Journal of Economic Studies

Journal of Economic Studies Volume 16 Number 1 1989

MCB
University Press

An Information Theory Approach to Measuring Industrial Diversification

by
Mohsen Attaran and Martin Zwick
California State University, Bakersfield, USA
and *Portland State University, Oregan, USA*

Introduction

Entropy as a measure of disorder, uncertainty, or homogeneity, has been used to analyse many different phenomena. In the physical sciences, it has been used to measure the irreversible increase of "unavailable energy". In the biological and behavioural sciences, entropy has been used as a measure of organisation (Nicolis and Prigogine, 1977; Rifkin, 1980). In communication theory, it quantifies the degree of uncertainty in a system (Shannon and Weaver, 1949). Taking the Shannon entropy as a measure of diversity yields:

$$D(P_1, P_2, \ldots P_n) = -\sum_{i=1}^{n} P_i \log_2 P_i$$

In the context of communication theory, where this measure gives the uncertainty H, P_i is the probability of some event i. As a measure of diversity, the P_i represents the proportion of some total quantity. For example, let K denote some quantity such as total employment, total output, foreign trade, or income, and K_i the amount contributed to this total by the ith entity, such as an industry in a region, a county, or a product traded. The value shares are $P_i = K_i/K$ and the entropy measure K gives the diversity or spread of the distribution (Horowitz and Horowitz, 1976). The maximum value of D is attained ($\log_2 n$) when all P_i are equal. If the ith entity is the only contributor to K, then $P_i = 1$, all other $P_i = 0$ and $D = 0$.

The entropy measure has been invoked in empirical studies in economics as well as in management, marketing, finance and accounting. In a marketing context, entropy can represent the distribution of consumer preferences for various brands. For example, Herniter (1973) uses entropy as a measure of uncertainty or disorder in the stochastic system that represents the consumer's preferences for special brands.

In the analysis of empirical data, entropy has also been used as a measure of dispersion, an alternative to the variance σ^2 (a measure of risk or uncertainty). For example, the use of entropy rather than variance as a measure of the risk of a securities portfolio whose components yield stochastic returns has been advocated by Philippatos and Wilson (1972; 1974). They concluded that since entropy can be estimated directly from variances (when the form of prior distribution is known) and can be computed from non-metric data, entropy is more general and better suited for the selection of portfolio than variance.

In the analysis of accounting data, entropy has been used to measure the loss of information from aggregation of items on financial statements, e.g. the balance sheet (Lev, 1968, 1970; Theil, 1969).

An extensive treatment of entropy-based measures in the analysis of economic data occurs in several studies (e.g. Bernhardt and Mackenzie, 1968; Georgescu-Roegen, 1971; Theil, 1967). For example, Theil (1967) discusses in detail the basic technical informational concepts and illustrates them with economic examples. Theil's books (1967, 1972) are primarily concerned with distributional issues and with decomposition analysis. In particular, he has argued that information concepts provide an appropriate measure which can be utilised in empirical studies in economics to answer such questions as: How is income distributed among the families of a nation or among the states of a nation? How are sales, total outputs or employment distributed among industries within a region and among regions? How is international trade distributed among countries?

In market structure analysis, entropy has been employed as a measure of the "competitiveness" of an industry. Here P_i represents the market shares of firms in the industry. As such, entropy varies inversely with the degree of industrial concentration. Using this measure, Horowitz and Horowitz (1968) analysed the concentration in the brewing industry between 1944 and 1964.

Using the decomposition property of entropy, some market structure researchers were able to analyse concentration, either within or between regions, or within brands of an individual company, and between companies (Bernhardt and Mackenzie, 1968; Horowitz and Horowitz, 1970; Theil, 1967). For example, Horowitz and Horowitz (1970) studied the industrial concentration in 26 two-digit manufacturing industries in the common market nations.

Along similar lines, Garrison and Paulson (1973) used entropy and a related measure to test the hypothesis that "labour-intensive" industries are less concentrated geographically than other types of industries. The results of the study supported the hypothesis.

