NBER WORKING PAPER SERIES

THE SURPRISING HYBRID PEDIGREE OF MEASURES OF DIVERSITY AND
ECONOMIC CONCENTRATION

Paolo M. Adajar
Ernst R. Berndt
Rena M. Conti

The Surprising Hybrid Pedigree of Measures of Diversity and Economic Concentration
Paolo M. Adajar, Ernst R. Berndt, and Rena M. Conti
NBER Working Paper No. 26512
November 2019
JEL No. B4,C43,D22,D4,I1,K21,L1,L4,L5

<u>ABSTRACT</u>

Measures of economic concentration, such as the k-firm concentration index and the Hirschman-Herfindahl Index (HHI) are commonly used to ascertain the competitiveness of a product market. Within a Cournot model industry equilibrium, it is known a relationship exists between the HHI and the gap between industry price and marginal cost, but the economic theory foundations and intuition underlying the HHI formula are seemingly arbitrary. Here we document that there are indeed powerful and intuitive theoretical foundations to the HHI, but those foundations emanate from outside economics, namely, ecology, where the HHI is known as Simpson's Diversity Index. We discuss the origins of the HHI and Simpson's Diversity Index, summarize other measures of concentration, and link them to common measures of inequality. Based on a priori reasoning, we conclude there is little on which to base a choice between the HHI and non-HHI measures of market concentration. We empirically illustrate the implementation of the HHI and other concentration indexes as the statin drug LipitorTM lost patent protection and faced generic competition in 2012; we find very similar empirical trends and high correlations among them. Our research provides support for the continued use of HHI as a measure of concentration, provided one recognizes its link to market power is equivocal.

Paolo M. Adajar
MIT Department of Economics
50 Memorial Drive
Cambridge, MA 02142
United States
padajar@mit.edu

Ernst R. Berndt
MIT Sloan School of Management
100 Main Street, E62-533
Cambridge, MA 02142
and NBER
eberndt@mit.edu

Rena M. Conti
Boston University
Questrom School of Business
Department of Markets, Public Policy and Law
595 Commonwealth Ave
Boston, MA 02215
rconti@bu.edu

Hybrid **n.** [L. *hybrid,* offspring of mixed parentage]. **1.** the offspring produced by crossing two individuals of unlike genetic constitution; specif., the offspring of two animals or plants of different races, varieties, species, etc.

Pedigree **n.** [ME. *pedigru, p de gre < MFr/pie de grue, lit., crane's foot <* L. *pes,* FOOT + *grus,* a crane: from the lines in the genealogical tree] **1.** a list of ancestors; record of ancestry; family tree

> *Webster's New World Dictionary of the American Language,*
> *New York: The World Publishing Company, 1970*

* * * * * * * * * * * * * * * *

I. INTRODUCTION

Students of industrial organization have long believed that market structure, market behavior and various market performance outcomes are interrelated phenomena, albeit with causality relationships among them considerably more complex than the classical "structure, conduct and performance" paradigm.[2]

One strand in the industrial organization literature has focused on characterizing and then using measures of market structure to ascertain how non-competitive is the industry, regardless of whether the extent of non-competitiveness is exogenous or endogenous. Since under competition price (P) equals marginal cost (MC); one way of evaluating any actual product market's departure from competition is to observe the size of the gap between P and MC. Specifically, when firms (with or without market power) choose output levels to maximize profits, Lerner [1934] showed that an index of market power is simply $\frac{P-MC}{P} = -1/\varepsilon$, where ε is the firm's (negative) own-price elasticity of demand.[3] In practice, it is often difficult to measure MC directly and accurately. The empirical literature commonly

---

[2] For a brief overview of the structure, conduct and performance paradigm, see Jean Tirole, *The Theory of Industrial Organization*, Cambridge, MA: The MIT Press, 1988, "Introduction: Why Should One Be Interested in Industrial Organization?", pp. 1-3.

[3] Abba P. Lerner, "The Concept of Monopoly and the Measurement of Monopoly Power", *Review of Economic Studies*, June 1934, 1:157-175.

addresses this challenge by reverse engineering the Lerner Index using econometric methods to estimate ϵ and thereby indirectly estimate gaps between P and MC.[4]

Are there other methods by which to quantify indirectly the gap between P and MC? At least since Stigler [1964], it is known that under certain assumptions involving Cournot behavior among oligopolists selling a homogeneous product at price (P), in industry equilibrium there is a positive association between industry variable profits (i.e., industry-wide P – MC) and a measure of industry concentration known as the Herfindahl-Hirschman (HHI) concentration index.[5] However, for firms and industries in which sunk or fixed costs are substantial and marginal costs are low (such as with many contemporary digital and R&D intensive products), competition may force P to be equal to average total costs, even as the gap between P and MC is large, thereby rendering HHI a problematic metric of market power.[6]

