Food, Processed

Organization and Performance of World Food Systems: NC-194

OCCASIONAL PAPER SERIES

The work reported herewithin contributes to the objectives of North Central Regional Project NC-194, a joint research project of state agricultural experiment stations and the U.S. Department of Agriculture

AN INTERNATIONAL COMPARISON OF CONCENTRATION AND DIVERSITY IN THE FOOD PROCESSING INDUSTRY

*Joseph G. Hirschberg

OP-46 JANUARY 1993

*THE AUTHOR IS AN ASSISTANT PROFESSOR, DEPARTMENT OF ECONOMICS, SOUTHERN METHODIST UNIVERSITY, AND VISITING PROFESSOR, DEPARTMENT OF AGRICULTURAL ECONOMICS, THE OHIO STATE UNIVERSITY

Abstract

International comparisons of market structure are complicated by a lack of comparable data. Although the U.S. Department of Commerce reports measures of industry concentration, they do not verify either the control of subsidiary firms or the possible multinational nature of their ownership nor are consistently based on sales. Other countries produce similar reports, but these studies are generally not comparable to U.S. values due, in part, to incompatible sector definitions. In addition, few government- sponsored studies provide firm-level detail or timely information. Furthermore, given the widespread multinational nature of many larger firms, an international analysis of ownership is necessary. This paper addresses the issues encountered in the construction of international market data from traditional financial report data and provides a set of methods for the comparison of measures of market concentration and industry diversity across countries.

Using 1991 financial data, a firm level data set is constructed and used to compute comparable measures of market concentration and industry diversity for the U.S. and E.C. food processing industries.

Introduction

The purpose of this paper is to make international comparisons of market concentration and diversity from multiproduct firm data. These comparisons are made for the food processing industry in the U.S. and in E.C. Due to problems of data compatibility, cross-country comparisons of industrial structure are far less common than interindustry studies (Yamawaki, Sleuwaegen and Weiss 1989). Furthermore, the majority of market structure studies employ establishment-level data that do not allow for the possibility of interrelated and multinational ownership that characterizes larger firms. This paper presents a strategy for using firm-level data to estimate market structure variables in order to make cross-country comparisons. This method provides not only point estimate comparisons of various market structure parameters but also a means for testing various hypotheses concerning the differences in market structure.

One method for avoiding the data incompatibilities of government produced concentration measures is to construct a firm-level data set, such as one collected by private investment information services or from direct contact with the firms. This approach was used by Sutton (1991) for the four largest food producing firms in France, West Germany, Italy, Japan, the United Kingdom and the U.S. The present study concentrates on the largest firms in all countries and unlike Sutton, it traces the impact of multinational ownership. Because this is a larger group of firms, it relies exclusively on the information available from private investor data.

This study employs a Monte Carlo method to simulate the sales and employment distributions for each industry by apportioning the firm's sales according to a specific statistical distribution. These generated characteristics are then used to compute measures of industrial concentration and diversity that are subsequently compared across industry and country.

Furthermore, given the global nature of the data, it is also possible to define markets among groups of countries, allowing comparisons between trade groups and specific nations. Due to data limitations, this study examines only major firms (with total sales in all lines of production over $150 million) with at least one product in the food processing industry (SIC 20) as reported in the Dun & Bradstreet computer data base. These data limitations also restrict the measures reported here to a class of Herfindahl-Hirschman shape measures as opposed to measures based on shares of the total market.

The paper proceeds as follows. In section 1 the steps in constructing the data set are outlined. Then a description of the simulation used in the analysis is given in section 2, along with a comparison to results obtained from a similar source of these data for the U.S. alone. Sample-size independent measures of industrial concentration and diversity are defined in section 3. Sections 4 and 5 respectively provide the results of market concentration and diversity comparisons between the E.C. and the U.S. by sector in the food processing industry.

1. Firm Data

A firm's sales and employment data by SIC are often difficult to obtain. The most readily available source for this information is the firm's annual report. However, there are at least three reasons why annual reports are frequently inadequate for this purpose:

> 1. Private firms and producer cooperatives often issue no report and a significant segment of many industries may be composed of these types of organizations.
>
> 2. Very rarely do annual reports provide a decomposition of sales or employment by product or country.
>
> 3. Annual reports often fail to identify the full set of subsidiary firms they hold, or if they are subsidiaries they may not identify any other similar subsidiaries or the parent firm.

Commercial investment data base vendors such as Dun & Bradstreet, Ward's Business Directory (Gale Research), and Trinet are examples of readily accessible sources that include information for both public and private firms (see Hirschberg, Dayton and Voros, 1992 for more details concerning these and other data sources). Some of these data series also identify firms that are subsidiaries to other firms and the level of investment involved. Furthermore, these data sets list a standardized set of product line classifications as well as sales and employment information.

The information from the Dun & Bradstreet's interactive computer data base service provide the most complete source for U.S. and foreign firms. However, these data do not indicate the intensity of production in any particular product line as do the Trinet data and Ward's data available for U.S. firms. In the *Who Owns Whom* published data source, Dun & Bradstreet provide a means for identifying both the U.S. and foreign parent firms for each data entry. Unfortunately the only market-specific information available is a ranking of up to six 4-digit SIC product codes. An example of the entries for the Dun & Bradstreet data is given in the Appendix. The information extracted from these data is:

Name of the firm.
Address of the firm.
Country where the firm is located.
Up to 6 4-digit SICs in order of importance.
The total annual sales (as of 9/91).
The total number of employees (as of 9/91).
Whether the firm imports and/or exports.
The name of the parent firm if the firm is a subsidiary.

A major element in construction of this data set is the identification of which firms are subsidiaries to other firms. Because of the incomplete nature of ownership correspondences in the electronic data sets, it was necessary to verify this information with two other sources: *Who Owns Whom* that contains the parent firm for subsidiaries located in North America, the United Kingdom, Ireland, and continental Europe, and the available annual reports that include information on subsidiaries. The data set contains information for 1,695 firms and subsidiaries that have total sales of over $150 million or more and that sell at least one product in a SIC 20 industry. In addition, because the entry for the parent firm may report the sales and employment total from a number of subsidiaries that also appear in the data set it was necessary to subtract the sales of the subsidiaries that do appear in the data from the parent firm to avoid possible double counting of the sales and employment values.

Table 1 shows the number of firms that are included from each country in the size group with combined sales of $150 million and that do business in at least one industry in the SIC=20 group. Thus, a firm may only do business in one such industry and have the majority of its sales in a sector other than food processing. **Table 2** provides an alternate view of the same data where the sales and employment are allocated by the country of the parent firms. Note that the ranking of the countries changes when total sales by ultimate ownership is used. In **Table 2** Switzerland moves from 9th to 5th, compared to **Table 1**. This shift is due mainly to the influence of the Nestle Company. The net difference in total sales reflects the net balance of foreign ownership in each country for firms in this size group. In the U.S., American-held firms account for 89 percent of domestic sales. For Switzerland, almost 2.5 times the domestic market sales are sold by Swiss held firms world-wide. The Netherlands is another net owner nation with a relative world market of almost 1.7 times domestic sales. Note that these are sales by firms owned in the parent country and do not reflect the export sales. The sales are allocated to the country in which the firm is headquartered, not by country in which the goods are sold. Thus, some small countries may have sales totals that are larger than their domestic markets. Also, these sales are for the firms in the data set, thus they include sales in industries other than SIC=20. In the next section, the method for allocating these sales by SIC will be described.

2. Simulated Diversity of Sales and Employment.

The model of firm diversity relies on the ordering of the SIC given for each firm or subsidiary, along with the assumption of a distribution for the shares of the sales and employment. Lacking detailed technological data, it is assumed that the distribution of a firms' sales is the same as that of employment. Although technical factors may differ by industry, scale and country, it is assumed that firms that produce similar products employ similar technologies. In order to generate a distribution of the sales or employment, a set of as many random numbers as reported SICs are drawn. These random values are generated according to a particular statistical distribution (the choice of distribution is discussed below) so that they form non-negative ordered weights that sum to one. The total sales and employment of the firm is then distributed by SIC.

The statistical distribution chosen to generate the random values will determine the form of the weights chosen. Five distributions were employed. Three distributions were used for sales in SIC; the uniform, the lognormal, and the Pareto. Two distributions were based on a particular data set describing the distribution of sales within a firm by SIC; an empirical distribution based solely on observations, and an estimated multivariate kernel density function. Each of these will result in a different characteristic pattern of the weights r_{1k} - r_{6k} (where k is the total number of SIC's listed for the firm). The construction of these weights is described below.

2.1 Distributions of Firms Sales.

If it is assumed that each firm's product is produced by an independent subsidiary then the distribution of sizes for each subsidiary within the firm will be related to the distribution of similar firms without regard to ownership. For example, a number of firms have purchased existing companies that produce a product that they had not previously sold in any corresponding market. Thus the size distribution of these firms that are owned by other firms are considered the same as the distribution of the firms without regard to ownership. In the following section a number of such distributions are proposed which have different implications for these distributions.

