Volume Title: Diversification and Integration in American Industry

Volume Author/Editor: Michael Gort

Volume Publisher: Greenwood Press

Volume ISBN: 0313-32442-5

Volume URL: http://www.nber.org/books/gort62-1

Publication Date: 1962

Chapter Title: Concepts and Methods

Chapter Author: Michael Gort

Chapter URL: http://www.nber.org/chapters/c1934

Chapter pages in book: (p. 8 - 26)

CHAPTER 2
Concepts and Methods

THIS chapter is concerned with developing definitions and measures of the concepts of diversification and integration, and with examining the technical characteristics of the data employed. It is written primarily for those interested in a close examination of the basis for conclusions reached in this study.

Concepts of Diversification

Diversification may be defined as an increase in the heterogeneity of *markets* served by an individual firm. Heterogeneity of *production* is distinct from diversification if it involves minor differences of essentially the same product, or if it takes the form of vertical integration. (The latter is discussed in a later section.)

In economic theory, it is usual to propose cross-elasticity of demand as a basis for identifying separate markets, and hence separate products for the purpose of measuring diversification. If cross-elasticity is high, the products are close substitutes and, hence, belong to the same market; if it is low, the products belong to separate markets. Unfortunately, there is little information on cross-elasticities of demand and, even were it available, one would still need to define the value of elasticity below which products may be considered as being separate—a definition that of necessity would be largely arbitrary.

Another solution to the problem is to identify separate markets on the basis of the immobility of productive resources. When resources can be shifted rapidly from one set of products to another, products may be deemed to fall in a common industry from the standpoint of the producer, though they are distinctly separate as seen by the consumer. Considerable mobility of resources leads to interdependence in supply for products. Changes in price (and hence in earnings) associated with one product lead to a shift in resources, and thus to changes in price and output of the other. Therefore, the phenomena of interdependence in prices, output, and earnings associated with high cross-elasticities of demand tend to be present also where a high degree of mobility of resources exists. Moreover, though there are many exceptions, the more common circumstances under which productive capacity can rapidly be diverted from one use to another are those involving products that are close substitutes. Thus the two sets of conditions may be considered complementary in identifying separable industries. Once again, however, progress in identifying diversification on the basis of interdependence in supply is constrained both by the absence

of appropriate information and by a lack of clear-cut definitions of limits.

In summary, diversification may be defined as an increase in the heterogeneity of output from the point of view of the number of markets served by that output. Two products may be specified as belonging to separate markets if their cross-elasticities of demand are low and if, in the short run, the necessary resources employed in the production and distribution of one cannot readily be shifted to the other. An increase in the number of regional markets served by an enterprise may, under some circumstances, fit the above definition of diversification. However, except when transportation costs are high, manufacturing resources are likely to be highly mobile between the demands of different regions for substantially similar products. (Mobility here simply entails selling the manufactured product in another region rather than moving plant and equipment.) Since limited short-run mobility is specified as a condition for distinguishing between markets, diversification through regional heterogeneity of output is far less likely to occur in manufacturing than in the distributive and service trades. Two quite similar products at opposite extremes of the relevant price range may also be characterized by very low cross-elasticity of demand. Immobility of resources between the two is, however, considerably less likely, though not inconceivable. In any event, our study is concerned neither with this last type of diversification nor with regional heterogeneity.

For this study, products were identified as belonging to separate markets if they could be classified into separate industries on the basis of the 1945 Standard Industrial Classification Code. As a practical matter, there was little choice since most of the data available were cast in the mold of the Classification Code. The industry code was developed mainly from product classes widely in use and therefore is largely based on differences and similarities in products. However, industries are sometimes also distinguished on the basis of production processes and raw materials employed. In most instances all three criteria lead to the same classifications, but there are many exceptions to this. In consequence, the implications of diversification, as measured on the basis of industry distinctions made in the Code, will vary to some extent depending on the classification criteria that had been employed. Generally, similarities in products are associated with high cross-elasticities of demand, and similarities in production processes with mobility of resources.

Measures of Diversification

How may quantitative differences in the heterogeneity of markets served by a firm be measured so as to distinguish more and less diversified enter-

9

prises? The choice among the several measures developed depends to some extent upon the problems one chooses to examine. Thus, if the problem is to predict differences in the response of earnings or output to cyclical fluctuations or to long-term trends in demand for the primary product of an enterprise, what is needed is a measure of concentration of output in the primary industry. That is, one would wish to know the extent to which the operations of the total enterprise are dependent upon the demand for its most important product. Another measure of diversification might be designed to throw light on the extent to which an enterprise is likely to enter activities unrelated to its primary operations in response, for example, to high earnings or growth prospects in other industries. For this purpose, an appropriate measure may simply be a count of the number of industries in which the enterprise produces goods or services.

