ENTROPY MEASURES OF SPATIAL CONCENTRATION IN POULTRY PROCESSING*

Thomas L. Sporleder

The theory of industrial organization provides a conceptual base for investigation of the market structure of specialized industries. A number of measures of concentration exist which have been applied to market structure investigations.[1] These measures are utilized to quantify seller concentration in an effort to classify individual markets with respect to relative competitiveness.

An important component of structure can be the spatial aspects of the industry under investigation. The purpose of this paper is to demonstrate how entropy, one measure of market concentration, is adaptable for investigating changes in long run spatial structure. Relative entropy, computed from absolute entropy, is suggested as a method appropriate for such spatial investigations. Comparison of a time series of relative entropies then allows documentation of temporal concentration propensities. Each entropy measure suggested is applied to the poultry processing industry (S.I.C. 2015)[2] for the last three available census years (1958, 1963, 1967).

BASIC ENTROPY MEASURES

Theil's entropy measure from information theory has been employed as an index of industrial concentration in several instances [1, 2, 3]. The basic entropy measure may be utilized whenever data are available on firm shares (either market shares, physical or dollar output shares, or ratio of individual firm employees to total employees). Given an N-firm industry with θ_i representing the share of the i^{th} firm in that industry, the entropy $H(\theta)$ is defined as [6, p. 24]:[3]

$$(1) \quad H(\theta) = \Sigma\, \theta_i \log_2 \theta_i^{-1} .$$

The entropy defined in equation (1) is regarded as an inverse measure of concentration since, if $\theta_i = 1$ for one i, zero otherwise, $H(\theta) = 0$ [6, p. 291]. Also, if all θ_i are equal, $H(\theta) = \log_2 N$. Thus, $0 \leqslant H(\theta) \leqslant \log_2 N$ where zero is the maximum degree of share concentration and $\log_2 N$ is the minimum degree of share concentration (maximum dispersion), given N. Of course, since θ_i represents a share, it is constrained such that:

$$(2) \quad \theta_i \geqslant 0 \qquad \text{for } i = 1, ..., N,$$

and

$$(3) \quad \Sigma\, \theta_i = 1 .$$

There are other properties of the entropy, $H(\theta)$, but they will not be repeated here since they are given by Horowitz [2, p. 463].

Thomas L. Sporleder is associate professor of agricultural economics at Texas A&M University.

*Texas A&M Agricultural Experiment Station Technical article number 10944. The author gratefully acknowledges partial financial support for this research from Texas Transportation Institute.

[1] Alternative measures include the Herfindahl-Hirschmann index, the Tall-Tideman index, the Gini coefficient, entropy, relative entropy, numbers equivalents, and the CCI (comprehensive concentration index). See Rose and Fraser [5] for an applied comparison of each measure except the CCI. For the CCI, see Horvath [4].

[2] The definition, from Census of Manufacturers, of S.I.C. 2015 is: "Establishments primarily engaged in slaughtering, dressing, packing, freezing, and canning poultry, rabbits and other small game for their own account or on a contract basis for the trade. This industry also includes the drying, freezing, and breaking of eggs."

[3] Logarithms to the base 2 are common in information theory, and when base 2 is used, the information content is said to be expressed in bits, short for "binary digits." Logarithms to the base 2 provide a convenient property for the entropy measure (see [6, pp. 4-5] for an explanation).

ADAPTION OF ENTROPY TO SPATIAL CONCENTRATION

Theil's basic entropy measure may be adapted to provide a spatial concentration measure by regarding θ_i as the share of the i^{th} geographic region. Thus, θ_i may be either the i^{th} region's share of total number of firms, share of total output, or share of total employees employed in the industry being studied. In the current instance, the total geographic area is the United States, and regions within this total follow the definitions from Census of Manufactures [7].

Spatial analysis obviously requires that the unit of investigation be defined by geographic boundaries. Since geographic regions are the basic unit of analysis and regions are of differing size, there is no a priori reason to expect equal shares among regions. This means that $H(\theta)$ for a particular point in time is without meaning. However, relative entropy measures over time provide a unique and useful means of investigating spatial concentration propensities.

