From Homogeneity to Heterogeneity: An Index of Diversification

Working Paper No. 129

Frank M. Gollop
James L. Monahan*

October 1986

The 1901 amalgamation of 180 previously independent iron and steel firms into U.S. Steel dominated the turn-of-the-century business press. The company's successful battle against Mobil for Marathon Oil claimed 1981 business headlines. Analysts are quick to observe that the shift in U.S. Steel strategy is not atypical. Philip Morris, for example, acquired Miller Brewing in 1970, Seven-Up in 1978, and then in 1985 leveled its sights on General Foods. U.S. Steel's name change to USX in 1986 barely raised eyebrows. Conglomerate expansion has replaced horizontal growth. While pure conglomerate mergers were relatively rare in the immediate postwar era, accounting for 3.2 percent of asset acquisitions in the 1948-53 period, they now account for more than 50 percent of merger activity.[1]

Economists have just begun investigating the properties of multiple-output production. Already, the notions of scale economies, natural monopoly, and productivity have been revised.[2] The applicability of marginal-cost pricing has had to be reinterpreted.[3] The prior universal condemnation of entry barriers has been questioned.[4] In spite of these provocative theoretical advances, the relative importance of diversification across industries and over time remains a mystery. There simply are no meaningful, publicly available statistics measuring the extent of establishment (plant), enterprise, or industry diversification.

While the Bureau of the Census annually collects extensive manufacturing data, the development of its published summary statistics reflects economic theory's partial equilibrium focus on homogeneous product markets. An establishment's entire output is assigned to a particular four-digit industry on the basis of the plant's primary product. There currently exists no measure indicating the extent of enterprise operations across industry lines. The Bureau's primary product specialization ratio notwithstanding, published Census of Manufactures data are silent on the extent of diversification within and across conventional industry lines.

The primary objectives of this paper are to develop a meaningful index of diversification and then, by applying the index to establishment and enterprise data currently collected by the Census Bureau, to describe the extent of product diversification within U.S. manufacturing. The properties required of a proper diversification index are described in section 1. An index satisfying these requirements is introduced in section 3 after the properties of previously suggested indexes are discussed briefly in section 2. Trends and inter-industry differences in manufacturing establishment and enterprise diversification are presented in section 4.

One interesting conclusion is the extent to which enterprise diversification is due to increasing diversification among rather than within establishments, suggesting that technical or within-plant economies of scope offer little explanation for the recent surge in diversification. If anything, the data make clear that the trend within establishments is for production to become more rather than less homogeneous.

1. Properties of a Diversification Index

There unfortunately is no formal model of production that leads unambiguously to a unique index of diversification. Still, the need for an index remains. As a result, this paper adopts an axiomatic rather than deductive approach. Fortunately, the diversification literature exhibits a clear concensus.[5] It posits that a well-designed index of diversification should have the following five properties:

(i) it should vary directly with the number of different products produced;

(ii) it should vary inversely with the increasingly unequal distribution of products across product lines;

(iii) it should vary directly with the dissimilarity or heterogeneity of the products;

(iv) it should have scope, applying equally well to plants, firms, and industries; and

(v) if possible, it should be bounded between zero and unity.

Properties (i) through (iii) form the critical subset. Only properties (ii) and (iii) require elaboration. The distribution property concerns the composition of production. Property (ii) posits that, holding all else constant, a firm producing n products with each having a 1/n share of the firm's total sales is more diversified than another firm producing the identical n products but with one representing, say, 90 percent of firm sales.[6] Holding all else constant, weighting production activity toward a subset of primary products reduces the diversified nature of the enterprise. The firm becomes more dependent upon fewer products. The heterogeneity property addresses product identity rather than composition. Property (iii) states that a firm producing distantly related products, ceteris paribus, is more diversified than another firm producing products which share technological and/or marketing similarities.[7] Property (iii) distinguishes between what the Federal Trade Commission labels pure conglomerates and market extension conglomerates.[8]

None of the diversification measures presented in the economics literature simultaneously satisfies the critical number, distribution, and heterogeneity properties. A brief survey, however, highlights guidelines for the development of an index that does.

2. Measuring Diversification

Gort (1962) generally is credited with having made the first quantitative examination of diversification in U.S. manufacturing. The familiar primary product specialization ratio (PPSR) forms the basis for his most popularly referenced measure:

$$(1) \quad d_1 \equiv 1 - \text{PPSR} = \sum_{j=2}^{n} s_j ,$$

where s_j, typically defined at the four-digit level, is the share of the jth secondary product in establishment, firm, or industry shipments.

The PPSR is a good measure of a producing unit's primary product concentration, but d_1 does not satisfactorily fulfill any of the critical properties of a well-defined index of diversification. It is silent regarding the number of products. It accounts only for the share distribution between the primary product and the aggregate of all secondary products.[9]

Only Utton (1977) and Goto (1981) offer a heterogeneity claim for d_1, arguing that applying the index at the aggregate two or three-digit level meaningfully measures product dissimilarity. This conclusion follows, the argument continues, because this form of d_1 would register only production activity outside the broadly defined industry aggregate housing the primary four-digit product. This notion of heterogeneity, however, is bivariate. Product heterogeneity is determined wholly by whether a product's four-digit SIC code falls within or outside the boundaries of the primary three- or two-digit group. Index d_1 treats heterogeneity as a discrete rather than continuous dimension. In spite of its weaknesses, however, but because of the ready availability of PPSR data in Census publications, d_1 remains a popularly used measure of diversification.[10]

Berry (1971) and McVey (1972) apparently independently propose and apply the Herfindahl index as a measure of corporate diversification. The index has the form $\sum s_i^2$, where s_i is the share of the ith product in total sales. McVey (1972) uses the index directly; Berry (1971) suggests the form

(2) $d_2 \equiv 1 - \sum s_i^2$.