The ratio of primary industry output to total output for the enterprise yields a measure of homogeneity—and the complement of this ratio, a measure of diversification—that is relevant for projections of the earnings of firms.[1] A limitation of this measure may be illustrated by a simple example. Assume that two firms show an identical primary industry specialization, but the nonprimary output of one company is evenly divided between five industries, whereas in the second company it is concentrated in a single industry. If diversification is measured as the complement of the ratio of primary industry output to total output, no differences between the two firms would be revealed, though for most purposes the first may be considered the more diversified. This could be corrected, to some extent, by adding the second largest industry to the numerator of the ratio. However, as the number of industries in the numerator is increased, more of the diversification to be measured is concealed. Indeed, if a sufficient number of industries were grouped with the primary industry in the ratio, all firms would appear to have a homogeneous output.

On the other hand, a measure of diversification based upon a simple count of industries would give undue weight to a wide dispersion over the industrial spectrum of activities that, in the aggregate, account for only a small proportion of the firm's total operations. While both the above measures were used individually, a composite measure which employs both the primary industry ratio and the number of industries in which operations are to be found was also used in this study. It was derived by multiplying the complement of primary industry specialization by the

---

[1] The complement of the ratio gives the relationship of nonprimary industry to total output.

number of industries in which the enterprise showed operations. A simple alternative designed to serve the same objectives would be to count the number of industries that account for a specified percentage of the total output or employment of the enterprise. The larger the number, the more diversified the firm. This method, however, proved insufficiently sensitive for our data in that it generated a large number of tied ranks in the sample of 111 large companies, ranked on the above basis.

A satisfactory direct measure of output was not available for nonmanufacturing activities. As a result, primary industry specialization was in most instances measured on the basis of employment data. For our sample of 111 firms, measures of diversification restricted to *manufacturing* activities were based on manufacturing payrolls. Because of variations in wage rates and in the number of hours worked, payrolls tend to approximate output somewhat more closely than does employment. For this sample of 111 companies further refinements were possible. Specifically, it was possible to separate and exclude payrolls associated with integration from the denominator of the primary industry specialization ratio.[2] Thus differences between companies in integration did not appreciably affect our measures of diversification for this sample. Again, for the manufacturing activities of the 111 firms, in measures of diversification based on a count of activities an attempt was made to exclude those which might be considered trivial in relation to the total size of the firm. For this purpose, all manufacturing activities within the firm with less than 1 per cent of manufacturing employment were excluded from the count. A 2 per cent exclusion rule was also tested.

Another variation was introduced in the measure of diversification by the choice of level of industry detail in identifying the primary industry. Thus, 2-, 3-, and 4-digit levels of detail were used at various points in the study. In summary, several measures with varying refinements were used. Some comparisons of the results using the various measures are shown later in this chapter and in Chapter 3.

Concepts and Measures of Integration

Integration may be defined as the act of combining two or more separable stages of production under common ownership. But how may a stage be identified? Within every establishment, the same productive functions

---

[2] In measuring diversification, it is conceptually best to combine integration employment with that in the industry to which it is auxiliary. This, however, did not prove feasible, partly because of limited information and partly, also, because operations in a given industry may be auxiliary to several rather than to one activity.

may be conceived of as a continuous process or, alternatively, subdivided into a vast number of separate operations, each of which may be identified as a separate stage in production. While in all industries certain productive processes are commonly regarded as being separate from others, these distinctions follow no uniform principle among the various industries. Stages of production may be identified as being separable if the productive processes could be performed successfully under separate ownership. Thus activities may be considered separable if, in fact, some firms successfully engage in them independently of the other activities in question. Therefore, if a firm combines two sets of successive operations though these operations are not combined under one ownership by other firms, the former may be said to have combined separable stages of production. A classification system based on this concept would not identify processes that are always combined within single plants as separable stages of production. Less obviously, if some manufacturing activities are never performed except in conjunction with the ownership of supplies of requisite raw materials, the extractive and manufacturing operations, even though performed in separate establishments, will not, in accordance with the above definition, be regarded as separable stages.

Changes in integration associated with the combination of two or more stages under common ownership normally entail separate investment decisions. Usually they will be made after explicit consideration of the alternative of purchasing the needed product or service outside the firm. Thus integration decisions, as defined above, are generally separable from decisions to manufacture the other products of the firm.