Entropy measures of geographical concentration have also been used by Garrison (1974) to examine the extent to which rural and small-town counties compete with urban areas for manufacturing employment in the Tennessee Valley region. Here P_i represents the relative ability of the ith county to attract manufacturing industries. Decomposition of entropy into its between-set and within-set components has also enabled Garrison to compare the low-wage and high-wage industries of the region with respect to the nature of their geographical dispersal over time.

Entropy has also been used to measure employment diversity by Hackbart and Anderson (1975). Within this context, the P_i represents the ith sector share of regional employment. Hackbart and Anderson illustrated the applicability of the entropy method by examining four river basin regions in Wyoming. They concluded that the entropy method "provides a direct means of comparing diversity in different regions or changes in diversity over time" (p. 378). However, they did not examine how their measure of diversity can be decomposed so that various patterns of inter-industry diversification within a region over time may be examined.

Entropy as a Measure of Industrial Diversification

The entropy method measures diversity of a region against a uniform distribution of employment where the norm is equiproportional employment in all industrial sectors. As it is applied to the United States estimate of employment data, the entropy measure of industrial diversity $D(E_1, E_2, \ldots E_n)$ is defined as follows:

$$D(E_1, E_2, \ldots E_n) = - \sum_{i=1}^{n} E_i \log_2 E_i \quad (1)$$

where n = the number of economic sectors,

E_i = the proportion of total employment of the region that is located in the ith sector.

The most important properties of the above measure are:

- the maximum value of D is attained when the E_i are all equal. This is the case where the region is totally diversified in the sense that all sectors contribute equally to the region's employment. Also, the greater the number of sectors sharing in the region's economic activity, the greater the value of D;
- $0 \leq D \leq \log_2 n$;
- $D = 0$ when only one of the $E_i = 1$ and the remainder are 0. This is an extreme case where the economic activity of a region is concentrated in only one sector; therefore, economic diversity is totally absent.

Attaran and Zwick (1987) argue that entropy is in all respects at least as good as any other measure of diversification and thus, due to its decompositional properties, it is superior to the others. The entropy measure is a more flexible and analytically powerful measure of economic diversity than the national average measure, portfolio measure and ogive approach.

The national average measure assesses the deviation of the regional distribution of economic activity from the national distribution. Since the norm (national distribution) changes over time, this measure does not determine whether the distribution of economic activity within a region itself has become more or less diversified over time. Because the uniform distribution is a comparative norm that is fixed, the entropy measure will accomplish the above objectives.

The portfolio measure is only indirectly (presumably inversely) related to diversity, in the context that diversity is measured by the ogive, national average, McLaughlin, or entropy. Also, the use of portfolio variance as a norm is clearly not a good measure of diversification.

The ogive and entropy measures are conceptually similar; both approaches compare actual distribution of employment to a hypothetical uniform distribution representing balanced industrial composition (equal percentage in each group). The entropy measure is more flexible than the ogive, however, as the entropy measure can be decomposed so that the various patterns of inter-industry diversification within a region over time may be examined. These patterns and changes may not be apparent from an examination of the single unit total diversity index of diversification.

The entropy measure as it is formulated in equation (1) can be disaggregated into its between-set and within-set aspects to express the extent and pattern of dispersal between and within different groups or subsets of industries. Consider, for example, that industries (sectors) are combined into G sets. The employment share of set S_g is then:

$$E_g = \sum_{i \epsilon S_g} E_i \qquad g = 1, \ldots G$$

The entropy index of diversity *within* each of the G sets can be measured by:

$$D \text{ within } s_g = -\sum_{i \epsilon S_g} (E_i/E_g) \log_2 (E_i/E_g) \tag{2}$$

The within-set measure merely represents the application of the entropy measure to different groups of industries treated independently. Representing each set's relative share of the total employment by E_g/E_t where E_t is total employment of the United States, the entropy measure of diversification between the G sets may then be expressed as:

$$D \text{ between} = -\sum_{g=1}^{G} (E_g/E_t) \log_2 (E_g/E_t) \tag{3}$$

The between-set measure identifies the extent to which employment is distributed equally between the G sets.

