INdustry and Company Analysis Program
Katholieke Universiteit Leuven
Departement voor Toegepaste Economische Wetenschappen
Dekenstraat 2
B-3000 Leuven ☎ 016/22.75.17

Onderzoeksrapport Nr. 8601

THE HERFINDAHL INDEX AND CONCENTRATION RATIOS REVISITED

By

Leo E. Sleuwaegen, Raymond R. De Bondt and Wim V. Dehandschutter

February 1986
Revised May 1986

Wettelijk depot D/1986/2376/2.

THE HERFINDAHL INDEX AND CONCENTRATION RATIOS REVISITED.

By

Leo E. Sleuwaegen, Raymond R. De Bondt and

Wim V. Dehandschutter.

Introduction

One issue of long standing in both economics (industrial organisation) and antitrust practice, concerns the correct and workable measurement of market power. Market power is intrinsically difficult to empirically grasp, if only because it is related to both structural (concentration) and behavioral (collusion) characteristics of the industry which in turn may be related to each other. In addition, nontrivial problems arise if market power measures are related to some attributes of economic welfare, such as allocative efficiency[1]. Notwithstanding a continuing debate on aspects of these issues (in the field of industrial organisation) antitrust laws are applied, relying among other things on measures of concentration.

1 Cf. efficiency vs. market power controversy. See e.g. Demsetz. Two systems of belief about monopoly, in Goldschmid, Mann and Weston (Eds.), Industrial concentration, the new learning, Boston, Little & Brown, (1974), 164-184; Clarke, Davies and Waterson, The profitability-concentration relation: markey power or efficiency?, The Journal of Industrial Economics, (1984), 32, 435-450.

The 1982 and 1984 modifications of the Justice department's merger guidelines[2] include a shift from the simple and intuitively "clear" four-firm concentration ratio (C_4) that equals the total market share of the four largest suppliers, to the somewhat more sophisticated Herfindahl (H) (or Herfindahl-Hirschman) index that equals a weighted average of all of the "revelant" firms' market shares, with each market share given a weight equal to its own value.

From a practical point of view it would appear worthwile to have a clear, operationally useful, understanding of the H index of concentration. A number of authors already made important progress towards "explaining" the H index, by relating it to C_4[3]. Through numerical examples and mathematical manipulation it has been shown that no guarantee exists for a one to one relationship between the H and C_4 indices. Two industries with different market share distributions can have the same H value. And with a given H number typically will correspond a range of C_4 values. It appears useful to report in this paper more fully on the characteristics and implications of the correspondence relation between the C_k (k firm concentration ratio) and the

2 See, for example Bronsteen, A review of the revised Marger Guidelines, The Antitrust Bulletin (1984), 29, 613-652.

3 For this Journal, see Weinstock, Using the Herfindahl Index to measure concentration, The Antitrust Bulletin (1982), 27, 285-301; Miller, The Herfindahl-Hirschman Index as a market structure variable: an exposition for antitrust practitioners, The Antitrust Bulletin (1982), 27, 593-618; Weinstock, Some little-known properties of the Herfindahl-Hirschman Index: problems of translation and specification, The Antitrust Bulletin (1984), 29, 705-717.

H index. Data analysis has revealed a horn shaped relationship between the H index and concentration ratios, with several characteristics that can contribute to a better understanding of these measures.

In the following section the boundaries of the C_k-H correspondence are discussed and illustrated with Belgian and U.S. data. The relevance for interpreting and valuating H values and for assessing economic performance is discussed next after which the relation with the U.S. Department of Justice's merger guidelines is clarified.

Relating the Herfindahl index to concentration ratios.

Suppose that a relevant industry is served by n firms, and that firms are ranked according to their size in terms of sales, with the largest firm being firm 1. The Herfindahl index of concentration H is

$$H = \sum_{i=1}^{n} s_i^2$$ with s_i denoting the market share of the i-th largest firm.

It can be verified that $0 \leq H \leq 1$ and

$$H = 1/n + \sigma^2 n$$

with σ^2 equalling the variance in the market shares, i.e

$(1/n)[\sum_{i=1}^{n} (s_i - 1/n)^2]$. All firms having equal market share means that $\sigma^2 = 0$ and then $H = 1/n$.