Relative entropy for any time t may be defined as:

$$(4)\quad R(\theta) = H(\theta) / \log_2 N.$$

Thus, relative entropy, $R(\theta)$, is the ratio of the estimated absolute entropy to the maximum entropy possible. As a result, $R(\theta)$ is an index where $0 \leqslant R(\theta) \leqslant 100$. If concentration is absolute, (i.e., $\theta_i = 1$ for one i, zero otherwise), $R(\theta) = 0$. When θ_i are equal for all i, $R(\theta) = 100$, or the case of greatest possible dispersion.

$R(\theta)$ is then a measure of the extent to which the industry under study is attaining the maximum possible geographic dispersion in firm or output shares given the number of geographic regions. Comparison of intertemporal changes in $R(\theta)$ provides information concerning spatial concentration propensities.

ENTROPY DISAGGREGATION

A convenient aspect of the total entropy, $H(\theta)$, is that it may be disaggregated into between-set and within-set entropies. As shown later, this is especially relevant when spatial concentration is considered. Following Theil [6, p. 291], suppose geographic regions are combined to form M sets of geographic regions. The share of set ψ_m is:

$$(5)\quad \psi_m = \sum_{i \epsilon \psi_M} \theta_i \qquad \text{for } m = 1, \ldots, M.$$

Disaggregated total entropy is:

$$(6)\quad H(\theta) = H_o(\theta) + \sum_{M=1}^{M} \psi_m H_m(\theta)$$

where

$$(7)\quad H_o(\theta) = \sum_{M=1}^{M} \psi_m \log_2 \psi_m^{-1}$$

and

$$(8)\quad H_m(\theta) = \sum_{i \epsilon \psi_M} [(\theta_i / \psi_m) \log_2 (\psi_m / \theta_i)]$$

$$\text{for } m = 1, \ldots, M.$$

Between-set entropy is defined by equation (7) while entropy within-set, ψ_m is defined by equation (8). Total within-set entropy is $\Sigma \psi_m H_m(\theta)$.

This disaggregation property is particularly useful when data are available by levels of aggregation such as in Census of Manufactures. Specifically, data for number of establishments and value of output are reported by state, division, and region within the United States.Sometimes data are not reported for all states within a division for disclosure reasons. As a consequence, the least disaggregation which may be consistently attained is divisions. In such a case, regions represent the set ψ_m, with divisions within each set. Total and disaggregated entropies may be computed using these data.

There are relative entropy measures, similar to $R(\theta)$, which may be constructed from the between-set entropy, $H_o(\theta)$, and the within-set entropy, $H_m(\theta)$. These are:

$$(9)\quad R_o(\theta) = H_o(\theta) / \log_2 M$$

and

$$(10)\quad R_m(\theta) = H_m(\theta) / \log_2 N.$$

Of course, both $R_o(\theta)$ and $R_m(\theta)$ are indices and have interpretations similar to $R(\theta)$. That is, $R_o(\theta)$ is a measure of the extent to which the industry under study is attaining maximum possible between-region geographic dispersion in shares given the number of sets, ψ_m. Also, $R_m(\theta)$ is a measure of the extent to which the industry is attaining maximum possible within-region geographic dispersion in shares given the number of divisions within regions.

POULTRY PROCESSING INDUSTRY ENTROPIES

As an illustration of the above methodology, both absolute and relative entropies are computed for the poultry processing industry. Spatial concentration propensities are investigated in terms of share of establishments and share of value of output by geographic areas.

Data on number of establishments and value of output for the poultry processing industry are reported by Census of Manufactures [7]. Shares of the United States total by divisions and regions are computed for each of the last three available census years, Table 1. Total and between-region absolute and relative entropies are computed from these shares regarding regions as four sets (m = 1, ...4), Table 2. Divisions within regions provide the base for relative within-region entropies, Table 3.

The only relevant aspect of the total and disaggregated absolute entropies is their change over time. Nevertheless, they are presented, primarily to illustrate the disaggregative property of $H(\theta)$. More important for analysis are the estimates of $R(\theta)$, $R_o(\theta)$, and $R_m(\theta)$.

