On the measurement of concentration : a cross-sectional analysis of alternative indices

Author: James B. Delaney

Boston College Electronic Thesis or Dissertation, 1974

ON THE MEASUREMENT OF CONCENTRATION: A CROSS-SECTIONAL ANALYSIS OF ALTERNATIVE INDICES

by

James B. Delaney

ACKNOWLEDGEMENT

I am sincerely grateful to the many faculty members at Boston College who guided the research in this dissertation. In particular, my thanks to H. Michael Mann who directed the general nature of this work and who provided the encouragement that was needed to complete it. Also, many helpful comments on the empirical chapters came from William Duffy. Finally, I acknowledge my wife Nancy whose general attitude was a necessary component in the completion of this work. Of course, any logical or technical errors are the sole responsibility of the author.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF ILLUSTRATIONS

CHAPTER ONE: OPTIMALITY AND THE GOAL IN THE MEASUREMENT OF CONCENTRATION

Recently a growing interest in the measurement of market concentration has emerged. Over the past quarter century the question has appeared in many of the professional journals. During this period a conventional wisdom seemed to have been established, in which the question of the choice of an index of concentration was treated as a moot point. It was argued that various indices were largely identical with respect to their predictive power, that there was no sufficient theoretical basis for chossing among alternative measures, and that, therefore, the Census concentration ratio was a superior index based on its availability and its computational ease.[1]

Yet, recent analysis has eroded the commonly accepted view.[2] Moreover, new measures of concentration have been proposed. These must be added to an already burgeoning list of concentration

---

[1] A sample of such statements may be found in any of the following: M. Hall and N. Tideman, "Measures of Concentration," JASA (March, 1967), p. 168; R.W. Kilpatrick, "The Choice Among Alternative Measures of Industrial Concentration," REStat (May, 1967), p. 260; D. Bailey and S.E. Boyle, "The Optimal Measure of Concentration," JASA (December, 1971), p. 706.

[2] An example of this current empirical work can be found in R.A. Miller, "Numbers Equivalents, Relative Entropy, and Concentration Ratios: A Comparison Using Market Performance," SEJ (July, 1972), pp. 107-12.

indices.[3] These trends are related to a growing concern over market concentration in conjunction with a general dissatisfaction with the indices used to measure it. This dissatisfaction stems from the low predictive power of traditional indices and from their general exclusion of asymmetrical share influences on behavior and market performance. These points will be expanded in the next chapter.

Thus, the economist faces an embarrassing situation. Which one out of the many concentration measures should be employed? Are the conclusions drawn from empirical investigation sensitive to the choice of an index? How can the economist evaluate the relative efficacy of alternative indices? This study will systematically examine these questions. But first consider how optimality is determined.

Optimality is not a categorical concept. Rather it is a condition or state in which the most favorable position possible has been achieved. In this context, favorableness is best described as the extent to which an index embodies the attributes necessitated by the ultimate goal in the measurement of concentration. But what is this goal?

The purpose of a concentration index should be to capture the extent to which the structure of an industry approximates the elements characteristic of competitive or monopolistic markets. The reason is that classical economic theory suggests that

---

[3]For an example, see J. Horvath, "Suggestions for a Comprehensive Measure of Concentration," SEJ (April, 1970), pp. 446-52.

monopolistic performance will be occasioned by the control of a large share of an industry's output in the hands of a few firms. "Few" is as small a number that results in a behavioral pattern which brings forth a monopolistic, as opposed to a competitive, outcome. Thus, the goal in the measurement of concentration is to provide an operational counterpart to the economic concept of fewness. A concentration index should impart knowledge about the likelihood that an industry's performance will be non-competitive. It is to accomplish this by characterizing the number and size distribution of firms into a one parameter index.[4]

The problem is that there exists a multitude of weighting schemes of the number and size distribution of firms consistent with the concept of fewness. The attributes that a measure of concentration should possess, as delineated by economic theory, are not sufficient to allow for an unambiguous selection among alternative concentration measures. The choice of an index of concentration essentially becomes an empirical proposition. Sufficient optimality conditions must be determined by empirical criteria.

The empirical criteria used in this analysis are twofold.

---

[4]Fellner expouses a similar view. He notes that a measure of concentration should tell the researcher something about the likelihood that oligopolistic behavior and performance will emerge. W. Fellner, Business Concentration and Price Policy, NBER (Princeton, 1955), p. 113.

Also, note that a one parameter index will not capture all the forces which reinforce or undermine non-competitive behavior, but number and size distribution serve as a first approximation. Some of these other factors include market growth, entry conditions, and product homogeneity.

The first is designed to answer the question of existence, i.e., whether or not there exist various indices which characterize diverse aspects of the structural dimension of concentration. This is accomplished by computing index intercorrelations. If these correlations are high, then once one index is known, additional information about the structural dimension of concentration from other indices is likely to be insignificant. Conversely, if these correlations are low, then the measures embody different elements of concentration and are not replicative with respect to their information content. In the latter situation, classification of markets along an atomistic-monopolistic continuum according to their respective index values is likely to be conflicting when different indices are used.

The second criterion confronts the question of identification, i.e., which particular index or indices possess better predictive capability. As previously noted, concentration is a structural indicator of market power. Its measurement is predicated on the relationship between market structure and market performance.[5]

---

[5]The structure-performance relation examined in this paper concerns industry concentration and industry profitability. It is the most testable hypothesis since classical theory suggests an unambiguous relation between them. The line of reasoning is straightforward: as the level of market concentration increases, the degree of recognized interdependence among existing firms increases; this results in a decrease in the pressures for price competition; tacit collusion on price is likely to emerge, accompanied by a restriction of industry output; the result is an increase in industry profitability. The reverse holds for a decrease in concentration. In other words, the expectation is that high profits, on the average, should be associated with highly concentrated markets as opposed to lowly concentrated markets.

In the following analysis, predictive power is measured by the adjusted coefficient of determination ($\bar{R}^2$). The rationale for selecting the equation specification with the highest $\bar{R}^2$ is based on the assumption that if there exists a correct specification, it will on the average exhibit a higher estimated $\bar{R}^2$ than any other specification. This rule is only suggestive since data and sample limitations can affect the magnitude of $\bar{R}^2$. All empirical research must face these handicaps. While not ignoring them, these limitations must be placed in their proper perspective. The $\bar{R}^2$ criterion is not a panacea, but is a reasonable rule upon which index predictive capability can be judged when applied to an imperfect world.

These remarks are meant to provide a framework within which the following analysis will take place. The major points will be amplified, particularly in the next chapter where a survey of alternative measures is presented. Let us now turn to that analysis.

CHAPTER TWO: AN EXAMINATION OF ALTERNATIVE INDICES OF INDUSTRIAL CONCENTRATION

As previously stated, the choice of an index of concentration must be consistent with the imperatives set forth by economic theory. This encompasses the development of an operational meaning to the concept of fewness. The economist reaches an impasse, however, since there are numerous representations which satisfy this theoretical requirement. In order to analyze alternative concentration measures, this chapter begins with a general discussion of the major categories of concentration indices. This background will allow for the evaluation of each index in a systematic fashion, emphasizing the theoretical relationships between them. The major conclusion of the chapter is a full recognition of the inability of economic theory alone to provide a yardstick by which to compare various indices. The choice of an index becomes primarily an empirical proposition.

RELATIVE VERSUS ABSOLUTE MEASURES OF CONCENTRATION: BACKGROUND TO THE CONTROVERSY

Two major categories of concentration measures are identifiable: absolute measures which focus on a subset of the firms in an industry, and relative measures which depend on the entire population of firms in an industry. The concept of fewness forms the focal point of analysis for absolute measures. For example, John Blair has commented that "it is the dominance of the few ... which tends to influence the market,"[1] while Morris Adelman has noted that "...fewness is the essential part of the study of competition and monopoly."[2] These represent characteristic statements of the proponents of absolute concentration measures. The empirical implementation of this theoretical emphasis has been to confine the analysis to a small number of leading firms in an industry, in which a "small number" has been translated into the largest four or eight firms.[3]

The advocates of absolute measures of concentration stress that discretionary power over price and output in a particular industry, when it exists, is held in the hands of a small number of dominant firms. It is the decisions of these firms which can result in non-competitive market performance. The potentialities lie there and not with the smaller fringe firms. Since the major goal of the measurement of concentration is to impart knowledge about the likelihood of the emergence of

---

[1]J.M. Blair, "Statistical Measures of Concentration in Business: Problems of Compiling and Interpretation," Bulletin of the Oxford University Institute of Statistics (November, 1956), pp. 355-56.

[2]M.A. Adelman, "Differential Rates and Changes in Concentration," REStat (February, 1959), pp. 68-69.

[3]One firm is considered larger than another usually by considering its value of shipments, although value-added and number of employees also are used. It should be noted that the question of how to measure firm size is logically prior to the question concerning the measurement of concentration and is not considered in this discussion.

monopolistic performance, the researcher should analyze its direct source, i.e., the dominant few.

The disadvantage of absolute measures is that interactions between fringe and core firms are ignored as well as those within each group itself. Proponents of relative measures of concentration stress that changes in the relative differentials among firm sizes can have important influences on competition in an industry even though the leading firms are unaffected. Disparity in firm sizes can significantly affect competition by altering the likelihood that the discretionary power held by the leading firms will ever exercised. Thus, the analysis should not be confined to a few dominant firms, but rather should encompass the entire population of firms in an industry.

The assertion that relative size differentials can affect the establishment and the maintenance of collusive behavioral patterns (and ultimately market performance) was promulgated by Stackelberg.[4] The postulated mechanism envisioned by Stackelberg was the substitution of follower reaction functions into the profit function of the leader, who then determines his optimal output. Follower firms then maximize their profits given the leader's output.

The leadership position, and thus the determination of of the follower firms, crucially depends on the disparity in

---

[4]Henrich von Stackelberg, The Theory of the Market Economy (London: Wm. Hodge & Co., Ltd., 1952). Also, see W.G. Shepherd, "On Appraising Evidence About Market Power," Antitrust Bulletin (Spring, 1967), pp. 65-72.

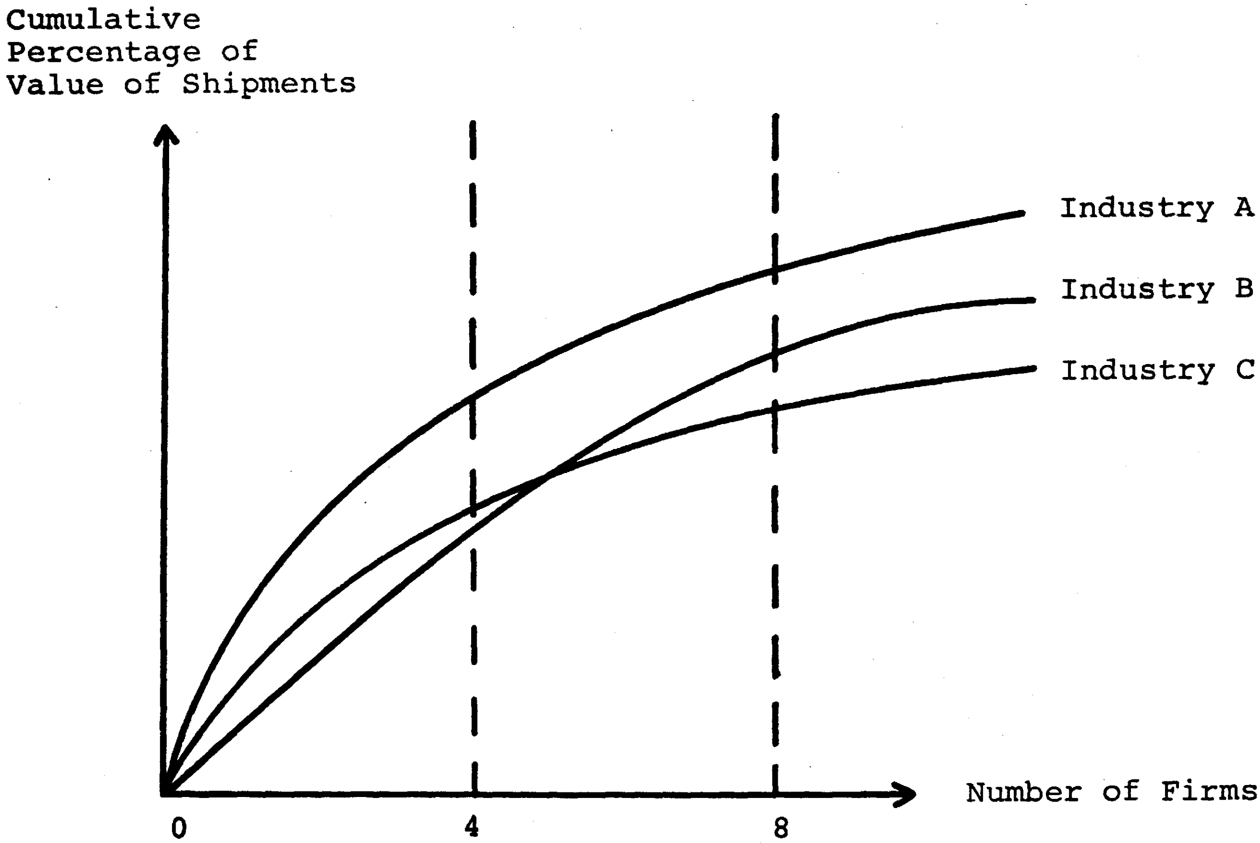
market shares. Equality in firm size leads to a contest for the leadership position, i.e., the familiar Stackelberg disequilibrium. This struggle is less likely to occur if one firm controls a large proportion of the market. The resulting interactions among the firms are similar to those described by the dominant-firm price leadership model. Stability is established because disparity dictates more obviously which firm will be the leader. Of course, stability is maintained only if the follower firms accept the leader and act accordingly.

The point to be emphasized is that the focus on the "dominant few" alone may be inappropriate. The concept of fewness involves not only numbers and relative size, but also the differentials in these sizes. Asymmetry in market shares may play a significant role. The proponents of relative measures of concentration contend that absolute measures are deficient in this latter area.

The controversy can be graphically interpreted by considering an industry's concentration curve. A concentration curve is a cumulative plot of the percentage of total industry value of shipments that is controlled by some number of firms, ranked from largest to smallest. Consider the cross-sectional comparison of three hypothetical industries in Diagram 2-A (p. 10).

Clearly, hypothetical industry A is relatively more concentrated than either industry B or C in a global sense since its concentration curve is everywhere above the other curves. However, the concentration curves for B and C intersect. It

DIAGRAM 2-A: A Comparison of Various Industry Concentration Curves

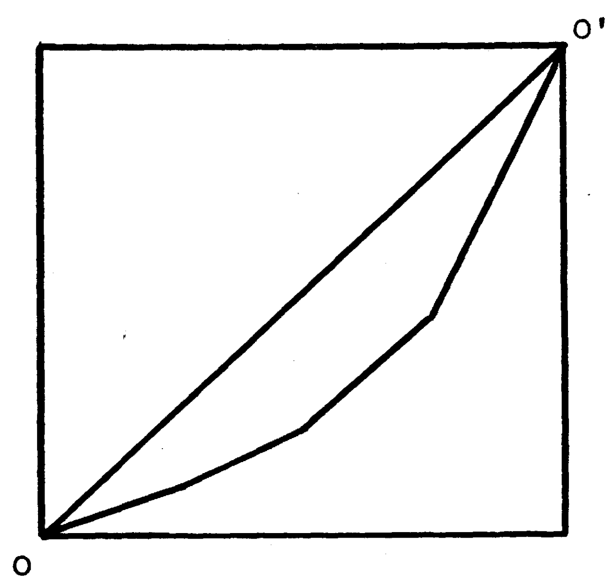


becomes obvious that one can select a neighborhood or a scheme of weights for the points on the curves that can make either industry appear more concentrated than the other. For example, consider just the top four firms. In this case, industry C appears more concentrated than industry B. Alternatively, when viewing the top eight firms, B is more concentrated than C. Thus, the choice of an index of concentration, even within a particular class of measures, can theoretically alter the results of empirical investigation. Indeed, Gideon Rosenbluth has noted that "the research worker using any one of these indexes will therefore want to know how much his results might be altered by the use of another index."[5] This specific problem will be extensively examined in later chapters.

Another graphical device that can be utilized is an industry's Lorenz curve. A Lorenz curve expresses the percentage of total industry value of shipments that is controlled by some proportion of firms in the industry, being cumulated from the smallest firm. Consider the Lorenz curve of a hypothetical industry as depicted in Diagram 2-B. An industry whose Lorenz

DIAGRAM 2-B: A Hypothetical Lorenz Curve

Cumulative Percentage of Value of Shipments



Percentage of Firms

[5] G. Rosenbluth, Business Concentration and Price Policy, NBER (Princeton, 1955), p. 64.

curve is coincident with the locus OO' is said to exhibit no concentration. The farther it diverges from this locus, the more concentrated the industry becomes. However, even though an industry is characterized by equally-distributed firm shares, it is possible that such firms could be few in number. The Lorenz curve cannot distinguish between an industry with only one firm and another with a thousand equally-sized firms.

Moreover, it is argued that the position of the Lorenz curve is extremely sensitive to the total number of firms classified into the industry. Blair comments that "there exists in most industries a considerable number of very small enterprises which exert little, if any, influence on the industry's behavior with respect to price.... Any measure of concentration which fluctuates with changes in the number of these tiny enterprises is not meaningful from an economic point of view."[6] Yet, judgments about monopolistic tendencies require a knowledge of the number of firms operating in that industry. Furthermore, an increase in the number of firms can result in a movement of the Lorenz curve in either direction depending on the effect on the dispersion of firm shares.[7,8]

---

[6]Blair, op. cit., p. 352.

[7]P.E. Hart and J.S. Prais, "The Analysis of Business Concentration: A Statistical Approach," Journal of the Royal Statistical Society, Series A, Part II (1956), pp. 152-53.

[8]A lucid development of these points can be found in E.M. Singer, Antitrust Economics: Selected Legal Cases and Economic Models (Prentice-Hall, 1968), pp. 141-44.

The question arises as to the relation between these two geometrical devices. The Lorenz curve measures the cumulative percentage of firms along the horizontal axis, while the concentration curve measures the cumulative number of firms. It is immediately clear that when the total number of firms is unknown, the Lorenz curve cannot be computed. But once the number of firms is known, differences in the Lorenz curve for two industries will be reflected in their respective concentration curves. For example, when the number of firms is fixed, increases in size inequalities will be associated with upward movements in the concentration curve.[9]

Given the previous graphical presentation, a convenient distinction can now be drawn. Concentration indices based on a weighting scheme of the points along a Lorenz curve are measures of relative concentration, while those based on a set of weights of the points along a concentration curve are measures of absolute concentration. The major problem arises when a specific measure of concentration is selected and applied to some predicted empirical relationship.

In general, the choice of an index is related to one's definition of concentration. Proponents of absolute measures view concentration as the degree of coercion that can be perpetrated by a few dominant firms on the market mechanism, while at the same time relegating the remaining firms to an

---

[9] Rosenbluth, op. cit., pp. 61-63, contains a more detailed discussion about the relation between the Lorenz and the concentration curves.

insignificant role. While not minimizing the role of leading firms in an industry, advocates of relative measures maintain that dispersion in or inequality of firm sizes has a significant impact on the propensity to implement such control. As such, relative measures of concentration impart more information about the degree of competition that is captured in the structural dimension of concentration.

More importantly, the controversy between absolute and relative measures serves to illustrate an important point: "...economic theory cannot offer much help in choosing among ... [measures of concentration]."[10] Are absolute or relative measures of concentration better able to characterize the market performance effects embodied in this structural dimension? And in a particular class, is there one index which is superior? Theory alone is unable to answer these questions because the concept of "few" lacks any precision with respect to numbers and to relative sizes.

The remainder of this chapter will be devoted to analyzing various measures of concentration on an individual level.

VARIOUS MEASURES OF MARKET CONCENTRATION

As previously noted, economic theory provides little, if any, direction in the search for the optimal measure of concentration. Many measures do possess economic significance

---

[10]T. Scitovsky, Business Concentration and Price Policy, NBER (Princeton, 1955), p. 112.

in the sense that they lend themselves to a representation of the structural characteristic of the number and size distribution of firms in an industry. The problem is that one cannot choose a priori among these various measures. The following section analyzes some proposed measures of industrial concentration, their computation, and their inter-relationships.

Concentration Ratio

The most widely used measure of concentration is the concentration ratio, which is simply the percentage of an industry's value of shipments, value-added, or employment that is controlled by a fixed number of leading firms. This number is usually four or eight, although other subsets of firms are occasionally used.

$$CR(j) = \sum_{i=1}^{j} x_i \qquad (1)$$

where x_i is the relative share of the i^{th} firm

The primary source of published ratios is the Bureau of the Census. These ratios are computed from categories defined in the Standard Industrial Classification Code. In order to ascertain the reliability of reported ratios, the Bureau also has defined the primary product specialization ratio and the coverage ratio: the former measures the extent to which plants classified into a particular industry specialize in the primary product of that industry, while the latter measures the proportion of a given product's shipments that originate from plants classified into a specific industry. When both of these ratios are high for an industry, the published concentration ratio can be used with more confidence.[11]

The computational ease and the availability of data probably account for the predominant use of the Census concentration ratio in empirical studies. However, this ratio does reflect a tenet of economic theory: when a small number of firms possess an inordinately large share of the market, mutually recognized interdependence is more likely; this can result in coordinated behavior and the exercise of discretionary market power. Thus, the concentration ratio represents an absolute measure of concentration.