Drawing from a uniform or rectangular distribution results in the least difference between the weights. The uniform can be used as the distribution that results when all firms in an industry have an equal probability of having a size between the limits of the distribution (a = lower bound and b = upper bound). The weights from this distribution are generated by:

$$r_{ik} = u_i \, / \sum_{i=1}^{k} u_i, \text{ where } u_i \sim U(a,b) \text{ for all } i \le k.$$

The limits of the distribution (a,b) do not effect the weights computed. In the simulations that follow a = 0 and b = 1. These values are then sorted in descending order. The number of sectors in which the firms sell is given as k. Thus if 4 SICs are given in the data entry, four values are drawn from the random number generator[2]. They are then sorted by size and divided by their total. The average values obtained from the uniform distribution are given in **Table 3**. From this table it can be seen that the average weights decline according to a linear relationship.

---

[2] All the uniform random numbers were generated by the SAS routine RANUNI.

Weights were also drawn from the lognormal distribution. The lognormal has been widely observed as a distribution of firm size (e.g., Quandt, 1966 and Silberman, 1967). These values were drawn in the same manner as the uniform, then sorted and weighted to form the sample weights by using $u_i \sim e^{N(0,1)}$.[3] The average weights from the lognormal distribution are listed in **Table 4.**

Another distribution that is also proposed for the distribution of firm sizes is the Pareto of the first kind (see Quandt, 1966). The cumulative distribution function of the Pareto is given by $F(x) = 1 - x^{-c}$, where $1 \leq x \leq \infty$ and $c > 0$. The weights generated by this distribution will depend on the shape parameter c. This parameter was estimated as .9124 using the data described below. A Pareto distributed pseudo-random value is generated by the following process $x_{ik} = (1/u_i)^{1/c}$, where $u_i \sim U(0,1)$. The resulting mean values of the weights are given in **Table 5.**

An alternative to the previous parametric statistical distributions is to use a nonparametric representation of the distribution. In order to use a nonparametric method, it is necessary to have observations on the proportion of sales in various markets. This differs from the generation methods described above that estimate a distribution of sales and then, by assuming the sales in each market are independent of each other, deriving the weights. The advantage of nonparametric representations is that they are based on observed behavior; the disadvantage is that these distributions will always reflect the data that were used to create them, thus anomalies in the data will be treated as information.

---

[3] The lognormal was generated by using the SAS RANNOR routine by raising e to the power of the pseudo-random value generated.

The data used in this estimation are from a market-oriented data base that lists the sales of the 50 largest firms by four digit SIC in the U.S. These data are for 1991 and are compiled by the Trinet Corporation (1991). To combine the sales of one firm across markets, a new data set was created where all the listings for a particular firm are combined with sales in each four digit SIC (using only SIC=20). To compute the weights implied by these values for firms that sell in two markets, all those firms that sell in at least two markets are used to estimate the weights; the same for those that sell in at least three, on up to six markets. In this way, the largest sample of firms in each category is used.

The first nonparametric technique employed is a multivariate kernel density estimate. This technique uses a weighting scheme (the kernel) to compute a continuous function as an estimate of the density function. This type of estimate may be viewed as a smoothed histogram with the possibility of tails that extend beyond the range of the data. In particular a modified multivariate Epanechnikov kernel estimator was used to generate a set of r_{ik}s, by values of k between 2 and 6 (see Silverman 1986). The modification was based on the two properties of the proportions that helps to simplify estimation: $\sum_{i=1}^{k} r_{ik} = 1$, and $r_{ik} \leq r_{jk}$, when $i<j$. Using these distributions a series of random numbers was generated using a look-up table of the cumulative density distribution and a uniform random number generator. The average weights are given in **Table 6.**

The second nonparametric method used for generating weights was an empirical random number generator based on the Trinet data. This is equivalent to randomly selecting the weights from the data. To construct this type of random number generator it was first necessary to replicate the weights from the data a number of times (the number depends on how many weights need to be generated) then to assure that each weight is independent from each other they are shuffled. The weights for each total number of SICs a firm sells (k) are constructed separately. **Table 7** reports the average values obtained from the Trinet data; they are the average weights generated from this procedure.

2.2 An Evaluation of the Statistical Distributions of Sales.

In order to evaluate these and other possible distributions from which weights could be drawn, the average weights computed from various distributions were compared to weights computed from the Trinet data set that estimates the largest (in sales) fifty U.S. firms by four digit SIC[4]. A modified goodness-of-fit statistic or distance measure (D) for discrete multivariate distributions (see Read and Cressie 1988), based on Kullback's (1958) concept of directed diversity, was calculated using the following formula:

$$D = \sum_{k=2}^{6} w_k \left(\sum_{i=1}^{k} \hat{r}_{ik} \ln(\hat{r}_{ik} / r_{ik}) + r_{ik} \ln(r_{ik} / \hat{r}_{ik}) \right)$$

where w_k is the proportion of the firms in the Dun & Bradstreet data that report selling in k markets (SICs) ($w_1 = .266$, $w_2 = .231$, $w_3 = .167$, $w_4 = .130$, $w_5 = .107$, $w_6 = .099$), $\hat{r}_{ik}$ is the average prediction of the proportion of sales in the ith largest market for a firm selling in k markets (as given in **Tables 3, 4, 5** and **6**), and r_{ik} is the average observed value from the Trinet data set (as given in **Table 7**). (Note that the case where k=1, $r_{11} \equiv \hat{r}_{11} \equiv 1$, the predicted and actual values of proportions are equal to one by definition, thus they are excluded from the computation of D.) The smaller the value of D, the smaller the distance and the greater the

[4] The Trinet data were created by a combination of market research and an economic model of firm diversity applied to aggregate firm data. Thus in using these distributions in this study it important to qualify the inferences drawn from these data.

similarity between the distributions that generated the weights. Comparing samples of size 1000, the following results were obtained from the mean and the median of the generated weights[5].

Distribution	D (Median)	D (Mean)
Uniform	.07649	.07906
Lognormal	.00664	.00659
Pareto (c = .9124)	.00497	.00756
Half-Normal	.03340	.03372
Normal^2	.04637	.05702
Normal^4	.22945	.31512
Multivariate Kernel Estimate	.00669	.00727

These values show that the kernel estimate and the lognormal distributions are very close, while the uniform distribution produces values that are an order of magnitude further away. The furthest distribution investigated was a normal raised to the fourth power; this is a highly skewed distribution. Note that all the candidate distributions were chosen so that they generate positive sales values and thus the normal was not used because it would require the assumption of a mean and a standard deviation that effects the distribution of the resulting weights. The lognormal, half-normal, and the normal distributions raised to even powers were functions of standard normals. As mentioned above, the Pareto was located with a shape parameter that minimized the value of D by estimating D under a series of values for c, thus, this value depends on the sample and is dependent to a small degree on the quality of the data.

Given the estimated nature of the Trinet data, it may be that the present analysis only serves to derive the distribution employed in the construction of the data. Thus, comparisons to

[5] The choice of seed for the random number generator and the size of the sample drawn for the simulation will result in slight variations in the orders listed here. Those values that are close to each other in magnitude may change order in a different experiment.

the Trinet data are not necessarily definitive, nor is the use of a distribution based on the Trinet data necessarily the best alternative. The uniform appears to make the most general assumption about sales distributions and may be considered the least restrictive. The lognormal and Pareto, besides appearing to fit the Trinet data the best, both coincide with a number of past studies of firm size distributions. The multivariate kernel estimate is intended to be more general than the empirical distribution because it allows for values that never occurred in the Trinet data while retaining the shape of the empirical distribution. However, as noted above, the empirical distribution may be too closely based on the Trinet sample. Yet, unlike any of the other distributions, it only reflects observed distributions (see Dagpunar, 1988). All the analysis that follows was performed with each of these five distributions so that any inferences to be drawn can be tempered by the choice of distribution.

The coverage of the Dun & Bradstreet International data allows the computation of comparable measures of firm concentration by SIC and by country or group of countries. In the remainder of this paper the combination of SIC and region will be used to define a market. Because the sample of firms chosen in this study is determined by a size factor and the allocation of sales is done via a random selection criteria, the analysis that follows concentrates on the differences between the parameters computed for various regions and SICs. This section proceeds as follows: first measures of concentration which are independent of sample size are defined, then comparisons are made between international regions.

3. Sample Size Independent Measures of Concentration.

A widely used index of market concentration is the Herfindahl-Hirschman index (H_{ij}) (Hirschman, 1945 and Herfindahl, 1950) for SIC j and region i:

$$H_{ij} = \sum_{k=1}^{n_{ij}} P_{ijk}^2$$

Where P_{ijk} is the proportion of total sales (or employment) in SIC i and region j for firm k, and n_{ij} is the number of firms in SIC i and region j (market ij). As the value of H_{ij} increases the level of concentration increases.

In order to make international comparisons of market concentration, it is necessary to construct H_{ij}s which are based on a comparable number of firms. Here n_{ij} varies by market, thus direct cross-market comparisons of H_{ij} will be contaminated by differences in sample size.