There exists, of course, no classification system that will uniformly identify separable stages of production in all the sectors of the economy. Nevertheless, the industry classification system used in this study largely achieves this objective. Productive processes associated with separate 4-digit SIC industries will, with only a few exceptions, constitute the entire scope of operations for at least some firms. Accordingly, a firm that combines two or more 4-digit activities, all of which are component processes associated with a single final product, may be considered more integrated than another whose operations are restricted to a single industry.

The measure of integration chosen for this study was based on the ratio of employment in all auxiliary activities to aggregate employment for the firm. An auxiliary product or service may either be an input (or a component part) of the "major" product or service or, alternatively, may constitute an operation that occurs at a later point in the production process (for example, a marketing outlet for a firm primarily engaged in

12

manufacturing). The distinction between "major" and auxiliary within the context of this study rests exclusively upon which of the two is the larger for the individual firm, with size measured by employment. One productive activity may be auxiliary to more than one major operation and, conversely, several activities may be auxiliary to the same major product or service. The specific technique employed in identifying auxiliary activities involved first grouping for each firm all the activities associated with common products or services. Second, within each group the largest in terms of employment was defined as major and the others were classified as auxiliary.

The attempt to measure integration was restricted to the sample of 111 large manufacturing enterprises. Distinctions between integration and diversification were made on the basis of qualitative information derived for each company from public records. In addition, individuals with expert knowledge of particular industries were consulted in some instances. Notwithstanding the care taken, arbitrary judgment is doubtless present in some of the allocations of activities between the categories of diversification and integration.[3] This is especially true for products or services designed to serve both purposes. Thus a manufacturer may maintain retail outlets both for the distribution of his products and for the sale of complementary products of other firms. Similarly, manufacturers frequently produce component parts for an external market as well as for their own use. In these circumstances, allocations were made on the basis of estimates of which category of uses was the larger in terms of sales.

Employment associated with sales and central administrative offices was classified as "integration" employment in computing a measure of integration. The major parts of sales and central office activities probably do not constitute separable stages in the sense defined earlier, since they are essential components of the operations of the firm. Yet the scope of activities of both sales and central offices varies considerably among companies. In some companies, these establishments perform services that elsewhere are purchased outside the firm. Examples are research and new product development, advertising, and various legal and engineering services. Our objective in classifying sales and central office employment as integration was to allow our measure of integration to reflect these differences among companies.

Integration (the ratio of auxiliary to total employment) is measured

[3] This was particularly a problem for activities that were very small relative to total firm size. The amount of information available was greater for the more important activities and thus permitted more reliable allocations.

on the basis of plant rather than product data, and consequently does not purport to measure differences in intraplant integration. It has been found, however, that establishments characterized by high primary product specialization ratios do not show a materially different ratio of value added to shipments than do establishments in the same industries characterized by lower specialization ratios.[4] Generally, a higher value-added to shipments ratio indicates greater integration. This would suggest that plant heterogeneity is not strongly related to integration. Stated in another way, for manufacturing enterprises on the whole, auxiliary operations that fall into distinguishable industries are reflected primarily in the variety of plants under common ownership rather than in product heterogeneity of the plants themselves.

Types of Data Used

Four sets of data were used for this study. First, the magnitude of primary relative to nonprimary activities for industry aggregates was derived from published census tabulations at roughly a 3-digit level of industry detail.[5] Second, for the relations between company size and diversification and company size and central office activities, unpublished data were prepared in a special census tabulation for 721 enterprises falling in 19 2-digit industries. Of these, 595 fell into thirteen manufacturing categories and 126 into six nonmanufacturing industries. The basic data were drawn from the 1954 Economic Censuses and consisted primarily of employment records for establishments, classified by industry of establishment. All of the employment of the relevant enterprises was included in these data, though out-of-census-scope records (that is, employment falling outside of manufacturing, minerals, trade, and services) were somewhat less reliable than those for in-census-scope industries. The enterprises comprised all of the multiestablishment companies with 2,500 and over employees that fell, on the basis of their primary industries, into the 19 2-digit industry categories.[6]

A major part of this study is based on a sample of 111 large enterprises.[7] Our third and fourth body of data relate to these enterprises.

[4] U.S. Bureau of the Census, *Working Paper Number 2*, 1956, Table II (Part II).

[5] The source of these data was U.S. Bureau of the Census, *Company Statistics: 1954 Censuses of Business, Manufacturing, Mineral Industries*, Washington, 1958. The Standard Industrial Classification Code was modified for these tables, but approximated the 3-digit level of detail.

[6] A narrower classification would not have generated a sufficient number of companies within individual categories to permit analysis of differences between companies within industries. Industries that had a negligible number of enterprises with 2,500 and over employees were omitted from the tabulation.