The k firm concentration ratios C_k, $k=1,2,3,4,...$ essentially give no weight to the market share of firms $j > k$. For example, $C_2 = s_1 + s_2$ and $C_4 = s_1 + s_2 + s_3 + s_4$, indicate respectively the two- and four-firm concentration ratio.

Both the C_k and H indices are a reflection of the market share distribution in the industry and hence it is not surprising that they somehow are related. Likewise it is clear, however, that in general there does not exist a functional (one to one) relation since the C_k is calculated from partial information on the size distribution of firms while the H index takes the complete distribution into account.

Using the extreme possibilities of market share distributions it is nevertheless possible to define a range that contains all possible H values for given C_k values[4]. This C_k-H range turns out to have the shape of a horn, since the range of possible H values increases as the magnitude of C_k (for a given k) becomes larger. The horn shaped correspondence is dependent on both the number of firms in

4 Or all possible C_k values for given H magnitudes.

the industry (n) and on the number of firms (k) included in the C_k computation[5].

For a "large" number of firms in the industry (strictly speaking for $n \to \infty$) the boundaries of the H-C_k relation are relatively simple, i.e.

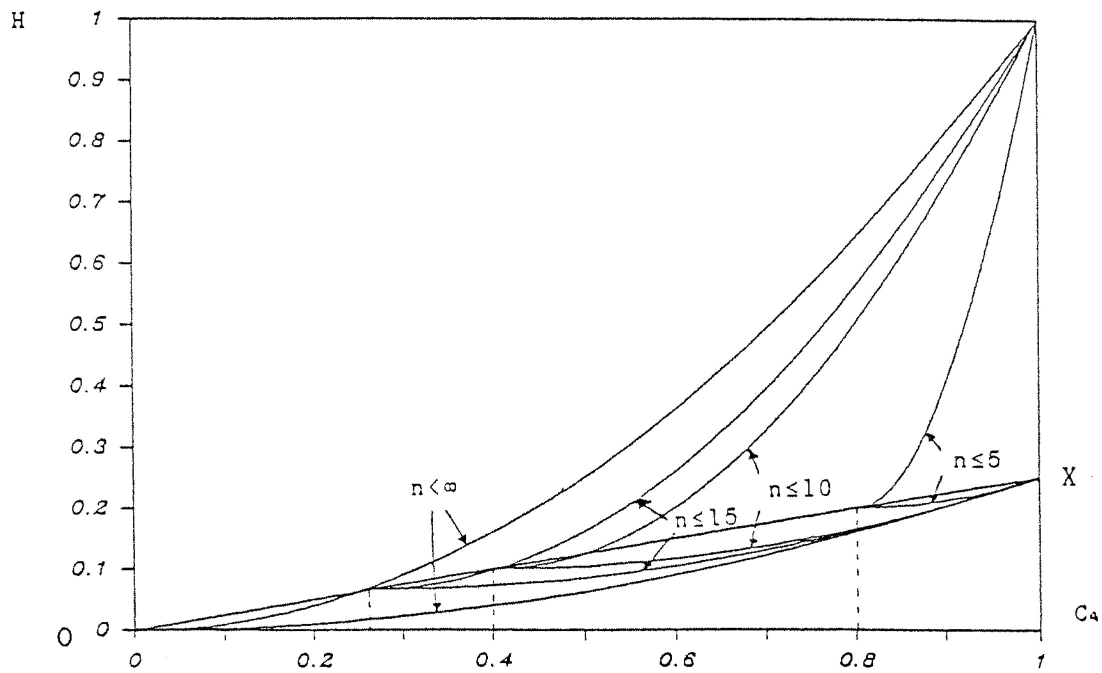

Fig.1. The horn shaped boundaries for the C_4-H relation

5 In general for n firms and a C_k ratio, which cannot be smaller than k/n, the boundaries are for $n > k$, min $H = C_k^2/k + (1-C_k)^2/(n-k)$; max $H = \max\{[1-(n-1)(1-C_k)/(n-k)]^2 + (n-1)[(1-C_k)/(n-k)]^2, C_k/k\}$, except for a fraction if the maximum is C_k/k and $k(1-C_k)/C_k$ is non-integer. A proof is given elsewhere, see Sleuwaegen and Dehandschutter, The critical choice between the concentration ration and the H index in assessing industry performance, forthcoming in The Journal of Industrial Economics, (1986).

$$\min H = C_k^2/4$$

and $\max H = C_k^2$ for $C_k \geq 1/k$

$\max H = C_k/k$ for $C_k \leq 1/k$ except for a fraction if k/C_k is not an integer number. This mathematical range turns out not to be very sensitive to the actual number of firms in the industry, except if there are only a few firms say $15 \leq n \leq 4$. This is illustrated for $k = 4$, $n \to \infty$, $n=15$, $n=10$, $n=5$ in the H-C_4 diagram of fig. 1.