The index reveals nothing about product heterogeneity, but it does satisfy both the number and distribution properties. This can be demonstrated by adding and subtracting 1/n to (2), where n is the number of distinct four-digit products:

(3) $d_2 = (1 - \frac{1}{n}) + \sum (\frac{1}{n^2} - s_i^2)$.

The first term in (3) varies directly with the number of products, thereby satisfying the number property. The second term can be interpreted as the difference between two Herfindahl indexes, one based on actual product shares (s_i) and a second calculated on assumed equal shares (1/n). The term satisfies the distribution property. It takes a zero value when the n products are distributed equally. It takes increasingly negative values as the distribution of production becomes more unequal.

Two other indexes have similar properties. Utton (1977) constructs a "cumulative diversification curve" and introduces an index based on the area above that curve:

(4) $d_3 \equiv 2 \sum i s_i - 1$,

where i is the shipments rank of the ith product and products are ranked in descending order. Jacquemin and Berry (1979) propose the entropy index:

(5) $d_4 = \Sigma s_i \ln(1/s_i)$.

Like d_2, d_3 and d_4 are sensitive to changes in both the number and distribution of products. In fact, each index can be decomposed into number and distribution components. Unlike their Herfindahl counterpart, however, neither d_3 nor d_4 is bounded from above by a constant; $1 \leq d_3 < n$ and $0 \leq d_4 < \ln(n)$. Another distinction is notable as well. As the simple number of products increases or the distribution of products becomes more equal, the Herfindahl-based index d_2 increases at a decreasing rate, d_3 increases at a constant rate, and the entropy index d_4 increases at a decreasing rate in product number but at an increasing rate in product distribution. The characterizations of diversification by the three indexes are quite different. In the end, however, none of the three indexes satisfies the heterogeneity property.

Pomfret and Shapiro (1980) introduce a diversification measure that does attempt to address the heterogeneity requirement:

(6) $d_5 \equiv \sum_i s_j v_{ij}$,

where i and j represent primary and secondary industries, respectively, and

v_{ij} = 0 if the producing agent operates in one 4-digit industry only,

= 1 if j is in the same 3-digit industry as i,

= 2 if j is in the same 2-digit industry as i, and

= 3 if i and j are in different 2-digit industries. Except for the case when n = 1, however, the index d_5 is not necessarily sensitive to variations in either the number or share distribution of products. The index is calculated by simply combining product shares s_j around common dummy variables v_{ij} identifying SIC boundaries. Changing the number or distribution of products assigned to any v_{ij} category leaves d_5 unchanged.

No single measure appearing in the literature meets all the requirements of a meaningful index of diversification but the cumulative history provides some useful insights. First, most empirical studies applying any of the above indexes treat the output of each establishment as a homogeneous product.[11] Regardless of its product mix, all shipments are assigned the plant's primary product code. The changing internal product diversification of establishments is masked. The resulting indexes are measures of industrial diversification, not product diversification. The latter is preferred. Second, while a number of studies acknowledge the importance of accounting for product heterogeneity, the few that formally incorporate some notion of product dissimilarity universally define distance between products wholly in terms of SIC codes. The Utton and Goto applications of d_1 and the Pomfret and Shapiro index d_5 are representative. Such indexes, however, register heterogeneity only as products cross SIC borders. As a result, the measured distance between chemical and petroleum products is assumed equal to the distance between chemical products and electrical machinery. A meaningful index of diversification requires a distance metric more sensitive to technical dissimilarities among products. Third, and most important, of the basic measures described above, the Herfindahl-based index appears the most promising candidate for further generalization. Index d_2 fails the heterogeneity requirement but otherwise satisfies the remaining four properties of a well-designed index.

The Utton (d_3) and entropy (d_4) indexes also satisfy the number and distribution properties but (1) neither is bounded in the appealing zero-one interval and (2) their second derivatives do not exhibit wholly desirable properties. The second derivatives of the entropy index with respect to the number of products and the extent of equality in product distribution are negative and positive, respectively. Unless strong arguments exist to the contrary, the derivatives should be signed equivalently. As for the Utton index, both derivatives equal zero. While reasonable people may disagree as to the most appropriate functional representation of diversification, the natural inclination of economists is to posit that the marginal contribution to diversification of a change from one to two or two to three distinct products exceeds the marginal contribution of a change from producing 80 to 81 products.

The Herfindahl-based index d_2 exhibits the desired negative second derivatives in both product number and distribution. Furthermore, as demonstrated by Davies (1979), for $n \leq 100$ the

Herfindahl index places equal weight on the number and distribution components identified in (3), a desirable property given the lack of prior information regarding which characteristic (if either) should dominate.

3. An Index of Diversification

The critical issue is how to generalize the Herfindahl-based index d_2 to make it sensitive to product heterogeneity. Fortunately, the Herfindahl index already has embedded in it a limiting treatment of heterogeneity. Two products are either identical or different. If identical, output units are simply aggregated and given a common product subscript. If different, output bundles are given distinct subscripts.