[7] The composition of this sample is discussed in a separate section later in this chapter.

Substantially the same type of information, but with a few additional statistics, was developed for this sample of companies from the 1954 Census as for the 721 companies mentioned above. In addition, similar information, but restricted to manufacturing industries, was derived from the 1947 Census of Manufactures for these 111 firms. Finally, in addition to census data on company employment and payrolls classified by industry, information on the products of the 111 firms was obtained from public records.

The information from public records consisted of all the products and services produced by these companies in 1954 and the changes in product composition for the companies that occurred in the periods 1929–39, 1939–50, and 1950–54.[8] The sources for these data consisted of information for individual companies in Moody's *Industrials,* in corporate annual reports to stockholders, and in *Thomas' Register of American Manufactures.* Products reported in 1954, as well as changes in product composition in each of the three periods indicated above, were classified into 4-digit SIC industries. Apart from conceptual advantages (discussed at a later point), the use of a 4-digit level of classification for product data served to eliminate the effects of variations in the detail with which various companies reported their product structures; that is, all companies reported their products at not less than a 4-digit level of detail, even though the Code was probably not directly considered when the reporting method was chosen. Products that were added and abandoned within the same interval of time were not recorded. Therefore, only changes in product structures between the initial and terminal dates of each period were entered. When a company added several products in the same 4-digit class during any one of the three periods, it was credited with only one addition. However, since each period was analyzed separately, if a product new to the company was added within one of the three periods, it was counted as an addition even though the company may have had other operations within the same 4-digit industry at the outset of the period. This procedure avoided a possible bias originating in the fact that since the total number of industry classes is limited, the more a firm diversifies in an earlier period, the fewer are the industries which remain to be entered at a later time.

The initial year 1929 was chosen because listings of products in public records are considerably less complete before the late 1920's. Nineteen hundred and fifty-four was chosen as the terminal year of the study because

[8] The periods began and ended in December of the initial and terminal years of each period.

an economic census was taken for that year. The three periods which subdivide the interval covered by the study were selected partly with a view to distinguishing diversification in the 1930's, a time primarily of low-level economic activity, from that associated with the high levels of output and investment in the 1950's. Also, diversification within the 1939–50 period probably reflects the impact on product structures of mobilization for World War II. While 1947 would probably have been preferable to 1950 as a point at which to divide the post-1939 interval, information available for 1950 from the Federal Trade Commission enabled us to check the accuracy of our data at the time of their preparation.[9] Moreover, since the initial and terminal dates of our periods had to be chosen in advance of the preparation of data, they do not necessarily mark off the points at which trends in diversification show marked changes in the rate of change over time.

To test the completeness of our product record, comparisons were made for 1954 between the product composition of companies as derived from public sources of information with that derived from census data for estab-

TABLE 1

MANUFACTURING INDUSTRIES APPEARING IN 1954 CENSUS ESTABLISHMENT DATA BUT NOT
SHOWN IN PUBLIC RECORDS, 111 COMPANIES

Primary Industry of Company	Number of Companies	Number of 4-Digit Industries Not Shown[a]	Average Ratio of Employment in Industries Not Shown to Total Manufacturing Employment[b]
Food products	12	6	.002
Tobacco manufactures	5	0	0
Textile mill products	4	1	.004
Paper products	8	0	0
Chemicals	14	10	.002
Petroleum	10	1	.002
Rubber products	5	0	0
Stone, clay, and glass products	7	1	.016
Primary metals	10	2	.001
Fabricated metal products	5	1	.002
Machinery	13	3	.015
Electrical machinery	5	2	.011
Transportation equipment	13	10	.007

SOURCE: Described in Chapter 2.

[a] The maximum number of 4-digit industries not shown for any one company was six.

[b] Unweighted average ratio for companies in each industry.

[9] Subsequently, however, when census data became available, the accuracy of information on products was checked against the latter source.

lishments. Since establishments frequently produce more than one product, it was to be expected that our product record would show a larger number of 4-digit industries than that revealed in census establishment data. The latter record only the primary product of an establishment. Though the degree of classification detail was the same for both, the 111 companies, in the aggregate, showed 1,709 manufacturing activities in the product record as compared with 1,073 in census establishment data. Of greater interest in testing the adequacy of the product record were instances in which census establishment data showed an industry that did not appear in the product record. These proved to be very few; for manufacturing activities there were only forty-seven such instances for the aggregate of 111 companies (as compared with 1,026 activities shown in census data which also appeared in our product record). Table 1 shows the distribution, by major industry of company, of instances in which industrial activities were revealed in the census but not in the public record. As may be seen from the table, in no industry group of companies was the unweighted average ratio of employment in omitted activities (those omitted from public records) to total manufacturing employment greater than 1.6 per cent. Employment in the omitted activities was measured on the basis of census establishment data. It is apparent that omissions of products that were primary to the operations of individual establishments proved to be of negligible importance.