The boundaries of the horn shaped correspondence shown in Fig. 1 define the range of possible H values that can be observed, because of purely mathematical reasons, with a given C_4 concentration ratio. The actual H and C_4 pattern for given industries may change, say from year to year or with the number of firms or with the level of aggregation, but in any case it will always lie within the horn boundaries corresponding to a large number of firms. For some cross sections the relation between the actual H and C_4 values may be approximately linear, but no guarantee whatsoever exists for this to be the typical case. This point is well underscored by Fig. 2 and 3. The Belgian horn pattern for 3 digit NACE sectors clearly widened from 1975 to 1981. The more disperse horn pattern in 1981 stayed nevertheless, as it should, within the boundaries indicated.

Published data for the U.S. 3 digit sectors (see Fig. 4) are less revealing and would appear to suggest a linear relation between the C_4 and H values. This, it should be clear now, is an artifact of the limited data set which

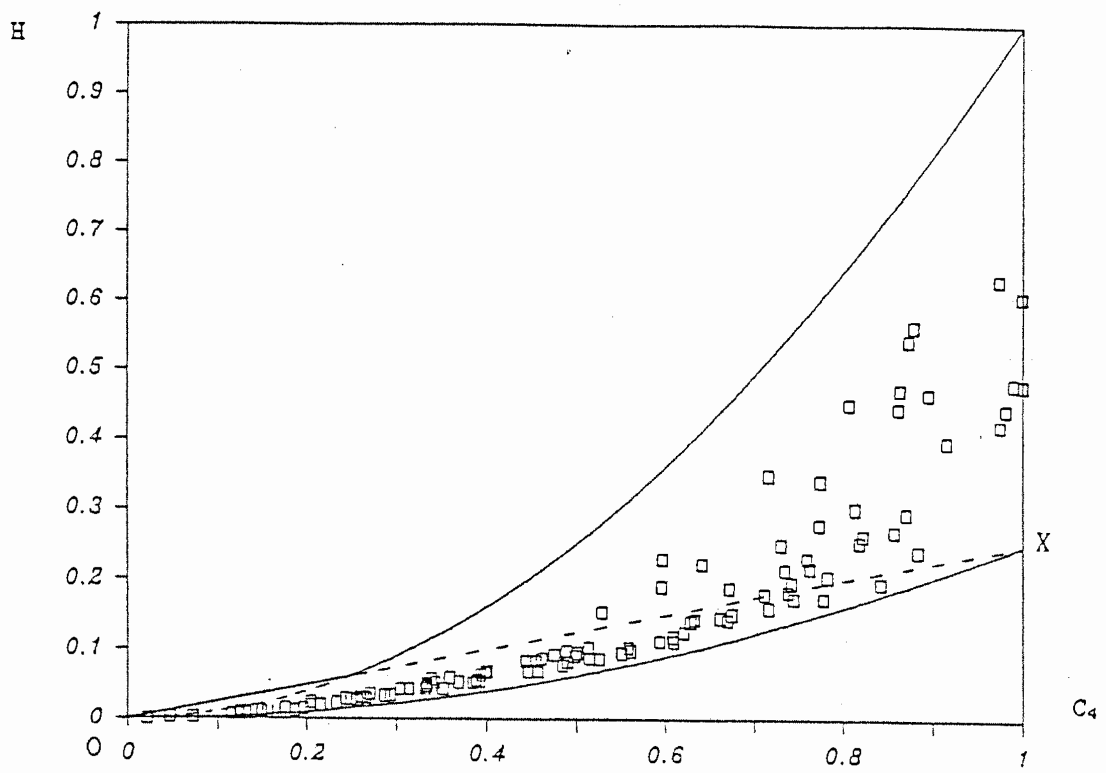

Fig.2. The horn relationship for Belgian industries (1975)

(source: unpublished statistics from the Belgian National Statistical Institute).