Making this bivariate notion of product heterogeneity explicit is the key to properly generalizing d_2. An example is instructive. Consider an enterprise that produces three different five-digit products y_i with shares s_i. The resulting value of d_2 is

$$(7) \quad d_2 = 1 - s_1^2 - s_2^2 - s_3^2$$
$$= (1 - \frac{1}{3}) - \sum (s_i^2 - \frac{1}{9}).$$

Contrast this with the scenario where y_2 and y_3, perhaps because they fall within the same four-digit product group, are treated as identical:

$$(8)\quad d_2 = 1 - s_1^2 - (s_2 + s_3)^2$$

$$= (1 - \frac{1}{3}) - \sum (s_i^2 - \frac{1}{9}) - 2s_2s_3.$$

The number and distribution components in (8) still are computed on the basis of the true number and distribution of five-digit products but, because the modeler has decided that y_2 and y_3 are to be treated as identical, the extra term $-2s_2s_3$ appears in (8). Generalizing the example to n products yields

$$(9)\quad d_2' \equiv 1 - \sum_i s_i^2 - \sum_i \sum_{k \neq i} s_i s_k z_{ik}$$

$$= (1 - \frac{1}{n}) - \sum_i (s_i^2 - \frac{1}{n^2}) - \sum_i \sum_{k \neq i} s_i s_k z_{ik} \quad ,$$

where $z_{ik} = \begin{cases} 1 \text{ if the ith and kth products are identical} \\ 0 \text{ if the ith and kth products are not identical.} \end{cases}$

As evident from (7) and (8), the last term in (9) serves as a corrective adjustment to the number and distribution components; $\sum\sum s_i s_k z_{ik}$ takes a non-zero value only when it is necessary to offset the upward bias resulting whenever two truly identical product bundles are improperly distinguished ($z_{ik}=1$). As such, $\sum\sum s_i s_k z_{ik}$ contains a primitive notion of product heterogeneity.

The problem with this formulation is that a zero value is assigned to z_{ik} whenever the ith and kth products are judged to be non-identical, without any regard to how minimal or substantial their differences may be. Note, this is perfectly acceptable given the limited "corrective" role $\sum\sum s_i s_k z_{ik}$ plays in (9). The number and share distribution of products should be sensitive only to the presence of distinct products, not to the extent of their dissimilarity. The heterogeneity property described in section 1, however, requires much more. The index d_2' must be generalized to include a continuous measure of product dissimilarity that makes an independent contribution to the final index of diversification.

The continuous variable σ_{ik} provides just that mechanism:

(10) $d_2'' \equiv 1 - \sum_i s_i^2 - \sum_i \sum_{k \neq i} s_i s_k (z_{ik} - \sigma_{ik})$,

where $\sigma_{ik} \equiv \left(\sum_j \left| \frac{w_{kj} - w_{ij}}{2} \right| \right)^{1/2}$ and $0 \leq \sigma_{ik} < 1$,

$w_{kj} \equiv$ input cost share of the jth input in the kth product, and

$w_{ij} \equiv$ input cost share of the jth input in the ith product.

Division by 2 ensures that σ_{ik} is bounded in the zero-one interval. The importance of the one-half power will be made clear below.

Stated formally, the input cost share sets $\{w_{kj}\}$ and $\{w_{ij}\}$ can be interpreted as estimates of the parameters $\{\beta_{kj}\}$ and $\{\beta_{ij}\}$ defining the Cobb-Douglas technologies $y_k = \Pi x_{kj}^{\beta_{kj}}$ and $y_i = \Pi x_{ij}^{\beta_{ij}}$ for the kth and ith products, respectively. Assuming competitive factor markets,

$$(11)\quad \beta_{\ell j} = \frac{\partial \ln y_\ell}{\partial \ell n x_{\ell j}} = \frac{p_{\ell j} x_{\ell j}}{\Sigma p_{\ell j} x_{\ell j}} \equiv w_{\ell j}\,, \qquad (\ell = k, i)$$

where $p_{\ell j}$ is the price of the jth input. The distance function σ_{ik} quantifies the difference between the technologies described by $\{\beta_{kj}\}$ and $\{\beta_{ij}\}$.

The definition of σ_{ik} is based on the simple premise that two products using the same inputs in nearly identical proportions are likely to be less heterogeneous than two products having significantly dissimilar input requirements. Holding all else constant, a well-designed index of diversification should identify a plant producing canned fruit and canned olives as less diversified than an establishment producing canned fruit and the cans themselves. Given the role of σ_{ik} in (10), the measured index d_2'' would be higher for the latter plant.

Three characteristics of the heterogeneity term $\sum\sum s_i s_k (z_{ik} - \sigma_{ik})$ are worth noting. First, the notion of product distance represented by the difference $|w_{kj} - w_{ij}|$ is consistent with the basic distance function underlying the Herfindahl distribution term $\sum(s_i^2 - 1/n^2)$. The latter is a simplified form of $(1/2n)\sum\sum(s_i - s_k)^2$. The distribution term therefore defines product distance in terms of output shares $(s_i - s_k)$; the heterogeneity term defines product distance in terms of input shares $(w_{kj} - w_{ij})$. Second, $\sum\sum s_i s_k(z_{ik} - \sigma_{ik})$ measures the heterogeneity of the firm's whole product mix and can be rewritten in a form revealing the heterogeneity contribution arising through each product:

$$(12) \sum_i \sum_{k \neq i} s_i s_k (z_{ik} - \sigma_{ik}) = \sum_i s_i \sigma_i ,$$

where $\sigma_i \equiv \sum_{k \neq i} s_k (z_{ik} - \sigma_{ik})$ represents the product heterogeneity specific to product i and $s_i \sigma_i$ identifies the contribution of that product's heterogeneity to overall firm diversification. Third, note that the product shares s_k play an important role in the definition of σ_i and therefore $\sum s_i \sigma_i$. The extent of product heterogeneity arising through σ_1 in, say, a three-product firm where $s_1 = .5$, s_2 or $s_3 = .49$, $\sigma_{12} = 1.0$, and $\sigma_{13} = .1$ depends on whether s_2 or s_3 has the .49 share. Intuition suggests that the measured heterogeneity of the firm's product mix should be higher (lower) if s_2 equals .49 (.01). Heterogeneity should be higher (lower) when it is the two dominant products that are very dissimilar (similar). The variables s_k ensure this result.