For the purpose of all data in this study, an enterprise was defined as the combination of all properties (including all corporate subsidiaries) falling under common ownership or control. Control was deemed to be exercised where the parent company held 50 per cent or more of the voting stock of a subsidiary, or where a parent company reported effective control to the Census Bureau even though it held less than 50 per cent of the voting stock. For purposes of non-census data, an indication of effective control with less than 50 per cent voting stock ownership was, on occasion, found in public records.[10] In short, the definition of a company or enterprise was substantially the same for both census data and data derived from public records.

For companies that had undergone large mergers during the period studied, the antecedent company of a merged one was deemed to be the

[10] A few instances in which two companies each held 50 per cent of the voting stock were encountered. In these cases, effective control was determined on the basis of qualitative information as to which of the parent companies performed the main supervisory functions. Companies that were related only in the sense that the same stockholder had a decisive voice in their management (that is, companies subject to the same personal, as distinct from corporate, control) were not combined in the data used.

largest individual component company, in terms of total assets, at the time of the merger.

The geographical scope of census data was defined by the boundaries of the United States. For data secured from public records, however, operations in Canada as well as the United States were incorporated in the study, but not those of subsidiaries in other countries.

Characteristics of the Sample of 111 Large Enterprises

The 111 large companies, on which much of our analysis is focused, were drawn from thirteen 2-digit manufacturing industries primarily on the basis of asset size. Appendix Table A-1 shows, for each company individually, total assets in 1929, 1939, and 1954, and the ratio of net income to net worth for the period 1947-54. The classification of companies according to industry was based on a 1951 Federal Trade Commission report which, in turn, classified companies on the basis of "the manufacturing activity which accounted for the largest percentage of total receipts in the post-war period."[11] In several instances, however, the industry in which a company was classified was altered on the basis of information derived from annual reports on the composition of the companies' output in 1954. The primary basis for selecting companies was inclusion in the list of 200 largest manufacturing firms.[12] The resources available required that the total sample be limited to between 100 and 120 companies. Accordingly, to secure broader industrial representation, it was necessary to limit the number of companies in a relatively homogeneous industry such as petroleum to ten, and the number in the considerably more heterogeneous chemical industry to fourteen. This was achieved by selecting the largest ten and fourteen firms in each of the two industries, respectively. Similar constraints were imposed on the number of companies selected for some of the other major industries. In a few instances, on the other hand, it was necessary to go outside the list of the 200 largest manufacturing companies to secure more adequate representation for the industry. Some companies were excluded because their histories, or public information on them, did not extend back to 1929. Industries which were not represented in the list of 150 largest companies, or which had fewer than four firms among the largest 500, were omitted

[11] U.S. Federal Trade Commission, *A List of 1,000 Large Manufacturing Companies, Their Subsidiaries and Affiliates, 1948*, June 1951, p. 2.

[12] The ranking of firms on the basis of total assets was primarily derived from "Directory of the 500 Largest U.S. Industrial Corporations," *Fortune Magazine*, Supplement, July 1956.

TABLE 2
TOTAL ASSETS OF 111 ENTERPRISES AND OF ALL CORPORATIONS IN
THIRTEEN MAJOR INDUSTRIES, 1954

Primary Industry of Company	Number of Companies	Total Assets of Companies in Sample ($000)	Total Assets of All Companies in Industry ($000)	Ratio of Assets in Sample to Assets of All Companies in Industry (per cent)
Food products	12	4,004,399	17,026,767	23.5
Tobacco manufactures	5	2,323,650	2,896,230	80.2
Textile mill products	4	543,527	8,623,017	6.3
Paper products	8	1,816,366	7,184,262	25.3
Chemicals	14	7,582,514	16,628,583	45.6
Petroleum	10	20,595,927	28,812,881	71.5
Rubber products	5	2,357,369	2,912,388	80.9
Stone, clay, and glass products	7	1,501,790	5,456,300	27.5
Primary metals	10	10,429,045	17,781,799	58.6
Fabricated metal products	5	1,315,766	8,149,487	16.1
Machinery	13	4,147,000	16,018,734	25.9
Electrical machinery	5	4,029,695	9,352,591	43.1
Transportation equipment	13	11,088,443	18,715,954	59.2
Total	111	71,735,490	159,558,993	45.0

SOURCE: For total assets of sample of companies, Moody's *Industrials*, 1954; for total assets of all corporations in industry, U.S. Internal Revenue Service, *Statistics of Income*, Preliminary Report, 1954.

altogether in selecting the sample. Table 2 shows that, for most of the thirteen major industries, the companies in our sample accounted for a substantial proportion of the total assets of all corporations that can be classified in those industries. The aggregate assets of the 111 companies accounted for 45 per cent of the total assets of all corporations in the thirteen industries in 1954.