The relative importance of each of the G sets determines the contribution of its diversification to the degree of industry diversification within the total economy. Weighting the result of equation (2) by the relative share of each set yields:

$$D \text{ within} = \sum_{g=1}^{G} (E_g/E_t) \left[-\sum_{i \epsilon S_g} (E_i/E_g) \log_2 (E_i/E_g) \right] \tag{4}$$

which is the average diversity within the G sets. The entropy measure of economic diversity for the entire economy can be obtained by summing equations (3) and (4):

$$D = -\sum_{g=1}^{G} (E_g/E_t) \log_2 (E_g/E_t) + \sum_{g=1}^{G} (E_g/E_t) \left[-\sum_{i \epsilon s_g} (E_i/E_g) \log_2 (E_i/E_g)\right] \quad (5)$$

This disaggregation of entropy into its between-set and within-set aspects, where $G = 2$ (manufacturing and non-manufacturing) is carried out for the United States and its results are presented in the following section.

Entropy Measure: An Illustration

Variation in Diversity within the USA

The United States is far from uniform in terms of the diversity of economic activity within its borders. It varies widely in climate, soils and vegetation, all of which affect the economic activities from area to area and thus the economic diversity of each region. To examine the changes in diversity, the diversification indices based on employment data were calculated for the 50 states and the District of Columbia for the ten-year period from 1972 to 1981. Calculation of the entropy measure is based on employment data from eight non-agricultural sectors. These sectors are: durable goods; non-durable goods; construction; transportation, communication and utilities; trade; finance, insurance and real estate; service and miscellaneous; and government.

The value E_i which measures the ith sector's relative share of employment for a given state, is calculated from the BLS series, *Employment and Earnings, States and Areas*, for a ten-year period from 1972 to 1981.

Since there are eight sectors, the maximum value of

$$D(E_1, E_2, \ldots E_8) = \log_2 8 = 3$$

and observed values may be directly interpreted in this scale. The diversification values would then range from 0 to 3, with a diversification value of 3 denoting the greatest diversification among the eight sectors of a state.

Table I gives the calculated diversity indices of each state for the ten-year study period. These indices were averaged, and grouped into four classes based on the level of diversity of the 51 areas.

Of the 51 study areas, roughly half showed high to moderate diversification, and none were distinguished as either highly diversified or highly specialised; however, there are patterns which may be identified. The west south central region can be considered a region of high diversity; the middle Atlantic clearly is a region of moderate diversity. Other highly diversified areas are scattered in the south and the northeast.

The most important areas of specialisation are: (1) the steel production area of Indiana and the auto-producing portion of Michigan in the eastern north central region; (2) North and South Dakota in the western north central region; (3) Montana and Nevada in the Mountain region, and (4) Alaska and Hawaii in the Pacific region. While the Mountain region is dominated by moderately diversified states, the Pacific and New England areas appear to be dominated by states having low diversity.

Thus, the south alone is mainly an area of high to moderate diversity while the west, central and eastern areas of the United States have moderate to low diversity.