It also is important to emphasize that the formulation of heterogeneity in (10) maintains that product distance is based wholly on technical considerations. Differences in distribution systems and/or marketing patterns are ignored. This is consistent with the historical focus of the Census of Manufactures on manufacturing activity at the production site.

Heterogeneity thus depends importantly on both z_{ik} and σ_{ik}. The discrete variable z_{ik} continues to play its corrective role, ensuring that the index of diversification is not upward biased by any disaggregation of a homogeneous output bundle into artificially distinct product groups. The continuous variable σ_{ik}, in contrast, captures the extent of dissimilarity between any two products for which $z_{ik} = 0$. By definition, $\sigma_{ik} = 0$ whenever $z_{ik} = 1$. For $z_{ik} = 0$, however, σ_{ik} takes some positive value in the zero-one interval, approaching unity as the products become increasingly dissimilar.

It follows that d_2'' is bounded in the zero-two interval. Dividing (10) by 2 bounds d_2'' in the desirable zero-one interval and produces the final form of the index:

$$(13)\quad D \equiv ½[1 - \sum_i s_i^2 - \sum_i \sum_{k \neq i} s_i s_k (z_{ik} - \sigma_{ik})].$$

The index takes a zero value for any entity that produces only a single product; $D = 0$ when $s_i = 1$ and z_{ik} and σ_{ik} are undefined. In contrast, D approaches unity as a producing unit approaches a state of "perfect diversification;" lim $D = 1$ as $s_i^2 \to 0$ and $\sigma_{ik} \to 1$ for all i,k.

One nice property of the index is that it can be decomposed into unique number, distribution, and heterogeneity components:

$$(14)\quad D = ½[(1 - \frac{1}{n}) + \sum_i (\frac{1}{n^2} - s_i^2) + \sum_i \sum_{k \neq i} s_i s_k (\sigma_{ik} - z_{ik})],$$

where $½(1 - \frac{1}{n}) \equiv$ the number component (N), $0 \leq N < .5$;

$½\sum_i (\frac{1}{n^2} - s_i^2) \equiv$ the distribution component (S),

$- ½(1 - \frac{1}{n}) < S \leq 0$; and

$$\frac{1}{2}\sum_i \sum_{k \neq i} s_i\, s_k(q_{ik} - z_{ik}) \equiv \text{the heterogeneity component (H)},$$

$$-\frac{1}{2}\left(1 - \sum_i s_i^2\right) \leq H \leq \frac{1}{2}\left(1 - \sum_i s_i^2\right).$$

The three components exhibit common behavior within their relevant ranges. The N, S, and H terms increase, respectively, with increases in the number of products, distributional equality, and product dissimilarity. Moreover, each increases at a decreasing rate. This can be demonstrated quite easily. Clearly,

$$\frac{\partial N}{\partial n} = \frac{1}{2n^2} > 0 \quad \text{and} \quad \frac{\partial^2 N}{\partial n^2} = -\frac{1}{n^3} < 0$$

since $n > 0$. The same properties hold for the distribution term. A change to a more equal product distribution can be represented by allowing a below average product share s_b ($s_b < 1/n$) to increase at the expense of an above average product share s_a ($s_a > 1/n$) holding constant all other product shares. As a result,

$$\frac{\partial S}{\partial s_b} = -s_b - s_a \frac{\partial s_a}{\partial s_b} = -s_b + s_a > 0 \quad \text{and}$$

$$\frac{\partial^2 S}{\partial s_b^2} = -1 + \frac{\partial s_a}{\partial s_b} = -2 < 0.$$

Finally, consider the heterogeneity term. The sum of differences in input requirements defined in (10), $\sum_j |x_{kj} - x_{ij}|/2$, is central to the notion of product heterogeneity. Consequently, an increase in H can be represented by an increase in this sum θ for any product pair (i,k) while, for convenience, holding constant the similar sums for all other product pairs:

$$\frac{\partial H}{\partial \theta_{ik}} = \frac{1}{4} s_i s_k \theta_{ik}^{-1/2} > 0 \text{ and } \frac{\partial^2 H}{\partial \theta_{ik}^2} = -\frac{1}{8} s_i s_k \theta_{ik}^{-3/2} < 0.$$

The importance of the one-half power in the definition of σ_{ik} in (10) becomes clear.

The important conclusion is that all three components defining D share common curvature properties, ensuring that the final index exhibits the usual marginal rate of substitution properties familiar to economists. The even more important conclusion is that the index D satisfies the five properties required of a meaningful index of diversification. The remaining task is to structure its application to Census of Manufactures data.

4. Diversification in U.S. Manufacturing

The index D is applied to the raw five-digit product data reported to the Census Bureau by each manufacturing establishment. Consequently, some simplifying restrictions can be imposed on D. In particular, it is reasonable to assume that products identified by a reporting plant as distinct five-digit products satisfy some minimum, however minimal, notion of product dissimilarity. It follows that all z_{ik} across five-digit product pairs take zero values. Symmetrically, it is reasonable to assume that all output units grouped by a responding establishment within a common five-digit code are sufficiently similar so as to be aggregated directly as in (10), thereby making unnecessary all corrective adjustments $-2s_i s_k z_{ik}$ where z_{ik} would take a unit value. In short, the length of the product vector reported by establishments is accepted, thereby reducing (13) to

$$(15)\ D = 1/2\,[1 - \sum_i s_i^2 + \sum_i \sum_{i \neq k} s_i s_k \sigma_{ik}].$$

An index of diversification is constructed for each establishment and each enterprise surveyed by the Census of Manufactures in 1963, 1967, 1972, 1977, and 1982. An establishment is defined as a manufacturing unit producing goods under one ownership and located at a single physical location. An enterprise is an aggregate of all establishments having common ownership. All necessary data are taken from the Census of Manufactures full establishment data file.