H_{ij} can be written as:

$$H_{ij} = \sum_{k=1}^{n_{ij}} X_{ijk}^2 \left(\sum_{k=1}^{n_{ij}} X_{ijk} \right)^{-2},$$

where X_{ijk} are the level of sales or employment and H_{ij} can be rewritten as:

$$H_{ij} = \frac{A_{ij}}{A_{ij} + 2B_{ij}}, \text{ where } A_{ij} = \sum_{k=1}^{n_{ij}} X_{ijk}^2 \text{ and } B_{ij} = \sum_{k=1}^{n_{ij}} \sum_{q=k+1}^{n_{ij}} X_{ijk} X_{ijq}.$$

Thus if H_{ij} is computed from data with the same distribution but from a different size sample (n_{ij}), H_{ij} will fall with increasing values of n_{ij} because, all else being equal, B_{ij} will be larger for larger n_{ij}. In order to compute a comparable H_{ij} it is necessary to define a sample size independent H_{ij}. One way of doing this is by weighing the average of A_{ij} and B_{ij} by sample size. Thus, one can compute:

$$\bar{A}_{ij} = \frac{A_{ij}}{n_{ij}}, \text{ and } \bar{B}_{ij} = \frac{B_{ij}}{(n_{ij}^2 - n_{ij})/2},$$

where $(n_{ij}^2 - n_{ij}) / 2$ is the number of terms in B_{ij}. These averages can then be weighted to compute an equivalent A_{ij} and B_{ij} for a hypothetical sample size using the formulae given below From these values we can compute a new H_{ij} based on these modified values of A_{ij} and B_{ij},

$$A_{ij}^{*} = \left(\frac{n_{ij}^{*}}{n_{ij}}\right) A_{ij}, \text{ and } B_{ij}^{*} = \left(\frac{(n_{ij}^{*2} - n_{ij}^{*})}{(n_{ij}^{2} - n_{ij})}\right) B_{ij}.$$

which will be referred to as the means equivalent (me) value of H_{ij}:

$$H_{ij}^{me} = \frac{(n_{ij} - 1)\, H_{ij}}{(n_{ij} - n_{ij}^{*})H_{ij} + n_{ij}^{*} - 1}.$$

For example, if n = 25, n^{*} = 50, and H = .2000 then H^{me} = .1091, the sample-size-compensated H is almost half the computed value.

An alternative method for creating a sample size compensated value of H_{ij} is to use the numbers equivalent (ne_{ij}) interpretation of the Herfindahl-Hirschman Index. For any H_{ij} it is possible to determine the number of equal size firms (ne_{ij}) that would have resulted in the same H_{ij}:

$$H_{ij} = \sum_{k=1}^{ne_{ij}} \left(\frac{1}{ne_{ij}}\right)^{2}, \text{ or } H_{ij} = \frac{1}{ne_{ij}}.$$

This number of firms can be compared to the sample size used to construct the H_{ij}. A relative equivalent number of firms can be defined as:

$$\alpha_{ij} = \frac{ne_{ij}}{n_{ij}}, \text{ or } = \frac{1}{H_{ij} n_{ij}}, \text{ where } 0 < \alpha_{ij} < 1.$$

H_{ij} written as a function of α_{ij} is $H_{ij} = 1 / n_{ij}\, \alpha_{ij}$. Under the assumption that α_{ij} is constant across sample size, a sample-size-equivalent H_{ij} can be defined that will be referred to as the numbers equivalent (ne) H_{ij}:

$$H_{ij}^{ne} = \frac{1}{n_{ij}^{*} \alpha_{ij}}, \text{ or } H_{ij}^{ne} = \left(\frac{n_{ij}}{n_{ij}^{*}}\right) H_{ij}.$$

Using the same example as above where n = 25, n^{*} = 50, and H = .2000 then H^{ne} = .1000.

Both transformations of H_{ij} result in smaller values for H_{ij} when the hypothetical sample is larger than the actual sample. Because the values of H_{ij} are often based on samples sizes of 50 (c.f. U. S. Census Bureau 1992) $n^* = 50$ was used in this paper to construct the sample-size-equivalent H_{ij}. In all but one SIC, the sample size observed was smaller than 50.

It must be cautioned that both of these methods attempt to capture the shape of a distribution based on either one parameter (α_{ij}), or two (A_{ij},B_{ij}). While the means equivalent transformation is based on more information, the numbers equivalent transformation has a more intuitive form. The numbers equivalent transformation will weight both H with the same values if both samples are of the equal size, while the means equivalent transformation for the same case would use weights that depend on the H as well. In most cases $H^{me} > H^{ne}$. This can be seen in **Figure 1** which is a contour plot of the percent difference (PDIFF, where PDIFF = $100\ (H^{me} - H^{ne}) / H^{ne}$) between the two sample-size-equivalent methods when $n^* = 50$. The percent increase of H^{me} to H^{ne} is at most 30 percent for a case in the lower right-hand corner where n = 10 and H = .3, but the percentage difference diminishes as H falls and n approaches 50. In the results discussed below both the "me" and the "ne" measures are reported.

4. A Comparison of Market Concentration in the U.S. and the E.C.

An index of concentration (H_{ij}) was computed by SIC and weighted using both the means and numbers equivalent methods. Two sets of differences between the H_{ij} for the E.C. and the H_{ij} for the U.S. were defined by

$$DH_i^{me} = H_{iEC}^{me} - H_{iUS}^{me}, \text{ and } DH_i^{ne} = H_{iEC}^{ne} - H_{iUS}^{ne}.$$

These differences were computed using the five distributions of the weights (uniform, lognormal, Pareto, multivariate kernel, and empirical) in a series of five hundred experiments each[6]. Five hundred values of the DHs were computed by SIC, distribution and equivalence method ("ne" or "me"). Those SICs in which 90 percent or more of all the DHs are either greater than or less than zero were are reported below. This amounts to a test under the assumption that the total sales or employment is correctly given but uncertainty existed as to the distribution of the weights for each firm's sales by SIC.

The table listed below summarizes the number of cases in which 90 percent or more of the DH_i are the same sign. The counts are based on cases where both the differences in employee and sales based concentration measures indicate significantly higher concentration in the U.S. or the E.C. The uniform distribution resulted in the largest number of significant differences and the empirical had the fewest. This dichotomy was especially pronounced for the DH^{ne} although overall the ne sample-size-equivalent methods appear to result in more SIC's with significant results than the me method.

Distribution	$H_{US} > H_{EC}$ (me)	$H_{EC} > H_{US}$ (me)	$H_{US} > H_{EC}$ (ne)	$H_{EC} > H_{US}$ (ne)
lognormal	13	5	14	9
Pareto	9	6	14	9
uniform	13	7	15	10
empirical	14	8	16	10
multi kernel	13	9	15	11
All	7	5	13	5

This table shows that the different distributions did not introduce any obvious pattern, except that the Pareto is a bit lower in the determination of higher U.S. concentration under the means equivalent method of comparison. However, the means equivalent method results in a

[6] In comparisons with experiments of differing size, little variation in the results were observed except for the experiments drawing from the empirical based distributions.

lower proportion of SIC's in which the E.C. has a higher concentration than the U.S. than is revealed by the numbers equivalent method. The row labeled "All" indicates the number of industries where over 90 percent of the cases generated by all the distributions share the same sign. Thus for the "$H_{US} > H_{EC}$ (me)" column over 90 percent of the differences calculated with any of the distributions share the same sign for seven industries. These results indicate that U.S. markets are more concentrated than the E.C. market.

Table 8 provides a summary of these results by SIC, of the count of cases in which 90 percent or more of the differences have the same sign. If $DH_{ij} < 0$ then the U.S. market is more concentrated than the E.C. and conversely if $DH_{ij} > 0$ the E.C. market has the higher concentration. Of the forty-nine 4-digit SICs in SIC=20, six had data for either only the E.C. or the U.S. and another five had no case in which 90 percent of the concentration indices were different between the U.S. and the E.C. For the remaining thirty-seven SICs in at least one case (combination of distribution and method for equivalence), over 90 percent of the EC/U.S. differences were of the same sign for either sales or employment. In a number of cases the signs were reversed for the differences computed from employment and those computed from sales. This table provides the count of cases in which there is a significant difference. Because there were no cases in which a significant result was observed for one sign of DH_{ij} under one measure of firm size and another measure of firm size resulted in an opposite sign, only the counts for one sign are reported.

In cases when the sample sizes are very disparate between the U.S. and the E.C. (especially when the number in one market is small, i.e., less than four) there is a marked difference between the results obtained from the two equivalency methods. This occurred for

SIC 2076 (vegetable oil mills), 2083 (malt) and 2091 (canned and cured fish). For this reason these cases are not referred to in the discussion of the SIC specific results that follows.

Table 8 reports a number of cases where a smaller number of firms in a market sample result in a lower market concentration than the market in the other region which has a larger sample size. This can be seen in SICs 2015, 2037, 2038, 2051, 2052, 2085, 2087. Obviously these results would not have been obtained without the application of an equivalence method.

The SICs which consistently indicate a higher concentration in the U.S. are 2024 (ice cream and frozen deserts), 2038 (frozen specialties), 2042 (cereal breakfast foods), 2045 (prepared flour mixes and doughs), 2052 (cookies and crackers), 2086 (bottled and caned soft drinks), 2087 (flavoring extracts and syrups), and 2095 (roasted coffee). The industries with less uniform results which indicate higher U.S. concentration are 2011 (meat packing plants), 2015 (poultry slaughtering), 2046 (wet corn milling), 2051 (bread cake and related products), and 2075 (soybean oil mills).

The E.C. only showed consistency higher concentration in the industries 2026 (fluid milk), 2062 (cane sugar refining), and 2064 (candy and other confection products). Less conclusive indications of higher E.C. concentration were found for 2063 (beet sugar) 2077 (animal and marine fats and oils), and 2085 (distilled and blended liquors).