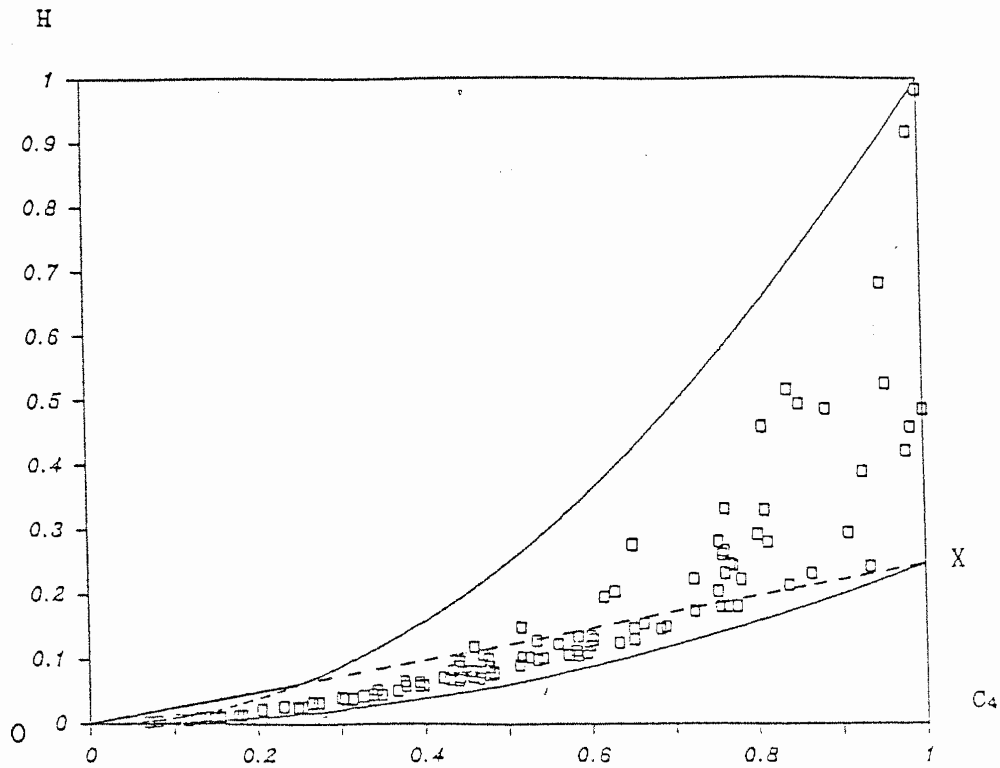

Fig. 3. The horn relationship for Belgian industries (1981)

(source: unpublished statistics from the Belgian National Statistical Institute)