A complete description of the data base, variable construction, and detailed data sources is presented in Monahan (1984). A brief summary is provided here, focusing attention on the output share (s_i) and input cost share (w_{ij}) data required by the index formula (15).

Outputs are distinguished at the five-digit product class level using 1972 SIC definitions, a classification system differentiating more than 1300 five-digit product classes. The product share variables s_i specific to each establishment and enterprise are measured using value of shipments data stated in current prices.

The definition of product heterogeneity requires a vector of input cost shares (w_{ij}) for each five-digit product. Ten input categories in the Census establishment files are distinguished: production workers, other labor, fuel, electricity, purchased services, agricultural materials, mineral inputs, nondurable materials, durable materials, and capital. The representative input mix for each five-digit product is defined as the unweighted mean of input cost shares in those establishments manufacturing only that product.[12] This procedure guarantees product-specific, Cobb-Douglas input requirement vectors.

Diversification indexes are constructed for each establishment and each enterprise and then are assigned to their appropriate two-digit industries.[13] Unweighted and shipments-weighted mean diversification indexes are calculated for each industry aggregate.

Table 1 reports weighted and unweighted average diversification indexes for establishments within each two-digit industry. Consider first the weighted indexes, the preferred measure of diversification within an industry. The 1982 indexes range from a low of .049 in tobacco to .428 in petroleum and coal products. Sixteen of twenty industries have 1982 indexes below .206, the manufacturing average. Decomposition of this average into its number, distribution, and heterogeneity components yields .244, -.101, and .063, respectively. Given the definitions in (14), the components reveal that establishments in 1982

Table 1

Diversification Indexes: Establishments

	Shipments-Weighted					Unweighted				
	1963	1967	1972	1977	1982	1963	1967	1972	1977	1982
Food and Kindred Products	.242	.240	.230	.217	.192	.106	.090	.088	.079	.080
Tobacco Manufactures	.040	.052	.055	.052	.049	.024	.027	.030	.025	.025
Textile Mill Products	.180	.169	.167	.156	.163	.071	.064	.062	.053	.060
Apparel	.118	.119	.121	.125	.137	.038	.033	.034	.037	.041
Lumber and Wood Products	.147	.130	.120	.132	.141	.054	.029	.030	.040	.044
Furniture and Fixtures	.202	.196	.180	.172	.185	.097	.066	.070	.060	.069
Paper and Allied Products	.184	.194	.177	.162	.155	.067	.070	.061	.049	.053
Printing and Publishing	.250	.237	.226	.230	.201	.155	.077	.072	.074	.064
Chemicals and Allied Products	.286	.297	.289	.282	.276	.133	.118	.112	.092	.093
Petroleum and Coal Products	.434	.486	.469	.464	.428	.108	.096	.090	.082	.082
Rubber and Plastics Products	.280	.244	.195	.176	.106	.080	.072	.056	.058	.034
Leather and Leather Products	.079	.072	.142	.128	.142	.041	.035	.054	.051	.053
Stone, Clay and Glass Products	.097	.086	.081	.079	.072	.050	.032	.029	.026	.029
Primary Metal Industries	.289	.286	.283	.272	.227	.100	.090	.084	.075	.069
Fabricated Metal Products	.155	.155	.136	.131	.116	.069	.059	.057	.049	.046
Machinery, except Electrical	.223	.229	.228	.217	.191	.060	.049	.045	.039	.040
Electrical Equipment and Supplies	.233	.224	.200	.165	.143	.074	.069	.060	.043	.043
Transportation Equipment	.200	.206	.161	.163	.188	.073	.064	.052	.044	.045
Instruments and Related Products	.227	.212	.217	.186	.169	.052	.047	.048	.038	.038
Miscellaneous Manufacturing	.114	.124	.120	.113	.112	.045	.034	.028	.029	.030
All Manufacturing	.223	.224	.208	.211	.206	.081	.059	.056	.052	.051

on average produced 1.95 fairly similar products in very unequal (86:14) proportions. On the index scale of zero to one, individual manufacturing plants, with the exception of those producing petroleum and coal products and chemicals, are not very diversified.[14]

The establishment-based indexes not only are small but also have decreased over the 1963-82 period. Seventeen of twenty industries exhibit declining diversification. The three exceptions are tobacco (an increase in the number of outputs, though each new product has a very small share in production), apparel (a slight increase in the number of dissimilar products), and leather and leather products (an increase in the number of products and a move to a more heterogeneous product mix).[15]

The ranking of industries, however, has remained relatively unchanged with only two notable exceptions: rubber and plastic products and leather. The former experienced the most severe decline. Its rank fell from fourth to eighteenth as its index declined steadily from .280 in 1963 to .106 in 1982, the result of the industry's move to a more limited number of products and to a more homogeneous product mix. The latter, leather and leather products, experienced precisely the opposite trend, as its rank climbed from nineteenth in 1963 to thirteenth in 1982.

The unweighted indexes, also reported in Table 1, generally are small and declining over time. The important observation, however, is that the weighted index is larger than its unweighted counterpart in every instance, suggesting that large establishments are more diversified than small ones. The 1982 all-manufacturing weighted average index is .206; the unweighted mean is only .051. The fact that the rate of decline in the unweighted index over the 1963-82 period (-37%) far exceeds the corresponding rate for the weighted index (-8%) indicates that the decline in diversification is especially severe among the smallest establishments.