For a number of SICs, sales and employment data imply contrary results. The most dramatic of these cases is SIC 2013 (sausages and other prepared meats) where the sales data infer higher concentration in the U.S. and the employment data infers a higher concentration for the E.C. In both 2079 (edible fats and oils) and 2084 (wines brandy and brandy spirits) these results are reversed. Although these differences may indicate a technological differences, it is more probable, given the inconsistency of the reporting for employment data, that these conflicts are data artifacts.

Of further interest are those markets where in both the E.C. and the U.S. the number of firms is about equal, and under no distribution or equivalence method are there significant differences between concentration measures. This observation appears to confirm the assumption of a common distribution for the same industry no matter where it is located. The industry with the greatest number of firms where this occurs is 2041 (flour and other grain mill products) followed by 2047 (dog and cat food), 2034 (dehydrated fruits, vegetables, soups), and 2098 (macaroni and spaghetti). In a number of markets only one analysis method led to significant differences, for example, 2024 (creamery butter), 2022 (cheese; natural and processed), 2023 (Dry, condensed evaporated products), 2032 (canned specialties), 2033 (canned fruits and specialties), 2035 (pickles, sauces, and salad dressing) and 2066 (chocolate and cocoa products).

Twenty tables of concentration measures were computed -- five distributions times two methods of equivalence, times two indicators of firm size (employment and sales). **Table 9** is a representative sample from this set of tables using the lognormal distribution, the means equivalent method of comparison and sales as the indicator of firm size. The lognormal is chosen because it fits the Trinet data well, while not being a function of that data. The means equivalent method is used because it incorporates more information in its value. Sales are used as the firm-size indicator because, although a subset of firms has no employment data, they all have sales data.

Table 9 provides the mean transformed concentration indices (times 1000) and a column that indicates a level of significance that is the percentage of the five hundred experiments in which the difference between the U.S. and the E.C. is the same sign. Also given are the equivalent numbers of firms as well as the total sales for each market. In particular, the total sales for the smaller sectors appear to be a bit high; and this is due to the model employed here which does not weight sales by SIC but by order in a firm's portfolio of SICs. Furthermore, the sales totals reflect the sample that is drawn. Thus, if an industry is made up of firms that will not be represented in this sample due to size, the total will grossly underestimate the total market sales. An interesting result from this table is the high proportion (thirty-nine of forty-two) of industries where over 90 percent of the differences between concentration indices are in the same direction. Also listed in this table is the equivalent number of firms computed as the average of the actual non-transformed reciprocal of H_{ij}. This means that they may not be made into an equivalent value and thus can not be compared between samples.

A correlation analysis was performed for the H_{ij} between the U.S. and the E.C. markets. A number of researchers (c.f. Bain 1966 and Sutton 1991) have noted the cross-economy relationship of industry concentration. The rationale for this phenomena is that the similarity of technology and tastes determines that a certain level of concentration will hold across countries. A test was performed to determine if this was true using the lognormal me transformed indices. The correlation between the concentration indices was computed for each experiment using both the Pearson correlation coefficient and the Spearman rank correlation coefficient. With both measures a positive correlation was found. For the Pearson correlation coefficient of the sales based concentration measure they ranged between .42 and .02 with a mean of .19 and a median of .18. The rank based Spearman coefficient ranged between .22 and -.06 (a lower 25 percent value at .0) with a mean of .05 and a median of .04. Both correlation analyses when applied to the employment based concentration measures did not result in statistically significant correlation coefficients. This was probably due to the poorer quality of the employment data. Thus, this statistical relationship between the concentration indices, at least in sales, is confirmed in these data.

5. Diversity Comparisons Between the U.S. and E.C.

The level of diversity of the firms selling in each market (SIC-region combination) is determined by the measure proposed by Berry (1971) for the study of diversity of production by a conglomerate firm. This value is defined as B_{ij}:

$$B_{ij} = 1 - \sum_{k=1}^{m_{ij}} Q_{ijk}^2$$

where Q_{ijk} is the proportion of total sales of all the firms in region j and SIC i that are in k industry SICs and m_{ij} is the number of these other markets where they operate; markets are defined both as different SICs and different regions of the world (U.S., E.C., rest of Europe, Mexico and Canada, the Far East, and the rest of the world). If all the firms in this market sold only in this market, B_{ij} would be equal to zero. As the firm sells an equal amount in a large number of other markets, this value would approach one. Thus, the greater the value of B_{ij}, the greater the diversity of the average firm in market ij.

Not all the SICs for each firm are SIC=20; they are all the other product markets in which the parent firm has an interest. This implies that B_{ij} will measure both the degree of vertical and horizontal integration. No classification of the SIC's was performed to differentiate these forms of diversity.

B_{ij} can be interpreted as 1 minus an appropriately defined Herfindahl- Hirschman index, thus the definitions of the sample size compensated versions (me and ne) of B_{ij} are identical to those defined for HH_{ij} given above. In the present case the ideal sample size of 250 other market/firm combinations is used to compare to m_{ij}. These measures are similar to the industry diversity measures proposed by Clarke and Davies (1983); however, in the present case, the individual diversity computations are not computed because in most cases they will be solely a function of the distribution assumed to compute them.

The statistical distributions described in **Section 4** are used to allocate each firm's sales and employment by SIC. Then the sales and employee values of firms that are subsidiaries to the same parent are aggregated into one large firm by SIC and country and/or region. Thus, via the activities of its subsidiary firms, an E.C. firm may sell in many different international regions and in more than 6 SICs. In this sample the largest number of SIC/region markets in which any firm sells is 46. However, more than 90 percent of the firms have 6 or fewer SIC/regions in which they sell; the average is 3.7 with a median of 3. Once this allocation has been made, sales and employment by SIC and region are summed to make an industry measure for comparison. As in the market concentration ratios compared above, the diversity indices are compared between the U.S. and the E.C.

The table given below shows the count, by distribution and sample-size-equivalency, of the number of SICs for which over 90 percent of the 100 experiments[7] result in differences in market diversity of the same sign. This table shows that the number of SICs in which a significant number of differences in diversity are recorded is greater than those for which concentration was high. Again, as with the comparison of concentration, it is the uniform distribution that results in the highest level of significant results compared to the other

[7] Due to the larger scale of the diversity computations only 100 experiments could be performed without the need for a completely restructured method for their computation.

distributions; the Pareto is the lowest. The E.C. markets appear to have firms that are less diverse than the comparable U.S. markets.

Distribution	$B_{US} > B_{EC}$ (me)	$B_{EC} > B_{US}$ (me)	$B_{US} > B_{EC}$ (ne)	$B_{EC} > B_{US}$ (ne)
lognormal	8	2	10	8
Pareto	3	2	3	8
uniform	18	3	17	12
empirical	14	6	14	4
multi kernel	6	5	8	10
All	3	2	3	2

Table 10 gives the breakdown of the diversity results by SIC. This table corresponds to **Table 8** for the concentration ratios in that it lists the counts of distributions in which 90 percent or more of the experiments lead to differences of the same sign. In this case a positive sign indicates that the E.C. market is more diverse than the U.S. market and a negative value indicates the opposite. Again, the first number is the comparison based on sales and the second is the number based on employment. The "number of firms" column lists the total number of other firms-markets that are sold to by the firms in the market over the number of firms in the market. Thus, for industry 2091 the 3 firms in this sample that sell in the E.C. have 45 other industries and regions in which they or their parent also sell.

From **Table 10** it can be seen that the U.S. markets are made up of more diverse firms than the corresponding E.C. markets, although there are only three industries in which the U.S. is more diverse under all comparisons; 2032 (canned specialties), 2038 (frozen specialties) and 2091 (canned and cured fish and seafoods). Definitive results were obtained for greater diversity in the E.C. in only 2037 (frozen fruits and vegetables) and 2046 (wet corn milling). A majority of the diversity differences are not significant and in general these results are not as strong as the concentration results; these results are based on only one hundred experiments versus the five hundred on concentration. Yet, the one hundred experiment results for the concentration values did not result in markedly fewer significant results.

Table 11 is the corresponding table to **Table 9**. These are the experimental results for the diversity computations using the lognormal firm diversity distribution, the mean equivalent comparison method and the sales as the indicator of size. The sales totals are for all the firms that sell in a particular market; this means that a firm's total sales may be included in both the E.C. and the U.S. total if it sells in both.

6. Conclusion

This research demonstrates the quality of the inferences available from a data set that is solely constructed from the information contained in typical financial reports that are supplemented with an ordering of market participation by importance. The concentration differences can be made a function of other variables that capture the taste and technological aspects of the SICs. The inclusion of all the countries in the E.C. in a single market may not be very reasonable for a number of industries -- such as 2051 (bread, cakes and related products), 2082 (malt beverages) and 2084 (wines, brandy, and brandy spirits) -- where individual E.C. countries have long histories of special tastes for these products. However, the E.C. is moving to develop true integration among these markets.

Furthermore, some of the E.C./U.S. comparisons may not be very meaningful due to the limit of the size of the firm included in the sample. In a number of cases the $150 million limit means that a large proportion of firms (especially for the E.C.) were excluded. This will result in an over-statement of the concentration. This may well be the reason for the high relative concentration of SIC=2082 (malt beverages) and SIC=2064 (candy and other confection products) of the E.C. over the equivalent U.S. data. The malt beverage concentration may reflect the presence of only the large UK brewers which dominate the market as constructed because the smaller firms in the German market are not included. Careful attention should be paid to many of the comparisons made here.