In spite of this uncertainty, there is a long tradition in economics of characterizing market structure by defining and constructing measures of industry or market concentration, and implicitly or explicitly relating them to indicators of market power such as profits. Syverson, for example, asserts "In the recent macroeconomic market power literature, the most frequently used measure of market power is concentration".[7] The most common measures are k-firm concentration ratios (typically for k = 4, 8, 20

---

[4] For recent examples, see Susanto Basu, "Are Price-Cost Markups Rising in the United States? A Discussion of the Evidence", *Journal of Economic Perspectives*, Summer 2019, 33(3):3-22; Chad Syverson, "Macroeconomics and Market Power: Context, Implications and Open Questions", *Journal of Economic Perspectives*, Summer 2019, 33(3):23-43, and Steven Berry, Martin Gaynor, and Fiona Scott Morton, "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization", *Journal of Economic Perspectives*, Summer 2019, 33(3):44-68..

[5] The proof of this relationship linking industry profitability to the HHI is essentially anticipated in George J. Stigler, "A Theory of Oligopoly", *Journal of Political Economy* [1964], 72(1), 44-61, specifically on page 55. For related proofs and discussions, see *inter alia*, Jean M. Tirole [1988], *op. cit*., pp. 221-223; F. M. Scherer and David Ross, *Industrial Market Structure and Economic Performance*, Third Edition, Boston, MA: Houghton Mifflin Company, 1990, pp.70-73; and Dennis W. Carlton and Jeffrey M. Perloff, *Modern Industrial Organization*, Fourth Edition, Boston, MA: Pearson Addison Wesley, 2005, Appendix 8A, p. 283.

[6] See, for example, John Sutton, *Sunk Costs and Market Structure: Price Competition, Advertising and the Evolution of Concentration*, Cambridge, MA: MIT Press, 1991; and Berry, Gaynor and Scott Morton [2019], *op. cit*., especially pp. 54-56.

[7] Syverson [2019], *op. cit.*, p. 25.

or 50 firms), or the HHI. Notably, the theoretical foundations of these measures of market concentration

and their linkages to market power are limited, leaving use of concentration measures potentially

vulnerable to Koopman's [1947] pointed criticism of the "the empirical approach" in his classic treatise,

"Measurement Without Theory"[8] – a critique that confounds antitrust enforcement issues to this day.[9]

Our contribution to this literature is that while we acknowledge continued ambiguity between

measures of market concentration and market power, we show in this paper that there are indeed

powerful and intuitive theoretical foundations to the widely used HHI metric of concentration which

emerge from outside economics. Thus, to the extent one wants to measure concentration, there are

good reasons for economists and other practitioners to continue to utilize the HHI metric.

Specifically, we argue that the characterization of stochastic outcome distributions is a common

operation in many disciplines and fields other than economics. We document that the pedigree of the

common HHI in economics involves a measure of diversity known in ecology as Simpson's Diversity

Index. Moreover, Simpson's Diversity Index is likely to be appealing to economists because of its elegant

and straightforward underlying probability intuition, thereby lending qualified support to the use of HHI

in economics. We discuss the origins of the HHI and Simpson's Diversity index, summarize other

measures of concentration and link them to common measures of inequality. We note the existing

economic literature on concentration reveals a preference for indexes that incorporate aspects of both

variability/inequality and absolute population size over such measures as the Gini index that only

incorporates relative inequality. Based on *a priori* reasoning, we conclude there is little on which to base

a choice between HHI and non-HHI measures of market concentration.

---

[8] Tjalling Koopmans, "Measurement Without Theory", *The Review of Economics and Statistics*, August 1947, 29(3):161-172.
[9] The summer 2019 issue of the *Journal of Economic Perspectives* (Vol. 33, No. 3) devotes a number of articles to the themes of measuring markups over marginal cost with implications for enforcement of antitrust statutes.

Finally, we empirically illustrate the implementation of the HHI and other concentration indexes in the context of a product whose manufacturer first experienced market exclusivity though patent protection and thus enjoyed a gap between P and MC, and then when that protection expired, encountered substantial competition, which likely reduced the markups and profits that new entrants could earn. In particular, we examine the statin prescription drug Lipitor™ (generic name atorvastatin), America's largest selling branded prescription drug earlier in this decade. We find very similar empirical trends and high correlations among them. Together, these considerations offer support and encouragement for the continued use of the HHI metric in measuring economic concentration, provided that users recognize that market structure as measured by concentration is only one of the various factors affecting market performance outcomes, and that market behavior and strategic considerations also play important roles.

