and vitality of a people; (2) on-the-job trainining, including old-style apprenticeship organized by firms; (3) formally organized education at the elementary, secondary, and higher levels; (4) study programs for adults that are not organized by firms, including extension programs for adults that are not organized by firms, including extension programs notably in agriculture; (5) migration of individuals and families to adjust to changing job opportunities.

T. W. Schultz, "Investment in Human Capital," *American Economic Review*, Vol. LI, (March 1961), p. 9. Though no systematic analysis was presented, Professor Sakamoto recently emphasized the importance of examining the question of "how the long-run (inter-regional) changes come about due to a balanced distribution of human capabilities in regions within a nation." J. Sakamoto, "Major Elements of Inter-regional Income Differentials," *Keizai Kenkyu*, Vol. 15 (July 1964), p. 221.

[20] John Meyer, "Regional Economics, A Survey," *American Economic Review*, Vol. LIII (March 1963), p. 47.

One such course might be, for example, to develop a model in which selected variables and policy parameters are lagged *vis-à-vis* Z scores in order to evaluate, isolate and possibly predict changes in Z scores. Not only some variables and parameters considered in this paper, but consumption patterns, types of investments, nature and source of income, etc., can be examined—in such forms as multivariate regression analysis or any other technique—against Z scores.

When these efforts are made, we may be able to discuss the policy questions at a more meaningful level than is done today. After all significant combinations of variables and parameters (including I's as raw data of Z's) are carefully examined, the noted weakness of the link existing between theoretical economics (unsupported by empirical studies) and regional economic development plans (often politically motivated) could be "rationalized."[21]

[21] One of the most favored words of the Japanese government officials and business community.