Marginal Concentration Ratio

For a given four-firm concentration ratio, the remaining industry output can be distributed among the other established firms in various ways, each of which could influence the price and output combinations observable in the marketplace. One proposed measure which attempts to characterize these diverse attributes of a firm size distribution is the marginal concentration ratio.[12] It is defined as the successive

---

[11]For a more detailed discussion, see E.M. Singer, op. cit., pp. 156-74. Mathematically, the number of existing firms is related to the minimum feasible value of various Census ratios. For example, if there are n firms in an industry, then

$CR(j)_{min} = j \times \frac{1.00}{n}$ is the minimum value of CR(j) where market shares are equally distributed. As n decreases, $CR(j)_{min}$ increases. See A.D.H. Kaplan, Big Enterprise in a Competitive System, Table 4-3, p. 80.

[12]R.A. Miller, "Marginal Concentration Ratio and Industrial Profit Rates: Some Empirical Results of Oligopoly Behavior," SEJ (October, 1967), pp. 259-67.

difference in published Census concentration ratios. For example, the marginal eight-firm concentration ratio is the relative share of industry output accounted by firms ranked 5 through 8, i.e.,

$$MCR(8) = (CR(8) - CR(4)) \tag{2}$$

If an industry has a substantial share of its output produced by firms in the "second-four-grouping," the result is likely to be a lower industry profit-rate.[13] The theory is similar to Stackelberg's already discussed. With a large secondary group of firms, it is difficult to maintain or to impose any collusive agreement. The result is an inability to realize a joint-profit maximizing solution. In other words, independent price rivalry is more likely when the second group of four firms produce a significant share of total industry output.

Although theoretically plausible, the empirical verification of such conjectures hinges on the reported negative sign for the coefficient of MCR(8). The statistical results crucially depend on the sample employed, determining the correlation between CR(4) and MCR(8) and the statistical significance of the latter in the regression analysis.[14]

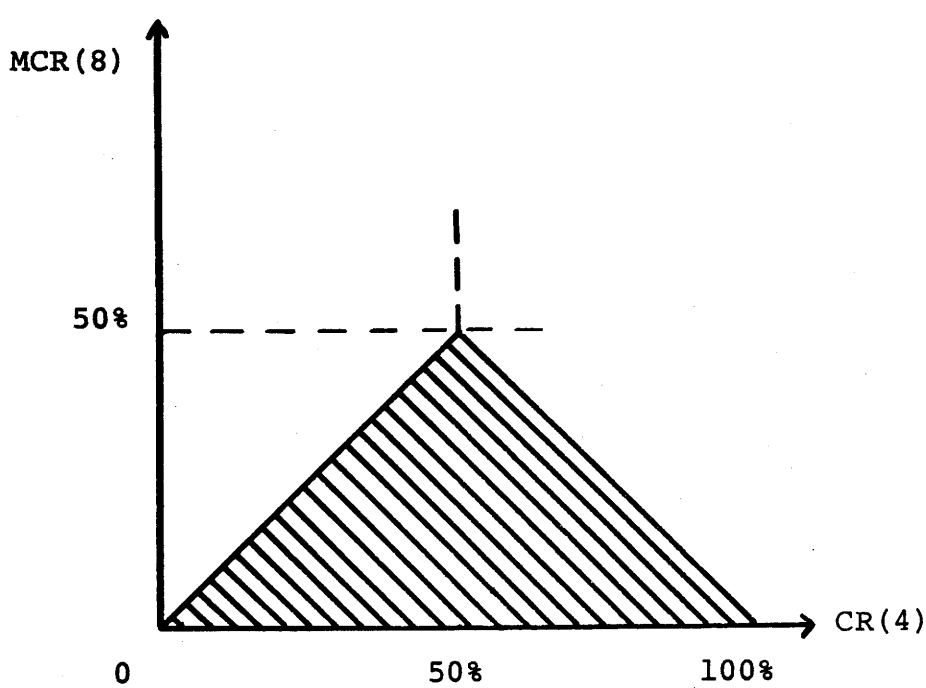
---

[13]Ibid., p. 264.

[14]Henning has noted that these variables are necessarily statistically dependent and has specified its nature. The following is largely adopted from his analysis. See J.A. Henning, "Marginal Concentration Ratio: Some Statistical Implications--Comment," SEJ (October, 1969), pp. 196-98.

For a given level of four-firm concentration, i.e., the relative market share of the leading four firms in the industry, the associated value of the eight-firm marginal concentration ratio can be neither greater than CR(4) nor greater than (1 - CR(4)). Thus, there is a range of possible values of the eight-firm marginal concentration ratio consistent with, and determined by, the four-firm concentration ratio. For values of CR(4) less than or equal to 50%, the binding constraint is the first; for values of CR(4) greater than 50%, the binding constraint is the second one. The feasible region for MCR(8) is depicted in Diagram 2-C.

DIAGRAM 2-C: Feasible Region of MCR(8) for Alternative Values of CR(4)



Consider a sample whose observed values of CR(4) are greater than 0.50. The sample correlation between the two variables may be negative regardless of any other relationship between them. Similarly, a positive correlation may arise for sample values of CR(4) less than 0.50.

In order to remove the restriction on the range of values for the eight-firm marginal concentration ratio imposed by its statistical dependence on the four-firm concentration ratio, it has been suggested that one should express MCR(8) relative to its maximum possible value:[15]

$$MCR(8)^{*} = \frac{MCR(8)}{CR(4)} \qquad \text{if } CR(4) \leq 0.50 \qquad (3a)$$

$$MCR(8)^{*} = \frac{MCR(8)}{(1 - CR(4))} \qquad \text{if } CR(4) > 0.50 \qquad (3b)$$

The transformed marginal concentration ratio indicates that the relation between MCR(8) and CR(4) is not linear over the entire range of values for the four-firm concentration ratio.[16] The point is that reported significance of marginal ratios in a regression model could be the result of its statistical dependence with Census ratios. The corrected marginal ratio allows the researcher to test for theoretical relations by removing statistical restrictions.

---

[15] Ibid., p. 198.

[16] Results of approximating this non-linear relationship and the resultant effects on performance as measured by profitability will be examined in a later chapter. In particular, a close look at the work of Collins and Preston and of Miller will be made.

A question arises as to the appropriateness of the marginal concentration ratio as a general measure of market concentration. It is a unique representation in the sense that it does not incorporate the leading firms directly. All other indices, whether absolute or relative, consider at least the top firms. By such exclusion, it seems inappropriate to use MCR(8) alone in the classification of markets as competitive or monopolistic or as a structural indicator of monopolistic performance. In this sense, MCR(8) is not a general measure of concentration. However, in a regression context, the importance of MCR(8) is enhanced. If MCR(8) is significant in an equation with CR(4) as a regressor, then MCR(8) captures different aspects of a firm size distribution than CR(4). This result would suggest that the measurement of concentration by CR(4) should be supplemented with MCR(8) (or $MCR(8)^*$) when estimating structure-performance relations. In this sense, MCR(8), as defined by successive points on an industry's concentration curve, represents a measure of concentration.

Disparity Index

Another aspect of a firm size distribution with a given four-firm concentration ratio is the differences in size within the leading firm group itself. One method to characterize this particular attribute is with an index of disparity, which is the relative mean absolute deviation of the value of shipments in a pre-specified category of leading firms. In the case of the four top firms,

$$DI(4) = \frac{1}{4} \sum_{i=1}^{4} |(X_i - \bar{X})| \div \bar{X}$$

where X_i: value of shipments of the i^{th} firm

$\bar{X}$: mean value of shipments of the top four firms

By focusing on the shares within the leading firm group, the disparity index can indicate potential patterns of price strategy adoption. For example, extremely large differentials in firm shares reveal the presence of one large firm within the top group who could establish a form of dominant firm price leadership. Conversely, very small values of the index might signal an inability to impose coordinated pricing policies by one firm on other members of the group.

It is not argued here that these pricing patterns will necessarily be established or even that a value of the index in any one year can convey the complexities of a multitude of decisions in past years which probably still influence the firm's present perspective. The point is that this index captures a dimension which is not measured by the Census ratio. An analysis of disparity of leading firm size can convey information about patterns of conduct and market performance that is not contained in the Census concentration ratio.

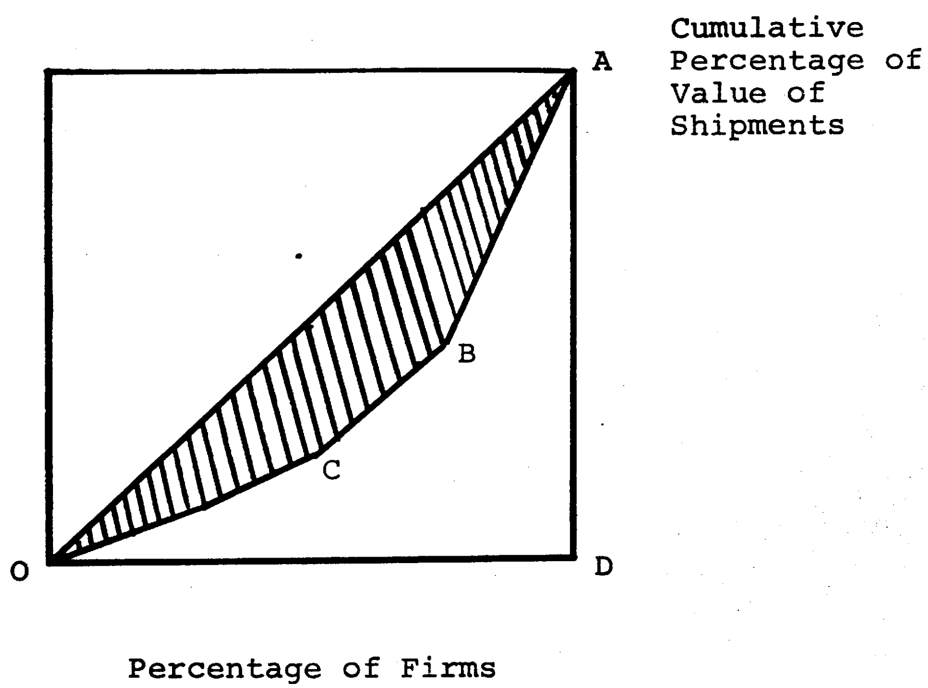
Gini Coefficient and Pietra Ratio

One common indicator of relative concentration is related to the extent to which the Lorenz curve deviates from the 45°

diagonal, which represents an equal-sized distribution.[17] The Gini coefficient for a particular industry can be geometrically represented as the ratio between the area between the diagonal and the Lorenz curve and the area below the diagonal. As shown in Diagram 2-D,

$$\text{GINI} = \frac{\text{Area of AOCB}}{\text{Area of AOD}}$$

DIAGRAM 2-D: The Gini Coefficient and the Lorenz Curve



The technique of "mean differences" is one method of computing Gini coefficients. The following relationships can

[17] A concise presentation of Gini coefficients and mean differences appears in E.M. Singer, op. cit., pp. 144-49.

be shown to exist:

$$\text{(a)} \quad MD(R) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |(X_i - X_j)|}{n^2}$$

$$\text{(b)} \quad MDR(R) = \frac{MD(R)}{(\sum_{i=1}^{n} X_i)/n} \qquad (5)$$

$$\text{(c)} \quad GINI = \frac{1}{2}(MDR(R))$$

where X_i: value of shipments of i^{th} firm

n: total number of firms in the industry

MD(R): mean difference with repetition[18]

MDR(R): relative mean difference with repetition

GINI: Gini coefficient

---

[18] Suppose an industry has 3 firms with the following sales vector, X = {30, 20, 5}. Let us calculate the matrix of pairwise differences, D, where each element of D is $|X_i - X_j|$.

$$D = \begin{pmatrix} 0 & 10 & 25 \\ 10 & 0 & 15 \\ 25 & 15 & 0 \end{pmatrix}$$

Differencing <u>with</u> repetition involves subtraction in both directions, i.e., it treats $|X_i - X_j|$ as distinct from $|X_j - X_i|$. For this hypothetical industry, there are $(3)^2$ "differences," or the elements of the matrix D. A simple calculation reveals that MD(R) = 100/9 = 11.11. Alternatively, differencing <u>without</u> repetition would produce only the 3 elements in the upper triangle of D. In this case, MD(not R) = 50/3 = 18.67.

The disadvantages of the Gini coefficient have been alluded to in the discussion of the Lorenz curve. It cannot distinguish between a large number of equally-sized firms and a monopoly. Moreover, the value of the coefficient is sensitive to the total number of firms in the industry. In general, the inclusion of marginal firms increases the degree of inequality registered by the index.

A related method of describing the dispersion of firm sizes is to compute the absolute differences between each established firm size and the average firm size, similar to the disparity index.

$$\text{(a)} \quad MD = \frac{\sum_{i=1}^{n} |(X_i - \bar{X})|}{n}$$

$$\text{(b)} \quad RMD = \frac{MD}{(\sum_{i=1}^{n} X_i)/n}$$

where X_i: value of shipments of i^{th} firm

$\bar{X}$: mean value of shipments of industry

n: total number of firms in industry

MD: mean deviation

RMD: relative mean deviation

Using this relative mean deviation measure, one can define the Pietra ratio:[19]

---

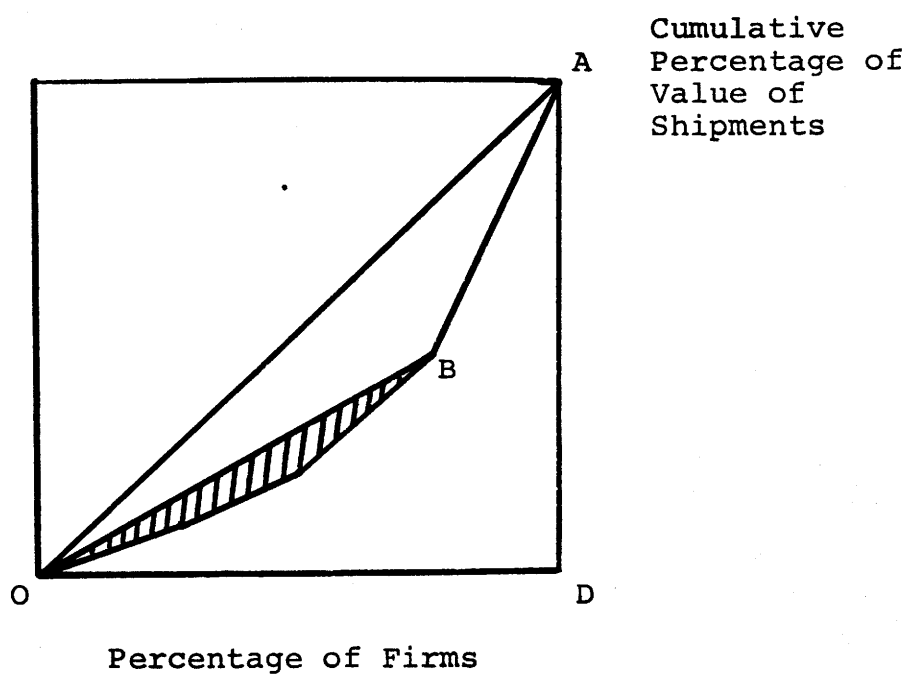
[19] <u>Ibid.</u>, pp. 149-52.

$$\text{PIETRA} = \frac{1}{2}(\text{RMD}) \quad (6)$$

In terms of the now familiar Lorenz curve, the Pietra ratio is the area of the maximum triangle that can be inscribed between the Lorenz curve and the line of equal distribution over the area under this diagonal. In terms of Diagram 2-E,

$$\text{PIETRA} = \frac{\text{Area of AOB}}{\text{Area of AOD}}$$

DIAGRAM 2-E: The Pietra Ratio and the Lorenz Curve



Hence, the Gini coefficient must be greater than or equal to the Pietra ratio. The shaded region is their difference in the above diagram.

Both of these measures represent unidimensional summary indices of points along the Lorenz curve and therefore are measures of relative concentration. It should be noted that the inability of the Gini coefficient to distinguish between equal size distributions across industries with different numbers of firms is not as damaging as some critics maintain. Similar problems arise with absolute concentration measures with respect to concentration curve intersections and the influences of intra-group size variations across different industries. The drawbacks of Gini and Pietra, especially when n is unknown, are not being minimized but being placed in the proper over-all perspective.

Herfindahl Index

One of the more commonly used measures of concentration is the Herfindahl index. It is defined as the sum of the squared relative firm shares across the industry, i.e.,

$$HERF = \sum_{i=1}^{n} \left(\frac{X_i}{X}\right)^2 \qquad (7a)$$

where X_i: i^{th} firm's value of shipments

X: industry's total value of shipments

Of more particular interest is the algebraic transformation of the above representation of the Herfindahl index into the following:

$$HERF = \frac{C^2 + 1}{n} \qquad (7b)$$

where C: coefficient of variation

The index can now be seen as a measure of dispersion. Note that when all the firms operating in an industry are the same size, C = 0 and HERF = (1/n); further, that when that industry is a monopoly, HERF = 1. The Herfindahl index can distinguish between various equal firm size distributions.

An interesting aspect of the index is its conversion into a numbers-equivalent.[20] With equal sized firms, HERF = (1/n), but also n = (1/HERF). The reciprocal of the Herfindahl measure can be viewed as the number of equal size firms required to generate a specific value of the index. If an industry has more firms than another but has a greater inequality among its firm shares, the question of which industry has a more competitive structure can be transformed into a comparison of the Herfindahl values converted into their respective numbers-equivalent.

Ranked Share Index

A variant of the Herfindahl index has been proposed by Hall and Tideman.[21] Note that the Herfindahl index implicitly weights each firm by its own relative share. An alternative that emphasizes the absolute number of firms in the industry is to premultiply each firm's relative share by its rank.

---

[20] See M.A. Adelman, "Comment on 'H' Concentration Measure as a Numbers-Equivalent," RESTat (February, 1969), pp. 99-101.

[21] M. Hall and N. Tideman, "Measures of Concentration," JASA (March, 1967), pp. 162-68. The reason for the specific form of the index results from a set of transformations which are required in order for the index to satisfy properties which the authors contend any concentration index should possess.

The inclusion of the absolute number of firms as rank weights in this measure of concentration eliminates a major source of concern with respect to the applicability of relative measures of concentration. The resultant index is:

$$THI = (2 \sum_{i=1}^{n} iX_i - 1)^{-1} \qquad (8)$$

where X_i: i^{th} firm's relative share

i: i^{th} firm's appropriate rank with the largest firm receiving a rank of 1

If a firm's relative share is measured in sales, then X_i represents the probability that a random dollar of sales will accrue to a firm of rank i. Therefore, iX_i represents the expected value of rank i and $\sum_{i=1}^{n} iX_i$ is the expected value of a random dollar of sales across all ranks in the industry. The greater this summation, the greater is the fluidity of consumer purchases and the lower is the concentration. So a transformed reciprocal is used. Also, note that in the case of equally-sized firms, the average rank is attached to each firm's share.

<u>Entropy Index</u>

In general, entropy is a measure of the expected information contained in a particular probability distribution. For example, consider a set of n events, $E_1, E_2, \ldots, E_n$, with the corresponding probabilities of occurrence, $p_1, p_2, \ldots, p_n$, one of which is certain to occur, or that $\Sigma p_i = 1$. Then the entropy contained in this probability distribution is $H(P) = \sum_{i=1}^{n} p_i \log (1/p_i)$, where

$0 \leq H(P) \leq \log n$. Entropy thus measures uncertainty since it attains its maximum value when all outcomes are equally likely.

Specifically, the entropy measure of concentration is a weighted average of the logarithms of the reciprocals of each firm's market share, i.e.,[22]

$$H(X) = \sum_{i=1}^{n} X_i \log(1/X_i) \qquad (9a)$$

where X_i: i^{th} firm's relative market share

If entropy is very high and firm size is measured by sales, then who captures a random dollar of sales is uncertain, which implies that concentration is very low. Conversely, as entropy approaches zero, uncertainty concerning the attainment of an additional dollar of sales is reduced and concentration is increasing. Thus, H(X) is an inverse measure of concentration.

Note that entropy is maximized when each firm has an equal share of the market. For any given industry with n number of firms, maximum entropy, H_{max}, is log n. Consider the following transformation:

$$H^* = \frac{H(X)}{H_{max}} \qquad (9b)$$

Here, actual entropy is expressed as a percentage of maximum entropy. The number of firms is directly embodied into the

---

[22]H. Theil, Economics and Information Theory, Ch. 8, contains the first proposed use of this measure. Also see J.L. Hexter and J.W. Snow, "An Entropy Measure of Relative Aggregate Concentration," SEJ (January, 1970), pp. 239-43.

index, given its relation to H_{max}. Cross-sectional comparisons can be more meaningfully undertaken using this revised measure. H^* has also been termed "relative entropy." It represents the extent to which an industry's sales are evenly distributed given the number of firms in that industry.