State	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981
Group 1: Region of High Diversity, 1972-1981										
Delaware	2.93	2.94	2.93	2.91	2.91	2.92	2.93	2.93	2.93	2.93
Kentucky	2.92	2.91	2.92	2.92	2.92	2.93	2.93	2.94	2.92	2.91
Louisiana	2.95	2.95	2.95	2.94	2.94	2.94	2.94	2.95	2.95	2.95
Oklahoma	2.90	2.90	2.90	2.89	2.89	2.89	2.90	2.91	2.92	2.92
Texas	2.92	2.92	2.92	2.91	2.91	2.92	2.93	2.93	2.94	2.94
West Virginia	3.01	2.90	2.99	2.99	2.99	3.00	2.99	2.98	2.97	2.94
Group 2: Region of Moderate Diversity, 1972-1981										
Alabama	2.85	2.86	2.87	2.86	2.86	2.86	2.86	2.86	2.86	2.85
Arizona	2.87	2.88	2.86	2.81	2.79	2.80	2.80	2.82	2.81	2.81
Arkansas	2.87	2.86	2.86	2.85	2.85	2.85	2.84	2.85	2.84	2.84
Colorado	2.84	2.85	2.84	2.81	2.81	2.82	2.84	2.86	2.85	2.86
Georgia	2.84	2.85	2.85	2.82	2.81	2.81	2.81	2.81	2.81	2.81
Idaho	2.82	2.83	2.83	2.82	2.81	2.82	2.83	2.84	2.83	2.83
Illinois	2.84	2.83	2.83	2.83	2.82	2.81	2.81	2.82	2.82	2.81
Kansas	2.82	2.83	2.84	2.83	2.83	2.84	2.84	2.85	2.85	2.85
Maine	2.80	2.80	2.82	2.81	2.82	2.82	2.81	2.82	2.83	2.83
Minnesota	2.84	2.84	2.83	2.81	2.81	2.80	2.81	2.81	2.80	2.79
Mississippi	2.83	2.84	2.85	2.84	2.83	2.83	2.83	2.84	2.84	2.85
Missouri	2.85	2.85	2.84	2.83	2.82	2.82	2.82	2.83	2.82	2.81
New Jersey	2.83	2.83	2.83	2.81	2.80	2.79	2.80	2.80	2.80	2.79
New York	2.84	2.84	2.83	2.80	2.80	2.79	2.79	2.79	2.78	2.77
Pennsylvania	2.86	2.86	2.85	2.85	2.85	2.85	2.85	2.85	2.84	2,83
Tennessee	2.83	2.83	2.84	2.83	2.82	2.82	2.83	2.83	2.82	2.82
Utah	2.81	2.82	2.83	2.82	2.83	2.84	2.84	2.85	2.84	2.84
Virginia	2.87	2.87	2.87	2.84	2.84	2.84	2.83	2.84	2.82	2.80
Wyoming	2.81	2.83	2.84	2.83	2.83	2.83	2.84	2.85	2.83	2.82
Group 3: Region of Low Diversity, 1972-1981										
California	2.80	2.80	2.79	2.77	2.77	2.77	2.78	2.79	2.78	2.78
Florida	2.81	2.82	2.80	2.75	2.74	2.75	2.75	2.76	2.76	2.76
Iowa	2.79	2.79	2.79	2.78	2.77	2.78	2.78	2.78	2.76	2.75
Maryland	2.80	2.80	2.80	2.75	2.75	2.75	2.74	2.74	2.72	2.71
Massachusetts	2.84	2.83	2.82	2.80	2.78	2.77	2.77	2.77	2.76	2.76
Nebraska	2.77	2.78	2.78	2.76	2.77	2.77	2.78	2.78	2.77	2.76
New Hampshire	2.81	2.81	2.80	2.78	2.78	2.78	2.78	2.78	2.77	2.77
New Mexico	2.76	2.76	2.77	2.75	2.75	2.77	2.77	2.78	2.77	2.77
North Carolina	2.78	2.79	2.79	2.79	2.77	2.78	2.79	2.79	2.79	2.80
Ohio	2.79	2.78	2.78	2.78	2.78	2.78	2.78	2.79	2.79	2.78
Oregon	2.79	2.79	2.78	2.76	2.75	2.76	2.77	2.77	2.76	2.74
Rhode Island	2.80	2.74	2.80	2.79	2.78	2.78	2.78	2.79	2.78	2.78
South Carolina	2.73	2.75	2.77	2.75	2.74	2.75	2.76	2.78	2.78	2.78
Vermont	2.82	2.82	2.80	2.77	2.76	2.75	2.76	2.77	2.75	2.75
Washington	2.77	2.77	2.76	2.75	2.76	2.76	2.77	2.78	2.77	2.76
Wisconsin	2.77	2.77	2.76	2.76	2.76	2.77	2.76	2.76	2.76	2.75

Table I.
(cont.)