Table 2 presents mean enterprise-based indexes for each industry. It has the same format as Table 1, but there the similarity stops. The 1982 weighted indexes range from .252 in lumber to .589 in petroleum and coal products. Six of the 20 industries have average enterprise-based diversification indexes above .50. Four have indexes greater than .55. These include paper, chemicals, petroleum and coal products, and transportation equipment. The .464 all-industry average contrasts with the establishment-based average of .206.

The number, distribution, and heterogeneity components share the limelight in Table 2. These enterprise-based contributions to the all-manufacturing index .464 are .390, -.083, and .157, respectively. The number component is the dominant force (enterprises, on average, produce 4.5 different five-digit products), but the heterogeneity contribution has a magnitude equal on average to more than 40 percent of the number component. Its contribution (.157) to enterprise diversification is two and one-half times its contribution (.063) to establishment diver-

Table 2

Diversification Indexes: Enterprises

	Shipments-Weighted					Unweighted				
	1963	1967	1972	1977	1982	1963	1967	1972	1977	1982
Food and Kindred Products	.400	.436	.479	.472	.428	.103	.081	.074	.062	.063
Tobacco Manufactures	.355	.410	.464	.477	.532	.041	.049	.063	.060	.074
Textile Mill Products	.428	.434	.468	.456	.448	.070	.062	.061	.050	.059
Apparel	.182	.224	.287	.266	.290	.040	.033	.035	.037	.040
Lumber and Wood Products	.259	.248	.352	.359	.252	.053	.028	.028	.038	.041
Furniture and Fixtures	.245	.262	.318	.263	.255	.095	.062	.066	.058	.065
Paper and Allied Products	.498	.524	.539	.541	.561	.070	.069	.067	.056	.064
Printing and Publishing	.316	.313	.328	.334	.332	.156	.074	.068	.069	.059
Chemicals and Allied Products	.491	.598	.590	.584	.563	.128	.102	.095	.070	.076
Petroleum and Coal Products	.545	.610	.631	.606	.589	.101	.090	.088	.091	.093
Rubber and Plastics Products	.480	.497	.487	.444	.373	.078	.070	.055	.057	.037
Leather and Leather Products	.216	.225	.331	.308	.266	.040	.034	.044	.041	.046
Stone, Clay and Glass Products	.314	.333	.412	.412	.294	.053	.034	.032	.030	.032
Primary Metal Industries	.553	.567	.576	.592	.542	.099	.091	.085	.075	.073
Fabricated Metal Products	.296	.346	.347	.355	.276	.069	.058	.056	.050	.048
Machinery, except Electrical	.350	.398	.476	.462	.409	.056	.044	.040	.035	.037
Electrical Equipment and Supplies	.529	.509	.494	.465	.454	.071	.065	.056	.039	.042
Transportation Equipment	.529	.563	.571	.567	.561	.073	.064	.052	.045	.046
Instruments and Related Products	.420	.440	.420	.424	.426	.052	.044	.045	.036	.038
Miscellaneous Manufacturing	.232	.237	.319	.247	.294	.046	.033	.028	.029	.030
All Manufacturing	.430	.462	.486	.488	.464	.080	.055	.051	.048	.047

sification. The contribution of product distribution is significant as well. Compared to its establishment counterpart (-.101), the enterprise-based distribution term (-.083) is smaller in both absolute and relative values.[16] The presence of an unequal product distribution offsets 40 percent of the number contribution in the establishment model but only 22 percent in the enterprise model.

The enterprise diversification ranking of industries is remarkably stable over the twenty-year period. The only outliers are rubber and plastic products and tobacco. The former's ranking dropped from seventh to twelfth between 1963 and 1982. The latter's ranking moved from eleventh to fifth and today, most assuredly, is higher given the 1985 acquisitions of Nabisco and General Foods by R.J. Reynolds and Philip Morris, respectively.

One of the more interesting insights surfaces from a comparison of weighted and unweighted enterprise indexes. Quite clearly, large enterprises are more diversified than small ones. The 1982 weighted all-manufacturing enterprise index is ten times its unweighted counterpart. But consider the trends. Weighted indexes increased over the 1963-82 period in 14 of 20 industries, while unweighted indexes increased in only two instances.[17] The opposite trends in the all-manufacturing indexes in Table 2 are striking. The underlying distribution of enterprise diversification has become bimodal. Enterprise numbers and shipments have clustered in opposite tails.

As equally important insight arises from a comparison of Tables 1 and 2. While 17 of 20 industries had declining establishment-based indexes between 1963 and 1982, enterprise-based diversification increased in 14 of 20 instances. There appears to be a shift over time in enterprise activity. Rising enterprise diversification and falling establishment diversification imply that firms appear to be moving toward more homogeneous production within establishments while holding a more diverse portfolio of establishments.

Table 3 investigates further the link between establishment and enterprise diversification. Since an enterprise is defined as the sum of its establishments, an enterprise's diversification must be a function of diversification within and among its establishments. Consider adding and subtracting a shipments-weighted average of diversification indexes for an enterprise's establishments to the right-hand side of an identity equating the enterprise's diversification index with itself:

$$(16)\quad D_i^{ent} = \sum_j a_{ij} D_{ij}^{est} + (D_i^{ent} - \sum_j a_{ij} D_{ij}^{est}),$$

where "ent" and "est" represent enterprise and establishment, respectively, and a_{ij} equals the shipments share of the jth plant in the ith firm.