The entropy index and its derivative forms act as measures of competition. This derives from the fact that entropy captures the degree of uncertainty of consumer purchases across firms. Moreover, its decomposition properties have led to suggestions for the development of entropy measures which take into account the effects of multimarkets and of buyer identities on seller concentration.[23] Thus, entropy is a richer relative measure of concentration because it can be thought of as representing not only dispersion, but also uncertainty concerning the stability of a firm's existing relative market position.

Comprehensive Concentration Index

A measure has been constructed that considers both the relative size of the largest firm as well as dispersion in firm sizes.[24] The so-called comprehensive index of concentration is defined as follows:

$$CCI = X_1 + \sum_{i=2}^{n} X_i^2(1 + (1 - X_i)) \qquad (10)$$

[23] I. Bernhardt and K.D. MacKenzie, "Measuring Seller Unconcentration, Segmentation and Product Differentiation," WEJ (December, 1968), pp. 395-403.

[24] J. Horvath, "Suggestion for a Comprehensive Index of Concentration," SEJ (April, 1970), pp. 446-52.

This index attempts to describe relative dispersion and absolute magnitude. The latter is reflected in the direct inclusion of the leading firm's relative market share. The relative aspects are represented by a Herfindahl-like summation. The range of values for the index is X_1 to 1. There has been some criticism of the index.[25] If the leading firm with X_1 percent of the market acquires the j^{th} firm in the industry, the change registered by the index is $\Delta CCI = X_j(1 - X_j)^2 > 0$. But if the leading firm captures only C% of the j^{th} firm's share, then $\Delta CCI = C[3X_j^2 - (3C + 4)X_j + (C + 1)^2]$, which is negative if $(3C + 4)X_j > 3X_j^2 + (C + 1)^2$. Thus, acquisition of additional sales by the leading firm can result in a decline in concentration as measured by CCI. In other words, the value registered by the index can actually decline when the largest firm captures part of the market share of one of its rivals.

A Digression on Fitting Functions

An alternative approach to the measurement of concentration has been to fit theoretical distributions to an observed size distribution of firms for particular industries. The fitted function is taken to represent a summary measure of concentration. The major problem with such analysis is that a particular theoretical distribution does not have general applicability across a large array of industries. For example, in a major

---

[25] For development of this criticism and a rejoinder see respectively A.R. Horowitz, "...Comment," SEJ (April, 1972), p. 602 and J. Horvath, "...Reply," Ibid., pp. 602-04.

study Irwin Silberman notes that "...the lognormal distribution does not provide us with a universal description of the size distribution of sellers."[26] Thus, cross-sectional analysis based on the estimated parameters of a given distribution that is limited in scope, i.e., to only actually observed distributions, is inappropriate.

Other problems arise. In many instances conventional statistical tests are not applicable. For example, the χ^2 goodness-of-fit test statistic contains arbitrary observation groupings when comparing actual versus expected frequencies. Moreover, small observed values of χ^2, indicating a good fit, may also result in significant differences when the cumulative distribution is examined. This also is a problem when the Kolmogorov-Smirnov test is used. The sparsity of industry data compounds these problems.

Many attempts have been limited to the description of the upper tail of some hypothesized cumulative distribution. However, Richard Quandt states that "...in many instances we shall be unable to test hypotheses about the distribution of the data ... becuase of the smallness of the number of observations."[27]

Although an interesting approach, fitting functions for the purposes of cross-sectional analysis is limited at the

---

[26] I.H. Silberman, "On Lognormality as a Summary Measure of Concentration," AER (September, 1967), p. 822.

[27] R.E. Quandt, "On the Size Distribution of Firms," AER (June, 1966), p. 432.

present time. Its limited applicability across industries, data limitations, and the lack of a standard, recognized statistical test hinder the development of a general measure of concentration when this method is employed.

ALTERNATIVE MEASURES IN RETROSPECT

The previous examination indicates that many varied, and at times conflicting, approaches have been undertaken in the development of the optimal measure of concentration. It seems clear that absolute measures of concentration ignore firm size disparities as a dimension of concentration. Even though a Census ratio can be supplemented by an index of dominant firm disparity or an eight-firm marginal concentration ratio, the question arises as to the arbitrary numbers game inherent in these measures. If important size disparity influences exist, then why not consider twenty firms or fifty?

At the same time, relative concentration measures are subject to criticism. An inequality index may reveal that 5% of the established firms control 50% of an industry's total value of shipments, but it does not reveal whether this 5% is one firm or a thousand firms. The likelihood of monopolistic conduct and performance crucially depend upon which is the case. Thus, absolute firm numbers and their respective control are important. The ranked-share index and the relative entropy index are attempts to include the influence of the absolute number of firms into a relative concentration index. How much of an improvement these measures are over the more traditional indices is an empirical question.

The burgeoning list of concentration indices illustrates that the establishment of an accepted measure of concentration has not been achieved. The alternative measures do have some inter-relationship. For example, CR(4) represents one point on an industry's concentration curve, while GINI summarizes the Lorenz curve. Yet the two indices are related when the number of existing firms is known; that is, directional changes in each are predictable. Each represents the size distribution of firms. Each measure is consistent with the vague concept of fewness. Furthermore, when other pairwise comparisons of alternative indices are made, a similar phenomenon is observed.

However, although the differences among the indices are not substantial, there is a basic distinction between the absolute and the relative concentration measures. It is the emphasis on the latter of all industry members as opposed to the dominant core. Neither class of concentration measure (or the various ways of measuring each) can be established as correct on an a priori basis. The selection of the optimal concentration index must be decided by empirical investigation. The aim is to ascertain the relative efficacy of alternative measures with respect to their explanatory power concerning the emergence of non-competitive market performance. A survey of past empirical studies regarding this question will be made in the next chapter.

CHAPTER THREE: EMPIRICAL STUDIES CONCERNING ALTERNATIVE MEASURES OF CONCENTRATION: A SURVEY

There does exist a growing body of empirical evidence suggesting that concentration indices are largely identical with respect to their predictive ability regarding various market performance dimensions. This chapter reviews these studies, specifying the data set and the nature of their conclusions. Generally, the employed procedure is to examine various correlation coefficients among a set of concentration measures. At times, these are reproduced for expository purposes. Also, only cross-sectional industry comparisons are considered. There has been a substantial amount of analysis concerning trends in industrial concentration and the measurement of aggregate concentration. However, these are not directly applicable to the question of ascertaining the efficacy of alternative concentration indices.

Rosenbluth[1]

One of the first cross-sectional industry comparisons was made by Rosenbluth. Measuring firm size by fixed assets,

---

[1] G. Rosenbluth, Business Concentration and Price Policy, NBER (Princeton, 1955), pp.

Rosenbluth calculated the leading-firm through the four-firm Census concentration ratios for 26 industries as reported in the Federal Trade Commission's Concentration of Production Facilities, 1947. The Spearman correlation coefficients were computed. As can be seen from Table 3-A, the rank correlations range from .91 to .98; not an unexpected result considering that successive ratios include the preceding one by definition.[2]

TABLE 3-A: Spearman Coefficients, Rosenbluth (1)

	CR(1)	CR(2)	CR(3)	CR(4)
CR(1)	1.0	.966	.924	.914
CR(2)		1.0	.961	.939
CR(3)			1.0	.984
CR(4)				1.0

A more interesting aspect of Rosenbluth's analysis was his comparison of concentration measures across different index classes for a group of 96 Canadian manufacturing industries. He calculated a three-firm Census ratio based on the employment size dimension, the number of firms required to account for 80% of industry employment, and a truncated Herfindahl based on employment. The Spearman coefficients

---

[2]Rosenbluth also compared CR(4) and CR(8) employment ratios for 135 industries using 1935 data. The Spearman rho was .989 as expected.

are presented in Table 3-B.

TABLE 3-B: Spearman Coefficients, Rosenbluth (2)

	CR(3)	NUM	H
CR(3)	1.0	.981	.979
NUM		1.0	.979
H			1.0

The extremely high magnitude of the correlation coefficients indicates a collinear dimension among these various measures, implying that the information embodied in each structural index is largely replicative. In other words, the dimensions spanned by each index vector largely overlap.

Rottenberg[3]

Supportive findings were reported by Ira Rottenberg. His analysis focused on 48, 4-digit industries drawn from the 1954 Census of Manufactures. Two four-firm ratios were computed, one based on employment and the other based on value of shipments; also studied were an eight-firm ratio based on shipments and a truncated Herfindahl index. The Spearman rank correlations appear in Table 3-C. An interesting anomaly

[3] I. Rottenberg, "New Statistics on Companies and on Concentration in Manufacturing from the 1954 Census," American Statistical Association: Business and Economic Statistics Section (Proceedings, 1957), pp. 216-27.

TABLE 3-C: Spearman Coefficients, Rottenberg (1)

	$CR(4)^s$	$CR(4)^e$	CR(8)	H
$CR(4)^s$	1.0	.971	.986	.989
$CR(4)^e$		1.0	*	.959
CR(8)			1.0	.982
H				1.0

was also reported. Considering only industries with "high" concentration, i.e., the 13 of the 48 sample industries that had CR(4) > 50%, the following correlations were obtained.

TABLE 3-D: Spearman Coefficients, Rottenberg (2)

	$CR(4)^s$	$CR(4)^e$	CR(8)	H
$CR(4)^s$	1.0	.891	.718	.885
$CR(4)^e$		1.0	*	*
CR(8)			1.0	.529
H				1.0

Of course, these results are only suggestive because of the sample size, but there have been significant decreases in the observed correlations for the restricted sample.

Kilpatrick[4]

Another cross-sectional approach was used by Kilpatrick. He compared the explanatory power of Census ratios and several other measures with respect to industry profit rates. A measure of concentration was considered a more accurate indicator of its structural dimension if it was more highly correlated with the performance variable which it theoretically should affect. The data group chosen was 111 minor manufacturing industries defined by the IRS, usually considered equivalent to the 3-digit SIC code. Thus, the measures of concentration for each IRS industry were a weighted average of the ratios for the component 1954 4-digit Census industries.

Robert Kilpatrick calculated the concentration indexes under varying bases, i.e., using different weighting schemes. For example, he distinguished between product shipments and industry shipments as well as 5-digit product class and 4-digit industry components. Also, imports were added to domestic shipments in order to account for any trade influence. The procedure was to compute the partial correlation coefficients for these alternative measures with respect to 1953-57 average profit rate levels and 1949-54 change in profit rates.

The results showed that the correlation coefficients did not differ by more than 0.07 and therefore indicated no significant differences in explanatory power. This was

---

[4] R.W. Kilpatrick, "The Choice Among Alternative Measures of Industrial Concentration," RESTAT (May, 1967), pp. 258-60.

occasioned by the high inter-correlations among the measures themselves. For example, the range of correlation for CR(4) with the remaining indices was from .929 to .998.

Kilpatrick's conclusion: "This investigation has failed to label any concentration measure as the best structural indicator of market power. The comparison of alternatives has, however, provided much evidence that the particular choice is not crucial."[5]

Hall and Tideman[6]

All of the previous studies have focused on measuring significant differences between the Census concentration ratio, the truncated Herfindahl, and several other derivative indices of concentration. Their results indicate very high inter-correlations within the index set and one reveals insignificant differences in explanatory power of industry profit rates. Another empirical investigation by Marshall Hall and Nicolaus Tideman examines similar questions but expands the index set to include the ranked-share index.

Using 446, 4-digit industries as reported in the 1958 Census of Manufactures, they calculated the following simple and rank correlations:

---

[5]*Ibid.*, p. 260. Kilpatrick also modified his measures for geographical market segmentation and considered a Kaysen-Turner classification. In both cases the results were similar to the unadjusted Census ratio.

[6]M. Hall and N. Tideman, "Measures of Concentration," *JASA* (March, 1967), pp. 162-68.

TABLE 3-E: Rank Correlations, Hall and Tideman

	CR(4)	H	TH
CR(4)	1.0	.995	.904
H		1.0	.933
TH			1.0

TABLE 3-F: Simple Correlations, Hall and Tideman

	CR(4)	H	TH
CR(4)	1.0	.976	.883
H		1.0	.947
TH			1.0

The Herfindahl and ranked-share index observations were minimum estimates since it was assumed that firms were of equal size within each Census concentration class, similar to the previous studies. All measures computed firm size by value of shipments.

The high correlation coefficients corroborate other empirical research. Hall and Tideman conclude that the Census concentration ratio gives cross-sectional rankings similar to the Herfindahl and ranked-share indexes. Indeed, "...if H or TH is the correct measure of concentration, then the concentration ratio is certainly a good proxy."[7]

One aspect of the previous study is the report of dispersion measures for each index.

TABLE 3-G: Dispersion Measures, Hall and Tideman

	Range min	max	Coefficient of Variation
CR(4)	.03	1.0	57.28
H	.0012	0.25	90.29
TH	.00008	0.25	127.01

The authors note that the greater dispersion of the ranked-share index implies that, *ceteris paribus*, TH is a more sensitive measure of concentration.

The question arises as to what Hall and Tideman are trying to convey. Does a more "sensitive" index mean that one can more accurately rank industries and compare respective values at a lower expected cost of an incorrect decision? Is it that a more sensitive measure is better able to detect differences in concentration across industries, i.e., is it a better discriminator? Clearly, an index with greater

---

[7] *Ibid.*, p. 168.

relative dispersion is more likely to generate values which are more widely scattered over the index's range. But this reveals nothing about whether respective index values are significantly different. In fact, classical statistical inference directly incorporates dispersion into its test statistics; deviations from a hypothesized expected value are measured in standardized units. The distinction between dispersion and sensitivity is not well-defined. Therefore, the ranked-share index cannot be interpreted as a superior measure on the grounds of having the largest coefficient of variation.

Bailey and Boyle[8]

An expanded analysis of Hall and Tideman was undertaken by Duncan Bailey and Stanley Boyle on 1963 Census value of shipment data for 417, 4-digit SIC industries. The concentration index set included a one-, four-, and eight-firm Census ratio, an over-all, eight-, and twenty-firm Herfindahl, as well as a ranked-share index.

The cumulative concentration measures were computed for varying firm size distribution assumptions. These included the mean-share assumption which produces minimum values of the summary measures, the linear-mean-share assumption, and

---

[8]D. Bailey and S.E. Boyle, "The Optimal Measure of Concentration," JASA (December, 1971), pp. 702-06.

the constrained-mean-share assumption. The first assumption is the most widely used. Each firm is assumed to have a size equal to the mean of its size class. The linear-mean-share assumption states that the top four firms are distributed so that the fourth firm has the same size as the mean of the next smallest size class, while all other firms below the top four are distributed by the mean-share assumption. The final assumption is more complicated. The largest firm is allowed to have a size up to, but not above, 50 percent of the industry total. The remaining three firms in the top grouping have equal sizes, but these cannot be smaller than the mean size of the next class. All other firm sizes are equal to the mean of their respective size class. A hypothetical calculation is presented in an appendix to this chapter. In any event, the correlations between the assumed firm size distributions were so high that the specific size distribution appeared to be irrelevant for all measures. The results re-produced will therefore confine themselves to the mean-share assumption.

The procedure was to compute the simple pairwise correlations for the various measures of concentration as defined by the 1963 Census data. The results were consistent with previous empirical studies. All variants of the Herfindahl index were highly inter-correlated. All variants of the Census ratio exhibited high correlations. Moreover, CR(1) and CR(4) had a range of correlation coefficients from .96 to .98 with all variants of the Herfindahl. Thus, there was little difference between variants of H and those of the Census ratio.

The authors conclude that "...it appears that on grounds of economic efficiency alone, the use of CR(4) concentration estimates seems to be called for in most studies which require a structural variable ... [T]he analytical results using this variable [are] equal to or superior to any other which might be suggested, in terms of predictive ability."[9]

This conclusion appears to be overstated. First, the superiority of CR(4) estimates' explanatory power was not tested. Moreover, pairwise correlations between the ranked-share index and the remaining measures, although high, were smaller in magnitude. They ranged from .86 to .94. The ranked-share index seems to possess a slightly different dimensionality from the other indices. Whether these differences are significant and are indicative of a greater predictive capability is an open question.

Miller[10]

The emergence of marginal concentration ratios as an important structural indicator of market performance began with Miller. His sample consisted of 118 IRS minor group industries for which two profit rate measures were computed:

---

[9] Ibid., p. 706.

[10] R.A. Miller, "Marginal Concentration Ratios and Industrial Profit Rates: Some Empirical Results of Oligopoly Behavior," SEJ (October, 1967), pp. 259-67.

$$PR_1 = \frac{\text{Net Income + Interest Paid}}{\text{Total Assets}}$$

$$PR_2 = \frac{\text{Net Income}}{\text{Net Worth}}$$

where $r_{12} = 0.974$.

The mean of the annual rate over the period 1958-59 through 1961-62 constituted the observed profit rate for each industry. Census ratios were calculated as the weighted means of the appropriate 4-digit SIC industries using value of shipments. These were found in <u>Concentration Ratios in Manufacturing Industries, 1958</u>.

Initial bivariate regressions with CR(4) and CR(8) resulted in the following adjusted coefficients of determination:

TABLE 3-H: $\bar{R}^2$ for Bivariate Regressions, Miller

	CR(4)	CR(8)
PR_1	.068	.040
PR_2	.064	.035

The inclusion of various marginal concentration ratios with CR(4) produced statistically negative coefficients for MCR(8) and raised $\bar{R}^2$.

TABLE 3-I: $\bar{R}^2$ for Multiple Regressions, Miller

	MCR(8)	MCR(8),(20)	MCR(8),(20),(50)
PR_1	.120	.138	.130
PR_2	.123	.130	.122

TABLE 3-J: Simple Correlations, Miller

	CR(4)	MCR(8)	MCR(20)	MCR(50)
CR(4)	1.0	.451	-.049	-.671
MCR(8)		1.0	.623	-.109
MCR(20)			1.0	.444
MCR(50)				1.0

One of the more important aspects of the multiple regression results is the simple correlation matrix among the independent variables. Particularly note the relatively low correlation between CR(4) and MCR(8). Its magnitude indicates that for this data set the two measures are not highly collinear. Indeed, the incremental increase in $\bar{R}^2$ after the insertion of MCR(8) into the bivariate analysis, along with the continued significant coefficient of CR(4), further supports this claim.

Because of their statistical dependence, the relation between MCR(8) and CR(4) is not linear over the entire range of feasible values. Initial results of approximating this non-linear relation by a quadratic equation specification produced a significant improvement over the simple linear relation, i.e., for the quadratic $\bar{R}^2 = .468$ as compared with $r^2 = (.451)^2 = .203$.[11] The revised regression incorporated this relationship.

$$PR = g(CR(4), MCR(8))$$
$$MCR(8) = f(CR(4), CR(4)^2)$$
$$\therefore \quad PR = h(CR(4), CR(4)^2)$$

However, estimating the function "h" did not support the quadratic relation between PR_1 and CR(4).

Another equation specification defined a trichotomy on the range of CR(4) values.[12] These subsample groupings were as follows:

Low	$0 < CR(4) \leq 30$
Intermediate	$30 < CR(4) \leq 60$
High	$60 < CR(4) \leq 100$

It was hypothesized that over these classes the sign of the MCR(8) coefficient should change sign. For concentrated

---

[11] R.A. Miller, "Marginal Concentration Ratios: Some Statistical Implications - Reply," SEJ (October, 1969), pp. 199-201.

[12] R.A. Miller, "Marginal Concentration Ratios as Market Structure Variables," REStat (August, 1971), pp. 289-93. Also see N.R. Collins and L.E. Preston, "Price-Cost Margins and Industry Structure," REStat (August, 1969), pp. 271-86.

industries, recognized interdependence is "tight" so that deviations from a joint profit maximizing strategy is unlikely. Increases in MCR(8) will complement this relationship. The coefficient of the marginal ratio should be greater than zero. For unconcentrated industries, competitive rivalry is likely to prevail. Increases in MCR(8) can not be expected to depress industry profits. The coefficient should be positive but close to zero and possibly insignificant. However, for intermediate values of CR(4) or "loose" oligopolies, collusive agreements are not likely to persist. Increases in MCR(8) give rise to independent price behavior to the extent that a joint profit maximizing strategy is not likely to be implemented. The coefficient should be negative in this range.

Regression results reported by Miller using this tri-classification scheme are generally consistent with the above. For the low range, the MCR(8) coefficient was not significantly different from zero. Over the intermediate range, both CR(4) and MCR(8) had significant coefficients with MCR(8) < 0. In fact, there was no apparent relation between the two in this range. In the high subsample MCR(8) did have a negative and statistically significant coefficient contrary to the predicted sign. The author offers the explanation that the negative correlation between CR(4) and MCR(8) signals a necessary redistribution of firm shares. The effect of increases in CR(4) on profit rates also includes decreases in MCR(8).