Table I. Diversity Indices

Group 4: Diversification Indices of Specialised Region, 1972-1981

State	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981
Alaska	2.60	2.63	2.70	2.73	2.75	2.75	2.75	2.73	2.74	2.74
Connecticut	2.61	2.60	2.58	2.59	2.57	2.57	2.58	2.58	2.58	2.58
Hawaii	2.76	2.77	2.76	2.75	2.71	2.70	2.70	2.71	2.71	2.70
Indiana	2.73	2.72	2.73	2.75	2.74	2.74	2.74	2.76	2.76	2.75
Michigan	2.70	2.69	2.70	2.70	2.69	2.69	2.69	2.70	2.70	2.69
Montana	2.77	2.78	2.77	2.71	2.72	2.72	2.72	2.74	2.73	2.73
Nevada	2.49	2.51	2.47	2.43	2.42	2.45	2.45	2.45	2.44	2.45
North Dakota	2.58	2.59	2.61	2.62	2.63	2.63	2.65	2.68	2.68	2.70
South Dakota	2.62	2.64	2.65	2.63	2.64	2.66	2.67	2.68	2.66	2.65
District of Columbia	2.21	2.21	2.20	2.16	2.11	2.11	2.10	2.10	2.09	2.08

Changes over Time in Diversity Patterns

To examine the changes in diversity over time, the diversity indices of employment data were calculated for the US for the 28 years from 1960 to 1987. The results are shown in Table II. Calculation of the entropy measure for the US is based on employment data of 56 sectors (two-digit SIC). The value E_i is calculated from the *Statistical Abstracts of the United States*, a national data book published annually by the US Department of Commerce, Bureau of the Census.

Since there are 56 sectors, the maximum value of $D = \log_2 56 = 5.8074$ and the diversification values would then range from 0 to 5.8074. There is evidence in Table II of a trend towards greater concentration in the overall US economy. [A t-test of the slope indicates a statistically significant relationship (computed value of $t = -4.79$ which is significantly beyond the 0.001 level)].

This single-unit total entropy measure does not, however, identify inter-industry diversification or concentration patterns and structural changes occurring within the entire economy. The entropy measure can be disaggregated into its between-set and within-set elements to express the extent and patterns of dispersal between and within different groups and subsets of industries. The results of this disaggregation analysis are presented in the next section.

Extent and Pattern of Diversification between and within Industries

The second section of this article discussed the decomposition properties of entropy in analysing the nature and extent of dispersal of economic diversity between and within different subsets of industries. Consider combining the 56 economic sectors (two-digit SIC) of the US into two separate groups or sets: S_g (g = 1, 2). S_1 is defined as manufacturing (20 sectors) and S_2 as non-manufacturing (the remaining 36 sectors).

Table II. Employment Diversification Indices for the US, 1960-87

Year	Diversity Index	Year	Diversity Index
1960	5.2405	1974	5.0538
1961	5.1898	1975	4.9993
1962	5.1921	1976	5.0092
1963	5.1786	1977	5.0077
1964	5.1881	1978	5.0093
1965	5.1720	1979	5.0109
1966	5.1514	1980	4.9940
1967	4.9840	1981	4.9951
1968	5.1113	1982	4.9733
1969	5.1199	1983	4.9527
1970	5.0960	1984	4.9796
1971	5.0810	1985	4.5769
1972	5.0604	1986	5.1046
1973	5.0678	1987	5.0930

The disaggregation of entropy for the above two groups is carried out using equations (2) to (5), and the results are presented in Table III, columns (2) to (7). The aggregated employment diversification indices initially presented in Table II are shown again in column (8). Table III enables a comparison of the manufacturing and non-manufacturing sets with respect to the nature of their economic dispersal over time.