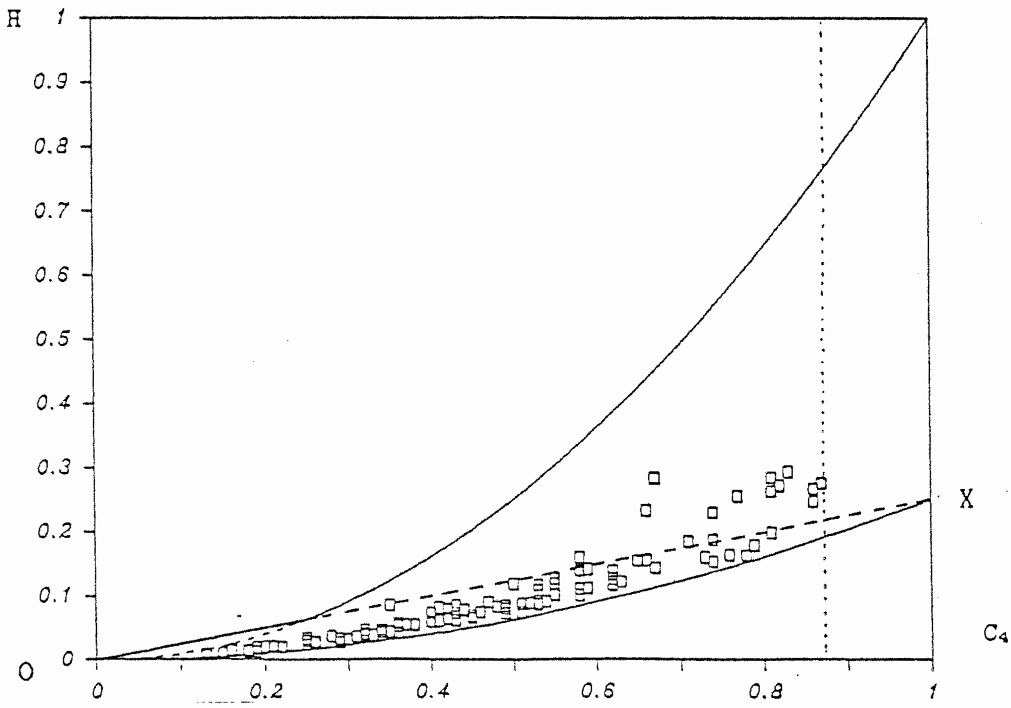

Fig. 4. The horn relationship for U.S. industries (1956)

(source: Nelson, R.L. (1963), Concentration in the Manufacturing Industries of the United States, New Haven, Yale Univ. Press)

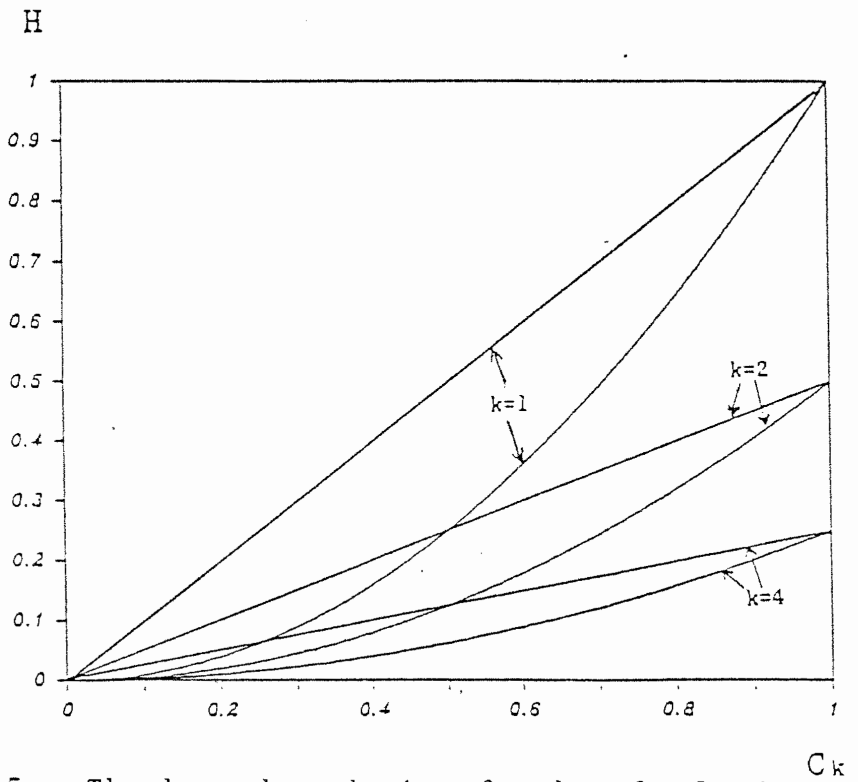

Fig. 5.: The horn boundaries for $k = 1, 2, 4$ and $n \to \infty$

stops at C_4 = .87 (presumably because of confidentiality reasons)[6].

The horn shaped correspondences between H and other C_k ratio's follow a pattern that is illustrated by Fig. 5. This figure indicates the boundaries for the H-C_1, H-C_2 and H-C_4 relationships in industries with a large number of firms. Within the corresponding set of boundaries the range for a small number of firms in the industry follows the pattern illustrated earlier in Fig. 1.

Interpreting H and C_k values

The horn shaped boundaries reveal some regularities than can be further interpreted. The horn boundaries for a finite number of firms all have their origin at a line through the origin (OX in Fig. 1). This line represents the minimum possible levels of H and C_k that can be observed in an industry with a number of firms equal to $n = 1/H = k/C_k$.