What does Miller's evidence reveal? Are the regression results using the subsample conclusive? At first glance, marginal ratios appear to have an important influence on industry profits in the intermediate range. However, note that if the subsample observations are clustered between CR(4) = 50 and 60, then the reported negative sign could be the result of a statistical dependence relation. A look at Miller's data shows that the mean value of CR(4) for this group is 41.3 with a standard deviation of 9.1. Moreover, no significant relation between CR(4) and MCR(8) was found in this range, so that the latter's influence is largely independent of CR(4) effects on profits. Intermediate subsample results are more convincing when these relations are also considered.

The results for the high subsample are not as convincing. A significant negative relation between CR(4) and MCR(8) is present. The reported mean value of CR(4) is 68.8 with a standard deviation of 10.1. The negative coefficient is not easily interpretable because of statistical dependence between the regressors. It is more appropriate to use Henning's corrected version of the marginal ratio in this group.

More recently, Miller has examined the relative predictive power of the four-firm Census ratio and several entropy measures.[14] Using 1958 price cost margin data as a

---

[14] R.A. Miller, "Numbers Equivalents, Relative Entropy, and Concentration Ratios: A Comparison Using Market Performance," SEJ (July, 1972), pp. 107-12.

measure of industry profitability, 25 4-digit industries were included in Miller's sample.[15] The zero order correlation between the four-firm Census ratio and relative entropy was -.244. Furthermore, regression results indicate that both measures have significant and independent effects on industry profits. These results, taken together, suggest that these measures of concentration capture diverse aspects of this structural dimension.

Summary

Past empirical investigation tends to indicate that Census concentration ratios and truncated Herfindahls are highly intercorrelated and are likely to affect industry performance in a similar fashion. In particular, they are likely to possess the same explanatory power with respect to industry profitability. Other studies have shown that the ranked-share index falls into the same category. However, there is some evidence suggesting that marginal concentration ratios may have independent influences on performance outside of Census ratios. The same can be said of relative entropy measures of concentration.

The next chapter will systematically analyze these questions using a data set that was not employed in any of

---

[15] The entropy data were obtained from I. Horowitz, "Numbers Equivalent in U.S. Manufacturing Industries: 1954, 1958, and 1963," SEJ (April, 1971), pp. 396-408. The performance variable was originally tabulated by Collins and Preston, Concentration and Price Cost Margins in Manufacturing Industries (University of California Press, 1968), Appendix A, pp. 119-48.

TABLE 3-K: Summary Listing of Empirical Studies

Study	Data Source	Sample	Measures Employed
Rosenbluth	FTC Report on Concentration of Production Facilities, 1947	26, 4-digit industries	CR(1), CR(2), CR(3), CR(4)
	Canadian Manufacturing, Dominion Bureau Stat.	96 Canadian industries	CR(3), H, 80% firm number
Rottenberg	Census of Manufactures, 1954	48, 4-digit industries	CR(4), CR(8), H
Kilpatrick	IRS Sourcebook of Income Stat. and Census of Manufacturers, 1954	111 minor group industries	CR(4), CR(8), CR(20), and several derivative measures
Hall and Tideman	Census of Manufactures, 1958	446, 4-digit industries	CR(4), H, TH
Bailey and Boyle	Census of Manufactures, 1963	417, 4-digit industries	CR(1), CR(4), CR(8), H, H(8), H(20), TH
Miller	IRS Sourcebook of Income Stat., 1958-1961; Census of Manufactures, 1958	118 minor	CR(4), CR(8), MCR(8), MCR(20), MCR(50)
	Collins and Preston Price-Cost Margin Data	25 4-digit industries	CR(4), and several forms of the entropy measure
	Horowitz Entropy Data		
	Census of Manufactures, 1958		

the aforementioned studies. Moreover, an expanded set of concentration indices will be considered with specific emphasis on their inter-relationships and a comparison to previous empirical research.

APPENDIX: Hypothetical Calculation of Firm Shares Under the Various Distribution Assumptions Used by Bailey and Boyle.

Suppose the following data for a hypothetical industry is reported by the Census: CR(4) = 0.60, CR(8) = 0.84, CR(20) = 1.00 and total industry value of shipments is 1000. From this data, generate the total value of shipments in each size class.

firm rank	1-4	5-8	9-20
value of shipments	600	240	160

The following individual sizes are obtained using the alternative assumptions.

Firm Rank	Value of Shipments Mean-Share	 Linear-Mean	 Constrained-Mean
1	150	240	480
2	150	180	60
3	150	120	60
4	150	60	60
5	60	60	60
6	60	60	60
7	60	60	60
8	60	60	60
9	13 1/3	13 1/3	13 1/3
10	13 1/3	13 1/3	13 1/3
11	13 1/3	13 1/3	13 1/3
⋮	⋮	⋮	⋮
20	13 1/3	13 1/3	13 1/3

CHAPTER FOUR: A CROSS SECTIONAL ANALYSIS OF SELECTED PRODUCT CLASSES AND OF ALTERNATIVE MEASURES OF CONCENTRATION

The choice of an index of concentration is an empirical question since economists have been unable to reach a consensus on the attributes that a measure of concentration should possess. All of the suggested indices discussed in Chapter II are representations of the structural characteristic of the number and size distribution of firms in an industry. Yet, there is no a priori basis upon which a selection among these alternative measures can be made. Moreover, many of the empirical studies presented in Chapter III conclude that indices of concentration are largely identical with respect to their predictive ability regarding various dimensions of market performance. These studies maintain that the Census concentration ratio is the index that should be utilized in empirical investigations, primarily because of its availability or its computational ease.

However, these studies are not all-encompassing since they examine a limited number of concentration indexes. Most consider only Census ratios and truncated Herfindahls. Others include marginal ratios or a ranked-share index. But there is no study that analyzes simultaneously the entire set of concentration measures that have been proposed on an identical data set. This chapter will rectify this situation by investigating a comprehensive set of alternative measures of concentration derived from recently published data by the Federal Trade Commission.

The data set and derivative sample listings are first presented. This is followed by an analysis of the selection procedure indicating possible bias. The next section sets forth the analytical procedure and enumerates the computed measures of concentration. An interpretation of these results follows, with a particular emphasis on comparisons to previous empirical research. Some of the questions to be considered are the relation between the Herfindahl index and the approximating truncated Herfindahl and the relation between Census ratios and marginal ratios with quadratic approximations. Of course the ultimate question concerns the measurement of the structural dimension of concentration. Are Herfindahls and Census ratios similar in their information content of this dimension so that their predictive abilities regarding performance are not significantly different? Do disparity indexes or entropy measures provide different aspects of market concentration? Can these various indices of concentration be viewed as acting in a complementary fashion, each measuring an aspect of concentration that another index lacks? Or does it make any significant difference?

The Data Source and Sample Definition

The basic data source for cross-sectional studies of concentration has been the Census of Manufactures. The 4-digit SIC level is the predominant category analyzed, although weighted averages with IRS minor group profit data are also common. No study has examined product class data, largely because of its non-availability. However, in January 1972, the Federal Trade Commission published a statistical report on the value of shipments by product class for the 1,000 largest manufacturing companies in 1950. This raw data constitutes the basic set that will be utilized in this analysis.

From this product class shipment data a core sample of 78 classes was selected, termed Sample I. Further refinements in sample definition produced three other samples. The selection criterion for the core sample was a pre-specified coverage ratio, i.e., the ratio of the sum of the company shipments allocated to a product class over the total value of shipments for that product class. Sample II increased this coverage ratio. These samples were then revised by using a minimum firm number criterion. Table 4-A summarizes these sample definitions, while Table 4-B lists the selected product classes.

A brief word about the selection criteria is necessary. In general, a coverage rule criterion is necessary since it

TABLE 4-A: Sample Definition and Coding

Sample I	80% Coverage
Sample II	90% Coverage
Sample I-A	80% Coverage and Num. > 8
Sample II-A	90% Coverage and Num. > 8

allows the inclusion of those product classes for which the available individual firm data largely describe the distribution of total shipments. This is particularly crucial for the accuracy of relative concentration measures. Sample estimates based on individual firms which account for a substantial proportion of product class shipments can be expected to minimize computational bias. However, restricting sample observations to those product classes that have a 100% coverage ratio is not necessarily desirable. The inclusion of "marginal" firms, those with a small market share, will not significantly alter the observed value of most indices. Moreover, these firms direct a measurable effect on market performance that is largely inconsequential. The exclusion of these firms will effect a minimal information loss. The 80%-rule, although arbitrary, is a reasonable criterion.

We now turn to a more detailed analysis of the selection procedure. Even though the selection criteria may be consistent with minimizing computational bias, it nonetheless is likely to

TABLE 4-B: Product Class Listing

Product Class Number	Name of Product Class	Firm Number	Coverage (%)	Sample Coverage
20111	Fresh Meats	14	93	All
20336	Canned Baby Food	6	95	I,II
20337	Canned Soups	5	87	I
20430	Cereal Foods	17	88	I,I-A
20730	Chewing Gum	4	92	I,II
20852	Distillers' Grain	9	86	I,I-A
20853	Bottled Liquor	9	82	I,I-A
21110	Cigarettes	6	99	I,II
22741	Linoleum	5	93	I,II
26113	Bleached Sulfate Pulp	18	96	All
26114	Unbleached Sulfate Pulp	25	81	I,I-A
26125	Container Board	40	96	All
26993	Sanitary Food Containers	16	94	All
28120	Alkalies and Chlorines	28	99	All
28251	Acetate Yarn	4	99	I,II
28252	Rayon Yarn	6	89	I
28421	Synthetic Organic Deter'ts	24	85	I,I-A
28250	Inorganic Color Pigments	20	86	I,I-A
28820	Linseed Oil Mill Products	7	87	I
28933	Dentifrices	8	84	I
28960	Compressed Gas	24	89	I,I-A
29111	Kerosine	41	95	All
29112	Distillate Fuel Oil	44	90	All
29113	Residual Fuel Oil	42	90	All
29118	Unfinished Petroleum Prod.	22	96	All
29119	Finished Petroleum Prod.	40	92	All

(Cont.)

TABLE 4-B: (Cont.)

Product Class Number	Name of Product Class	Firm Number	Coverage (%)	Sample Coverage
29321	Coke, Screenings, Breeze	29	90	All
29323	Coke-Oven Products	28	95	All
29924	Lubricating Oils	31	97	All
29925	Lubricating Oil-Base Stocks	22	99	All
30110	Tire and Tubes	12	97	All
32112	Window Glass	4	84	I
32311	Laminated Glass	4	90	I,II
32720	Gypsum Products	7	95	I,II
32925	Asbestos-Cement Shingles	7	86	I
33112	Pig Iron	29	99	All
33120	Steel Ingots	34	87	I,I-A
33121	Semifinished Steel Shapes	37	96	All
33122	Steel Plates	22	99	All
33123	Hot-Rolled Sheet and Strip	33	81	I,I-A
33124	Tin, Terneplate, Bulk Plate	12	97	All
33125	Hot-Rolled Bars, Bar Shapes	27	90	All
33126	Structural Shapes and Piling	10	98	All
33128	Steel Mill Transfers	39	97	All
33526	Aluminum Plates	6	90	I,II
33527	Rolled Aluminum	9	95	All
33927	Steel Wire	26	82	I,I-A
33937	Welded Rivet Pipe	21	84	I,I-A
33993	Cold-Rolled Sheet and Strip	29	92	All
33994	Non-Ferrous Forgings	11	83	I,I-A
34111	Metal Cans	18	86	I,I-A
34212	Razor and Razor Blades	4	99	I,II

(Cont.)

TABLE 4-B: (Cont.)

Product Class Number	Name of Product Class	Firm Number	Coverage (%)	Sample Coverage
34415	Lath, Other Building Tools	19	92	All
34612	Other Enamel Products	10	98	All
34718	Incandescent Street Lighting	5	91	I,II
34893	Wire Fencing, Fence Gates	14	81	I,I-A
34914	Steel Barrels, Drums	17	80	I,I-A
35112	Steel and Hydraulic Turbines	5	98	I,II
35192	Deisel Engines(ex.Bus or Tur)	20	86	I,I-A
35193	Gas Engines	6	87	I
35211	Wheel-Type Tractors	14	94	All
35212	Track-Type Tractors	6	98	I,II
35720	Typewriters	6	98	I,II
35811	Household Washing Machines	19	90	All
35851	Household Refrigerators	15	98	All
35852	Home and Farm Freezers	12	83	I,I-A
35855	Condensing Units	12	95	All
36153	Transformer Parts, etc.	7	81	I
36172	Arc Welding Electrodes	8	99	I,II
36217	Electrical Appliance Parts	24	98	All
36612	House-Radio Receivers	21	81	I,I-A
36640	Telephone-Telegraph Equip.	7	93	I,II
37171	Passenger Cars	9	99	All
37172	Truck Tractors and Chassis	15	95	All
37318	Self-Propelled Ships(non-mil)	6	88	I
37411	Locomotives-Railroad Type	5	99	I,II
37412	Locomotives-Switching Type	5	99	I,II
38613	Photographic Film	6	97	I,II

be subject to other biases. The following section examines this possibility, analyzing survey standard errors of total shipment estimates, coverage rates and the tendency to include large average firm-sized classes, and the resultant effects on the range of index values.

Analysis of the Selection of Product Classes

As previously stated, the selection of product classes was initially based on varying coverage ratios. For example, Sample I was limited to those classes in the 1972 FTC report that could account for at least 80% of the total value of shipments of that product class. The latter data was obtained from Report of the Federal Trade Commission on Industrial Concentration and Product Diversification in the 1,000 Largest Manufacturing Companies, 1950.[1] These constituted only estimates of the total value of shipments. The sampling procedure was to survey one-sixth of all manufacturing establishments, including all those known to employ more than 250 persons.[2] In most cases a corresponding standard error of estimate is reported indicating the reliability of the estimated value. An effort was made to screen carefully product classes when these errors were available prior to the implementation of the coverage criterion.

[1]This report was published in January 1957. The relevant section of the report is Appendix C - Table 1, pp. 152-254.

[2]See Appendix A - Technical Note 3 of the aforementioned report, pp. 128-30.

TABLE 4-C: Frequency of Standard Errors for Total Value of Shipment Data

	a	b	c	1	2	3	4	5	6	7	15
Sample I	4	2	14	25	7	7	10	5	2	1	1
Sample II	3	2	10	17	6	4	6	1	-	1	1
Sample I-A	4	2	13	13	5	4	6	1	2	1	-
Sample II-A	3	2	9	9	4	2	3	-	-	1	-

[a]Standard error not available

[b]Aggregate value reported to FTC exceeded mean estimate of Census Bureau for product class or one of its components

[c]Estimate obtained from special survey believed to cover <u>all</u> producers

As can be seen from Table 4-C, one-half of the included product classes for Sample I (39/78) have standard errors less than or equal to 1%, while almost nine-tenths of the sample has estimates whose errors do not exceed 5%. Of interest is Arc Welding Electrodes, product class number 36172. It has a reported standard error of 15% but is included since the surveyed firms virtually account for the entire output for the product class. Thus, the probability that the true, total value of shipments would fall below an 80% coverage under the normality assumption is less than 10%.

For Sample II, a similar frequency distribution is observed. Over half of the included product classes have estimates for which the standard errors are not greater than 1%; almost 90% of the sample is in the not-more-than-5% category. Table 4-D summarizes these characteristics for all four samples.

TABLE 4-D: Cumulative Frequency of Standard Errors for the Total Value of Shipment Data

	≤1%	≤5%
Sample I	50	87
Sample II	53	87
Sample I-A	50	82
Sample II-A	54	82

Thus, the estimates of the total value of shipments for the selected product classes are fairly accurate.

After screening total value data, coverage ratios were calculated for the companies classified in each product class, and included on the 80% rule. The coverage rule criterion is similar to the selection procedure used in Ralph Nelson's study.[3] In that study a 4-digit SIC industry is selected if 90% of its shipments are accounted for by component 5-digit product classes that are "well-covered." To be well-covered, a product class must have establishments of at least 100 employees in which more than 70% of the total value of

[3]R.L. Nelson, Concentration in Manufacturing Industries of the United States (New Haven: Yale University Press, 1963), pp. 26-27.

shipments is produced. Sample II conforms with the Nelson criterion, although it applies directly to the 5-digit SIC level. Sample I, which uses 80% as a minimum coverage rule, is also consistent with Nelson since at the 5-digit level his coverage rate is only 70%.

In conjunction with these two coverage rules, 80 and 90 percent, the number of firms within each product class constitutes the basis for the remaining derivative samples, i.e., exclude those product classes in Sample I and Sample II that do not have more than eight component companies. The resultant sets of observations represent Sample I-A and II-A respectively. The minimum firm number criterion reflects the computational requirements for the eight-firm marginal concentration ratio, one measure of concentration that has shown independent influences on performance in previous empirical studies. Thus, the selection criteria is not arbitrary. Rather, it is the necessary consequence of the computational process of various measures of concentration.

A further look at the distribution of firm numbers is presented in Table 4-E. Twenty-seven product classes are deleted from Sample I, while eighteen observations are eliminated from Sample II. In both cases, approximately one-third of the original data set is omitted.

TABLE 4-E: Interval Distribution of Firm Number in Each Product Class

	Total	1-4	5-8	9-20	21-50
Sample I	78	5	22	27	24
Sample II	51	4	14	16	17
Sample I-A	51	-	-	27	24
Sample II-A	33	-	-	16	17

However, using coverage rules and firm number as criteria for selection can produce certain biases. For example, consider a comparison of all 1950 product classes and sample product classes by major industry group frequency. Examining Table 4-F, one can see that a disproportionate share of the sample product classes comes from Petroleum and Coal Products (29) and Primary Metal Industries (33). These two major industry groups account for 30.7%, 43.4%, 37.3%, and 54.5% of the total across each sample as compared to 8.6% of the overall 1950 Annual Survey of Manufacturers.

This pattern of coverage primarily results from the data source itself. Considering the 1,000 largest manufacturing companies themselves, major industry groups such as Apparel, Printing, and Leather Products are almost necessarily excluded by the coverage criterion. In fact, major groups 23, 24, 25, 27, 31, and 39 have no observations in any sample. Table 4-G shows the percentage distribution of shipments and firm number for

TABLE 4-F: Pattern of Coverage by Major Industry Group

	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
Sample I	7	1	1	-	-	-	4	-	8	9	1	-	4	15	7	10	5	5	1	-
Sample II	4	-	-	-	-	-	4	-	5	9	1	-	-	14	5	6	3	2	-	-
Sample I-A	3	1	1	-	-	-	3	-	1	9	1	-	2	10	4	7	3	3	1	-
Sample II-A	1	-	-	-	-	-	3	-	1	9	1	-	-	9	2	4	2	1	-	-
Overall	65	5	42	13	22	13	36	7	43	20	8	13	33	31	58	86	45	22	15	20

TABLE 4-G: Percentage Distribution of Value of Shipments and Firm Number by Major Industry Group for the 1,000 Largest Manufacturing Companies in 1950

	20	21	22	23	24	25	26	27	28	29
Value of Shipments	14.2	1.9	4.6	0.8	0.8	0.5	4.1	0.9	8.4	7.8
Firm Number	14.7	2.0	4.4	0.6	0.7	0.3	3.7	1.0	8.1	7.2

	30	31	32	33	34	35	36	37	38	39
Value of Shipments	2.0	0.8	2.0	14.5	4.1	6.8	5.4	18.7	1.0	0.8
Firm Number	2.5	0.7	2.0	15.7	2.9	5.9	5.8	20.4	0.9	0.5

these 1,000 companies across each major industry group.[4] Each major group that is omitted from sample coverage accounts for less than 1% of the total value of shipments for the 1,000 largest companies. A similar pattern emerges for their number. Because of such small percentages, the likelihood of selecting product classes from these major industry groups is extremely small.

From these various tables it can be seen that the major group frequency distribution is similar to the one for the 1950 product class data set, but is different from the frequency distribution for all product classes defined in the 1950 Annual Survey of Manufacturers. The disparities are inherent in the basic data set itself, although the selection procedure does have a tendency to magnify them. The resultant effect is a movement toward 2-digit groupings that have a large average company size. For example, Apparel, Printing, and Furniture are known to have unimportant scale economies. Moreover, small firms in these groupings have had a good earnings record.[5] Yet

---

[4]See Appendix D, Table 1, of the 1957 FTC report previously cited.

[5]See H.O. Stekler, Profitability and Size of Firm (Berkeley: University of California Institute of Business and Economic Research, 1963). Also, M. Hall and L.W. Weiss, "Firm Size and Profitability," REStat (August, 1967), pp. 319-31. The issue of firm size and profitability is far from settled. The major problems concern a measure of profitability and the roles of absolute size versus relative size of firms. Many conditions needed to generate behavioral patterns that result in a significant relation between absolute size and profitability are traceable to relative size implications. The point is that the 2-digit groupings excluded in this study are those with lower average company "size," however measured, than those groupings that are included in the sample.

all of these 2-digit groupings are entirely excluded from the core sample.