However, under the objective to study the potential for U.S. firms competing abroad and for E.C. firms competing in the U.S., the limitation to only large firms may prove to be very useful. If an argument can be made that scale economy is needed to consider competition in foreign markets, then limiting the analysis to large firms may be reasonable. However, the argument that concentration translates into potential ability to compete abroad may not be a viable argument, especially in light of the highly concentrated U.S. car market and the relatively low propensity for U.S. food producers to export (see Handy and Henderson 1992).

Future directions for this research include the verification of these results using simulated data for smaller firms that would be sampled under the $150 million sales level. Another future topic would be to differentiate the diversity measures to account for upward and downward vertical integration as well as other horizontal integration by region. Furthermore, the simulations used here could be extended to include simulations of data used in a second level econometric analysis. This could involve the use of the simulated data along with other information in regression analysis. A first step in this direction was the interregional correlation of the concentration measures.

References

Bain, Joe S., 1966, *International Differences in Industrial Structure*, Yale University Press, New Haven.

Berry, Charles H., 1971, "Corporate Growth and Diversification," *Journal of Law and Economics*, October, 371-384.

Clarke, R. and S. W. Davies, 1983, "Aggregate Concentration, Market Concentration and Diversification", *The Economic Journal*, 182-192.

Dagpunar, John, 1988, *Principles of Random Variate Generation*, Clarendon Press, Oxford, UK.

Dun & Bradstreet International, Dun's Marketing Services, Parsippany, NJ. Dun & Bradstreet, 299 Park Ave., New York, NY.

Handy, Charles R. and Dennis R. Henderson, 1992, "Foreign Investment in Food Manufacturing", OP-41, Organization and Performance of World Food Systems: NC-194, Department of Agricultural Economics, Ohio State University.

Hirschberg, Joseph G., James Dayton and Peter Voros, 1992, "Firm Level Data: A Compendium of International Data Sources for the Food Processing Industries", OP-34, Organization and Performance of World Food Systems: NC-194, Department of Agricultural Economics, Ohio State University.

Herfindahl, Orris C., 1950, "Concentration in the U.S. Steel Industry", Unpublished doctoral dissertation, Columbia University.

Hirschman, Albert O., 1945, *National Power and the Structure of Foreign Trade*, Berkeley.

Kullback, S., 1959, *Information Theory and Statistics*, John Wiley, New York.

Quandt, Richard E., 1966, "On the Size Distribution of Firms", *American Economic Review*, 56, 416-432.

Read, T. R. C. and N. A. C. Cressie, 1988, *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer-Verlag, New York.

Silberman, Irwin H., 1967, "On Lognormality as a Summary Measure of Concentration", *American Economic Review*, 57, 807-31.

Silverman, Bernard W., 1986, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York, NY.

Sutton, John, 1991, *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, MIT Press, Cambridge, MA.

Trinet Line of Business Database, Trinet Inc., Parsippany, NJ. Trinet Inc., (Contacts Influential Inc.), Parsippany, NJ.

U. S. Bureau of the Census, 1992, *Concentration Ratios In Manufacturing*, Census of Manufacturing (1987), Subject Series, MC87-S-6, GPO, Washington. DC.

Utton, M. A., 1977, "Large Firm Diversification in British Manufacturing Industry", *The Economic Journal*, 87, 96-113.

Who Owns Whom, Dun's Market Services, Parsippany, NJ.

Yamawaki, Hideki, Leo Sleuwaegen and Leonard W. Weiss, 1989, "Industry Competition and the Formation of the European Common Market", in *Concentration and Price*, ed. by Leonard W. Weiss, MIT Press, Cambridge MA.

Appendix

The following are fabricated examples of the data Collected from the Dun & Bradstreet computer data base. A U.S. firm would appear as:

```
BIG FOOD COMPANY
ONE E DESOTO STREET
CHICAGO, UNITED STATES

TELEPHONE: 3125554000
STATE/PROVINCE: IL

BUSINESS:  CANNED FRTS,VGTBLS

PRIMARY SIC:   2033  CANNED FRUITS AND SPECIALTIES
SECONDARY SIC: 2079  EDIBLE FATS AND OILS
SECONDARY SIC: 2015  POULTRY SLAUGHTERING AND PROCESSING
SECONDARY SIC: 2013  SAUSAGES AND OTHER PREPARED MEATS
SECONDARY SIC: 2022  CHEESE; NATURAL AND PROCESSED
SECONDARY SIC: 2099  FOOD PREPARATIONS, NEC

YEAR STARTED:                1990

EMPLOYEES TOTAL:              16,900
SALES (LOCAL CURRENCY):       4,560,000,000
SALES (U.S. CURRENCY):        4,560,000,000
THIS IS:
    A SUBSIDIARY

DUNS NUMBER:           14-468-2555

PARENT NAME:           BIGGER INC
PARENT DUNS:           00-527-9000
PARENT CITY:           LOS ANGELES
PARENT STATE/PROVINCE: CA
PARENT COUNTRY:        UNITED STATES

        Copyright 1991 Dun & Bradstreet, Inc.
```

A firm in the European Community might appear as:

```
GROSSE-BRAUEREI
JOSEFSBERGSTR 25
XBURG, GERMANY FED REP OF

TELEPHONE: 5555 1111
TELEX: 4444444
STATE/PROVINCE: SAARLAND

BUSINESS:  MALT BEVERAGES

PRIMARY SIC:   2082  MALT BEVERAGES
SECONDARY SIC: 2086  BOTTLED AND CANNED SOFT DRINKS
SECONDARY SIC: 2024  ICE CREAM AND FROZEN DESERTS
SECONDARY SIC: 5181  BEER AND ALE

EMPLOYEES TOTAL:              1,855
SALES (LOCAL CURRENCY):       604,800,000
SALES (U.S. CURRENCY):        345,000,000

THIS IS:
    IN THE EUROPEAN ECONOMIC COMMUNITY (EEC)
DUNS NUMBER:           31-555-5555

CHIEF EXECUTIVE:       KARL SCMIDT,KOMPLEMENTAR

        Copyright 1991 Dun & Bradstreet, Inc.
```

Table 1. Firms by country of sales.

Country	Number of firms	Sales in Millions of $ US	Number of Employees
UNITED STATES	600	617004	3323686
UK	136	173853	1424119
JAPAN	111	100645	203048
NETHERLANDS	76	76354	111972
AUSTRALIA	53	49243	107359
FRANCE	94	45020	218136
WEST GERMANY	104	43456	121434
NEW ZEALAND	17	26637	42984
SWITZERLAND	6	22376	87010
SWEDEN	7	19533	104486
ITALY	54	18378	42516
CANADA	11	11322	71333
SPAIN	37	11095	48677
DENMARK	17	7429	34283
IRELAND	15	6316	21352
HONG KONG	2	5870	27500
BELGIUM	18	5430	16011
FINLAND	11	5113	19590
AUSTRIA	4	5111	19492
KOREA, REP OF	13	4459	38325
ISRAEL	6	4384	4110
MEXICO	5	1052	13057
SINGAPORE	2	492	4270
PORTUGAL	2	357	4740
NORWAY	1	208	1890
GREECE	1	170	2250
TOTAL	1403	1261307	6113630

Table 2. Parent firms by country.

Parent Country	Number of Parent Firms	Sales in Millions of $ US	Number of Employees
UNITED STATES	357	549410	2977714
UK	78	178298	1249673
NETHERLANDS	42	129125	449981
JAPAN	100	97497	213698
SWITZERLAND	6	55779	252519
AUSTRALIA	29	54069	102803
FRANCE	58	42124	214208
WEST GERMANY	82	36308	113090
SWEDEN	6	19533	104486
ITALY	39	16165	30729
CANADA	13	15919	147517
NEW ZEALAND	13	15464	35644
DENMARK	14	7683	36239
SPAIN	25	7134	32110
IRELAND	9	6539	16392
HONG KONG	2	6070	28600
FINLAND	11	5379	22790
AUSTRIA	4	5111	19492
ISRAEL	6	4384	4110
KOREA, REP OF	11	3981	29705
BELGIUM	10	3293	9383
MEXICO	4	815	9597
SINGAPORE	2	492	4270
PORTUGAL	2	357	4740
NORWAY	1	208	1890
GREECE	1	170	2250
TOTAL	925	1261307	6113630

Table 3. The Average weights from a Uniform distribution.

Number of SICs (k)	r_1	r_2	r_3	r_4	r_5	r_6
2	.697	.303				
3	.522	.321	.157			
4	.418	.300	.190	.092		
5	.348	.270	.194	.126	.062	
6	.296	.239	.188	.139	.092	.045

Table 4. The Average weights from a Lognormal Distribution.

Number of SICs (k)	r_1	r_2	r_3	r_4	r_5	r_6
2	.719	.281				
3	.596	.273	.131			
4	.520	.256	.147	.078		
5	.461	.239	.150	.096	.053	
6	.424	.225	.145	.010	.067	.039

Table 5. The Average weights from a Pareto Distribution (c = .9124).

Number of SICs (k)	r_1	r_2	r_3	r_4	r_5	r_6
2	.704	.296				
3	.614	.237	.149			
4	.564	.213	.129	.094		
5	.506	.201	.127	.094	.072	
6	.509	.183	.113	.081	.063	.051

Table 6. The Average weights from a Multivariate Estimated Kernel Distribution.