The selection of firms that were largest in their respective industries in 1954 may have introduced a small bias in our data. The firms that are largest on the terminal date of a fairly long interval of time are likely, as a group, to have grown faster than average in the course of the period. Since diversification is a form of growth, they are also likely to have diversified somewhat more than average. However, within the sample (as Appendix Tables A–2 and A–3 show) there was no significant concentration of rapidly growing companies in particular 2-digit industries. Appendix Table A–4 shows that companies in petroleum and in primary metals were, on the average, larger in 1954 than those in the sample as a whole, while companies in textile mill products, paper products, and stone, clay, and glass products tended to fall in the lower asset-size classes.

Conceptual Problems in the Data Used

An important problem arises from the use in this study of both product and plant data. When census (plant) data were used, *all* of the payrolls of an establishment were classified into the single largest (primary) industry irrespective of the product mix of the plant. The use of plant data opens the possibility that measures of diversification are biased to the extent that certain groups of enterprises have more homogeneous plants than others. For example, it is possible that larger plants will tend to be more heterogeneous than smaller ones, or that certain classes of enterprises tend to segregate varied industrial activities into separate plants while others combine varied activities in the same plants. Since the definition of an establishment itself contains a margin of ambiguity, the proportion of a company's activities falling within the scope of a single establishment is at times arbitrary. If it were true that companies with a larger variety of plants tended to have plants that internally were more homogeneous, measures of diversification based on plant data would be seriously inadequate.

Does internal homogeneity of plants systematically offset the greater diversity of plants owned by some firms? To resolve this problem, companies were divided into deciles on the basis of a measure of diversification (hereafter referred to as measure D_3) derived from establishment data. The measure was based on 1954 census information and consisted of the complement of the ratio of primary 4-digit industry payrolls to total company payrolls in all manufacturing activities, multiplied by the number of manufacturing industries in which the company maintained activities. The number of activities included only those which were not associated with integration and which accounted for at least 1 per cent of a company's total manufacturing employment. Manufacturing payrolls associated with integration were subtracted from the denominator of the ratio given above. For the group of companies in each decile, a ratio was computed of the total number of 4-digit manufacturing products produced in 1954 (derived from the product record described earlier in this chapter) to the sum of 4-digit manufacturing industries shown for the companies in the 1954 census establishment record. The higher the ratio, the greater the heterogeneity of products within plants, and the greater the tendency of establishment data to understate diversity of operations.

As may be seen in Table 3, the ratio of products to industries of establishments shows a reasonable degree of stability among companies grouped by the above measure of diversification. For eight of the deciles, the relevant ratio fell within the range of 1.4 to 1.7. Though the decile associated

20

TABLE 3
DIVERSIFICATION AND RATIO OF NUMBER OF INDUSTRIES BASED ON
PRODUCT RECORD TO THOSE BASED ON ESTABLISHMENT DATA

Diversification Decile[a] (ascending order)	Number of Companies[b]	Ratio of Number of 4-Digit Industries in Product Record to Those in Establishment Data[c]
1	13	3.1
2	10	1.7
3	11	1.5
4	11	1.4
5	11	1.7
6	11	1.5
7	11	1.7
8	11	1.6
9	11	2.1
10	11	1.4
All companies		1.6

SOURCE: Special census tabulation based on 1954 Census of Manufactures and product data for 1954 developed from public information by methods described in this chapter.

[a] The diversification measure was based on the 1954 Census of Manufactures and consisted of two components. First, the complement of the ratio of primary 4-digit-industry payrolls to manufacturing payrolls, minus manufacturing payrolls associated with integration, was computed. Second, the resultant statistic was multiplied by the number of manufacturing industries shown in census establishment data each of which accounted for at least 1 per cent of manufacturing employment for the company concerned. Activities characterized as integration were excluded.

[b] There are thirteen companies in the first decile because of tied ranks. Since ninety-eight companies remained for deciles 2–10, one decile was randomly selected to consist of ten companies.