The first term on the right-hand side of (16) reflects the contribution of within-establishment diversification to overall D_i^{ent}. Each D_{ij}^{est} is calculated wholly in terms of the product mix and input requirements of the jth establishment. The second

Table 3

Relation Between Establishment and Enterprise Diversification

	Shipments Weighted			Unweighted		
	Enterprise Diversification	Contribution: Diversification Within Establishments	Contribution: Diversification Among Establishments	Enterprise Diversification	Contribution: Diversification Within Establishments	Contribution: Diversification Among Establishments
All Manufacturing:						
1963	.430	.038	.392	.080	.069	.011
1967	.462	.037	.425	.055	.045	.010
1972	.486	.035	.451	.051	.039	.012
1977	.488	.023	.465	.048	.036	.012
1982	.464	.014	.450	.047	.033	.014
Two-Digit Industries, 1982:						
Food and Kindred Products	.428	.018	.410	.063	.035	.028
Tobacco Manufactures	.532	.000	.532	.074	.009	.065
Textile Mill Products	.448	.013	.436	.059	.027	.032
Apparel	.290	.036	.254	.040	.027	.012
Lumber and Wood Products	.252	.049	.203	.041	.034	.007
Furniture and Fixtures	.255	.051	.204	.065	.050	.015
Paper and Allied Products	.562	.006	.556	.064	.029	.036
Printing and Publishing	.332	.045	.288	.059	.050	.009
Chemicals and Allied Products	.563	.009	.555	.076	.040	.036
Petroleum and Coal Products	.590	.004	.586	.093	.028	.065
Rubber and Plastics Products	.373	.016	.357	.037	.022	.015
Leather and Leather Products	.266	.033	.233	.046	.027	.019
Stone, Clay and Glass Products	.294	.011	.283	.033	.020	.013
Primary Metal Industries	.542	.007	.535	.073	.038	.035
Fabricated Metal Products	.276	.029	.247	.048	.033	.015
Machinery, except Electrical	.409	.016	.393	.037	.026	.011
Electrical Equipment and Supplies	.454	.007	.447	.042	.022	.019
Transportation Equipment	.561	.003	.559	.046	.028	.019
Instruments and Related Products	.426	.009	.418	.038	.022	.016
Miscellaneous Manufacturing	.294	.020	.273	.030	.023	.007

term recognizes that differences in product mix and input combinations across plants are captured in the enterprise measure D_i^{ent} but not in the individual D_{ij}^{est}. The difference between D_i^{ent} and the weighted establishment sum $\sum_j a_{ij} D_{ij}^{est}$ quantifies the contribution of diversification among an enterprise's plants.

Both terms must be greater than or equal to zero. That this is true for the first term follows from the constraints $a_{ij} \geq 0$ and $D_{ij}^{est} \geq 0$ for all i and j. The proposition also holds for the second term since $D_i^{ent} \geq \sum_j a_{ij} D_{ij}^{est}$. An enterprise cannot be less diversified than the weighted sum of its plants' diversification indexes.

Equation (16) holds for any enterprise. Summary statistics for two-digit industries can be formed by calculating shipments-weighted and unweighted versions of (16) over all enterprises within each two-digit industry. Results are reported in Table 3. As before, attention focuses on the shipments-weighted results.

The important conclusion is immediately apparent. Enterprise diversification comes from diversification among and not within establishments. Column 3 dominates column 2 in all cases. The within and among plant contributions to all-manufacturing diversification in 1982 are .014 and .450, respectively. The latter is more than 32 times the former. Within-establishment contributions in 1982 have the largest absolute and relative values in the lumber and furniture sectors, but even here the "within" contributions are only a fourth of the corresponding diversification "among" contributions. At the other extreme, enterprises in tobacco derive their diversification almost wholly from diversification among plants. The within-establishment contribution in 1982 is only .00002. The corresponding among-plant contribution is .532.

A comparison of the weighted and unweighted columns in Table 3 also reveals an interesting phenomenon. The unweighted within-establishment contribution is greater than its weighted counterpart in the all-manufacturing sector in every census year and in 16 of 20 industries in 1982, in spite of the fact that the weighted diversification index exceeds the unweighted index in every industry in every year and never by less than a factor of 4. The measures suggest that large enterprises, on average, have their individual plants producing a more homogeneous product mix than do much smaller firms, those typically having single plants. Perhaps this reflects a lesser degree of flexibility available to those smallest firms constrained to operate a single plant. Whatever the interpretation, it illustrates another dimension of important differences between large and small enterprises.[18]

5. Conclusion

A number of detailed conclusions are reported above but one stands out in importance. Enterprises appear to be restructuring their production activity, moving toward more homogeneous production within establishments while holding a more diversified portfolio of establishments. While all-manufacturing enterprise diversification has increased over the 1963-82 period, the contribution of diversification within establishments has declined markedly, from an all-manufacturing average of .038 to .014. In contrast, the contribution to enterprise diversification from diversification among establishments has increased from .392 to .450. Moreover, by 1982 the latter contribution to enterprise diversification was more than 32 times the former.

An important implication follows directly. Contrary to the popular argument, technical economies of scope appear to play little role in explaining the measured increase in enterprise diversification. If such technical economies exist and are actively pursued by firms, the economies most likely would be realized at the establishment level, but it is precisely here that diversification has decreased. Between 1963 and 1982 establishment diversification has declined in 17 of 20 industries. On average, establishments were producing fewer and less dissimilar five-digit products in 1982 than in 1963. In fact, over this 20-year interval wholly undiversified (D=0) plants in 13 of 20 industries increased both their proportion of industry establishments and their share of industry shipments. If anything, firms appear at the establishment level to be seeking more conventional scale rather than scope economies.

Notes

*Boston College and U.S. Bureau of the Census, respectively. We wish to thank J.R. Norsworthy for his helpful comments and the U.S. Department of Commerce for research support under Grant No. 3396. The views expressed in this paper do not necessarily reflect those of the Census Bureau.

1. These data are taken from Federal Trade Commission (1969), p. 673 and FTC (1981), Table 20.

2. See Baumol (1977).

3. See Baumol (1976).

4. See Panzar and Willig (1977) and Sharkey (1981).

5. See the discussion in section 2.

6. The importance of the distribution property has been recognized previously. See, for example, Berry (1971) and McVey (1972).