To see this, note that, for example for the H-C_4 combinations, the line OX represents all H and C_4 values for which $H = C_4/4$. The corresponding concentration levels all

6 A referee pointed out that unpublished material prepared by the U.S. Department of Justice contains similar horn shaped correspondence for U.S. industries. The investigation of the boundaries and their implications, however, appears to have not been explored previously.

have the same numbers equivalent. The numbers equivalent n_e is the number of equally sized firms that corresponds with a given concentration measure value. It is the lowest possible number of firms for that given concentration measure value. For example, the numbers equivalent n_e^H for $H = 0.2$ satisfies $0.2 = 1/n_e^H$ or $n_e^H = 5$, while for $C_4 = 0.4$ the numbers equivalent $n_e^C = 4/C_4 = 10$. In general n_e^H and n_e^C solve:

$$H = 1/n_e^H \text{ and}$$

$$C_4 = 4/n_e^C$$

Those values of H and C_4 for which the n_e values are equal, satisfy

$$n_e^H = 1/H = n_e^C = 4/C_4$$

or $H = C_4/4$ and $n_e^H = n_e^C = 4/C_4$

When the actual number of firms equals the calculated numbers equivalent, then the market is equally distributed and consequently the concentration is minimal for that given n.

Both for the U.S. and Belgium the majority of 3 digit industries are characterized by C_4 and H values that lie below the line $H = C_4/4$ (below OX in Fig. 2, 3 and 4). H values corresponding with a C_4 lower than or equal to $C_4/4$ can be considered "low" while H values larger than the $C_4/4$ value of the industry can intuitively be judged to be "high". Indeed, if $H < C_4/4$ it means that given the market share of the four largest firms, the inequality between the firms in the industry is not very large as the following example illustrates.

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	C_4	H	σ^2
.28	.20	.19	.13	.07	.04	.03	.02	.02	.02	.80	.18	.008
.34	.21	.14	.11	.07	.04	.03	.02	.02	.02	.80	.20	.010
.44	.16	.11	.09	.08	.055	.025	.02	.01	.01	.80	.25	.015

More precisely, if $n = 10$ it is clear that a concentration level of $H = 0.18$ is larger than its minimum value ($1/n = 1/10 = 0.10$), hence, firms in the industry have unequal market shares ($\sigma^2 > 0$). The inequality between the firms, with the four largest firms having a share of 0.80, is however not very large in the precise sense that the positive variance (σ^2) is not able to tilt the H value above 0.20 (0.80/4). A five firm industry with a $C_4 = 0.80$ would have a minimum $H = 0.20$. Now there is a 10 firm industry also with a $C_4 = 0.80$ and unequal market shares. But precisely because there is a small variance in the market shares the larger number of firms lowers the H value below 0.20. If and only if the variance in market shares is very small will an increase in n lower the H value and if and only if this variance is large enough will the H index in the 10 firm industry with $C_4 = 0.80$ be larger than the minimum H index in the 5 firm industry with $C_4 = 0.80$[7].

7 The reasoning is based on $dH/dn = \sigma^2 - 1/n^2$.

Calculating H values from C_k

The horn correspondences also allow to calculate H values from a limited knowledge of market shares. The intuition that the H index somehow gives a more "complete" picture of the industry, because it takes into account all market shares, needs to be qualified. The spread for example of possible H values that can accompany C_2 values in an industry is very low (see Fig. 5)[8]. Somewhat crudely one could state: if you know the market shares of the two largest firms, you don't need to know the other shares, since the possible H values lie within a narrow bound that can be calculated.

The range of possible H values that can accompany C_k is dependent on k and this can be explored to approximate H. If one is interested in obtaining the H index for a particular industry and allows for a maximum error of 1 percent, it

8 The maximum spread of possible H values is $\Delta \leq \max H(n\to\infty) - \min H(n\to\infty)$ with $\max H(n\to\infty)$ and $\min H(n\to\infty)$ defined on p. 5. For $C_k < 1/n$ $\Delta \leq C_k/k - C_k^2$ approximately. This range increases as k gets larger but is small for k = 2 anyway. For $C_k \geq k$, $k \geq 2$, $\Delta \leq ((k-1)/k)\, C_k^2$ and the maximum spread of H values is minimal for k = 2!

suffices to sum the squares of the k largest firms and an amount HR given by the following expression[9].

$$HR = 0.5\ [(1-C_k)\ (s_k + (1-C_k)/(n-k)]$$

where, as before, $C_k = s_1 + s_2 + \ldots + s_k$.

The number of firms k required in the computation is implicitly given by the condition:

$$s_k(1-C_k) - HR \leq 0.01 \qquad (*)$$

This implies that in many cases the need for information on market shares may be very small making the computation of H indices very easy. For example, take an industry composed of 50 firms but where $s_1 = 0.40$, $s_2 = 0.30$, $s_3 = 0.10$, $s_4 = 0.07$, $s_5 = \ldots$, $s_{50} = \ldots$ To obtain an almost perfect approximation of the H-index it suffices to take into account only the market shares of the three largest firms.