The within-set component of the entropy measures for the manufacturing and non-manufacturing groups produced from applying equation (2) is presented in columns (2) and (3), respectively. The within-set measure represents the application of the entropy measure to two industry groups treated independently. There is evidence in column (3) of a trend towards increasing concentration within the non-manufacturing set over the 28-year period (t = −3.16, significant beyond 0.01). However there is no evidence in column (2) of any trend revealed by the within-set entropy of the manufacturing set (t = −0.11, not significant).

The weighted within-set entropy measures of the two groups appearing in columns (4) and (5) reflect each group's contribution to the degree of economic diversification within the total economy. In applying equation (4), the weighted within-set measures for the two groups are summed to yield the total weighted within-set entropy measure shown in column (6). There is evidence of a trend toward increasing concentration revealed by the total weighted within-set entropy measure (t = −2.05, significant beyond 0.05). However, examination of its component revealed a trend toward decreasing contribution of manufacturing (t = −19.89, significant beyond 0.001) and increasing contribution of non-manufacturing (t = 6.61, significant beyond 0.001).

As discussed above, there is a decreasing contribution of manufacturing and an increasing contribution of non-manufacturing to the degree of economic diversification within the total economy. The consequence of increasing concentration within non-manufacturing sector, then, is a trend toward increasing concentration within the total economy revealed by the total weighted within-set entropy measure.

Table III. A Disaggregated Entropy Measure of Employment Diversity for the US, 1960-1984

(1) Year	(2) Within-set Entropy (Mfg)	(3) Within-set Entropy (Non-mfg)	(4) Weighted Within-set Entropy (Mfg)	(5) Weighted Within-set Entropy (Non-mfg)	(6) Total Weighted Within-set Entropy	(7) Between-set Entropy	(8) Total Entropy
1960	4.0506	4.4309	1.4772	2.8172	4.2944	0.9461	5.2405
1961	4.0454	4.3812	1.3622	2.9059	4.2681	0.9217	5.1898
1962	4.0491	4.3605	1.4655	2.7823	4.2478	0.9443	5.1921
1963	4.0495	4.3440	1.4424	2.7966	4.2391	0.9395	5.1786
1964	4.0503	4.3586	1.4414	2.8074	4.2489	0.9392	5.1881
1965	4.0431	4.3439	1.4105	2.8285	4.2390	0.9330	5.1720
1966	4.0338	4.3158	1.4135	2.8035	4.2169	0.9344	5.1514
1967	3.5994	4.2907	1.4702	2.5381	4.0083	0.9757	4.9840
1968	4.0300	4.2697	1.3606	2.8281	4.1888	0.9225	5.1113
1969	4.0274	4.2868	1.3494	2.8505	4.1999	0.9200	5.1199
1970	4.0391	4.2584	1.3084	2.8790	4.1874	0.9086	5.0960
1971	4.0440	4.2507	1.2585	2.9278	4.1864	0.8946	5.0810
1972	4.0494	4.2224	1.2492	2.9198	4.1690	0.8914	5.0604
1973	4.0408	4.2378	1.2437	2.9335	4.1772	0.8906	5.0678
1974	4.0357	4.2288	1.2171	2.9534	4.1705	0.8832	5.0538
1975	4.0380	4.1753	1.1542	2.9818	4.1361	0.8633	4.9993
1976	4.0418	4.1970	1.1324	3.0211	4.1535	0.8557	5.0092
1977	4.0423	4.1831	1.1610	2.9817	4.1426	0.8651	5.0077
1978	4.0345	4.1910	1.1523	2.9940	4.1463	0.8630	5.0093
1979	4.0221	4.2037	1.1351	3.0173	4.1524	0.8585	5.0109
1980	4.0150	4.2026	1.0860	3.0659	4.1519	0.8422	4.9940
1981	4.0154	4.2092	1.0737	3.0837	4.1574	0.8377	4.9951
1982	4.0170	4.2035	1.0192	3.1369	4.1562	0.8171	4.9733
1983	4.0100	4.1927	0.9866	3.1611	4.1477	0.8049	4.9527
1984	4.0106	4.2169	1.0114	3.1535	4.1649	0.8147	4.9796
1985	3.9988	3.7123	0.9613	2.8199	3.7812	0.7957	4.5769
1986	3.9989	4.3764	1.0380	3.2402	4.2784	0.8262	5.1046
1987	4.0014	4.3699	1.0134	3.2632	4.2766	0.8164	5.0930