[c] The product record was based on 1954 data developed from public records. Establishment data were derived from the 1954 Census of Manufactures.

with lowest diversification showed the highest ratio, there was no consistent tendency for companies that were more diversified on the basis of establishment data to show a lower ratio of number of products to number of industries of establishments. In short, the more diversified companies derived their greater diversity mainly from the greater heterogeneity of their establishments. These results tend to show that even had we, in later analysis, used a diversification measure based on product rather than on plant data, the resultant measures would have been substantially similar. Table 4 was developed as a further test of this conclusion. The table shows that when companies were grouped into deciles on the same basis as in Table 3, the average number of products per company was roughly thirty-seven for the highest decile as compared with only about ten for the lowest. Similarly, the average number of *manufacturing* products per company was roughly thirty for the highest decile and only five for the lowest. The table shows a relatively steady rise, from one decile to

TABLE 4

THE RELATION OF AVERAGE NUMBER OF PRODUCTS PER COMPANY TO DIVERSIFICATION
DECILES BASED ON MEASURE D_3 FOR 111 LARGE COMPANIES

Deciles on Basis of D_3	Number of Companies[a]	Average Number of Products per Company, 1954	Average Number of Products in Manufacturing Per Company, 1954
1	13	10.5	5.2
2	11	14.1	8.4
3	10	14.1	8.4
4	11	17.9	12.1
5	11	18.5	12.1
6	11	21.8	16.7
7	11	21.0	15.7
8	11	29.2	21.8
9	11	30.9	26.0
10	11	37.2	30.5

SOURCE: Number of products per company based on product record described in this chapter. D_3 based on census establishment data for 1954. The latter measure is defined on p. 58.

[a] The numbers of companies in the deciles were uneven because of tied ranks and uneven total number of companies.

the next, in the number of products that companies produced in 1954.

Several qualifications, however, need to be appended. First, there appear to be differences in plant heterogeneity between industry groupings of companies, so that not all of the variations between individual companies are random. Second, as compared with the cross-section of products in 1954, a somewhat larger proportion of products added since 1939 were produced in plants classified in industries other than those of the added products. Thus recent increases in diversification tend to be understated in census plant data. This results from the lesser likelihood that new activities, which have not as yet attained maturity, will account for the major part of the output of multiproduct plants. In addition, the Standard Industrial Classification Code (the industry classification system used for the study) is unlikely to incorporate industries of very recent origin (such as nuclear products). Under these circumstances, products that fall into the newest industries tend to be classified into previously existing categories and thus may not contribute to a measure of change in diversification. To the extent that these factors are present, there will be a relative understatement in the measure of diversification in 1954 for firms that have diversified more recently.

Some might also raise the question whether the directions of diversification (that is, the types of industries into which firms most frequently

diversify), as reflected in plant data, are the same as those revealed in product data. On this point, however, the answer given in Chapter 7 is decisively in the affirmative.

Another possible bias arising from the nature of the data may be found in some other unsatisfactory aspects of the Classification Code. The division between industries in the Code does contain elements of arbitrary judgment, and industries are not equidistant from each other either in terms of substitution or mobility of resources between products. Nevertheless, if the errors in the Code as judged by an ideally designed classification system are generally random with respect to both the dependent variable (diversification) and the explanatory variables, these errors need not do much violence to broad generalizations.

Because the Code has been developed at several levels of industry detail, its use necessitates selection of the most appropriate level. The problem is made more difficult by the fact that if our objective is to segregate groupings of products defined by market boundaries, the most appropriate level given in the Code is not uniform for all industries. In general, the 4-digit level of detail represents the best compromise between, on the one hand, arbitrary breakdowns between close substitute products and, on the other, common classifications for heterogeneous products. Nevertheless, we have tried wherever possible to analyze the results at more than one level of industry detail. To the extent that varying classifications lead to substantially similar findings, the problem of choosing among the several alternatives becomes less critical.

The Relation Among Diversification Measures

Because of the variety of diversification measures used in this study, it is well to recapitulate them before analyzing further how much the use of one, rather than another, affects the findings. Fundamentally, three types were used: first, those based on primary industry concentration; second, those based on number of products produced or number of industries in which the firm was producing; third, measures that were composites of these two. In addition, the above measures were used at differing levels of industry detail, with as well as without adjustment for integration, and on the basis of differing bodies of data. The following is a summary of the measures used and tested.