In summary, coverage rates and minimum firm number rules constitute selection criteria that are required in order to calculate various proposed indices of concentration. The application of such criteria on the 1950 product class data has produced sets of observations which reflect generally this data source, but which also exclude various major industry groups completely. These groupings possess insignificant scale economies and, therefore, are likely to be smaller company-sized than those that are included. This can potentially restrict the range of various indexes, because industry groups mostly comprising industries with large firm numbers and low concentration are not as equally represented as classes generally encompassing industries with small numbers and higher concentration (see Table 4-I below). Such exclusions limit the interpretability of cross-sectional analysis and should always be kept in mind.

Analytical Procedure

One major goal of this investigation is to determine if an optimal measure of concentration exists. The choice of an index can be predicated on its predictive ability of market performance dimensions. One approach that has been widely utilized is to compute correlation matrices for a set of alternative concentration measures. If these correlations are high, then the indexes embody or contain similar information. Although possible, their predictive accuracy is not likely to be significantly different. If these correlations are low, then the indexes characterize different aspects of the structural dimension of concentration; a significant difference in predictive power is likely. Alternatively, the indices embody the same characteristics, but also contain noise superimposed onto their information signal. In this case, predictive power will not vary significantly even though low correlations are observed. Of course, high pairwise correlations do not necessarily imply equal predictive capability. Likewise, low intercorrelations do not necessarily imply different predictive ability. The examination of index correlations can only address the existence question, i.e., whether or not the various indexes are diverse characterizations of the concentration dimension, and if so, imply a likelihood about unequal predictive capabilities. A simple listing of the index set follows:

TABLE 4-H: Concentration Index Listing and Coding

CR(4)	Census Four-Firm Ratio
MCR(8)*	Corrected Eight-Firm Marginal Ratio
DI(4)	Four-Firm Disparity Index
GINI	Gini Coefficient
PIETRA	Pietra Ratio
HERF	Herfindahl Index
THI	Ranked Share Index
ENTRO	Relative Entropy Index
CCI	Comprehensive Index
MCR(8)	Uncorrected Eight-Firm Ratio
THERF	Truncated Herfindahl Index

These measures have been previously defined in Chapter II. For comparison purposes, the disparity index has been transformed so that its range is between 0 and 1 inclusive. The necessary transformation is

$$DI(4) = [\frac{1}{4} \sum_{i=1}^{4} (X_i - \bar{X})/\bar{X}]/1.5, \text{ where } DI(4)_{max} = 1.5.$$

All other measures are calculated by the appropriate formulas presented in Chapter II.

The resultant numerical values of each index for the sample product classes appear in Table 4-I. One of the more striking characteristics of the observed values is the relatively high magnitude of CR(4). In fact, over half of the product classes in the core sample have a four-firm Census ratio that is greater than 70%. This is a consequence of the selection procedure and its application to this particular data set. For example, by examining Table 4-J, note that increasing the coverage ratio raises the mean value of CR(4), although not significantly, while for a given coverage ratio increasing minimum firm number reduces the mean value of CR(4). In any event, the mean value of CR(4) across all samples is inordinately high and further tends to indicate a bias to relatively larger firms, as alluded to in the previous section.

TABLE 4-I: Estimated Values for Each Concentration Index

Product Class Number	Measures of Concentration								
	CR(4)	MCR(8)*[a]	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
20111	.6969	.5226	.3431	.6005	.4739	.1629	.1960	.7211	.4320
20336	.8657		.3123	.4276	.3373	.2464	.3122	.8141	.5560
20337	.8705		.8445	.7300	.6832	.5961	.9559	.3228	.7784
20430	.7269	.4249	.1317	.6958	.5933	.1429	.2266	.6336	.3974
20730	.9190		.4353	.4273	.3265	.3525	.4940	.7326	.6598
20852	.6818	.4782	.0960	.4277	.3565	.1292	.2361	.7846	.3640
20853	.6581	.4139	.1999	.4356	.3569	.1246	.2504	.7766	.3611
21110	.8790		.2011	.3304	.2731	.2207	.2521	.8932	.5324
22741	.9015		.4044	.4704	.3843	.3092	.4223	.7485	.6171
26113	.6014	.5367	.1969	.5562	.4104	.1162	.1305	.7935	.3440
26114	.4620	.2989	.2881	.5937	.4310	.0768	.1255	.6929	.2736
26125	.4413	.3647	.3920	.6011	.4338	.0846	.0653	.7868	.2926
26993	.5811	.5926	.1167	.5204	.3944	.1059	.1398	.7996	.3174
28120	.5993	.6564	.1767	.7208	.5883	.1182	.1299	.6973	.3442
28251	.9984		.3327	.3021	.2496	.3388	.3590	.8821	.6614
28252	.8399		.3165	.4644	.3611	.2334	.3626	.7668	.5346
28421	.7448	.2249	.5593	.8454	.7267	.2800	.3331	.4353	.5571
28520	.6468	.3834	.3958	.7146	.5549	.1619	.2100	.6144	.4236
28820	.8125		.3957	.5568	.4568	.2316	.3881	.6839	.5376
28933	.8228		.5094	.7049	.5585	.3068	.5512	.5180	.5880

(Cont.)

TABLE 4-I: (Cont.)

Product Class Number	CR(4)	MCR(8)*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
	Measures of Concentration								
28960	.8211	.2482	.2921	.8400	.7617	.2063	.3048	.4794	.5019
29111	.4747	.4169	.2246	.6786	.5265	.0784	.0801	.7452	.2636
29112	.3832	.5550	.1895	.6705	.5419	.0592	.0770	.7317	.2212
29113	.3641	.5944	.0840	.6437	.5075	.0525	.0750	.7469	.1863
29118	.5734	.5319	.2146	.6161	.4727	.1089	.1237	.7653	.3318
29119	.5142	.3475	.2896	.7060	.5523	.0947	.0941	.6904	.3106
29321	.5630	.3513	.4062	.6793	.5231	.1289	.1213	.6780	.3725
29323	.6448	.4143	.4175	.7067	.5498	.1663	.1285	.6727	.4292
29924	.4144	.5023	.2256	.5393	.4031	.0671	.0727	.8327	.2332
29925	.5387	.5834	.0795	.5722	.4633	.0974	.1078	.8082	.2986
30110	.7812	.6059	.1182	.5428	.4738	.1650	.1895	.7785	.4370
32112	.8333		.3703	.3550	.2778	.2442	.5043	.7877	.5562
32311	.9039		.4443	.4409	.3333	.3387	.5194	.7378	.6485
32720	.9171		.4825	.6672	.5530	.3753	.4678	.5691	.6734
32925	.6851		.3218	.3687	.3079	.1595	.2750	.8214	.4280
33112	.6621	.4370	.4021	.6909	.5568	.1726	.1173	.6974	.4400
33120	.4486	.4104	.0491	.6466	.5058	.0647	.0993	.7195	.2168
33121	.7068	.3930	.5869	.8037	.6529	.2595	.1447	.5520	.5294
33122	.7148	.5879	.5815	.7224	.5572	.2672	.1645	.6311	.5420
33123	.4619	.3933	.3394	.7065	.5532	.0847	.1299	.6367	.2917

(Cont.)

TABLE 4-I: (Cont.)

Product Class Number	CR(4)	MCR(8)*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
	Measures of Concentration								
33124	.7504	.7706	.3954	.6204	.4717	.2209	.2092	.7112	.5092
33125	.9273	.6529	.5406	.7321	.6598	.3630	.3857	.5236	.7025
33126	.6299	.3557	.2413	.7148	.5731	.1244	.1411	.6542	.3637
33128	.6681	.4134	.4459	.7907	.6264	.1883	.1270	.6075	.4586
33526	.8943		.3344	.5680	.4782	.2952	.4463	.6240	.6167
33527	.9310	.2844	.6283	.7893	.6938	.4929	.5703	.4044	.7475
33927	.5087	.2638	.4791	.6582	.4746	.1189	.1411	.6394	.3540
33937	.5298	.4370	.2997	.6454	.5017	.1026	.1644	.6773	.3270
33993	.4781	.4535	.2099	.6263	.4788	.0797	.1007	.7528	.2618
33994	.7004	.4327	.3747	.6323	.5044	.1777	.3120	.6415	.4523
34111	.8112	.1820	.5639	.8409	.7568	.2922	.4288	.4155	.6061
34212	.9882		.4623	.4076	.3468	.4085	.4292	.7815	.7012
34415	.4872	.5762	.1654	.5317	.4124	.0884	.1232	.7922	.2848
34612	.7158	.9050	.2994	.4636	.3258	.1815	.1895	.8244	.4645
34718	.8646		.2776	.3392	.2714	.2273	.3418	.8540	.5376
34893	.5228	.4307	.2793	.5262	.3838	.1109	.1942	.7139	.3428
34914	.6208	.3565	.1703	.6740	.5409	.1130	.2363	.6274	.3408
35112	.9734		.5652	.5842	.5148	.4431	.4934	.5996	.7574
35192	.5609	.3734	.1951	.5953	.4596	.0983	.1465	.7236	.3149
35193	.8571		.4650	.6275	.5040	.3229	.5523	.5766	.6221

(Cont.)

TABLE 4-I: (Cont.)

Product Class Number	CR(4)	MCR(8)*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
			Measures of Concentration						
35211	.6320	.6959	.2188	.5481	.4153	.1339	.1692	.7680	.3854
35212	.9715		.4627	.6138	.1705	.4063	.4421	.0643	.7066
35720	.8326		.1743	.3066	.2045	.2103	.2490	.8615	.5177
35811	.5508	.5547	.1539	.5984	.4651	.0989	.1473	.7372	.3050
35851	.5894	.6732	.1909	.4965	.3635	.1188	.1349	.8197	.3448
35852	.5429	.4596	.1439	.4260	.3243	.0918	.1815	.7865	.2874
35855	.8375	.6131	.4301	.6915	.5582	.2567	.2884	.6217	.5696
36153	.7577		.3753	.5611	.4809	.1965	.4422	.6516	.4856
36172	.8601		.2866	.5035	.4164	.2477	.2517	.7807	.5599
36217	.8098	.4158	.5401	.7952	.6691	.3276	.2080	.5248	.6145
36612	.4462	.4127	.0895	.5229	.4057	.0637	.1258	.7412	.2252
36640	.9289		.8180	.8110	.7205	.6550	.8636	.2582	.8241
37171	.9009	.9080	.3612	.6223	.5345	.2989	.2960	.6750	.6165
37172	.7666	.5620	.3745	.6807	.5435	.2115	.2233	.6569	.5022
37318	.8449		.4353	.5562	.4233	.2921	.4464	.6707	.5881
37411	.9625		.5973	.5680	.4741	.4910	.4654	.6263	.7451
37412	.9785		.3078	.4395	.3405	.3337	.3568	.7718	.6692
38613	.9784		.7009	.7217	.6087	.5997	.6199	.3937	.8039

[a]MCR(8)* is the corrected marginal concentration ratio. The transformation factor is the Henning adjustment discussed in Chapter II.

TABLE 4-J: Mean and Standard Deviation for Four-Firm Census Ratio

	Mean	Standard Deviation
Sample I	0.71	0.170
Sample II	0.74	0.185
Sample I-A	0.63	0.140
Sample II-A	0.64	0.150

With these deficiencies apparent, the simple pairwise and rank correlations among the various measures of concentration are listed in the following tables for each sample. Their interpretation comprises the next and core section of this chapter.

TABLE 4-K(1): Simple Correlations, Sample I

	CR(4)	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.580	-.150	.001	.857	.780	-.310	.953
DI(4)		1.00	.418	.481	.824	.710	-.756	.783
GINI			1.00	.972	.124	.008	-.795	.032
PIETRA				1.00	.260	.159	-.861	.171
HERF					1.00	.895	-.612	.960
THI						1.00	-.577	.856
ENTRO							1.00	-.509
CCI								1.00

TABLE 4-K(2): Simple Correlations, Sample II

	CR(4)	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.603	-.251	-.100	.867	.833	-.330	.961
DI(4)		1.00	.424	.501	.838	.699	-.804	.786
GINI			1.00	.975	.069	-.072	-.767	-.052
PIETRA				1.00	.215	.077	-.840	.096
HERF					1.00	.940	-.663	.961
THI						1.00	-.566	.899
ENTRO							1.00	-.538
CCI								1.00

TABLE 4-K(3): Simple Correlations, Sample I-A

	CR(4)	MCR*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.143	.557	.377	.509	.891	.810	-.603	.947
MCR*		1.00	-.252	-.448	-.418	.012	-.134	.509	.055
DI(4)			1.00	.645	.602	.771	.509	-.739	.775
GINI				1.00	.967	.537	.335	-.862	.518
PIETRA					1.00	.627	.488	-.904	.607
HERF						1.00	.827	-.750	.976
THI							1.00	-.708	.808
ENTRO								1.00	-.724
CCI									1.00

TABLE 4-K(4): Simple Correlations, Sample II-A

	CR(4)	MCR*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.237	.648	.336	.466	.904	.841	-.661	.963
MCR*		1.00	-.269	-.549	-.455	.025	.117	.367	.112
DI(4)			1.00	.690	.667	.811	.562	-.821	.809
GINI				1.00	.965	.525	.300	-.874	.477
PIETRA					1.00	.635	.456	-.920	.582
HERF						1.00	.899	-.841	.975
THI							1.00	-.697	.857
ENTRO								1.00	-.786
CCI									1.00

TABLE 4-L(1): Rank Correlations, Sample I

	CR(4)	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.578	-.117	-.017	.956	.888	-.260	.964
DI(4)		1.00	.420	.427	.764	.588	-.673	.580
GINI			1.00	.977	.074	-.010	-.828	.751
PIETRA				1.00	.156	-.001	-.859	.134
HERF					1.00	.877	-.439	.997
THI						1.00	-.372	.884
ENTRO							1.00	-.420
CCI								1.00

TABLE 4-L(2): Rank Correlations, Sample II

	CR(4)	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.638	-.191	-.070	.959	.944	-.288	.972
DI(4)		1.00	.415	.451	.796	.599	-.721	.766
GINI			1.00	.971	.013	-.185	-.816	.024
PIETRA				1.00	.120	-.648	-.865	.089
HERF					1.00	.930	-.479	.994
THI						1.00	-.331	.940
ENTRO							1.00	-.499
CCI								1.00

TABLE 4-L(3): Rank Correlations, Sample I-A

	CR(4)	MCR*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.119	.498	.397	.471	.956	.827	-.528	.953
MCR*		1.00	-.300	-.429	-.354	.044	-.021	.493	.039
DI(4)			1.00	.633	.561	.693	.359	-.685	.708
GINI				1.00	.978	.495	.211	-.878	.494
PIETRA					1.00	.528	.276	-.874	.529
HERF						1.00	.762	-.611	.997
THI							1.00	-.524	.759
ENTRO								1.00	-.614
CCI									1.00

TABLE 4-L(4): Rank Correlations, Sample II-A

	CR(4)	MCR*	DI(4)	GINI	PIETRA	HERF	THI	ENTRO	CCI
CR(4)	1.00	.232	.618	.361	.454	.963	.918	-.586	.971
MCR*		1.00	-.351	-.466	-.384	.122	.373	.329	.137
DI(4)			1.00	.690	.639	.754	.422	-.759	.751
GINI				1.00	.968	.486	.180	-.925	.468
PIETRA					1.00	.535	.285	-.939	.525
HERF						1.00	.850	-.682	.996
THI							1.00	-.473	.859
ENTRO								1.00	-.671
CCI									1.00

Interpretations of Index Intercorrelations

First, consider the simple pairwise correlations among the concentration index set. These tables indicate that the Gini coefficient and the Pietra ratio are highly intercorrelated, as expected given their mathematical relationship. The simple correlations range from .965 to .979. Both of these relative measures appear equally able to summarize an industry's Lorenz curve. However, most of the correlations with the remaining measures of concentration are in an intermediate range and are not interpretable without further analysis.

The Herfindahl index and the four-firm Census ratio have high intercorrelations, ranging from .86 to .90. This is consistent with all previous studies, although generally of a lower magnitude which could be an effect of the data's comparative level of disaggregation, i.e., 5-digit product class data.[6] Also, note that the ranked-share index and the Herfindahl index correlations range from .90 to .94 for three of the four samples. Tideman and Hall, and Bailey and Boyle report .95 and .94 intercorrelations respectively. This result is therefore similar to prior empirical research and indicates that a scheme of rank weights on a firm's relative market share produces an index similar to the standard Herfindahl index.

---

[6]Rottenberg, Tideman and Hall, and Bailey and Boyle report a correlation span of .96 to .99. Of interest is the Rottenberg finding that when CR(4) > 50%, the correlation was only .88. Combining this with the mean and standard deviation of CR(4) in each sample makes these results more consistent with previous empirical work.

One area where no extensive empirical analysis has been undertaken is with regard to the comprehensive concentration index proposed by J. Horvath. Its simple correlations with the four-firm Census ratio are .95 or .96. Furthermore, the correlations with the Herfindahl index are from .96 to .97. Thus, the comprehensive index, which includes both relative firm size and their dispersion, is coincident with the CR(4) and HERF vectors. This suggests that the comprehensive index provides little, if any, information about an industry's structural dimension of concentration that is not also revealed by either the four-firm Census ratio or the Herfindahl index.

The four-firm disparity index has intermediate correlations with CR(4), ranging from .56 to .65. This indicates that the indices are neither overlapping nor orthogonal in their information content. But of more interest is DI(4)'s relation with the corrected eight-firm marginal concentration ratio. For the restricted firm number samples, the pairwise correlations are low and negative (-.25 and -.27).[7] The magnitude shows that the strength of the relation is small, but it is also significant. Lower disparity ratios tend to be associated with less dominance by the top four firms, as measured by CR(4),

[7]These correlations are -.63 and -.77 when using the uncorrected marginal ratio. The comparison of MCR(8)* and other indices is confined to Samples I-A and II-A because the other samples include product classes where the individual firm number is less than eight so that an eight-firm marginal ratio is meaningless in these instances. Also, see a following section that examines the relation between MCR(8)* and MCR(8).

since there is a mild positive correlation between these two indices. This implies an increase in the dominance of the second four-firm grouping. Thus, DI(4) and MCR(8)* are negatively related.

One of the more obscure measures of concentration is the entropy index. How does it compare with the remaining indices? Its correlation range with the Pietra ratio is from -.84 to -.92; with the Gini coefficient from -.77 to -.87. Of course the negative signs are expected since entropy is an inverse concentration measure. In both cases, the upper end of these ranges is realized in the restricted firm number samples, where GINI and PIETRA are likely to have better resolution toward their respective industry values. In this sense, the entropy measure of concentration is approximately coincident with summary index vectors of the Lorenz curve.

The magnitudes of the rank correlation coefficients are similar to the simple pairwise correlations. The range of intercorrelation between the Gini coefficient and the Pietra ratio is from .97 to .98, while both have an upper endpoint of correlation of -.93 or -.94 with the entropy measure. The Herfindahl index and the Census ratio have a correlation of .96. At the same time, HERF and CCI have a rank correlation of .99 for all samples, while CR(4) and CCI range of correlation is from .95 to .97. The disparity index and the corrected marginal ratio have low to intermediate coefficients, almost identical to the simple pairwise correlations. Thus, all of the conclusions regarding these particular pairs of indices in the preceding paragraphs are supported by the rank correlations.

In summary, what can we conclude from the foregoing analysis with respect to the choice of an index of concentration? Previously, it was mentioned that the computation of correlation matrices and their examination can resolve the existence question, i.e., whether or not there exist various indices which characterize diverse aspects of the structural dimension of concentration and thus are likely to exhibit different predictive abilities. The question of which specific index is better able to predict market performance is not answerable using the prior analysis. With this firmly in mind, let us return to the existence question.

From the estimated correlations, it is strongly suggested that CR(4), HERF, and CCI are not likely to possess significant differences in predictive power. All can be easily interchanged. The susceptibility of empirical results to the choice of any of these indices is minimal. Furthermore, the ranked-share index can be placed in this class, although its pairwise correlations are lower in magnitude but still high with the other indexes. A second grouping emerges, largely orthogonal to the first. This group includes the Gini coefficient, the Pietra ratio, and the entropy measure of concentration. Note that in the first grouping both absolute and relative measures are included. It is surprising that HERF, a relative index, has such low pairwise correlations with the other relative measures, which are the components of the second class. Finally, a third group can be found. These indices have low correlations with the other concentration measures, including between themselves. In this class are the disparity index and the corrected marginal concentration ratio.

TABLE 4-M: A Tentative Index Grouping

Group I:	CR(4), HERF, CCI, THI
Group II:	GINI, PIETRA, ENTRO
Group III:	DI(4), MCR(8)*

Since all are not highly intercorrelated, there is reason to believe that some of the indices are likely to vary in their predictive capabilities. We should note that the major findings of previous empirical studies with respect to a limited index set, largely Group I, have been supported by this analysis. A larger set, Groups II and III, considered here for the first time, shows the emergence of some low or intermediate correlations.