Table IV.
Actual Entropy as Percentage of Maximum Entropy

Year	Within-set Entropy (Mfg)	Within-set Entropy (Non-mfg)	Weighted within-set Entropy (Mfg)	Weighted within-set Entropy (Non-mfg)	Total weighted within-set Entropy	Between-set Entropy	Total Entropy
1960	93.89	85.70	34.20	54.49	90.41	94.61	90.24
1987	92.58	84.53	23.45	63.12	90.03	81.64	78.81
Percentage changes 1960-87	−1.31	−1.17	−10.75	+8.63	−0.38	−12.97	−11.43

The between-set entropy measure which results from applying equation (3) is presented in column (7). The between-set measure merely identifies the extent to which US employment is distributed equally between the manufacturing and non-manufacturing sets. There is a trend towards greater between-set concentration over the 25-year period ($t = -14.76$, significant beyond 0.001).

The total weighted within-set measure is added to the between-set measure to yield the total entropy measure as formulated by equation (5). (This is the same as the aggregated entropy measure of industry diversification initially presented in Table II.)

As mentioned earlier, the maximum value of entropy equals 5.8074. For the years 1960 and 1987, actual entropy as a percentage of maximum entropy was calculated. This provides a better understanding of what has been happening to relative concentration (diversification). These results are given in Table IV.

In interpreting Table IV, it must be kept in mind that when actual entropy as a percentage of total entropy falls, relative concentration rises. Examination of Table IV reveals very little change in the within-set entropy of manufacturing and non-manufacturing over the period 1960-1987. The total weighted within-set measure shows a change of 0.38 per cent. Among its components, manufacturing shows the highest change (decrease of 10.8 per cent to the degree of diversification within the total economy) followed by non-manufacturing (increase of 8.6 per cent). With regard to the between-set diversification, there has been a 13 per cent decline. This indicates that during the period of the study, there has been an increase in size disparity between these two major groups of the economy. In the case of the 56 sectors as a whole, there has been an 11.43 per cent reduction in relative diversification.

It should be noted that the US economy is relatively diverse in terms of employment, achieving approximately 85 per cent of total possible entropy during the study period. Thus, in terms of employment, the US economy seems to be fairly equally distributed in size. Yet this study also showed a shift towards an increased disparity in size for the total US economy.

In 1960, the percentage employment shares of manufacturing and non-manufacturing were about 36 and 64 per cent, respectively. During the 1970s non-manufacturing percentage shares generally increased to 71 per cent, and in 1987 they stood at 75 per cent. Because the non-manufacturing group's percentage share initially was considerably higher than 50 per cent (comparative norm for the greatest diversity) the increase in the group's percentage share over the 28-year period accounts for the decrease in between-set entropy, i.e. the greater concentration of the US economy in one of the two aggregated sets.

As discussed earlier, there was a trend toward increasing industry concentration within the non-manufacturing group. The consequence of increasing between-set concentration toward non-manufacturing and away from manufacturing, then, could be greater concentration and less diversification in the total economy. This finding may lend perspective to interpreting the structural changes (concentration) occurring within the entire US economy.

