A Decomposition of the Herfindahl Index of Concentration

de Gioia, Giacomo

26 November 2017

A Decomposition
of the Herfindahl Index of Concentration

Giacomo de Gioia

Abstract

The Herfindahl index is one of the most known indices used to measure the concentration of a variable distributed over a certain number of units, and tipically to measure the degree of concentration of business in a market. Its worth is the sensitivity both to the dimensional variability of these units and to their numerical consistency. In this note a decomposition of the H-index into these two terms is offered.

Keywords: Herfindahl, Decomposition of the Herfindahl index, Industrial concentration, Metric index of concentration.

JEL Classification: C43, C49, D40, L11.

Introduction

The reasons why the Herfindahl index is widely used to measure the degree of concentration in a certain market are essentially two:

(i) it is very simple to be calculated; and

(ii) it allows to capture the two main features of the industrial concentration that a good index has to take into account, that is the dimensional inequality across firms in the observed market and their number.

Here we want to isolate and quantify these two essential components.

A metric definition of concentration

Suppose we have to measure the degree of concentration of a positive variable X distributed over n units according to x_i, $i = 1, 2, ..., n$. To speak concretely, we can think, for example, of n firms of an industry and their respective sales recorded in a certain period.

So we can consider the $\mathbb{R}^n_+$-vector:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

together with the n-dimensional vector having all the coordinates equal to $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$:

$$\bar{\mathbf{x}} = (\bar{x}, \bar{x}, \dots, \bar{x})$$

with a clear meaning of both of them: $\mathbf{x}$ is the point of the space $\mathbb{R}^n_+$ actually (empirically) observed, whereas $\bar{\mathbf{x}}$ is the ideal point representing the equidistribution state of the variable X across the units under consideration (in our example, the firms of an industry). Therefore one of the most natural way to define the concentration of X is to determine how close to or how far from $\bar{\mathbf{x}}$ the actual point $\mathbf{x}$ is. For this we can compute the (*Euclidean*) *distance* between the two points (see also Ricci, 1975, p. 42):

$$d(\mathbf{x}, \bar{\mathbf{x}}) = \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sqrt{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2} \qquad (1)$$

Such a measure can be normalized taking the ratio to the greatest virtually observable distance in the mentioned space of the $\mathbb{R}^n_+$-vectors, consistent with our actual data set on the firms' sales. Naturally this maximum distance is the distance between $\bar{\mathbf{x}}$ and a maximum-concentration point in which the total volume of the market is thought as attached to a single firm, the others having an output equal to zero, say $\mathbf{x}^*$ (this is the *concentration-maximizing distribution*) [see Appendix (a)]. Hence we have:

$$\max_{\mathbf{x}} d(\mathbf{x}, \bar{\mathbf{x}}) = d(\mathbf{x}^*, \bar{\mathbf{x}}) =$$

$$= \sqrt{\left(\sum_i x_i - \bar{x}\right)^2 + (n-1)(0 - \bar{x})^2}$$

$$= \sqrt{\left(\sum_i x_i\right)^2 - n\bar{x}^2} = \sqrt{(n-1)n\bar{x}^2} \qquad (2)$$

and a relative index of concentration (with $n > 1$) is:

$$\tau = \frac{d(\mathbf{x}, \bar{\mathbf{x}})}{d(\mathbf{x}^*, \bar{\mathbf{x}})} = \sqrt{\frac{\sum_i x_i^2 - n\bar{x}^2}{(n-1)n\bar{x}^2}} \tag{3}$$

A first immediate remark is that $\tau = \sigma_X/\sigma_{X^*}$, namely the ratio of the standard deviations of the actual distribution to the concentration-maximizing distribution of X. Moreover we see that $\tau \in [0,1]$: if $\tau = 0$, it means that $\mathbf{x}$ coincides with the equidistribution point, where by definition the degree of concentration is null; if $\tau = 1$, $\mathbf{x}$ coincides with one of the max-concentration points; and, finally, if $0 < \tau < 1$, the empirical point $\mathbf{x}$ lies in the space $\mathbb{R}_+^n$ in an intermediate position, and more exactly at a distance from $\bar{\mathbf{x}}$ equal to the fraction τ of the maximum observable distance (2): the closer to (resp. farther from) $\bar{\mathbf{x}}$ the actual point is, the more τ tends to 0 (resp. to 1), thereby providing a measure of the degree of concentration of X.

A decomposition of the Herfindahl index

The metric index τ has two main advantages: first, it is very easy to compute, and, second, it has a clear geometric, intuitive meaning. However, in spite of these advantages, this index might not be completely satisfactory especially as regards the analysis of industrial concentration phenomena, because it takes into account only the aspect of the dimensional heterogeneity of firms, not being affected by their number, in the same sense that it could be said about other concentration ratios, like the Gini index for instance. More exactly, any distribution of X with n equal-sized firms always corresponds to null concentration for τ, and so $\tau = 0$ for any market with 2 or 200 or 20,000 identical firms. It is quite evident that this is a crucial issue from the economic point of view, because it is not acceptable that an industry could be examined with reference to the features of concentration regardless of the number of players. The number of firms has relevant effects on the firms' behaviour itself, on the degree of competition, in short on the market conditions; in fact, it is very hard to think that a market with two firms only is the same that one with tens or hundreds or more. For this reason a good index of industrial concentration should be able to embed both of the main aspects related to (a) the dimensional inequality across firms and (b) the number of firms within the industry, according to a wider definition of concentration.