---

9 The H index, $H = \sum_{i=1}^{n} s_i^2$, can be written as $H = H_k + H_{n-k}$, where $H_k = \sum_{i=1}^{k} s_i^2$ and $H_{n-k} = \sum_{i=k+1}^{n} s_i^2$. If s_k (the market share of the k largest firm) is known, straightforward application of the transfer-principle (see Hannah and Kay, p. 22) implies:
$\max H_{n-k} = (1-C_k)s_k + (\mu^2 - \mu)s_k^2$, where μ is the fraction implied by the division in the first term.
Similarly, $\min H_{n-k} = (1-C_k)^2/(n-k)$.
The possible spread of H values is $\max H_{n-k} - \min H_{n-k}$.
The midpoint is $(\max H_{n-k} + \min H_{n-k})/2$. The condition (*) ignores the small fraction μ and takes a 0.01 % deviation upwards and downwards from this midpoint.

Hannah and Kay, Concentration in modern industry, Theory, Measurement and the UK Experience, London, The MacMillan Press, (1977).

Using the method outlined above, computing HR for the three largest firms yields:

$$HR = 1/2\ [(1-0.80)(0.10 + 0.004)] = 0.0104$$

implying that the precision condition is met:

$$(1-0.8)0.1 - HR \leq 0.01$$

The corresponding H index approximately equals:

$$H = (0.4)^2 + (0.3)^2 + (0.1)^2 + 0.0104$$

$$= 0.2704$$

The horn correspondence and economic performance

Assessment of possible welfare losses in an industry on the basis of one concentration measure may be misleading, in view of the horn correspondence and established economic theory. Static misallocations and excess profits because of market power, will tend to be larger as the H index increases, even if the firms in the industry act independently[10]. This does not mean, however, that an industry with a low H necessarily will be accompanied with

10 See Dansby and Willig, Industry Performance Gradient Indexes, American Economic Review, (June 1979), 69, 249 -260.

small welfare losses[11]. Indeed a low H may be accompanied with a higher C_4 and if the leading four firms were to collude, excess profits will tend to be high as well[12]. If a low H were to be accompanied with a low C_4 (C_k) there would be no need to look beyond H.

To achieve an operational interpretation of H values without reference or knowledge of the C_4, the several regions of the horn correspondence can be helpful to achieve guidelines for antitrust analysis.

The U.S. Department of Justice's new merger guidelines

One of the most noted modifications that resulted from the 1982 revision of the merger guidelines was the adoption of the H index instead of the four-firm concentration ratio as a first indicator of possible exercise of market power. The primary argument for this change was the observation that the H index reflects better than the C_4 the actual distribution of firm sizes in the industry. Another important argument brought forward by the Justice Department

11 As would be suggested, for example, by the contribution of Stigler, A theory of oligopoly, Journal of Political Economy, (Febr. 1964), 72.

12 See also Clarke, Davies and Waterson, opus. cit. and Sleuwaegen, On the nature and significance of collusive price-leadership, International Journal of Industrial Organisation, (1986), 4 (in press).

is the greater predictability of its decisions since the decision rules concerning whether or not to challenge a merger are clearly specified in terms of the post merger H-index and the resulting increase of the H-index. The report states :

> The Department divides the spectrum of market concentration as measured by the HHI into three regions that can be broadly characterised as unconcentrated (HHI below 1000), moderately concentrated (HHI between 1000 and 1800) and highly concentrated (HHI above 1800). An empirical study by the Department of the size dispersion of firms within markets indicates that the critical HHI thresholds at 1000 and 1800 correspond roughly to four-firm concentration ratios of 50 % and 70 %, respectively. Although the resulting regions provide a useful format for merger analysis, the numerical divisions suggest greater precision than is possible with the available economic tools and information. Other things being equal, cases falling just above and just below a threshold present comparable competitive concerns".

Comments on these changes have argued that the use of the H-index has a probusiness benefit : a non-competitive industry according to the C_4 ratio is classified more competitive when using the H-index the more the largest firms' shares are equal[13].