Number of SICs (k)	r_1	r_2	r_3	r_4	r_5	r_6
2	.665	.335				
3	.571	.281	.148			
4	.549	.252	.131	.068		
5	.533	.235	.125	.070	.038	
6	.517	.235	.127	.067	.035	.018

Table 7. The Average weights as Estimated from the Trinet data for U.S. Food Processing Firms.

Number of SICs (k)	r_1	r_2	r_3	r_4	r_5	r_6
2	.706	.294				
3	.592	.267	.141			
4	.540	.245	.136	.079		
5	.510	.228	.130	.079	.053	
6	.478	.222	.127	.081	.055	.038

Table 8. The Counts of cases where 90 percent or more of the experiments result in interregional differences of market concentration (DH_i) of the same sign.

SIC	Description	me*	ne	# of US firms	# of EC firms
2011	MEAT PACKING PLANTS	-3, -3	-4, +1	42	47
2013	SAUSAGES AND OTHER PREPARED MEATS	-5, +5	-4, +3	30	40
2015	POULTRY SLAUGHTERING AND PROCESSING	-3, -3	-4, +1	37	14
2021	CREAMERY BUTTER	+1, -1	+2, -1	19	44
2022	CHEESE; NATURAL AND PROCESSED	+1, +1	+1, +1	34	41
2023	DRY, CONDENSED, EVAPORATED PRODUCTS	+1, +1	+1, +1	24	32
2024	ICE CREAM AND FROZEN DESERTS	-5, -5	-4, +1	38	10
2026	FLUID MILK	+5, +5	+5, +5	59	63
2032	CANNED SPECIALTIES	+1, +1	+1, +1	15	12
2033	CANNED FRUITS AND SPECIALTIES	----	+1, +1	42	19
2034	DEHYDRATED FRUITS, VEGETABLES, SOUP	----	----	15	9
2035	PICKLES, SAUCES, AND SALAD DRESSING	+1, -1	----	20	11
2037	FROZEN FRUITS AND VEGETABLES	-2, -2	-5, +1	35	6
2038	FROZEN SPECIALTIES, NEC	-5, -5	-5, -5	25	16
2041	FLOUR AND OTHER GRAIN MILL PRODUCTS	----	----	22	22
2043	CEREAL BREAKFAST FOODS	-4, -4	-3, -3	10	7
2044	RICE MILLING	-1, +1	-3, +3	6	2
2045	PREPARED FLOUR MIXES AND DOUGHS	-5, -5	-5, -5	9	8
2046	WET CORN MILLING	-3, -3	-4, -4	10	5
2047	DOG AND CAT FOOD	----	----	15	15
2048	PREPARED FEEDS, NEC	+1, +1	+1, +1	40	50
2051	BREAD, CAKE, AND RELATED PRODUCTS	-4, -4	-5, +1	23	12
2052	COOKIES AND CRACKERS	-5, -5	-5, -5	19	13
2053	FROZEN BAKERY PRODUCTS, EXPT BREAD	----	----	6	0
2061	RAW CANE SUGAR	----	----	1	5
2062	CANE SUGAR REFINING	+5, +5	+5, +5	7	8
2063	BEET SUGAR	+5, +5	+5, -1	7	18
2064	CANDY AND OTHER CONFECTION PRODUCT	+5, +5	+5, +5	21	23
2066	CHOCOLATE AND COCOA PRODUCTS	----	+1, -1	8	15
2067	CHEWING GUM	----	----	5	0
2068	SALTED AND ROASTED NUTS AND SEEDS	----	----	2	0
2074	COTTONSEED OIL MILLS	----	----	2	0
2075	SOYBEAN OIL MILLS	-4, -4	-5, -5	13	7
2076	VEGETABLE OIL MILLS, NEC	----	+5, -1	3	9
2077	ANIMAL AND MARINE FATS AND OILS	+5, +5	+4, +4	5	10
2079	EDIBLE FATS AND OILS	+5, -5	+3, -3	17	13
2082	MALT BEVERAGES	+1, -1	+5, -1	6	33
2083	MALT	----	+4, -1	2	8
2084	WINES, BRANDY, AND BRANDY SPIRITS	+4, -4	+4, -3	7	21
2085	DISTILLED AND BLENDED LIQUORS	+3, +3	+2, +2	12	18
2086	BOTTLED AND CANNED SOFT DRINKS	-5, -5	-5, -5	31	27
2087	FLAVORING EXTRACTS AND SYRUPS, NEC	-5, -5	-5, -5	22	14
2091	CANNED AND CURED FISH AND SEAFOODS	-1, -1	-5, -4	10	3
2092	FRESH OR FROZEN PACKAGED FISH	----	----	15	5
2095	ROASTED COFFEE	-5, -5	-5, -5	10	9
2096	POTATO CHIPS AND SIMILAR SNACKS	----	----	10	0
2097	MANUFACTURED ICE	----	----	0	2
2098	MACARONI AND SPAGHETTI	----	----	4	8
2099	FOOD PREPARATIONS, NEC	-3, +3	-1, +1	42	53

*The first number is the number of indices based on sales and the second number is the index based on employment. A plus sign indicates that the E.C. industry is more concentrated than the U.S. industry and a minus sign indicates that the U.S. industry is more concentrated than the E.C. industry.

Table 9. Concentration experiment results for the lognormal distribution using the means equivalence.

SIC	Description	HHI US (sales)	HHI EC (sales)	Level of Signif.	Est N of Firms US (sales)	Est N of Firms EC (sales)	Num of US Firms	Num of Firms EC	US Sales (Mill $)	EC Sales (Mill $)
2011	MEAT PACKING PLANTS	1186.42	390.44	0.990	7.2343	24.1289	42	47	41693.60	12926.61
2013	SAUSAGES AND OTHER PREPARED MEATS	791.98	2917.34	0.990	8.1662	2.9399	30	40	12561.57	17806.62
2015	POULTRY SLAUGHTERING AND PROCESSING	668.22	375.93	0.990	11.3023	7.8348	37	14	13550.15	1325.30
2021	CREAMERY BUTTER	743.30	643.14	.	5.9162	13.9374	19	44	2576.94	10734.47
2022	CHEESE; NATURAL AND PROCESSED	647.07	588.04	.	11.0147	14.0830	34	41	10849.12	6320.68
2023	DRY, CONDENSED, EVAPORATED PRODUCTS	1254.43	1455.70	.	4.4552	4.7981	24	32	9607.94	8420.56
2024	ICE CREAM AND FROZEN DESERTS	875.92	392.14	0.990	9.1891	5.5220	38	10	11125.91	3047.41
2026	FLUID MILK	441.08	625.37	0.990	26.8407	19.9964	59	63	18263.70	22372.91
2032	CANNED SPECIALTIES	521.83	613.93	.	6.2749	4.5356	15	12	7705.43	3571.79
2033	CANNED FRUITS AND SPECIALTIES	1081.28	1205.17	.	8.3159	3.7159	42	19	24057.54	8185.58
2034	DEHYDRATED FRUITS, VEGETABLES, SOUP	595.02	582.24	.	5.7837	3.6482	15	9	5396.37	5445.94
2035	PICKLES, SAUCES, AND SALAD DRESSING	818.19	496.18	0.950	5.5681	5.1235	20	11	6524.89	2616.08
2037	FROZEN FRUITS AND VEGETABLES	535.40	410.58	0.950	13.3282	3.4188	35	6	7069.89	1620.73
2038	FROZEN SPECIALTIES, NEC	979.38	347.76	0.990	5.6226	9.5636	25	16	8857.75	2134.62
2041	FLOUR AND OTHER GRAIN MILL PRODUCTS	813.26	498.53	0.950	6.1657	9.1765	22	22	8819.09	6826.99
2043	CEREAL BREAKFAST FOODS	406.08	329.83	0.900	5.4109	4.6225	10	7	8739.95	2406.52
2044	RICE MILLING	245.58	718.86	0.990	5.0512	1.3102	6	2	1733.76	194.47
2045	PREPARED FLOUR MIXES AND DOUGHS	963.73	513.35	0.950	2.6064	3.8148	9	8	2475.82	2072.20
2046	WET CORN MILLING	756.82	334.12	0.990	3.3887	3.3877	10	5	8675.49	1091.39
2047	DOG AND CAT FOOD	480.65	502.40	.	6.6977	6.4911	15	15	4079.42	3358.53
2048	PREPARED FEEDS, NEC	1187.94	1898.13	.	7.2563	6.6083	40	50	10246.06	19147.64
2051	BREAD, CAKE, AND RELATED PRODUCTS	697.29	497.71	0.950	7.0831	5.3732	23	12	11092.42	5220.11
2052	COOKIES AND CRACKERS	1075.70	335.90	0.990	4.1847	8.1429	19	13	8896.47	3773.94
2053	FROZEN BAKERY PRODUCTS, EXPT BREAD	558.81	.	0.990	2.7901	.	6	.	2292.32	.
2061	RAW CANE SUGAR	.	631.95	0.990	.	2.2668	.	5	.	2281.49
2062	CANE SUGAR REFINING	311.33	689.27	0.990	4.8480	2.9828	7	8	2390.55	4792.99
2063	BEET SUGAR	229.28	363.65	0.990	6.2179	10.1968	7	18	3121.63	6921.76
2064	CANDY AND OTHER CONFECTION PRODUCT	481.34	1078.66	0.990	9.2914	4.7389	21	23	3845.64	9330.75
2066	CHOCOLATE AND COCOA PRODUCTS	700.57	565.27	0.900	2.9322	5.8016	8	15	3423.92	6044.02
2067	CHEWING GUM	246.89	.	0.990	4.2598	.	5	.	1960.71	.
2068	SALTED AND ROASTED NUTS AND SEEDS	1988.53	.	0.990	1.0986	.	2	.	1494.87	.
2074	COTTONSEED OIL MILLS	10000.00	.	0.990	1.0000	.	2	.	1718.13	.
2075	SOYBEAN OIL MILLS	1053.04	220.53	0.990	3.1818	6.4284	13	7	12983.43	1216.09
2076	VEGETABLE OIL MILLS, NEC	381.38	425.66	.	2.1461	4.7018	3	9	785.39	2991.78
2077	ANIMAL AND MARINE FATS AND OILS	302.62	507.22	0.990	3.6512	4.5838	5	10	358.96	2023.55
2079	EDIBLE FATS AND OILS	840.39	630.02	0.950	4.6186	4.6640	17	13	9026.01	3914.05
2082	MALT BEVERAGES	518.54	488.52	.	2.8669	13.7398	6	33	14911.26	20786.31
2083	MALT	499.95	502.35	.	1.4638	3.8909	2	8	948.75	1586.93
2084	WINES, BRANDY, AND BRANDY SPIRITS	592.16	372.59	0.975	3.0299	11.7079	7	21	1181.02	5933.45
2085	DISTILLED AND BLENDED LIQUORS	451.65	572.43	.	5.9189	6.7713	12	18	4681.99	9809.87
2086	BOTTLED AND CANNED SOFT DRINKS	1014.35	370.58	0.990	6.4484	14.8173	31	27	16030.70	6275.97
2087	FLAVORING EXTRACTS AND SYRUPS, NEC	1816.46	527.31	0.990	3.0032	5.8548	22	14	11230.86	2134.34
2091	CANNED AND CURED FISH AND SEAFOODS	558.61	358.87	0.900	4.1843	2.2118	10	3	2414.52	842.61
2092	FRESH OR FROZEN PACKAGED FISH	411.36	430.47	.	7.8982	2.8792	15	5	2004.35	585.40
2095	ROASTED COFFEE	1446.44	458.02	0.975	2.3106	4.4443	10	9	4099.67	5913.36
2096	POTATO CHIPS AND SIMILAR SNACKS	554.68	.	0.990	4.2229	.	10	.	5654.24	.
2097	MANUFACTURED ICE	.	362.52	0.990	.	1.7263	.	2	.	400.74
2098	MACARONI AND SPAGHETTI	394.29	560.84	.	2.5880	3.4789	4	8	757.81	706.45
2099	FOOD PREPARATIONS, NEC	1204.60	1490.59	0.900	7.2935	7.0699	42	53	18760.52	23626.78