Of course, these index groupings are subjective in the sense that they are only subjectively based on the magnitude of pairwise intragroup correlation estimates versus pairwise intergroup correlations. This is purely an arbitrary descriptive tool. A more objective method of index groupings can be achieved by employing the method of principal components. This is examined extensively in the next chapter.

Truncated Herfindahls as Proxy Variables for Standard Herfindahls

Because of the non-availability of the entire firm size distribution, empirical studies which have employed a Herfindahl index have had to make certain assumptions about the industry's concentration or Lorenz curve. Census of Manufactures data only provide a few points along the concentration curve. The most common assumption can be termed the mean-share assumption. In fact, the empirical studies cited in Chapter III made this assumption in the computation of the Herfindahl index. The mean-share assumption can be stated as follows: each firm is assumed to have a size equal to the mean of its appropriate size class. These size classes usually are defined between known points on the concentration curve, e.g., firms ranked 1 to 4, 5 to 8, 9 to 20, and 21 to 50. The resultant index is termed a truncated Herfindahl. This particular assumption minimizes the variance in firm shares for given Census ratio data. Thus, the procedure produces minimum Herfindahl estimates (see equation (7b) in Chapter II).

However, with the 1950 product class data the Herfindahl index can be computed without the mean-share assumption. This allows us to consider the question of whether or not truncated

Herfindahls are accurate indicators of the entire concentration curve. Are previous empirical results based on truncated Herfindahls susceptible to the computational assumption embodied in the index?

In order to investigate this issue, simple pairwise correlations were calculated across all samples. If these correlations are extremely high, a truncated Herfindahl is a good proxy for the standard Herfindahl index. Conversely, if these correlations are low, then the use of truncated Herfindahls in empirical research can be called into question.

TABLE 4-N: Simple Correlations Between Standard and Truncated Herfindahl Indices

Sample I	0.870
Sample II	0.879
Sample I-A	0.907
Sample II-A	0.912

From the above table, note that the simple correlations range from .87 to .91. The implication is that a truncated Herfindahl closely approximates an industry's Herfindahl index and, therefore, can be used as a proxy variable when individual firm data is not available.

Marginal Concentration Ratios and Non-Linearity Relations with the Census Ratio

One of the more perplexing areas in the measurement of concentration is the theoretical role of marginal concentration ratios and their impact on market performance. The necessary statistical dependence between marginal ratios and Census ratios has led to non-linear equation specifications and theoretical conjectures concerning the coefficient sign of MCR(8) over the range of CR(4) values.[8] The first question that arises is what are the effects on this statistical dependence of Henning's proposed transformation of the marginal concentration ratio, the form of the index that was employed in the previous analysis. MCR(8)* expresses MCR(8) relative to the maximum feasible value that is imposed by the level of the four-firm Census ratio. From Table 4-O, note that the

TABLE 4-O: Simple Correlations of Marginal Ratios with the Census Ratio

	MCR(8)	MCR(8)*	Mean of CR(4)	Standard Deviation CR(4)
Sample I-A	-.67	.14	.63	.14
Sample II-A	-.70	.24	.64	.15

[8]Previous references have been made in Chapter II concerning this point. Particularly, consider the works of Henning and Miller.

uncorrected marginal ratio is negatively correlated with the four-firm Census ratio. This reflects their statistical dependence when the corresponding sample mean and standard deviation values of CR(4) are taken into account.[9] After applying the Henning transformation, the correlations substantially decrease and change sign. The statistical dependence has been reduced, if not eliminated. In essence, this transformation implies that the relation between the marginal concentration ratio and the four-firm Census ratio is not linear over the entire range of the latter's values. Two non-linear equation specifications are examined below: a quadratic approximation and a dichotomized sample over the range of CR(4) values.

Table 4-P(1) presents regression results on two overall samples and subsamples within each that were defined by Miller. The "intermediate" subsample contains those product classes for which $30 < CR(4) \leq 60$, while the "high" subsample has observations where $60 < CR(4) \leq 100$. Significant linear and quadratic equation specifications are found for both overall samples with an adjusted coefficient of determination for the quadratic relation consistent with Miller's results, i.e., Miller reports $\bar{R}^2 = .468$. Also similar to Miller is the insignificant linear relation in the intermediate subsample and the significant one found in the high subsample.

---

[9]See Diagram 2-C and following pages in Chapter II.

TABLE 4-P(1): Regression Analysis, MCR(8) as the Dependent Variable

		Const.	CR(4)	$CR(4)^2$	$\bar{R}^2$
Sample I-A	Overall	.357	-.307 (6.31)[a]		.437 (39.8)*,b
		.038	.731 (1.99)	-.806 (2.84)	.508 (26.8)*
	Intermediate	.105	.202 (1.51)		.053 (2.29)
	High	.451	-.440 (4.75)		.453 (22.6)*
Sample II-A	Overall	.377	-.313 (5.39)		.467 (29.1)*
		.004	.896 (2.30)	-.929 (3.14)	.585 (23.6)*
	Intermediate	.125	.196 (1.41)		.065 (1.97)
	High	.480	-.457 (4.47)		.527 (19.9)*

*Significant at the 1% level

[a]T-ratios in parentheses

[b]F-ratio in parentheses

Turning now to Table 4-P(2), no significant linear or quadratic relations exist for the overall sample or the derivative subsamples in either Sample I-A or II-A product class data when the corrected marginal concentration ratio is the dependent variable. It appears that those relations found in Table 4-P(1) have been eliminated.

TABLE 4-P(2): Regression Analysis, MCR(8)* as the Dependent Variable

		Const.	CR(4)	$CR(4)^2$	$\bar{R}^2$
Sample I-A	Overall	.382	.156 (1.01)		0.00 (1.02)
		.542	-.369 (0.30)	.407 (0.42)	0.00 (0.59)
	Intermediate	.295	.265 (0.68)		0.00 (0.45)
	High	.298	.340 (0.95)		0.00 (0.89)
Sample II-A	Overall	.391	.235 (1.36)		0.03 (1.85)
		.439	.082 (0.06)	.117 (0.12)	0.00 (0.90)
	Intermediate	.363	.306 (0.82)		0.00 (0.65)
	High	.301	.348 (0.83)		0.00 (0.67)

TABLE 4-Q: Covariance Analysis Using Miller's Classification

	Dependent Variable	Residual Sum of Squares			Computed F-Value
		Overall	Intermed.	High	
Sample I-A	MCR(8)	0.117	0.038	0.049	7.89*
	MCR(8)*	1.168	0.278	0.881	0.18
Sample II-A	MCR(8)	0.077	0.021	0.030	7.22*
	MCR(8)*	0.678	0.153	0.521	0.09

*Significant at the 1% level

Further investigation using covariance analysis shows that the intermediate-high classification scheme is not meaningful when MCR(8)* is the dependent variable, a not so surprising result. However, a significant point of discontinuity is indicated when MCR(8) is the dependent variable.

Collectively, these results imply that the utilization of standard marginal concentration ratios in an attempt to ascertain its effects on performance variables independent of the four-firm Census ratio influences is inappropriate. However, when these ratios are adjusted according to their maximum feasible values, marginal ratios may indicate influences on performance that emanate from the structural dimension of concentration independent of the Census ratio. In fact, no significant relation is found with CR(4) after such adjustment. Thus, the use of MCR(8)* in the preceding analysis is consistent with the goal of this paper and it represents the preferred form of the eight-firm marginal concentration ratio.

Conclusion

One of the principal reasons for analyzing alternative measures of concentration is the concern over the susceptibility of conclusions from empirical research founded on a particular index of concentration. More specifically, does the choice of an index affect empirical results? If one selects a different measure, will the conclusions be altered? Essentially, this is the overriding issue. Of course, the answer lies in empirical investigation itself. This chapter has set forth a comprehensive set of proposed measures of concentration and has presented a compilation of index values for a data set that has not been previously analyzed. The approach was to investigate the likelihood of diverse predictive abilities within the index set, i.e., to examine the existence question regarding the "optimal" measure of concentration.

Many of the results in this chapter are supportive of previous empirical studies. This is particularly the case with respect to the identical predictive powers (or its likelihood) between the four-firm Census ratio and the Herfindahl index. Similar findings also appeared with the ranked-share index and the comprehensive concentration index. Moreover, it was confirmed that the use of truncated Herfindahls as proxy variables for standard Herfindahls is not likely to bias empirical research because of their high intercorrelation.

However, some of the results indicate that the choice of an index may be crucial. This is particularly true when you consider Group I correlations with the Gini coefficient, the Pietra ratio, or the entropy measure. Furthermore, using the disparity index or the marginal concentration ratio in conjunction with another index is likely to increase predictive capability. This stems from their low range of correlation with the remaining concentration measures.

In all cases, each of the proposed concentration indices characterizes certain attributes of the number and size distribution of firms within an industry. Some appear to embody similar, if not identical, aspects of this structural dimension. Others describe different aspects. This suggests that no single measure is capable of summarizing all of the information content in this structural dimension. As we have seen, groupings of indices have been found, largely independent of each other. One question that arises is the distinctive character of each index group. Recalling the controversy between absolute and relative measures of concentration discussed in Chapter II, note that Group II has all relative measures, while Group I contains absolute and relative measures. Thus, Group II stresses dispersion in firm size. Group I indexes are a conglomerate of both dispersion influences and dominance of the leading firm group. The appearance of HERF and THI, both relative measures, in Group I possibly results from the predominance of the leading firm in many of these industries, so that the change in their observed values as more firms are included is extremely small.

Another question that arises is which index within each group should be selected. The measures in a grouping are likely to possess the same predictive power. One solution is to base selection on availability or computational ease. But if all the indices within a group are available, then what should be the selection criterion?

Another question concerns the between group selection. Once a measure has been selected from each grouping, how can we determine which one of these is optimal? This uniqueness question can be answered by employing the predictive power criterion. Of course, we are put into a position of justifying the consideration of uniqueness at all. Why not use the selected measures in a complementary fashion? Assuming that the goal in the measurement of concentration is to reveal something about the emergence of monopolistic performance, then why not characterize this structural dimension by more than one index? The predictive capability is likely to improve, particularly if the included indexes are largely orthogonal.

The latter question will be examined in a later chapter. But first, let us turn to a discussion of the principal component transformation as a method of establishing objective index groupings as well as its applicability within the groupings found in this chapter.

CHAPTER FIVE: A PRINCIPAL COMPONENT ANALYSIS OF CONCENTRATION INDEX GROUPINGS

In the concluding remarks of the previous chapter, it was noted that no single measure of concentration can completely describe firm-size distributions. Yet, various index groupings were found, within which predictive capability is likely to be the same for each concentration measure. This chapter will examine one solution to the problem of index selection within these groupings. The selection criterion developed in the foregoing analysis seeks to avoid arbitrary index selections. Even though the indices within a specific group are highly intercorrelated, an empirical selection procedure should incorporate all of their representations of the structural dimension of concentration. In other words, the problem is one of index construction of the various measures of concentration per se, where each index is not constrained to have a zero weight. The technique that will be applied to these index groupings is the principal component analysis.

Once a general index of each grouping has been ascertained, the logical question to consider is their relationship. If the group indices are largely orthogonal, then their joint utilization in a regression analysis is likely to enhance the predictive power of the equation, while at the same time allowing for unambiguous coefficient interpretations. If the general indexes are largely interdependent, then the relative importance of the group indices in the explanation of some performance variable is not as clearcut. In the latter situation, an alternative procedure is required.

The first section of this chapter contains some introductory comments on principal componant analysis, presents some previous applications of the technique, and examines the applicability of principal components in the measurement of concentration. This is followed by a formal derivation of principal component estimators. Some important component properties are presented with an emphasis on their empirical significance. Finally, the technique is applied to the index groupings defined in Chapter IV and to the entire index set.

One of the major conclusions of this chapter is the general desirability of the principal component transformation in the construction of economic indexes. It is asserted that the predictive power of a component index, which reflects <u>all</u> variables in a multivariate system, will generally exceed that of any <u>one</u>, individual variable in that system. This represents the major advantage of principal component analysis when the problem is to devise an index of concentration.

The Applicability of Principal Component Transformations to the Analysis of Alternative Concentration Measures

One of the major problems encountered in the measurement of concentration occurs when an industry is ranked differently by alternative indices. In such cases the determination of which industry is "more concentrated" depends on the choice of an index. Assuming that there exists no a priori basis for selection among the various measures and that arbitrary selection is undesirable, how can this problem be surmounted? One solution is to construct an index of the various measures of concentration themselves, i.e., to represent a set of multidimensional vectors in a space of fewer dimensions. In the case of index construction, the reduced space is typically of one dimension. Consider the properties of this new index. Any index is essentially a function that defines a scheme of weights to be applied to a set of observations on some variables. It seems desirable that the linear combination of the concentration measures should preserve or capture most of the sample variability in order to retain maximal information about the movements of the index values across different industries. The restriction to linear weighting schemes is not a necessity. However, if the admissable weighting functions are limited, then a principal component transformation satisfies the above requirement.

In general, the principal component technique is an orthogonal transformation which decomposes the cumulative variance of a set of jointly distributed variables into smaller and smaller proportions. If the first few components account for most of this variance, and if such components can be interpreted, then the original system is better described by these components.

The technique has been successfully used in the past. For example, Kendall examined the yields of ten crops in forty-eight English counties.[1] From the correlation matrix, the largest characteristic root was extracted and the corresponding principal component computed. Each coefficient (of the characteristic vector) was of equal magnitude, indicating that each crop was equally correlated with the associated principal component. Therefore, this new variate was identified as a measure of productivity and counties were ranked and grouped according to the value of this index. Here the problem can be viewed as one that involved the reduction of a ten-dimensional system into a one-dimensional system.

Another example can be found in a later work undertaken by Dhrymes, who attempted to determine the relation between price and various characteristics of automobiles using standard least-squares estimation.[2] In order to reduce the

---

[1] Kendall, M.G., "The Geographical Distribution of Crop Productivity in England," J. Roy. Stat. Soc. (1939), pp. 21 ff.

[2] P.J. Dhrymes, "On the Measurement of Price and Quantity Changes in Some Consumer Capital Goods," Discussion Paper 67, Economic Research Unit, University of Pennsylvania (1967).

dimensionality of and the multicollinearity in the independent variable set, he extracted characteristic roots and vectors. It was found that the first two components accounted for 97% of the sample variance. Thus, these components were substituted for the original variables in the regression model.

There have been more applications of the principal component transformation in economic research.[3] Of course, the technique has certain drawbacks. For example, changes in the scale of measurement will alter the characteristic vectors and, therefore, the form of the principal components. To avoid this problem, most extraction procedures are performed on the correlation matrix for a given set of jointly distributed variates, i.e., when all the observation vectors are measured in standardized units. More importantly, there exists no rule for deciding when a sufficient proportion of the sample variance has been accounted for by the components. In practice, a large percentage is arbitrarily specified. This proportion is usually contained in the first few components. If some of the original variates are not highly correlated with these components, then these variables in their original form can be used with a subset of the components in further analysis. However, even if these variates are highly correlated with later components, the associated characteristic roots are so small that considerable computational

---

[3]See J.R.N. Stone, "The Interdependence of Blocks of Transactions," J. Roy. Stat. Soc., Supplement (1947), pp. 1-32.

error is likely to occur. In general, the procedure is to select initially an arbitrary large percentage of the cumulative sample variance that is to be captured. Then, the components are calculated up to this point. If some variables are unrelated to this subset of the extracted components, and they are crucial to the model, then they are included in their original form, along with the components, in further analysis such as least-square estimation.

With these introductory remarks aside, let us consider the application of principal components to the problem of selecting an index of concentration. In the previous chapter several groupings of concentration measures were identified. The question arose as to which measure within each group should be selected. Assuming availability, there appears to be no theoretical basis upon which a selection can be made. The problem is similar to the empirical question examined by Kendall, i.e., to reduce a multidimensional system into a one-dimensional representation. Kendall's result was a measure of productivity. In this analysis, the goal is to devise an overall index of concentration. The principal component transformation has potential in this context. The degree of its applicability depends on the percentage of total variance accounted for by the first or second component, and whether these components can be interpreted meaningfully. If these conditions are satisfied, then industry rankings can be made on a component basis.

Principal Component Estimation and Some Component Properties

Suppose we have a set of T observations on the variables $x_1, x_2, \ldots, x_k$ from which the sample covariance matrix, S, is computed. S summarizes the dependent structure among these K variates. The problem is index construction, i.e., to represent these K-dimensional vectors in a reduced space. It is desirable to preserve the sample's variability, so the problem can be restated as finding a linear function of the sample vectors that maximizes the variance of the resultant index. The principal component technique can define such linear functions.[4]

Formally, the first principal component is that linear combination of the original variables

$$P_1 = Xa_1 \tag{1}$$

whose variance, $s_1^2 = a_1'Sa_1$, is maximized for all coefficient vectors, a_1, subject to the normalization constraint that $a_1'a_1 = 1$. The latter condition is to avoid the trivial indeterminacy implied by variance maximization. Formulating the Lagrangian function, we obtain

$$L_1 = a_1'Sa_1 + \lambda_1(1 - a_1'a_1) \tag{2}$$

Differentiating with respect to a_1,

---

[4]The following analysis is largely adopted from D.F. Morrison, Multivariate Statistical Methods (New York: McGraw Hill), pp. 221-58.

$$\frac{dL_1}{da_1} = 2Sa_1 - \lambda_1 2a_1 = 0$$

$$2(S - \lambda_1 I)a_1 = 0 \qquad (3)$$

$$(S - \lambda_1 I)a_1 = 0$$

Disregarding the trivial solution to this set of K homogeneous equations, (3) requires that $\det[S - \lambda_1 I] = 0$. But this is exactly the condition for extracting characteristic roots from a real symmetric matrix. Hence, λ_1 is a characteristic root of S and a_1 is its associated characteristic vector. Moreover, λ_1 is the largest root of S since it represents the variance of P_1 which was to be maximized. This is easily demonstrated by premultiplying (3) by a_1' and employing the normalization constraint.

$$a_1'Sa_1 - \lambda_1 a_1'a_1 = 0$$

$$a_1'Sa_1 = \lambda_1 \qquad (4)$$

$$s_1^2 = \lambda_1$$

In order to determine the second principal component, $P_2 = Xa_2$, the coefficients of the vector a_2 are chosen so as to maximize the variance of P_2, subject not only to a normalization condition, but also to an orthogonality constraint between the vectors a_1 and a_2, i.e., $a_1'a_2 = 0$. Thus,

$$L_2 = a_2'Sa_2 + \lambda_2(1 - a_2'a_2) + \gamma(a_1'a_2) \qquad (5)$$

Differentiating with respect to a_2,

$$\frac{dL_2}{da_2} = 2Sa_2 - \lambda_2 2a_2 + \gamma a_1 = 0 \qquad (6)$$

In order to simplify this first order condition, premultiply (6) by a_1' and use both the normalization and orthogonality conditions.

$$2a_1'Sa_2 - \lambda_2 2a_1'a_2 + \gamma a_1'a_1 = 0$$
$$2a_1'Sa_2 + \gamma = 0 \qquad (7)$$

Returning to equation (3), premultiply by a_2', so that

$$2a_2'Sa_1 - \lambda_1 2a_2'a_1 = 0$$

or $$2a_1'Sa_2 - \lambda_1 2a_1'a_2 = 0 \qquad (8)$$

$$2a_1'Sa_2 = 0$$

Substituting (8) into (7), note that $\gamma = 0$ so that (6) becomes

$$2Sa_2 - \lambda_2 2a_2 = 0$$

or $$(S - \lambda_2 I)a_2 = 0 \qquad (9)$$

Therefore, the coefficients of a_2 satisfying equation (9) are the elements of the characteristic vector corresponding to the second largest characteristic root.

The remaining K-2 principal components are generated by replicating this procedure. Note that the orthogonality condition implies that the variance of the successive components sum to the total variance of the original variates, i.e., $\lambda_1 + \lambda_2 + \ldots + \lambda_k = tr(S)$. Also, note that in the preceding derivation the components were successively extracted from the sample covariance matrix, S. In most applications the sample correlation matrix, R, is used in order to standardize the units of measurement. The components obtained from R are different from those of S, although the extraction technique is identical.

How do we characterize the importance of each component and of each original variate in terms of the sample's variance? Recall that the successful application of the principal component transformation is conditioned upon the capability of the first few components to account for a large proportion of the total sample variance. Thus, one way of ascertaining the importance of each component is to define the ratio of the component's variance, i.e., its associated characteristic root, over the total variance in the original multivariate system, tr(R). If the first or second components sum to some prespecified percentage of the tr(R), then the technique is well conditioned on the original sample data.

Of course the components convey a meaningful representation of the original variables only if we can ascertain their degree of association. The sign and the magnitude of each element in the characteristic vector indicate the importance of the corresponding original variate in the resultant principal component. Also, if the components have been extracted from the correlation matrix, then the correlation of the original variables with the i^{th} principal component is given by

$\sqrt{\lambda_i}\, a_{ij}$, j = 1,2,...,K.[5] With these correlations, we can measure the degree of association between the components and the original multivariate system. Such associations provide input into the interpretation of the principal components themselves.