Establishment Shares

Examining the intertemporal change in $R(\theta)$ reveals that geographic concentration in terms of establishments shares increased from 1958 to 1967, but the rate of change has not been substantial. Less than a 2 percent decline in $R(\theta)$ over a 10-year period substantiates how slow this change in geographic concentration has been. Also, the magnitude of $R(\theta)$ suggests that the industry was about 93-95 percent of maximum possible dispersion during this period.

Turning to the intertemporal change in $R_o(\theta)$, about the same propensity toward concentration of establishments between regions is revealed as for the total. The rate of change in $R_o(\theta)$ over the 10-year period is just at 2 percent. The rate of change in $R_o(\theta)$ did accelerate from 1963 to 1967 compared to the previous five-year change.

Intertemporal change in the relative within-region entropy, $R_m(\theta)$, shows that the greatest propensity toward within-region concentration over the 10-year period occurred in the Northeast region, followed by the West. During this same period the South region actually became more dispersed among divisions in terms of establishments, while the North

Table 1. **SHARES OF NUMBER OF ESTABLISHMENTS AND VALUE OF OUTPUT FOR POULTRY AND EGG PROCESSING BY DIVISIONS AND REGIONS, SELECTED YEARS***

Geographic Area by Regions and Divisions	Shares[a/]					
	1958		1963		1967	
	Establishment Shares	Output Shares	Establishment Shares	Output Shares	Establishment Shares	Output Shares
			percent[b/]			
United States	100.0	100.0	100.0	100.0	100.0	100.0
Northeast Region	14.0	12.1	13.4	10.0	12.4	8.7
New England Division	4.5	5.5	2.9	4.1	2.0	2.7
Middle Atlantic Division	9.6	6.6	10.5	5.9	10.4	6.0
North Central Region	34.4	30.6	32.2	22.8	31.4	20.2
East North Central Division	18.4	12.3	18.1	7.5	16.8	6.4
West North Central Division	16.0	18.3	14.1	15.3	14.6	13.8
South Region	36.5	48.1	38.7	57.0	41.4	61.6
South Atlantic Division	17.3	27.0	18.7	29.8	20.0	32.9
East South Central Division	6.3	9.0	8.0	12.5	8.5	12.2
West South Central Division	12.9	12.1	12.0	14.7	12.9	16.5
West Region	15.2	9.2	15.7	10.1	14.7	9.5
Mountain Division	3.2	1.1	3.4	1.1	2.6	0.9
Pacific Division	12.0	8.1	12.3	9.0	12.1	8.6

*Source: [7].

[a]Includes all establishments regardless of size.

[b]Region or division percentages may not add exactly due to rounding error.

Table 2. ENTROPY MEASURES FOR NUMBER OF ESTABLISHMENTS AND VALUE OF OUTPUT SHARES BY CENSUS YEARS*

Entropy Measure	Census Year 1958 Establishment Shares	1958 Output Shares	1963 Establishment Shares	1963 Output Shares	1967 Establishment Shares	1967 Output Shares
Total entropy, $H(\theta)$	2.9945	2.8658	2.9830	2.8108	2.9435	2.7248
Between-region entropy, $H_0(\theta)$	1.8705	1.7160	1.8644	1.6148	1.8315	1.5258
Total within-region entropy, $\Sigma\ \psi_m H_m(\theta)$	1.1240	1.1498	1.1186	1.1960	1.1120	1.1990
Relative total geographic dispersion, $R(\theta)$	94.5	90.4	94.1	88.7	92.9	86.0
Relative between-region geographic dispersion, $R_0(\theta)$	93.5	85.8	93.2	80.7	91.6	76.3

*Source: Computed from Table 1.