This division of the spectrum does not, contrary to common interpretation, necessarily imply a loosening or a non-change of the guidelines in judging mergers. Figure 6 indicates the "winners" (those industries which are considered less concentrated, the dotted line zones) and the "losers" (the more concentrated, the dashed line zones)

13 See, for example, A Loosening of Merger Rules, Business Week, (May 17, 1982).

resulting from the switch from C_4 thresholds to the new H-index thresholds. Comparing the highly concentrated industries defined according to the C_4 or the H upper tresholds, it should indeed be observed that the

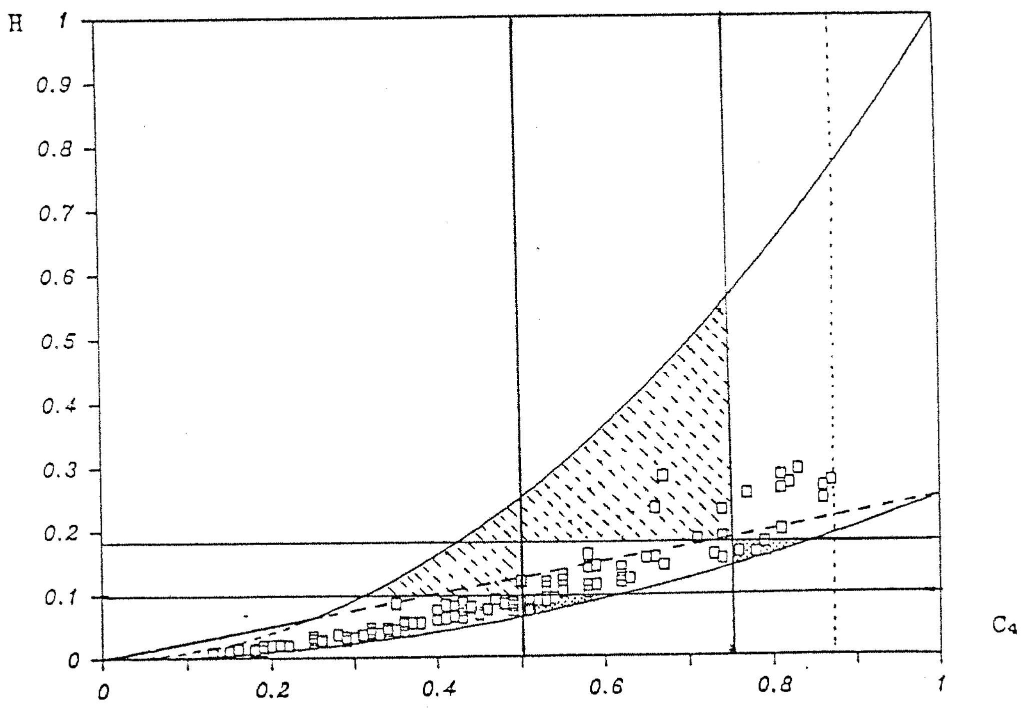

Fig. 6 : The horn relationship and the tresholds of the Justice Dept. for both the C_4 (1968 guidelines) and the H index (1982 guidelines) illustrated with U.S. 1956 data (Nelson (1963)).

actual (1958 VS) distribution of firms would imply a larger number of losers. But as distributions change this need not necessarily be the case (anymore).

However, in view of the new evidence on welfare losses and concentration measures, the adoption of H-index tresholds constitutes some progress. The upper H index treshold clearly delineates a region of industries characterized by high H index, as well as high C_k values, related under either collusive or non-collusive behaviour to unwanted possible allocative inefficiencies.

Conclusions

There exists a horn shaped correspondence relation between the Herfindahl index and the k firm concentration ratios that allows to :

- define the range of possible H values that can exist in an industry with given C_k ratios (or alternatively the range of possible C_k's with a given H);

- interpret better changes in H and C_k, since high (low) H values may be equated with magnitudes larger (smaller) than C_k/k;

- possibly approximate the range of possible H values if partial observation on say market shares of leading firms are available;

- interpret the new merger guidelines, which in view of the horn shaped correspondence do not, in principle, mean a "loosening" of merger policy, but are, better than before, in line with recent evidence on the relationship between welfare losses and structural concentration measures.

The discussion of the horn shaped correspondence suggested that in fact it may make little sense to restrict the focus on H. Given existing insights on welfare losses associated with either collusion or independent behavior, it makes sense to relate H to the corresponding C_4. In fact the boundaries employed in the new merger guidelines implicitly appear to do precisely this.