One of the more interesting properties of principal components is revealed by the geometric interpretation of the transformation. The matrix R can be viewed as specifying a K-dimensional ellipsoid, or a scatter of T observations in K space. Assuming R is not diagonal, this ellipsoid will exhibit a certain orientation. What are the characteristic vectors? Essentially, the linear combination of the X's that define the principal components collectively represent a "change-of-basis" transformation. This re-referencing of the original variables is done so as to transform the sample correlation matrix into its canonical form. Since R is a real, positive-definite, symmetric matrix, there must exist an orthogonal matrix, Q, that diagonalizes R. In this case, the matrix Q is the matrix of normalized characteristic vectors of R, whose canonical form is the diagonal matrix of characteristic roots. As previously noted, these roots are the variances of the principal components which are pairwise orthogonal. Thus, the principal component transformation is a rigid rotation of the original system into a coordinate orientation along the principal axes of the ellipsoid defined by R.

---

[5]As an aside, note that the sum of the squared correlations of the original variables with the i^{th} component is the component variance.

From this geometric interpretation, it has been shown that the coefficients of the characteristic vectors define linear combinations of the original variates that minimize the sum of the squared deviations between the X's and the corresponding principal component. In other words, the new coordinate system corresponds to successive least-square solutions.[6]

Let us summarize the salient points that have been presented in this section. Given a set of T observations on K variables and a sample correlation matrix R,

(1) $P = XA$ where A is the matrix of normalized characteristic vectors of R

(2) $\sum_P = \Lambda$ where Λ diagonal, whose elements are the characteristic roots of R

(3) $tr(R) = \sum_{j=1}^{k} \lambda_j$ or cumulative variance equality

(4) $\gamma_{P_2,X_j} = \sqrt{\lambda_i}\ a_{ij}$ correlation between component and original variables

Using this general knowledge of the principal component transformation, the technique can now be applied to the product class data. Recall that in Chapter IV substantial inter-correlations between the alternative measures of concentration were found. However, the high magnitude of these coefficients did not permeate the entire index set. Groupings based on pairwise correlations were identified. Industry rankings are

---

[6]Morrison has an excellent discussion of these points. In particular, see Morrison, op. cit., pp. 230-33.

likely to be different for alternative indices across these groups. By employing principal component analysis, an index of the alternative measures can be computed. This index ideally will reflect all of the diverse aspects of the structural dimension of concentration contained in the entire index set or in various index groupings or subsets.

Principal Component Analysis of the Concentration Index Groupings

One of the pre-estimation problems in the application of the principal component transformation is the specification of a large proportion of sample variance which will terminate the extraction of additional characteristic roots and vectors from the sample correlation matrix.[7] Arbitrarily, 90% is chosen as the minimum proportion of sample variance to be explained in the foregoing analysis. This percentage must be accounted for by the first or second components. Otherwise, the principal component transformation has little, or no, applicability for the purposes of index construction. Ultimately, the goal is to reduce the dimensionality of the index set into one or at most two dimensions.

Consider the Group I indices defined in Chapter IV. Tables 5-A(1) and 5-A(2) present the composite results of

---

[7]In actual practice, all roots and vectors may be calculated. But the cumulative variance criterion will signal the inclusion of the subset of components that possess information about the significant dimensionalities in the data set. The excluded components will be along principal axes that are small in length and the computed characteristic vectors are likely to be imprecise.

TABLE 5-A(1): Characteristic Root and Vector of the First Component for Group I Indices

	Root	Associated Vector			
		CR(4)	HERF	CCI	THI
Sample I	3.65	-.492	-.509	-.516	-.483
Sample II	3.73	-.491	-.505	-.512	-.492
Sample I-A	3.63	-.503	-.509	-.515	-.472
Sample II-A	3.72	-.499	-.508	-.510	-.483

TABLE 5-A(2): Correlations of the First Component with Group I Indices

	Percent of Total Var.	Correlations with Group I Indices			
		CR(4)	HERF	CCI	THI
Sample I	91%	-.939	-.972	-.986	-.923
Sample II	93%	-.948	-.976	-.989	-.950
Sample I-A	91%	-.959	-.970	-.981	-.900
Sample II-A	93%	-.963	-.980	-.984	-.932

performing the principal component transformation on this data set. First, note that the various tabulated estimates are almost identical for all four samples. This suggests an insensitivity to varying selection criteria. The magnitudes of the largest characteristic root indicate that 91 to 93 percent of the cumulative variance can be expressed along the major axis of the ellipsoid associated with the correlation matrix. This axis is represented by the associated characteristic vector. Because of the relative equality of the vector coefficients, the correlations between the first component and the Group I indexes are almost identical. What is striking about these correlations is their magnitudes! The first principal component appears to be a more parsimonious representation of the Group I index set. It allows the avoidance of arbitrary selection among a group of highly intercorrelated variables. The signs of the index correlations imply that the first principal component can be viewed as an overall index of <u>inverse</u> concentration for Group I measures. Industry rankings according to this index will provide structural information about the relative degrees of competition.

Similar results are obtained for Group II indices. The first component captures 91 to 95 percent of the total sample variance; the coefficients of the associated vector are almost identical, so that Group II index correlations with the component are relatively equal. Finally, these correlations

TABLE 5-B(1): Characteristic Root and Vector of the First Component for Group II Indices

	Root	Associated Vector GINI	PIETRA	ENTRO
Sample I	2.75	-.581	-.595	.556
Sample II	2.73	-.583	-.598	.551
Sample I-A	2.82	-.579	-.588	.565
Sample II-A	2.84	-.577	-.587	.568

TABLE 5-B(2): Correlations of the First Component with Group II Indices

	Percent of Total Var.	Correlations with Group II Indices GINI	PIETRA	ENTRO
Sample I	92%	-.964	-.988	.923
Sample II	91%	-.962	-.987	.909
Sample I-A	94%	-.973	-.987	.950
Sample II-A	95%	-.970	-.986	.954

are extremely high and their signs indicate that the first component represents an overall index of inverse concentration.

The question now arises as to the relationship between the first component of each index group. Successive components extracted from a given sample correlation matrix are necessarily orthogonal. But component orthogonality across index groupings is not guaranteed. Table 5-C shows that the group components for Samples I and II are not highly intercorrelated. However,

TABLE 5-C: Correlation Between Group Principal Components

Sample I	0.224
Sample II	0.169
Sample I-A	0.605
Sample II-A	0.600

for Samples I-A and II-A, the component intercorrelations are significantly higher.[8] In the case of Samples I and II, the structural dimension of concentration, as measured by Group I and II indexes, can be adequately characterized by a dual-index set. Each index grouping has been reduced to a one-dimensional representation. Using the extracted components simultaneously

[8]The Wilks-Bartlett test for independence is accepted for Samples I and II, but is rejected for Samples I-A and II-A. See M.S. Bartlett, "Note on Multiplying Factors for Various Chi-Square Approximation," J. Roy. Stat. Soc. (1954).

is likely to enhance predictive capability since the components are largely orthogonal. However, the intermediate range of correlation between the group components for Samples I-A and II-A suggests that the gain in predictive power is likely to be smaller than for Samples I and II.

As an alternative to component extraction from individual index groupings, a "merge" procedure can be employed for Samples I-A and II-A. By considering the indices of Groups I and II jointly, we hope to characterize the residual interdependence between these groups that remained after separate component extraction and representation by the first component index.

Table 5-D presents the results of the application of the principal component transformation to the respective correlation matrices for Samples I-A and II-A. The first component for Sample I-A accounts for 75% of the total sample variance, while the second captures 18%. These percentages meet the minimum variance criterion when cumulated. Of critical importance is their meaningful interpretability. The first component can be considered a general index of inverse concentration, similar to the component index defined on the separate index groupings. The magnitudes of the index correlation with this component are high across the entire merged index set. But what does the second component reveal? A close examination of the signs of index correlation suggests that the second principal component represents a comparison of Group I versus Group II indices.

TABLE 5-D: Correlations of Components with Merged Index Groups

	Component Number	Percent of Total Var.	Correlation Between Indices and Components CR(4)	HERF	CCI	THI	GINI	PIETRA	ENTRO
Sample I-A	1	75%	-.857	-.933	-.929	-.829	-.748	-.834	.913
	2	18%	.432	.264	.303	.382	-.651	-.530	.312
Sample II-A	1	76%	-.862	-.959	-.937	-.842	-.718	-.812	.944
	2	20%	.430	.236	.299	.420	-.689	-.566	.293

Group I indexes have changed the sign of their correlations, while those of Group II indices have remained the same. In other words, the second component captures disparities between the groupings; the first component describes their similarities.[9]

The results for Sample II-A are almost identical. The cumulative variance explained by the first two components is 96%, 76% for the first component, and 20% for the second. The component correlations with the original indices are the same in magnitude and sign. Therefore, the components lend themselves to the same interpretation.

As mentioned in Chapter IV, the index groupings were formed using a subjective criterion, one based on arbitrary magnitudes of pairwise correlations. Yet, the principal component transformation is an objective tool. In the following, the transformation is applied to the entire index set, the purpose of which is to avoid subjective groupings. The correlations between the original variables and the components is presented in Table 5-E.

The results are in general agreement with those that were found for the merged group in the previous section. The addition of the disparity index for Samples I and II produces no significantly novel results. DI(4) is highly correlated with the first principal component, but not with the second component. GINI and PIETRA are highly correlated with the latter. For Samples I-A and II-A, the corrected marginal concentration ratio is also added to the index set. Here, DI(4)

---

[9]This interpretation of the pattern of signs is common. Illustrative examples may be found in Morrison, op. cit., pp. 229, 243.

TABLE 5-E: Correlations of Components Over the Entire Index Set

	Component Number	Percent of Total Var.	CR(4)	HERF	CCI	THI	GINI	PIETRA	ENTRO	DI(4)	MCR(8)*
Sample I	1	62%	-.773	-.942	-.909	-.863	-.395	-.519	.813	-.903	
	2	31%	.539	.274	.380	.332	-.908	-.833	.543	-.083	
Sample II	1	62%	-.796	-.964	-.924	-.884	-.299	-.433	.813	-.914	
	2	33%	.537	.225	.355	.340	-.949	-.886	.559	-.166	
Sample I-A	1	66%	-.820	-.924	-.919	-.801	-.774	-.840	.931	-.821	.265
	2	20%	.495	.327	.374	.307	-.521	-.413	.310	-.055	.809
	3	7%	-.050	.008	.037	-.449	.331	.219	.077	.144	.485
Sample II-A	1	67%	-.833	-.950	-.927	-.807	-.751	-.824	.957	-.873	.188
	2	23%	.486	.272	.349	.424	-.605	-.470	.243	-.106	.863
	3	4%	-.009	.080	.033	.178	-.253	-.275	.051	.160	-.443

is associated with the first component, while MCR(8)* is correlated with the second principal component. The third component represents Group II indices and MCR(8)*.

Conclusion

The application of the principal component transformation on the index groups defined in Chapter IV, individually and collectively, has allowed us to represent a multidimensional system in terms of one or two dimensions. In all cases, the first component captured the influences of all the indices included in the sample, although in varying degrees. In some instances, the second principal component embodied residual influences, where those indexes which were weakly associated with the first component emerging dominant. At other times, the second component represented a comparison of some indexes versus others, as revealed by the sign pattern of the correlations. In general, the results suggest that there exist diverse dimensionalities within the entire set of alternative concentration measures. However, the degree of disparity appears small. The actual identification of which index, if any, is superior is predicated on relative predictive capability in the estimation of a hypothesized structure-performance relation. The gain associated with principal components is not in predictive power, but is in the representation of the predictive power of the index set in a reduced space.

CHAPTER SIX: A STRUCTURE-PERFORMANCE TEST OF ALTERNATIVE MEASURES OF CONCENTRATION

The ultimate selection criterion for a measure of concentration must be predicated on its ability to predict market performance. This results from the fact that there is no a priori grounds upon which a selection can be made. In this chapter an empirical examination of the concentration-profit hypothesis is made using the various indices of concentration that have been examined.

Recall that an index of market concentration should reflect something about the likelihood that oligopolistic performance will emerge in a market. To the extent that potential discretionary power over price is actually utilized, increases in concentration should be associated with increases in profitability. Classical economic theory suggests that given similar market demand and cost structures, prices are higher and profits are greater in monopolistic markets than in competitive markets. Measures of concentration allow us to classify markets into such categories. In other words, as industry output increasingly falls under the control of a few firms, deviations from the equality of price and marginal cost are more likely in the long-run. Industry profits, on the average, are greater because of increased recognized interdependence. Conversely, as industry output becomes more widely distributed among existing firms or new firms, more rigorous price competition is fostered and industry profits are reduced. The classical hypothesis involves the relation between profits and relative size distributions, while measures of concentration are summary indicators of the firm size distribution. Thus, the theory suggests that there exists a direct connection between concentration and profitability.

In the following pages, the concentration-profit hypothesis is estimated using the 1950 product class data. The measure of profitability is 1958 price-cost margin data. The primary concern is to determine differences in the explanatory power of the alternative indices. Previous analysis suggested disparities were likely to exist. The hope is to identify the concentration index or indexes which have higher predictive capability. Many questions arise in this context. Do small fringe firms exert pressure on industry profitability? How do these effects differ from core firm influences that persist in monopolistic markets? Are these effects independent? The answers to these questions provide the necessary input required in the selection of a measure of concentration.

Note that the profit-concentration hypothesis is only a single aspect of the overall structure-performance relationship. One index may emerge as superior with regards to predicting profitability. However, it should not be concluded that this index is universally a better discriminator and predictor of performance. Ideally, one should examine all the dimensions of the structure-performance relation, e.g., the relation between concentration and the quality of product. Unfortunately, economic theory's predictions about market structure's effects are conflicting and ambiguous. Empirical work has not advanced our knowledge very much because the conditions which need to be satisfied in order to interpret the results meaningfully are generally lacking.[1] The emphasis in this analysis is to determine whether one index is a better predictor of profitability, whether various indices have independent effects on profitability, and whether the choice of an index can seriously prejudice empirical investigation. The conclusions relate solely to the hypothesis under consideration.

The Four-Digit Industry Sample

As previously noted, the choice of an index of concentration can be based on its predictive power as exhibited in some hypothesized structure-performance relationship. The most testable hypothesis is between industry profits and industry concentration. The first problem encountered was the index data itself.

---

[1]J. Bain, Industrial Organization (John Wiley & Sons, 1968), pp. 418-25.

All of the prior analysis was conducted on 5-digit product class data. Unfortunately, profitability data at this level of disaggregation are non-existent. This necessitated "aggregating" the 5-digit data to the 4-digit SIC industry level. Selection of 4-digit industries was based on a simple coverage rule: an industry was included in the sample if the value of shipments of the component product classes totaled at least 70% of the value of shipments attributed to the 4-digit industry. In some instances, the component product classes numbered only one, e.g., cigarettes and cereals. Here, the 5-digit product class and the 4-digit industry were identical so that no aggregation was required. In other cases, there were 4-digit industries that satisfied the 70% coverage criterion for which there were more than one component product class.

Leading firms in each product class were not necessarily the same. Moreover, many firms did not operate in every product class. The generation of the 4-digit firm size distribution was accomplished by adding each firm's value of shipments in those component classes in which it did operate. Then firms were re-ranked, highest to lowest, according to the combined value of shipments. Finally, concentration indices were calculated on this new firm size distribution. An illustrative computation is presented in the appendix to this chapter.

The resulting 4-digit industries that met the coverage criterion are presented in Table 6-A. An examination of this table reveals that the sample is biased toward "concentrated"

TABLE 6-A: Sample of Four-Digit Industries and Corresponding Index Values

SIC Number	CR(4)	DI(4)	HERF	GINI	PIETRA	THI	ENTRO	CCI
2043	.7269	.1317	.1429	.6958	.5933	.2266	.6336	.3974
2073	.9190	.4353	.3525	.4273	.3265	.4940	.7326	.6598
2085	.6552	.2081	.1248	.5444	.4402	.2564	.7108	.3614
2094	.6496	.3879	.1595	.5856	.4202	.3887	.6262	.4200
2111	.8790	.2011	.2207	.3304	.2731	.2521	.8932	.5324
2131	.5251	.2022	.0938	.2832	.2071	.2431	.8069	.3047
2812	.5993	.1707	.1182	.7208	.5883	.1299	.6973	.3442
2852	.6468	.3958	.1619	.7146	.5549	.2100	.6144	.4236
2896	.8211	.2921	.2063	.8400	.7617	.3048	.4794	.5019
2992	.3855	.2160	.0585	.6110	.4596	.0769	.7659	.2105
3011	.7812	.1182	.1650	.5428	.4738	.1895	.7785	.4370
3221	.6578	.4987	.1940	.5610	.4292	.4111	.6426	.4593
3229	.6700	.3807	.1683	.6295	.4794	.4513	.5983	.4327
3272	.9171	.4825	.3753	.6672	.5530	.4678	.5691	.6734
3297	.6940	.3294	.1594	.5845	.4629	.2641	.6971	.4268
3411	.7942	.5678	.2811	.8512	.7653	.3963	.4087	.5925
3511	.8219	.5286	.3022	.7259	.6315	.4319	.5297	.6055
3572	.8326	.1743	.2103	.3066	.2045	.2490	.8615	.5177
3581	.4879	.1354	.0782	.5972	.4679	.1670	.6895	.2547
3584	.7187	.4986	.2382	.6646	.5594	.4544	.5732	.5170
3593	.6341	.3737	.1401	.6037	.4492	.2729	.6598	.3908
3615	.7486	.4037	.1984	.6748	.5642	.4034	.5828	.4869
3617	.5778	.4294	.1373	.6321	.4898	.3414	.5967	.3842
3715	.7276	.4827	.2331	.7785	.6377	.6380	.4524	.5099
3741	.8868	.5896	.4087	.7984	.6638	.4089	.4625	.6737
3861	.6515	.6370	.2371	.8572	.7217	.6225	.3692	.4978

industries, as measured by the four-firm Census ratio. This is similar to the 5-digit samples analyzed in previous chapters. The effect is to consider only a part of the concentration curve for U.S. manufacturing as a whole.

TABLE 6-B: Descriptive Statistics for Four-Digit Sample

Index	Mean	Standard Deviation	Range
CR(4)	.71	0.13	.38 - .92
DI(4)	.36	0.15	.12 - .64
HERF	.20	0.09	.06 - .41
GINI	.62	0.15	.28 - .86
PIETRA	.51	0.15	.21 - .72
THI	.34	0.14	.08 - .64
ENTRO	.63	0.13	.37 - .89
CCI	.46	0.12	.21 - .67

Table 6-C indicates that the pairwise index correlations for the 4-digit sample are similar to those that were obtained for the various derivative 5-digit samples. The Census ratio is highly intercorrelated with the Herfindahl index and Horvath's comprehensive index. A similar pattern is observed for the Gini coefficient, the Pietra ratio, and the relative entropy index. One difference does appear. Pairwise correlations

TABLE 6-C: Product-Moment Index Correlations for Four-Digit Sample

	CR(4)	DI(4)	HERF	GINI	PIETRA	THI	ENTRO	CCI
CR(4)	1.00	.310	.848	.037	.135	.441	-.164	.922
DI(4)		1.00	.674	.541	.508	.817	-.774	.634
HERF			1.00	.257	.309	.657	-.447	.974
GINI				1.00	.982	.352	-.901	.209
PIETRA					1.00	.347	-.889	.276
THI						1.00	-.666	.661
ENTRO							1.00	-.406
CCI								1.00

between the ranked-share index and all other indices, except the four-firm disparity index, are lower in magnitude. However, the pattern of index correlations are remarkably similar. The aggregated industry sample, although small in number, is consistent with the analysis performed at the product class level.

Price-Cost Margins as a Measure of Profits

Price-cost margins were initially used as a measure of profitability by Collins and Preston.[2] The margin is defined as the difference between value added and various variable

[2] Collins, N.R. and Preston, L.E., Concentration and Price-Cost Margins in Manufacturing Industries (1968), Appendix A.

costs expressed as a percentage of value of shipments. Since this measure is the ratio of profit plus fixed costs (e.g., depreciation) and thus excludes capital costs and advertising, the industry capital output ratio was included in Collins and Preston's regression analysis. Moreover, geographic market segmentation was accounted for by a dispersion index. The inclusion of these variables in the regression model was necessitated by the use of a price-cost margin. Interindustry differences in margins could be more appropriately attributable to differences in concentration after their inclusion.

Many alternative measures of industry profits exist. Interestingly, price-cost margins are indicators of monopoly power, particularly the Lerner index, and are directly related to the gap between price and marginal cost. The choice of price-cost margin measures was largely based on its availability at the 4-digit industry level for a large group of industries. The data used was that reported in Collins and Preston for 1958. The 4-digit sample has already been screened to account for changes in classification or definition by the Bureau of Census.