Table 3. WITHIN-REGION ENTROPIES FOR NUMBER OF ESTABLISHMENTS AND VALUE OF OUTPUT SHARES BY CENSUS YEARS*

Geographic Region	Entropy within separate regions, $R_m(\theta)$: 1958 Establishment Shares	1958 Output Shares	1963 Establishment Shares	1963 Output Shares	1967 Establishment Shares	1967 Output Shares
Northeast	90.0	99.4	75.4	97.7	63.7	89.4
North Central	99.7	97.2	98.9	91.4	99.7	90.1
South	93.3	89.6	94.7	93.0	94.7	91.8
West	74.3	52.8	75.4	49.7	67.3	45.2

*Source: Computed from Table 1.

Central dispersion among divisions remained constant. The estimates or $R_m(\theta)$ also suggest that for the most recent year, the North Central region is near maximum dispersion within the region while the Northeast and West regions are 60-70 percent of maximum.

Output Shares

Entropy measures based on output shares are more important than establishments shares for some purposes, since size of establishments is reflected in the former shares but not in the latter.

Propensity toward geographic concentration for output shares has been more pronounced than for establishments shares, Table 2. The change in relative total entropy, $R(\theta)$, was about 5 percent from 1958 to 1967, based on output shares. This compares to less than 2 percent change in $R(\theta)$ over the same period based on establishments shares.

An even more striking difference exists in the intertemporal changes in between-region entropy, $R_0(\theta)$, for the two share types. For the 10-year period, output shares between region entropy, $R_0(\theta)$, declined slightly over 11 percent, compared to 2 percent for the comparable statistic based on establishments shares. This indicates that geographic concentration in terms of size of establishment occurred substantially faster than in number of establishments. The rate of change in output concentration between regions did slow slightly from

1963 to 1967, compared to the previous 5 years. This is contrary to the rate of change in $R_o(\theta)$ based on establishments shares which increased from 1963 to 1967.

The within-region relative entropies, $R_m(\theta)$, show marked changes over the 3 census years, Table 3. The greatest within-region concentration increase was the West, followed by the Northeast and the North Central regions. The South, as with $R_m(\theta)$ based on establishments, was more geographically dispersed among divisions in 1967 than in 1958.

Output shares concentration is greater within-region in the South and West than the concentration of establishments shares. This is especially pronounced comparing $R_m(\theta)$ in 1967 for the two shares. In the North Central and South regions the within-region entropies are similar for either share type.

CONCLUSIONS

Spatial concentration may be quantified by adapting the entropy measure of information theory. Intertemporal comparisons of entropy allow concentration propensities to be investigated. Relative entropy is more useful for spatial analysis than absolute entropy, since regions are of different size and equal shares are not expected.

The disaggregation property of total entropy into between-set and within-set entropies is particularly useful for analysis of data reported by divisions and regions. Entropies for the poultry processing industry document a slight propensity toward increased concentration between regions and a relatively rapid propensity toward within-region concentration for the Northeast and West. In general, the relative between- and within-region entropies reveal that this industry is tending toward spatial concentration, especially on an output shares basis. This suggests that spatial concentration is occurring more rapidly in size of establishments than in numbers of establishments.

REFERENCES

[1] Horowitz, Ann, and Ira Horowitz, "Entropy, Markov Processes and Competition in the Brewing Industry." *Journal of Industrial Economics,* 16:196-211, July 1968.

[2] Horowitz, Ira, "Employment Concentration in the Common Market: An Entropy Approach," *Journal of the Royal Statistical Society,* A, Part 3:463-479, 1970.

[3] ____, "Numbers - Equivalents in U.S. Manufacturing Industries: 1954, 1958, and 1963," *Southern Economic Journal,* 37:396-408, April 1971.

[4] Horvath, Janos, "Suggestion for a Comprehensive Measure of Concentration," *Southern Economic Journal,* 36:446-452, April 1970.

[5] Rose, P. S., and D. R. Fraser, "The Prices of Bank Services and Alternative Measures of Market Structure," paper presented at the Southern Economics Association meetings, Houston, Tex., 1973.

[6] Theil, Henri, *Economics and Information Theory.* Amsterdam: North-Holland Publishing Co., 1967.

[7] U.S. Department of Commerce, *Census of Manfactures,* Vol. II, Industry Statistics, 1958, 1963, 1967.