7. The importance of product heterogeneity as a determinant of diversification also has been discussed in the literature. Both Berry (1971) and Gorecki (1975) acknowledge its importance. With various degrees of formalism, Utton (1977), Jacquemin and Berry (1979), Pomfret and Shapiro (1980), and Goto (1981) explicitly account for it in their research.

8. A conglomerate formed through market extension merges noncompeting products with related marketing channels or production processes.

9. Gort (1962) designs an alternative measure of diversification $n\sum_{j=2}^{n} s_j$ in an attempt to satisfy both the number and distribution requirements. Like d_1, however, this measure is sensitive to neither the distribution nor heterogeneity of secondary products.

10. The measure d_1 appears frequently in the literature. See, for example, Rhoades (1973), Gorecki (1974, 1975), Greening (1980), and Caves (1981).

11. Examples include Berry (1971, 1974), McVey (1972), Rhoades (1973), Gorecki (1974, 1975), Utton (1977), Jacquemin and Berry (1979), Greening (1980), and Pomfret and Shapiro (1980). Goto (1981) is an exception. He considers the commodity composition of the 124 largest Japanese firms in the 1963-75 period.

12. Increasing the number of input categories to twenty increases the establishment and enterprise diversification indexes, on average, by approximately .01 points. The results of this and other sensitivity tests are reported in Monahan (1984).

13. The 1963 Census of Manufactures data file contains 296,033 establishments and 265,778 enterprises. There are 348,385 establishments and 294,397 enterprises in the 1982 Census file.

14. The 1982 number, distribution, and heterogeneity components take .413, -.122, and .137 values, respectively, in petroleum and coal products and .297, -.111, and .090 values, respectively, in chemicals.

15. The parenthetical explanations are derived from observed movements in the number, distribution, and heterogeneity components for each industry's index.

16. Recall, an increasingly equal product distribution reduces the negative distribution term in absolute value.

17. Three out of four industries exhibit a decline in enterprise diversification between 1977 and 1982. This may be more a statistical artifact than an indication of a trend given the poor state of the economy in 1982, the census year.

18. That the weighted among-establishment contribution exceeds the unweighted contribution in Table 3 follows automatically from the dominance of single-plant firms in every industry. For single-establishment firms, the among-establishment term $D_i^{ent} - \sum_j a_{ij} D_{ij}^{est}$ definitionally equals zero, a condition applying to nearly 80 percent of enterprises in 1982. These firms collectively have little effect on the weighted analysis. Each firm, however, whether single or multi-plant, has an equal weight in forming the unweighted measures.

References

Baumol, William J. (1976). "Scale Economies, Average Cost, and the Profitability of Marginal Cost Pricing," in Public and Urban Economics, Ronald E. Grieson (ed.), Lexington, MA: Lexington Books, 43-57.

__________ (1977). "On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry," American Economic Review, 67 (December), 809-22.

Berry, Charles H. (1971). "Corporate Growth and Diversification," Journal of Law and Economics, 14 (October), 371-83.

__________ (1974) "Corporate Diversification and Market Structure," Bell Journal, 5 (Spring), 196-204.

Caves, Richard E. (1981). "Diversification and Seller Concentration: Evidence from Changes," Review of Economics and Statistics, 63 (May), 289-93.

Davies, Stephen (1979). "Choosing Between Concentration Indices: The Iso-Concentration Curve," Economica, 46 (February), 67-75.

Federal Trade Commission (1969). Economic Report on Corporate Mergers. Washington, D.C.: U.S. Government Printing Office.

__________ (1981). Statistical Report on Mergers and Acquisitions. Washington, D.C.: U.S. Government Printing Office.

Gorecki, Paul K. (1974). "The Measurement of Enterprise Diversification," Review of Economics and Statistics, 56 (August), 399-401.

__________ (1975). "An Inter-Industry Analysis of Diversification in the United Kingdom Manufacturing Sector," Journal of Industrial Economics, 24 (December), 131-46.

Gort, Michael (1962). Diversification and Integration in American Industry. Princeton: Princeton University Press.

Goto, Akita (1981). "Statistical Evidence on the Diversification of Japanese Large Firms," Journal of Industrial Economics, 29 (March), 271-8.

Greening, Timothy (1980). "Diversification, Strategic Groups and the Structure-Conduct-Performance Relationship: A Synthesis," Review of Economics and Statistics, 62 (August), 475-77.

Jacquemin, Alexis and Charles Berry (1979). "Entropy Measure of Diversification and Corporate Growth," The Journal of Industrial Economics, 27 (June), 359-69.

McVey, J.S. (1972). "The Industrial Diversification of Multi-Establishment Manufacturing Firms: A Developmental Study," Canadian Statistical Review, 47 (July), 4, 6, 112-117.

Monahan, James L. (1984). "The Bureau of the Census Index of Diversification: A Technical Description of Its Construction," Bureau of the Census Technical Notes (September).

Panzar, John C. and Robert D. Willig (1977). "Free Entry and the Sustainability of Natural Monopoly," Bell Journal 8 (Spring), 1-22.

Pomfret, Richard and Daniel Shapiro (1980). "Firm Size, Diversification, and Profitability of Large Corporations in Canada," Journal of Economic Studies, 7 (November), 140-50.

Rhoades, Stephen A. (1973). "The Effect of Diversification on Industry Profit Performance in 241 Manufacturing Industries: 1963," Review of Economics and Statistics, 55 (May), 146-55.

Sharkey, William W. (1981). "Existence of Sustainable Prices for Natural Monopoly Outputs," Bell Journal, 12 (Spring), 144-54.

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