One of the most common indices holding such a good property is the Herfindahl index, which has in reality a very simple structure:

$$H = \sum_{i=1}^{n} s_i^2 \tag{4}$$

where s_i represents the market share of the i-th firm, that is $x_i/\sum_i x_i$. While the maximum of H is again 1, in case of equidistribution we have $s_i = 1/n$ for all i's, and $H = 1/n$, which is therefore the minimum of H. It follows that the Herfindahl index allows to distinguish among uniform distributions according to the number of the firms in the market, in the sense that H increases as n decreases (as indeed we expect from an economic point of view). Thus it is usually said that the Herfindahl index is able to capture the effect on concentration by side not only of the dimensional variability of firms but also of their number. Our purpose is to isolate clearly these two effects [see Appendix (b)].

In this regard, first of all it is necessary to express τ in terms of H, which is not difficult to obtain if we divide by $(\sum_i x_i)^2$ both numerator and denominator of the ratio under the square root in (3) (that in turn is the same as considering s_i's in place of x_i's, meaning that τ is exempt from scaling problems). Then, by virtue of (4), we get:

$$\tau = \sqrt{\frac{n}{n-1}\left(H - \frac{1}{n}\right)} = \sqrt{\frac{nH-1}{n-1}} \tag{5}$$

which represents a sequence of functions whose limit is $\sqrt{H}$ as $n \to +\infty$ [see Appendix (c)].

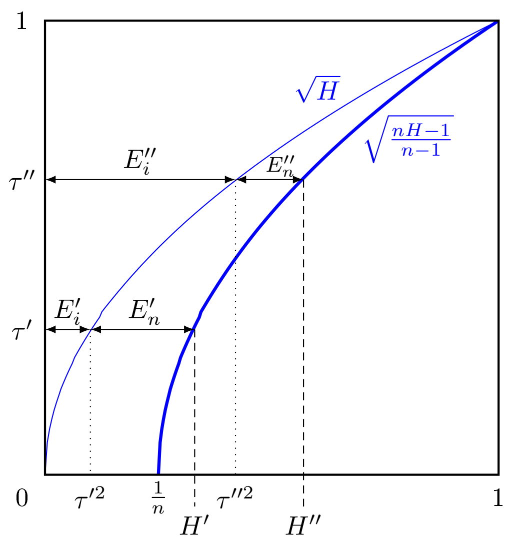

Figure 1. An H-index decomposition: for each level of H (e.g. H' or H'' on the horizontal axis) and of τ (e.g. τ' or τ'' on the vertical axis), the decomposition of H into E_i and E_n, respectively the inequality and the n effects on the concentration.

In Figure 1 expression (5) is depicted together with $\sqrt{H}$ in such a way that we can visualize the horizontal distance between the two curves. The point is that this distance reflects the difference between two market conditions: one being a limit case with an infinite number of firms – on the curve $\sqrt{H}$ taken as a term of comparison –, and the other one with the actual (finite) number of firms, *in correspondence of the same degree of metric concentration* τ. Hence, this distance can be assumed as a measure of the pure n-effect on the Herfindahl index, the complement being related to the remaining determinant that we can call inequality effect. Notice that, as wanted, the n-effect on H is unambiguous: as n increases, H certainly decreases for each level of τ, and vice versa H certainly rises as n falls.

Some developments and remarks:

A) for any $n > 1$ there exists a one-to-one correspondence between H and τ. For any n and for each level of H (or, that is the same, for each level of τ), H can be decomposed into two parts: the above-mentioned inequality effect (E_i) and the n-effect (E_n). See Figure 1 for an exemplification. It is easy to realize that:

$$E_i = \tau^2 \qquad \left(= \frac{nH-1}{n-1} \right) \tag{6}$$

$$E_n = H - \tau^2 \qquad \left(= \frac{1-H}{n-1} \right) \tag{7}$$

B) Moreover, by virtue of (5), we have:

$$H = \frac{(n-1)\,\tau^2 + 1}{n} \tag{8}$$

so that the (first-order) differences of H with respect to n (τ constant) are:

$$\Delta H_n = \frac{\tau^2 - 1}{n(n+1)} \leq 0$$

which expresses the fact that, *given whatever* τ, there exists an inverse relation between H and n (and hence between E_n and n).