Table 10. The Counts of cases where 90 percent or more of the experiments result in interregional differences of market diversity (DB_i) of the same sign.

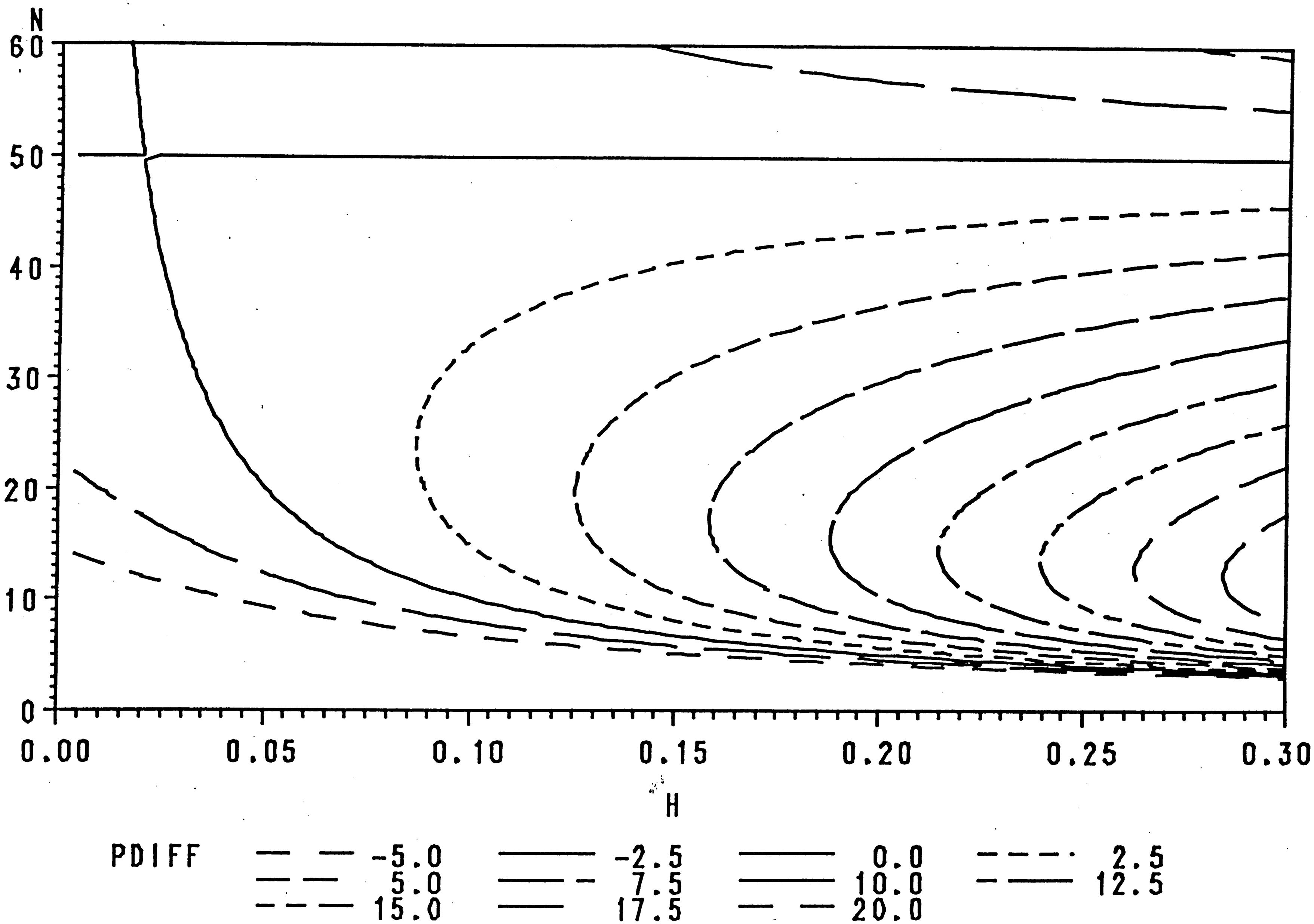
SIC	Description	me*	ne	# of US firms	# of EC firms
2011	MEAT PACKING PLANTS	-1, +1	-1, +1	208/42	198/47
2013	SAUSAGES AND OTHER PREPARED MEATS	-2, -2	-3, -3	224/30	285/40
2015	POULTRY SLAUGHTERING AND PROCESSING	-1, +1	-4, +1	228/37	134/14
2021	CREAMERY BUTTER	-2, -2	-4, -4	88/19	339/44
2022	CHEESE; NATURAL AND PROCESSED	-1, -1	-1, -1	209/34	294/41
2023	DRY, CONDENSED, EVAPORATED PRODUCTS	----	----	160/24	217/32
2024	ICE CREAM AND FROZEN DESERTS	-2, -2	-1, -1	294/38	100/10
2026	FLUID MILK	----	----	307/59	301/63
2032	CANNED SPECIALTIES	-5, -5	-5, -5	173/15	116/12
2033	CANNED FRUITS AND SPECIALTIES	-3, -3	-3, -3	355/42	131/19
2034	DEHYDRATED FRUITS, VEGETABLES, SOUP	+2, +2	+2, +2	105/15	103/9
2035	PICKLES, SAUCES, AND SALAD DRESSING	-4, +4	-4, +4	200/20	135/11
2037	FROZEN FRUITS AND VEGETABLES	+5, +5	+5, +5	256/35	35/6
2038	FROZEN SPECIALTIES, NEC	-5, -5	-5, -5	257/25	104/16
2041	FLOUR AND OTHER GRAIN MILL PRODUCTS	----	----	213/22	228/22
2043	CEREAL BREAKFAST FOODS	+1, -1	+1, -1	145/10	114/7
2044	RICE MILLING	+1, +1	+1, +1	39/6	27/2
2045	PREPARED FLOUR MIXES AND DOUGHS	----	----	80/9	58/8
2046	WET CORN MILLING	+5, +5	+5, +5	104/10	15/5
2047	DOG AND CAT FOOD	-3, -3	-3, -3	120/15	76/15
2048	PREPARED FEEDS, NEC	-4, -4	-4, -4	238/40	240/50
2051	BREAD, CAKE, AND RELATED PRODUCTS	+2, +2	----	159/23	109/12
2052	COOKIES AND CRACKERS	+2, +2	+2, +2	125/19	112/13
2053	FROZEN BAKERY PRODUCTS, EXPT BREAD	----	----	62/6	0
2061	RAW CANE SUGAR	-2, -2	-3, -3	6/1	33/5
2062	CANE SUGAR REFINING	-2, +2	-2, +2	26/7	35/8
2063	BEET SUGAR	-1, +1	-1, +1	38/7	66/18
2064	CANDY AND OTHER CONFECTION PRODUCT	-1, -1	-2, -2	168/21	181/23
2066	CHOCOLATE AND COCOA PRODUCTS	-2, -2	-2, -2	91/8	139/15
2067	CHEWING GUM	----	----	26/5	0
2068	SALTED AND ROASTED NUTS AND SEEDS	----	----	15/2	0
2074	COTTONSEED OIL MILLS	----	----	11/2	0
2075	SOYBEAN OIL MILLS	-2, -2	-2, -2	113/13	48/7
2076	VEGETABLE OIL MILLS, NEC	-2, -2	-3, -3	34/3	76/9
2077	ANIMAL AND MARINE FATS AND OILS	-1, +1	-1, +3	24/5	122/10
2079	EDIBLE FATS AND OILS	-2, -2	-1, -1	155/17	116/13
2082	MALT BEVERAGES	----	-1, +1	63/6	164/33
2083	MALT	----	-3, +3	21/2	82/8
2084	WINES, BRANDY, AND BRANDY SPIRITS	-2, -2	-2, -2	30/7	95/21
2085	DISTILLED AND BLENDED LIQUORS	-1, +1	----	89/12	107/18
2086	BOTTLED AND CANNED SOFT DRINKS	+1, +1	+1, +1	150/31	184/27
2087	FLAVORING EXTRACTS AND SYRUPS, NEC	----	----	212/22	151/14
2091	CANNED AND CURED FISH AND SEAFOODS	-5, -5	-5, -5	45/10	45/3
2092	FRESH OR FROZEN PACKAGED FISH	----	----	61/15	17/5
2095	ROASTED COFFEE	----	----	84/10	69/9
2096	POTATO CHIPS AND SIMILAR SNACKS	----	----	72/10	0
2097	MANUFACTURED ICE	----	----	0	6/2
2098	MACARONI AND SPAGHETTI	-1, -1	-1, -1	46/4	124/8
2099	FOOD PREPARATIONS, NEC	-2, -2	-2, -2	380/42	446/53