Conclusions

The proposed entropy measure provides a flexible and analytically powerful measure of industrial diversity. The rectangular distribution (uniform distribution) of economic activities used as a comparative norm with the entropy measure is objective and conceptually consistent with the intuitive notion of diversification as the absence of concentration. Furthermore, the entropy measure can be decomposed to allow for identification of some important inter-industry diversification patterns which may not be at all apparent merely from examining the single-unit total entropy measure of diversity.

This technique was useful, not only in providing an overall index of diversity over time for the US but also, through its decomposition properties, in analysing the nature of such a dispersal. The decomposition properties have permitted the analysis of economic concentration and structural changes, both within and between groups of sectors, which appeared to offer some useful extension of regional analysis.

References

Attaran, M. and Zwick, M. (1987), "Entropy and Other Measures of Industrial Diversification", *Quarterly Journal of Business and Economics*, Vol. 26 No. 4, pp. 17-34.

Bernhardt, I. and MacKenzie, K.D. (1968), "Measuring Seller Unconcentration, Segmentation and Product Differentiation", *Western Economic Journal*, Vol. 6, pp. 395-403.

Garrison, C. (1974), "Industrial Growth in the Tennessee Valley Region, 1959 to 1966", *American Agricultural Economics*, Vol. 56, February, pp. 50-60.

Garrison, C. and Paulson, A. (1973), "An Entropy Measure of the Georgraphic Concentration of Economic Activity", *Economic Geography*, Vol. 49, pp. 319-24.

Georgescu-Roegen, N. (1971), *The Entropy Law and the Economic Process*, Harvard University Press, Cambridge, MA.

Hackbart, M.M. and Anderson, D.A. (1975), "On Measuring Economic Diversification", *Land Economics*, Vol. 51, pp. 374-8.

Herniter, J.D. (1973), "An Entropy Model of Brand Purchase Behavior", *Journal of Marketing Research*, Vol. X, November, pp. 361-75.

Horowitz, A. and Horowitz, I. (1976), "The Real and Illusory Virtue of Entropy-based Measures for Business and Economic Analysis", *Decision Science*, Vol. 7, pp. 121-35

Horowitz, A. and Horowitz, I. (1968), "Entropy, Markov Processes and Competition in the Brewing Industry", *Journal of Industrial Economics*, Vol. 16, pp. 196-211.

Horowitz, I. and Horowitz, A.R. (1970), "Structural Changes in the Brewing Industry", *Applied Economics*, Vol. 2, May, pp. 1-13.

Horowitz, I. (1970), "Employment Concentration in the Common Market: An Entropy Approach", *Journal of the Royal Statistical Society*, Vol. 133 Part 3, pp. 463-79.

Lev, B. (1968), "The Aggregation Problem in Financial Statements: An Information Approach", *Journal of Accounting Research*, Vol. 6, Autumn, pp. 247-61.

Lev, B. (1970), "The Informational Approach to Aggregation in Financial Statements: Extensions", *Journal of Accounting Research*", Vol. 8, Spring, pp. 78-94.

Nicolis, G. and Prigogine, I. (1977), *Self Organization in Non-Equilibrium Systems*, Wiley, New York.

Philippatos, G.C. and Wilson, C.J. (1972), "Entropy, Market Risk, and the Selection of Efficient Portfolios", *Applied Economics*, Vol. 4, pp. 209-20.

Philippatos, G.C. and Wilson, C.J. (1974), "Entropy, Market Risk and the Selection of Efficient Portfolios: Reply", *Applied Economics*, (1980), Vol. 6, pp. 77-81.

Rifkin, J. (1980), *Entropy: A New World View*, Viking, New York.

Shannon, C.E. and Weaver, W. (1949), *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana.

Theil, H. (1967), *Economics and Information Theory*, North Holland Publishing Company, Amsterdam.

Theil, H. (1969), "On the Use of Information Theory Concepts in the Analysis of Financial Statements", *Management Science*, Vol. 15, May, pp. 459-80.

Theil, H. (1972), *Statistical Decomposition Analysis*, American Elsevier Publishing Company, New York.