C) The most interesting point is the one which mainly emerges just from looking at Figure 1: as we pass from minimum levels of dimensional inequality, measured by τ (or E_i), to higher ones, the effect of the number of firms decreases; in other words, the firms' crowding on the market is bound to have a decreasing-in-importance role in the determination of H as the players' inequality raises. And this seems to be reasonable: in a neighbourhood of the maximum of τ the market tends to a monopoly whatever the number of firms (think about the substantial equivalence of two industries both dominated by one very big company, but in a case flanked by just one very small firm and, in the other, by a large number of little co-players); on the other side, in a neighbourhood of the minimum of τ, where players are all fairly comparable, their number is rather crucial in discriminating even very different situations: a market with few quite equal firms (= duopoly, oligopoly) and that one with a great many firms (= competitive market). In particular:

$$\text{if } \tau = 1 \quad \Rightarrow \quad H = E_i \text{ and } E_n = 0$$

$$\text{if } \tau = 0 \quad \Rightarrow \quad H = E_n \text{ and } E_i = 0$$

As a last remark we must note that sometimes, in some applications, normalizations of the Herfindahl index are used for the claimed purpose to make markets with different n comparable and hence to neutralize the n-effect. It is not difficult to conclude that this is in principle incorrect, because, for all said thus far, the distinctive function of H is precisely to catch that effect. Indeed we can see that the standard normalization, for example, obtained as a composition ratio with respect to the range of variation, i.e.:

$$\frac{H - 1/n}{1 - 1/n}$$

corresponds exactly to E_i we derived above at (6), which discloses the real meaning of this kind of operation. The result is therefore and simply to reduce the index to just one component of two, with consequent remarkable loss of information (unless, obviously, E_n is unimportant). On the normalization of H, see, for instance, the recent work of Cracau and Lima (2016).

Conclusions and examples

If studying the concentration of business in a certain market at a given point in time our aim is to answer questions like: "How much of it is explained by the factor α rather than the factor β?", we need to partition the measure of the phenomenon into measure of α + measure of β, both positive. Here, in particular, our intention has been to decompose the Herfindahl index of concentration (H) into two main components, one (E_i) expressing the (squared) relative distance of the actual state of the market with respect to the ideal equidistribution state (same number of firms, same volume of the market, but equidistributed), and the other one (E_n), a sort of distance from that other ideal state of the market in which the number of firms tends to infinity, so-called perfect competition.

This decomposition hence allows to get precise measure of situations as follows: let $n = 100$ and consider $\sum_i x_i = 1000$ distributed at 60% within the 40%-top units (for the sake of simplicity suppose uniformely, in and out of this 40% range); then we obtain:

H	τ	E_i	E_n
0.01167 (100%)	*0.04103*	0.00168 (14.4%)	0.00998 (85.6%)

Now let $n = 50$ and the same volume of the market at 80% within the 20%-top units:

H	τ	E_i	E_n
0.06500 (100%)	*0.21429*	0.04592 (70.6%)	0.01908 (29.4%)

where we see both components increased, easy to guess also without decomposition, but with an inversion in weights, less easy to ascertain without decomposition.

Appendix

(a) Note that $(\sum_i x_i)^2 \geq \sum_i x_i^2$ for the assumption $x_i \geq 0$, and so $(2) \geq (1)$.

(b) Some representations of H in terms of main determinants are well known. For instance, in Hay and Morris (1984, p. 142):

$$H = n\sigma^2 + \frac{1}{n}$$

where σ is the standard deviation of market shares, but in this espression H is lacking of a clear one-way dependence by n. In Scognamiglio Pasini (2013, p. 138):

$$H = \frac{c^2 + 1}{n}$$

where c is the coefficient of variation, that is the ratio of the standard deviation of x_i's to $\bar{x}$, which in fact remedies to the previous inconvenience, but it seems to be not suitable for a decomposition of H into separate pure positive additive components.

(c) Precisely we can think of (5) as $\tau_n : [1/n, 1] \to [0, 1]$ functions of H, extendable to $\tilde{\tau}_n : [0, 1] \to [0, 1]$ functions of t as follows:

$$\tilde{\tau}_n(t) = \begin{cases} 0 & \text{if } 0 \leq t < 1/n \\ \tau_n(t) & \text{otherwise} \end{cases}$$

or simply see below (8) and notice that it converges to τ^2, $\tau \in [0, 1]$, as $n \to +\infty$.

References

Cracau, D. and Lima, J. E. D. (2016), "On the Normalized Herfindahl-Hirschman Index: A Technical Note", *International Journal on Food System Dynamics* 7 (4), 382-386.

de Gioia, G. (2000), "Measures of concentration and probability. A decomposition of the Herfindahl index with an application to Italian data", Research report at Master in Economics–Coripe Piemonte, Moncalieri (Turin); unpublished.

Hay, D. A. and Morris, D. J. (1984), *Economia industriale. Teoria e verifiche empiriche*, Il Mulino, Bologna. (Orig. ed.: *Industrial Economics. Theory and Evidence*, Oxford University Press, Oxford, 1979.)

Ricci, F. (1975), *Statistica ed elaborazione statistica delle informazioni*, Zanichelli, Bologna.

Scognamiglio Pasini, C. (2013), *Economia industriale. Economia dei mercati imperfetti*, Luiss University Press, Roma.