*The first number is the number of indices based on sales and the second number is the index based on employment. A plus sign indicates that the E.C. industry is has a higher degree of diversity than the U.S. industry and a minus sign that the U.S. industry is the more diverse than the E.C. industry.

Table 11. Diversity experiment results for the lognormal distribution using the means equivalence.

SIC	Description	B US (sales)	B EC (sales)	Level of Signif.	US N of Inds (sales)	EC N of Inds (sales)	Num of US Firms	Num of EC Firms	US Sales (Mill $) all inds	EC Sales (Mill $) all inds
2011	MEAT PACKING PLANTS	0.97433	0.97580	.	208	212	42	47	146167.73	102061.22
2013	SAUSAGES AND OTHER PREPARED MEATS	0.98162	0.97944	0.990	236	309	30	40	133262.59	169969.76
2015	POULTRY SLAUGHTERING AND PROCESSING	0.97821	0.97953	.	238	148	37	14	129895.84	138194.32
2021	CREAMERY BUTTER	0.98738	0.97936	0.990	88	363	19	44	18304.63	191194.86
2022	CHEESE; NATURAL AND PROCESSED	0.97834	0.97761	.	219	318	34	41	110265.13	164137.75
2023	DRY, CONDENSED, EVAPORATED PRODUCTS	0.98442	0.98528	.	170	227	24	32	73962.19	89108.43
2024	ICE CREAM AND FROZEN DESERTS	0.98106	0.97845	0.900	317	104	38	10	174156.07	68102.97
2026	FLUID MILK	0.98434	0.98440	.	320	311	59	63	114280.33	105292.19
2032	CANNED SPECIALTIES	0.98586	0.98055	0.990	190	130	15	12	79779.14	89284.60
2033	CANNED FRUITS AND SPECIALTIES	0.97873	0.97344	0.950	385	135	42	19	222763.12	70207.77
2034	DEHYDRATED FRUITS, VEGETABLES, SOUP	0.97753	0.97932	0.990	109	107	15	9	71547.62	72340.73
2035	PICKLES, SAUCES, AND SALAD DRESSING	0.98015	0.98149	0.950	220	149	20	11	116744.11	113344.94
2037	FROZEN FRUITS AND VEGETABLES	0.98023	0.99185	0.990	280	35	35	6	117867.32	9902.77
2038	FROZEN SPECIALTIES, NEC	0.98311	0.96547	0.990	289	108	25	16	158979.62	52558.93
2041	FLOUR AND OTHER GRAIN MILL PRODUCTS	0.97590	0.97932	0.990	216	252	22	22	136534.49	163335.12
2043	CEREAL BREAKFAST FOODS	0.98614	0.98518	0.950	172	138	10	7	123004.53	95437.36
2044	RICE MILLING	0.95728	0.98412	0.975	43	27	6	2	18936.41	5560.31
2045	PREPARED FLOUR MIXES AND DOUGHS	0.96740	0.97440	.	80	58	9	8	50993.34	53428.26
2046	WET CORN MILLING	0.97785	0.99344	0.990	108	15	10	5	101729.96	3085.00
2047	DOG AND CAT FOOD	0.98813	0.98199	0.990	127	80	15	15	67273.75	17920.47
2048	PREPARED FEEDS, NEC	0.96600	0.93033	0.990	243	244	40	50	105069.02	140999.97
2051	BREAD, CAKE, AND RELATED PRODUCTS	0.98133	0.98248	.	184	119	23	12	105644.51	66613.60
2052	COOKIES AND CRACKERS	0.98471	0.98626	.	146	122	19	13	95830.20	58240.31
2053	FROZEN BAKERY PRODUCTS, EXPT BREAD	0.98702	.	0.990	66	.	6	.	29044.68	.
2061	RAW CANE SUGAR	0.99208	0.98765	.	6	33	1	5	665.83	10254.02
2062	CANE SUGAR REFINING	0.98677	0.98758	.	26	35	7	8	9214.22	12956.08
2063	BEET SUGAR	0.98891	0.98935	.	38	66	7	18	13527.87	19499.83
2064	CANDY AND OTHER CONFECTION PRODUCT	0.98396	0.98329	0.900	195	207	21	23	119650.87	115194.01
2066	CHOCOLATE AND COCOA PRODUCTS	0.98575	0.98373	0.900	105	165	8	15	55823.01	102113.75
2067	CHEWING GUM	0.98912	.	0.990	29	.	5	.	22027.04	.
2068	SALTED AND ROASTED NUTS AND SEEDS	0.98834	.	0.990	18	.	2	.	16012.81	.
2074	COTTONSEED OIL MILLS	0.98334	.	0.990	11	.	2	.	8420.34	.
2075	SOYBEAN OIL MILLS	0.97883	0.97419	0.990	113	48	13	7	96874.10	48490.32
2076	VEGETABLE OIL MILLS, NEC	0.97977	0.97790	.	34	80	3	9	28045.22	85936.34
2077	ANIMAL AND MARINE FATS AND OILS	0.97134	0.98005	0.975	24	136	5	10	4838.48	133838.13
2079	EDIBLE FATS AND OILS	0.98085	0.97920	.	168	130	17	13	130876.79	96956.69
2082	MALT BEVERAGES	0.98283	0.98782	0.990	73	164	6	33	62819.74	63687.71
2083	MALT	0.97710	0.98793	0.990	21	82	2	8	19116.42	33096.47
2084	WINES, BRANDY, AND BRANDY SPIRITS	0.98692	0.98551	.	39	107	7	21	11669.66	29205.85
2085	DISTILLED AND BLENDED LIQUORS	0.98374	0.98472	.	94	112	12	18	46167.18	52359.48
2086	BOTTLED AND CANNED SOFT DRINKS	0.98237	0.98359	.	172	202	31	27	74427.33	108959.82
2087	FLAVORING EXTRACTS AND SYRUPS, NEC	0.98349	0.98300	.	240	175	22	14	145062.67	144941.01
2091	CANNED AND CURED FISH AND SEAFOODS	0.98931	0.97551	0.990	53	49	10	3	15073.40	42467.67
2092	FRESH OR FROZEN PACKAGED FISH	0.98520	0.98745	.	75	17	15	5	20125.72	1880.79
2095	ROASTED COFFEE	0.98015	0.98679	0.975	94	79	10	9	77390.68	68326.71
2096	POTATO CHIPS AND SIMILAR SNACKS	0.98593	.	0.990	75	.	10	.	54573.22	.
2097	MANUFACTURED ICE	.	0.99271	0.990	.	6	.	2	.	2900.00
2098	MACARONI AND SPAGHETTI	0.98517	0.98407	.	46	144	4	8	11759.47	94539.56
2099	FOOD PREPARATIONS, NEC	0.98005	0.97780	.	407	479	42	53	263343.59	266320.52

Figure 1 The % difference between H^{me} and H^{ne} ($n^* = 50$).



This material is based in part on work supported by the U.S. Department of Agriculture, Cooperative State Research Service, under Agreement No. 89-34210-04238 and successor(s).

Additional information on NC-194 and a complete list of project publications can be obtained from:

Executive Director, NC-194
Department of Agricultural Economics
The Ohio State University
2120 Fyffe Road
Columbus, Ohio 43210-1099
(614)292-2194