The question immediately arises as to the compatibility of 1950 concentration data and 1958 margin data. It would, of course, be desirable for the concentration and price-cost data to be coincident with respect to time. This is not possible. This lack of correspondence is not detrimental if concentration is a stable structural characteristic, so that a highly "concentrated" industry in 1950 is very likely to be so characterized in 1958. Since the theory claims that price-cost margins should vary positively with concentration, 1958 margins should vary positively with the 1950 concentration data if these are appropriate measures of the 1958 situation. The data from the Bureau of Census indicate that concentration, as measured by the Census ratio, tends to be very stable through time.[3] One suggestive result is the correlation between each index series over time. For the sample used in this analysis the correlation between the 1950 CR(4) estimates and the reported 1958 CR(4) ratios is 0.83. Unfortunately, the four-firm Census ratio is the only 1958 series available. However, the pattern of index intercorrelations for 1950 is similar to those that have been found for the 1958 data.[4] Although the correlation between the four-firm ratios is not perfect and the stability of the remaining measures can only be examined indirectly by the pattern of index intercorrelation, it is reasonable to claim that the noncontemporaneous nature of the data does not do serious damage to an attempt to determine the relative predictive power of various measures of concentration.

In summary, it has been argued that price-cost margins are indicators of an industry's profitability, that the non-compatible time span of the data is not damaging when structural stability is a reasonable expectation, and that

---

[3]J.S Bain, "Comparative Stability of Market Structure," Industrial Organization and Economic Development (1970), pp. 38-46.

[4]See Table 6-C in conjunction with R.A. Miller, "Numbers Equivalent, Relative Entropy, and Concentration Ratios," SEJ (July, 1972), p. 110.

for the 4-digit sample suggestive correlations indicate at least relative stability. With the price-cost margin and aggregated concentration data, an empirical test of the concentration-proft hypothesis was conducted. The results are reported in the following section.

Regression Results

Simple bivariate equation specifications indicate that the various alternative measures of concentration are not identical with respect to their relative predictive power. Emerging as superior to the other measures are the Gini coefficient and the relative entropy index. In fact, only these indices had significant linear relations at the 10% level. The inclusion of capital-output ratios as a regressor did not produce a significant change in any equation, although $\bar{R}^2$ did increase. A similar coefficient pattern was observed when the geographical dispersion index was added. These variables were included to account for segmented markets and differing capital intensities across industries.

These expanded equation specifications are similar to those that were estimated by Miller.[5] A comparison with Miller's work shows that there is a general complementarity. In both cases, relative entropy appears superior to the four-firm Census ratio. Furthermore, in the 4-digit sample the

[5] *Ibid.*, p. 111.

Gini coefficient equation had $\bar{R}^2 = 0.12$, similar to the entropy measure and larger than the traditional Census-ratio index. However, these conclusions must be tempered by the realization that the estimated relations, i.e., those that included the capital-output ratio and the geographical dispersion measure, are not statistically significant at the 10% level. This is largely a degrees-of-freedom problem. The contribution of the capital-output ratio and the dispersion index is not enough to compensate for the loss in degrees-of-freedom.

Of interest are the signs of the coefficients; negative for the Gini coefficient and positive for the entropy measure.[6] Some after-the-fact theorizing may explain the unexpected signs obtained by the two measures. Recall that GINI was calculated using relative mean differences.[7] Consider two industries, A and B, in which the average firm size is the same but the average difference in firm size is higher for Industry A. Greater average size differentials indicate greater asymmetry among the firms in A, increasing the likelihood of dominant firm pricing with the smaller firms accepting the established price as their marginal revenue curve.[8]

There may be some critical value associated with this

---

[6]See Chapter Appendix B for a compilation of the estimated equations.

[7]See the discussion in Chapter II on pages 21-24.

[8]The role of asymmetric market shares in influencing the effectiveness of oligopolistic coordination has not been given much theoretical or empirical attention, except for the dominant-firm model. W.G. Shepherd points out: "...most discussions of oligopoly now routinely assume it to be largely a grouping of equals; and virtually nothing in oligopoly theories or general discussions yet recognizes or analyzes asymmetry as a (cont.)

difference, but the point here is that Industry A is likely to have higher profits than Industry B. The coefficient sign should be positive.

Alternatively, consider two industries in which the average difference in firm size is the same, but the average firm size is larger in Industry B. Further, suppose that the larger average size in B is attributable not to market size (or sales) but to a smaller number of existing firms. Because firms are smaller in number, they are more likely to establish tacit agreements on price and develop coordinated behavior patterns. This may also induce a perceived increase in the limit price function, deterring the threat of new entry. In any event, Industry B is likely to have higher profits than Industry A. The coefficient sign should be negative, since average firm size appears in the denominator of GINI. It seems that this latter consideration is dominant in this sample.[9]

A similar line of reasoning can be developed for the relative entropy measure. This is presented by Miller.[10] Briefly, suppose two industries have the same number of firms but different entropy. Higher values of ENTRO, therefore,

---

major element of market structure." W.G. Shepherd, "On Appraising Evidence About Market Power," <u>Antitrust Bulletin</u> (Spring, 1967).

[9]The key to this line of argumentation is whether or not increases in average firm size are associated with decreases in firm number, wholly or partially. The simple correlation between average and firm number for this sample is -0.08.

[10]Miller, <u>op</u>. <u>cit</u>., p. 109.

reflect greater uncertainty with respect to consumer purchases and presumably lower profits. This implies a negative coefficient sign. Larger numbers of firms with identical entropy also imply lower profits. But since numbers are in the denominator, the coefficient sign should be positive. The regression results, consistent with Miller's, suggest that numbers dominate.

Thus, although the strengths of relationships were not large, traditional measures of concentration, e.g., Census ratio or Herfindahl index, were generally inferior to the Gini coefficient and the relative entropy measure.

Because the predictive power of the alternative indices differed, a more extensive examination was conducted. Simple equation specifications or ones that include geographical market segmentation and capital intensive effects on profitability across different industries cannot separate direct and indirect influences of alternative concentration measures on industry profitability. For example, suppose the four-firm Census ratio increases for a certain industry. There is a predicted direct effect on industry profits. However, there may be indirect effects on profits that are channeled through the dispersion of leading firm market shares, i.e., increases in CR(4) may imply changes in π and changes in DI(4) whose alteration also implies changes in π. Symbolically,

$$\frac{d\pi}{dCR(4)} = \frac{\partial\pi}{\partial CR(4)} + \frac{\partial\pi}{\partial DI(4)} \cdot \frac{\partial DI(4)}{\partial CR(4)}$$

where $\pi = \pi[CR(4), DI(4)]$

Equations were estimated that paired the four-firm Census ratio with other alternative measures of concentration.

TABLE 6-D: Regressions with Census Ratio Paired with Other Concentration Indexes

Const.	CR(4)	DI(4)	GINI	PIETRA	ENTRO	CCI	HERF	$\bar{R}^2$
15.35	28.3 (1.29)	-21.85 (2.08)						**0.18
25.95	21.25 (1.82)		-21.47 (2.15)					**0.19
21.54	23.5 (1.96)			-20.85 (1.95)				*0.16
-6.80	24.72 (2.11)				26.73 (2.29)			**0.20
5.07	74.27 (2.42)					-65.03 (1.91)		*0.15
0.28	56.73 (2.52)						-64.67 (1.91)	*0.15

*significant at 10%

**significant at 5%

t-ratio in parentheses

As indicated in Table 6-D, CR(4) becomes significant when paired with various other concentration indices. Pairings with DI(4), GINI, and ENTRO are significant at the 5% level, while the remaining equations are significant at the 10% level. The only insignificant pairing was with the ranked-share index and it is not reported. The inclusion of K/O and GD as regressors did not add any strength to the overall relationship.[11]

In summary, the regression results suggest that the interrelationship between profitability and concentration is far more complex than conventional theory implies. The sole reliance on Census ratios in estimating this relation appears to be insufficient. Other measures of concentration have significant, independent effects on profitability as measured by price-cost margins. The conclusions are surely tentative when we consider the sample size, its make-up, and the non-contemporaneous nature of the data. However, when viewed in conjunction with the previous research done by Miller, the case for dispersion measures of concentration seems strong. It is time that the appropriate government agencies collect and publish

---

[11]One equation of this form is of interest for comparison purposes with Miller's work.

Const.	CR(4)	ENTRO	K/O	GD	$\bar{R}^2$
-12.59 (0.96)	22.74 (1.91)	29.61 (2.41)	0.097 (1.23)	0.022 (0.46)	0.19*

The estimated relation is similar to Miller's.

data which more richly describes firm-size distributions than do the discrete four- and eight-firm concentration ratios.

APPENDIX A

Illustrative Calculation of Four-Digit Data Using Five-Digit Product Class Value of Shipments

Suppose hypothetical industry 0000 has a total value of shipments of 2000 which are distributed among four component product classes with the following firm size distribution:

Class Number	00001		00002		00003		00004	
Total Class Sales	500		800		400		300	
Reported Sales	500		400		300		200	
Coverage Ratio	100%		50%		75%		67%	
Firm Sales*	B	100	A	180	B	90	B	60
	C	75	C	50	C	70	A	50
	A	70	E	40	A	50	C	30
	D	60	F	35	F	25	D	25
	E	50	H	30	G	25	J	15
	F	35	B	25	E	20	L	15
	H	30	D	15	D	15	G	5
	G	25	G	15	J	5		
	I	25	J	5				
	K	20	L	5				
	M	10						

*Letters indicate firm code

Under the coverage rule for 5-digit product classes, only 00001 would be included in the analysis of Chapters IV and V. However, the reported value of shipments of the firms in the component classes sum to 1400, i.e., 70% of the 4-digit industry total. Thus, it would be included in the 4-digit sample. Also, there are 13 different firms in the component classes but only a few operate in all.

The aggregated firm size distribution is simply the summation of each firm's value of shipments in each class.

Firm Code	Sales	Firm Code	Sales
A	350	H	60
B	275	I	25
C	225	J	25
D	115	K	20
E	110	L	20
F	95	M	10
G	70	Total Sales	1400

From this aggregated firm size distribution, the index values are computed. For example, CR(4) equals 69%. Note that a weighted average of the CR(4)'s in the component classes, where the weights are the proportion of the total reported value of shipments accounted for each product class, produces a CR(4) ≈ 72%. This results from the fact that the leading four firms in the 4-digit industry are not the leading firms in every product class.

APPENDIX B

Compilation of Regression Results

Const.	Geo	K/O	CR(4)	ENTRO	GINI	PIETRA	DI(4)	HERF	CCI	THI	$\bar{R}^2$
13.2			20.3								0.06
13.3				22.7*							0.09*
40.6					-20.8*						0.11*
36.7						-18.1					0.06
32.7							-14.4				0.03
26.0								7.86			0.00
22.5									11.0		0.00
28.9										-4.09	0.00
9.22	.036	.064	18.9								0.01
4.29	.027	0.12		26.5**							0.09
37.6	.015	0.12			-25.4**						0.12
33.2	.015	0.11				-21.6*					0.05
26.9	.036	.097					-15.5				0.00
20.3	.040	.082						6.83			0.00
17.4	.039	.077							9.73		0.00
23.0	.040	.084								-4.14	0.00
12.5		.036	19.3								0.03
6.40		.098		27.3*							0.12*
39.4		.112			-26.2**						0.15*
35.1		.098				-22.6*					0.09
30.5		.069					-15.9				0.02
24.2		.051						6.77			0.00
21.2		.047							9.84		0.00
26.8		.054								-4.20	0.00

APPENDIX B
(Cont.)

Const.	Geo	K/O	CR(4)	ENTRO	GINI	PIETRA	DI(4)	HERF	CCI	THI	$\bar{R}^2$
-11.2		.081	23.1*	30.3**							0.21**
26.6		.094	18.9		-25.9**						0.21**
21.3		.081	21.8*			-24.4**					0.17*
14.4		.051	27.2**				-22.6**				0.16*
43.9		.022	55.4**					-63.6*			0.12
45.3		.030	72.9**						-64.4*		0.12
12.6		.030	26.6*							-15.2	0.04

*significant at the 10% level

**significant at the 5% level

other results reported in Chapter

CHAPTER SEVEN: MEASUREMENT OF CONCENTRATION IN RETROSPECT

The general purpose of my inquiry has been to elucidate the issues concerned with the measurement of market concentration, the key variable in any investigation focused upon the causes and consequences of oligopolistic market structures. Although many alternative measures exist, there seems to be a consensus that one index is as "good" as another. A careful examination of this proposition does not provide such an ardent conclusion. The various measures do capture different characteristics of the firm-size distributions and they are not perfectly substitutable, as indicated by the low correlations between some of the indexes.

Unfortunately, a decision about which measure is appropriate in such circumstances would have to be arbitrary, since economic theory provides no a priori basis for claiming superiority of one index as opposed to another. On the other hand, principal component analysis permits a researcher to capture objectively the diverse elements of structure embodied in the various indices, thereby obtaining a richer measure of the firm-size distribution. Whether or not this artificial representation is meaningful depends largely on its interpretability in terms of the original indexes. If the data were available, a principal component representation of an enlarged index set would be superior to the reliance on only one statistic of concentration, the four-firm concentration ratio of the Bureau of the Census.

The inferiority of sole dependence on Census ratios as <u>the</u> measure of concentration also is seen from the structure-performance test in that some measures predicted "better" than others, i.e., had greater explanatory power as measured by the adjusted coefficient of determination. The relative entropy index and the Gini coefficient emerged as superior to other traditional measures of concentration. Although hampered by a small sample and non-contemporaneous data, the regression results suggest that a richer data set, like the one utilized in the foregoing analysis, is needed if our understanding of the causes and effects of oligopolistic markets is to increase.

Thus, this study has demonstrated that a reassessment of the measurement of market concentration is required. The renewed debate will enhance our knowledge of market dominance, its measurement and its interpretability in terms of the causes and the effects of oligopolistic markets.

REFERENCES

Adelman, I. and Preston, L.E. "A Note on Changes in Industry Structure." RESTat (February, 1960).

Adelman, M.A. "The Measurement of Industrial Concentration." REStat (November, 1951).

Adelman, M.A. "Rejoinder on 'The Measurement of Industrial Concentration.'" REStat (May, 1952).

Adelman, M.A. "The Measurement of Industrial Concentration: Rejoinder." REStat (November, 1952).

Adelman, M.A. "Differential Rates and Changes in Concentration: REStat (February, 1959).

Adelman, M.A. "Comment on 'H' Concentration Measure as a Numbers-Equivalent." REStat (February, 1969).

Bailey, D. and Boyle, S.E. "The Optimal Measure of Concentration." JASA (December, 1971).

Bain, J. Industrial Organization. John Wiley & Sons, 1968.

Bain, J. "The Comparative Stability of Market Structures." Industrial Organization and Economic Development. Edited by Markham and Papanek. Houghton Mifflin, 1970.

Bartlett, M.S. "Note on Multiplying Factors for Various Chi-Square Approximations." J. Roy. Stat. Soc. (1954).

Berle, Jr., A.A. "Comment on 'The Measurement of Industrial Concentration.'" REStat (May, 1952).

Bernhardt, I. and MacKenzie, K.D. "Measuring Seller Unconcentration, Segmentation and Product Differentiation." WEJ (December, 1968).

Blair, J.M. "The Measurement of Industrial Concentration: Reply." _RESTat_ (November, 1952).

Blair, J.M. "Statistical Measures of Concentration in Business: Problems of Compiling and Interpretation." _Bulletin of the Oxford University Institute of Statistics_ (November, 1956).

Collins, N.R. and Preston, L.E. _Concentration and Price-Cost Margins in Manufacturing Industries_. University of California Press, 1968.

Dhrymes, P.J. "On the Measurement of Price and Quantity Changes in Some Consumer Capital Goods." University of Pennsylvania, Economic Research Unit, Discussion Paper 67 (1967).

Edwards, C.D. "Comment on 'The Measurement of Industrial Concentration.'" _RESTat_ (May, 1952).

Fellner, William. _Business Concentration and Price Policy_. Princeton: National Bureau of Economic Research, Princeton University Press, 1955.

Finkelstein, M.O. and Friedberg, R.M. "The Application of an Entropy Theory of Concentration to the Clayton Act." _Yale Law J._ (March, 1967).

George, E.B. "Comment on 'The Measurement of Industrial Concentration.'" _RESTat_ (May, 1952).

Grossack, I.M. "Towards an Integration of Static and Dynamic Measures of Industry Concentration." _RESTat_ (August, 1965).

Hall, M. and Tideman, N. "Measures of Concentration." _JASA_ (March, 1967).

Hall, M. and Weiss, L.W. "Firm Size and Profitability." _RESTat_ (August, 1967).

Hart, P.E. "Statistical Measures of Concentration vs. Concentration Ratios." RESTAT (February, 1961).

Hart, P.E. and Prais, J.S. "The Analysis of Business Concentration: A Statistical Approach." J. Roy. Stat. Soc., Series A, Part II (1956).

Henning, J.A. "Marginal Concentration Ratio: Some Statistical Implications - Comment." SEJ (October, 1969).

Hexter, J.L. and Snow, J.W. "An Entropy Measure of Relative Aggregate Concentration." SEJ (January, 1970).

Hexter, J.L. and Snow, J.W. "An Entropy Measure of Relative Aggregate Concentration - Reply." SEJ (July, 1971).

Horowitz, A. "Suggestion for a Comprehensive Measure of Concentration: Comment." SEJ (April, 1972).

Horowitz, A. and I. "Entropy, Markov Processes and Competition in the Brewing Industry." SEJ (July, 1968).

Horowitz, I. "Numbers Equivalents in U.S. Manufacturing Industries: 1954, 1958, and 1963." SEJ (April, 1971).

Horvath, J. "Suggestions for a Comprehensive Measure of Concentration." SEJ (April, 1970).

Horvath, J. "Suggestions for a Comprehensive Measure of Concentration: Reply." SEJ (April, 1972).

Joskow, J. "Structural Indicies: Rank-Shift Analysis as a Supplement to Concentration Ratios." RESTAT (February, 1960).

Kaplan, A.D.H. Big Enterprise in a Competitive System. Washington: 1954.

Kendall, M.G. "The Geographical Distribution of Crop Productivity in England." J. Roy. Stat. Soc. (1939).

Kilpatrick, R.W. "The Choice Among Alternative Measures of Industrial Concentration." REStat (May, 1962).

Kottke, F.J. "An Entropy Measure of Relative Aggregate Concentration - Comment." SEJ (July, 1971).

Linter, L. and Butters, J.K. "The Measurement of Industrial Concentration: Further Rejoinder." REStat (November, 1952).

Miller, R.A. "Marginal Concentration Ratio and Industrial Profit Rates: Some Empirical Results of Oligopoly Behavior." SEJ (October, 1967).

Miller, R.A. "Marginal Concentration Ratios: Some Statistical Implications - Reply." SEJ (October, 1969).

Miller, R.A. "Marginal Concentration Ratios as Market Structure Variables." REStat (August, 1971).

Miller, R.A. "Numbers Equivalent, Relative Entropy, and Concentration Ratios." SEJ (July, 1972).

Morrison, D.F. Multivariate Statistical Methods. McGraw-Hill, 1967.

Nelson, R.L. Concentration in the Manufacturing Industries of the United States. New Haven: Yale University Press, 1963.

Pashigian, P. "Market Concentration in the United States and Great Britain." J. of Law and Economics (October, 1968).

Prais, S.J. "The Statistical Conditions for a Change in Business Concentration." REStat (August, 1958).

Quandt, R.E. "On the Size Distribution of Firms." AER (January, 1966).

Rosenbluth, G. Business Concentration and Price Policy. NBER, Princeton University Press, 1955.

Rottenberg, I. "New Statistics on Companies and on Concentration in Manufacturing from the 1954 Census." Amer. Stat. Assoc.: Bus. and Econ. Section (Proceedings, 1957).

Saving, T.R. "Concentration Ratio and the Degree of Monopoly." IER (February, 1970).

Scherer, F.M. Industrial Market Structure and Economic Performance. Rand McNally, 1970.

Scitovsky, T. Business Concentration and Price Policy. NBER, Princeton University Press, 1955.

Shepherd, W.G. "On Appraising Evidence About Market Power." Antitrust Bulletin (Spring, 1967).

Silberman, I.H. "On Lognormality as a Summary Measure of Concentration." AER (September, 1967).

Simon, H. and Bonini, C. "The Size Distribution of Business Firms." AER (September, 1958).

Singer, E.M. Antitrust Economics: Selected Legal Cases and Economic Models. Prentice Hall, 1968.

Stackelberg, Henrich von. The Theory of the Market Economy. London: Wm. Hodge & Co., Ltd., 1952.

Stekler, H.O. Profitability and Size of Firm. University of California Institute of Business and Economic Research (1963).

Stocking, G.W. "Comment on 'The Measurement of Industrial Concentration.'" REStat (May, 1952).

Stone, J.R.N. "The Interdependence of Blocks of Transactions." J. Roy. Stat. Soc. (1947).

Theil, Henri. Economics and Information Theory. Rand McNally, 1967.

Weiss, L.W. "Average Concentration Ratios and Industrial Performance." JIE (July, 1963).