$$D_1 = 1 - \frac{P_r}{M_r - I_r}$$

23

$$D_2 \quad = \quad 1 - \quad \frac{P_r + S_r}{M_r - I_r}$$

$$D_3 \quad = \quad N_1\left(1 - \frac{P_r}{M_r - I_r} \right)$$

$$D_4 \quad = \quad 1 - \frac{G_e}{V_e - A_e}$$

$$D_5 \quad = \quad N_2\left(1 - \frac{G_e}{V_e - A_e} \right)$$

$$D_6 \quad = \quad 1 - \frac{P_e}{M_e}$$

$$D_7 \quad = \quad 1 - \frac{L_e}{V_e}$$

$$D_8 \quad = \quad N_3$$

P_r = primary 4-digit industry payrolls; M_r = total manufacturing payrolls; I_r = manufacturing payrolls associated with integration; S_r = manufacturing payrolls in the second largest 4-digit industry; N_1 = number of 4-digit manufacturing industries, excluding those which account for less than 1 per cent of manufacturing employment and those classified as integration; N_2 = the total number of 2-digit industries; G_e = employment in the primary 2-digit industry; V_e = total company employment; A_e = employment in 2-digit activities associated primarily with integration; P_e = primary 4-digit industry employment; M_e = total company employment in manufacturing; L_e = primary industry employment at the modified 3-digit level of detail; and N_3 = number of 3-digit industries.[13]

The first five measures were used exclusively for the sample of 111 firms because the requisite information was not available for the other sets of data. The only difference between measures D_1 and D_2 is the addition in the latter of the second largest activity in the numerator of the ratio. The Spearman coefficient of rank correlation between these two measures

[13] The modified 3-digit level of detail was that employed in *Company Statistics*.

was .94. Measures D_1 and D_3 differ in the addition to the latter, as a multiplier, of number of industries engaged in. The correlation here was also .94. Measures D_4 and D_5 are the same as D_1 and D_3 except that they employ the 2-digit level of detail and include nonmanufacturing activities; they were also highly correlated (.85). High correlation for the three sets of measures was, of course, to be expected since in each case one of the two measures being correlated is a component of the other.

To determine the effect of changing the level of industry detail, rank correlation coefficients were computed between measures D_1, D_2, and D_3, on the one hand, and D_4 and D_5, on the other. The results were as follows: the Spearman coefficient of rank correlation for diversification measures D_1 and D_4 was .62, for measures D_2 and D_4 it was .66, and for measures D_3 and D_5 it was .75. Differences in the ranks assigned to companies on the basis of the 4- and 2-digit measures did not arise merely from differences in the level of industry detail. The 2-digit measures incorporated diversification in nonmanufacturing industries, whereas those at the 4-digit level were restricted to manufacturing activities. Further, the latter were based on data for payrolls, while the former were based on employment data (information on payrolls being unavailable for out-of-census-scope activities). Therefore, if both 2- and 4-digit measures had been computed on the same basis, the resultant rank correlations would very probably have been higher. In short, while differences obviously arise as one changes the level of industry detail, the coefficients were sufficiently high to indicate that the measures of diversification used were not merely the accidental results of a single arbitrary breakdown of industries.

Greater confidence can be placed in a measure of diversification that yields roughly similar results for several levels of industry detail than in one that is highly unstable with regard to the system of classification used. The greater stability of measure D_3 than of measure D_1 (when compared with similar measures at the 2-digit level) makes it preferable, on that basis, as a general measure of diversification. However, all measures were used to a degree, though some only for comparative purposes.

Measures D_6 and D_7 were developed for the broader sample of 595 manufacturing enterprises in thirteen industry groups. The companies were ranked within each group on the basis of the two sets of ratios, and Spearman coefficients of rank correlation were computed for each group. As Table 5 shows, for nine of the thirteen groups the coefficients exceeded .8. Thus, once again, alternative levels of classification generate measures that lead to substantially similar rankings of companies.

TABLE 5
SPEARMAN COEFFICIENTS OF RANK CORRELATION FOR NONPRIMARY
EMPLOYMENT RATIOS AT TWO LEVELS OF INDUSTRY DETAIL[a]

Primary Industry of Company	Number of Companies	Rank Correlation[b]
Food products	78	.642
Textile mill products	66	.841
Paper products	39	.825
Chemicals	60	.810
Petroleum	27	.790
Rubber products	12	.934
Stone, clay, and glass products	22	.764
Primary metals	52	.649
Fabricated metal products	27	.858
Machinery	88	.866
Electrical machinery	52	.834
Transportation equipment	54	.899
Instruments	18	.944

SOURCE: Data used for computing correlations were derived from a special census tabulation based on the 1954 Censuses.

[a] The ratios correlated were nonprimary employment over *total manufacturing* employment at the 4-digit level and nonprimary employment over *total* employment at a modified three-digit level of detail.

[b] All coefficients were significant at the .01 level.