II. A SHORT HISTORY OF ECONOMIC CONCENTRATION DESCRIPTIVE MEASUREMENTS

Economists and other analysts of industrial organization have long used descriptive measures to depict and quantify industry concentration. One of the earliest and simplest descriptive metrics is the k-firm Concentration Ratio (CR_k), defined as

$$CR_k \equiv \sum_{i=1}^{k} S_i, \qquad \text{(Eqn. 1)}$$

where in an N-firm market, S_i is the ratio of the ith firm's revenues (typically, its value of shipments) relative to that of revenues in the entire N-firm market, and k is the set of k largest ranked firms in that market, $i \in k$.[10] Instead of revenue shares, some analysts have utilized unit or quantity shares (particularly when firms' quantities are measured in common units), or shares of output capacity, value added or employment. Historically, common choices for k have been 4, 8, 20 and 50.[11] Virtues of CR_k

---

[10] Issues concerning market definition remain controversial, and although they are discussed briefly below, they are beyond the scope of this research.
[11] See, for example, F.M. Scherer and David Ross [1990], *op. cit.*, pp. 71-73.

include its simplicity, its ability to highlight concentration by focusing on the largest k firms, and its computational feasibility even in the presence of incomplete data about each firm in the particular N-firm market. Drawbacks of the CR$_k$ metric include its sensitivity to the arbitrary choice of k, and its inability to provide any additional potentially useful information concerning the distinct contributions among all firms sized between 1 and k and between $k + 1$ and N. Moreover, its value is invariant to whether N is, say, 10 or 10,000 as long as the share sum of the k largest firms is identical.

Arguably the most common descriptive economic concentration metric that combines elements of both firm numbers and firm variability among *all* N firms is the so-called Herfindahl-Hirschman Index ("HHI"), defined as

$$HHI \equiv \sum_{i=1}^{N} S_i^2 \,. \qquad \text{(Eqn. 2)}$$

Note that unlike the CR$_k$ metric, HHI explicitly incorporates information from each of the N firms in the market. HHI ranges from 0 to 1 with larger values signaling markets with increased concentration. When an industry contains only one firm (i.e., a pure monopolist), HHI attains its maximum value of 1. By squaring market shares, the HHI weights more (less) heavily the values for larger (smaller) firms. As a result, an attractive feature of the HHI is that when data on shares of very small firms are unavailable, it can still be reasonably informative. However, because of the HHI's disproportionate weighting of large firms, it is crucial that the largest firms' shares be measured accurately. Moreover, the lower weight assigned to smaller firms in the HHI implies that if small firms play a critical role in disciplining larger firms' behavior (as may occur in some market segments), that market disciplining role may not be captured by the HHI concentration metric.

Defining lower case s_i as the deviation of upper-case S_i from its arithmetic mean, one can rewrite HHI as[12]

$$HHI \equiv \sum_{i=1}^{N} S_i^2 = \sum_{i=1}^{N} s_i^2 + \frac{1}{N}. \qquad \text{(Eqn. 3)}$$

Hence, given firms' variability (measured as the sum of squared deviations from the arithmetic mean), the value of HHI declines with increases in the number of firms N, but increases with rising variability among any given number of firms N.[13] Consequently, the HHI captures two dimensions of concentration – the variation between firms, and the absolute number of firms in an industry. According to Hirschman "…the use of the index is indicated when concentration is a function of both unequal distribution *and fewness*.[14] Other measures of concentration, generally devised in connection with income distribution and the Lorenz curve, instead focus entirely on *relative* shares and inequality of distribution rather than on fewness and the possible disciplining of large by small firms (more on this in Section IV below where we discuss the Gini coefficient).

A particularly interesting counterfactual interpretation of (Eqn. 3) was introduced by Adelman [1969].[15] Suppose all firms were equal in size with market share S_i so that all s_i in (Eqn. 3) were zero. In that perfect equality scenario, $HHI = 1/N$ in (Eqn. 3) – i.e., in that case HHI collapses to $1/N$, and it only captures the number of firms in an industry. For a given HHI, therefore, Adelman defined the "*numbers equivalent*" N_e in the market as the number of equally sized firms that would generate the same HHI, i.e.,

---

[12] *Ibid*, fn. 45, p. 73.

[13] Rosenbluth [1955] has noted that the HHI can also be rewritten as $(c^2 + 1)/N$ where c is the coefficient of variation, i.e. $c \equiv$ standard deviation/mean. See Gideon Rosenbluth, "Measures of Concentration", in *Business Concentration and Price Policy*, Princeton, NJ: Princeton University Press, 1955, fn. 16a, p. 62.

[14] Albert O. Hirschman, "The Paternity of an Index", *American Economic Review*, 1964, 54(5): 761. Italics in original text.

[15] Morris A. Adelman. "Comment on the 'H' Concentration Measure as a Numbers-Equivalent", *Review of Economics and Statistics*, 1969, 51(1):99-101.

$$N_e = \frac{1}{HHI}.$$ (Eqn. 4)

The HHI has come to play a prominent role in public policy evaluations, particularly involving possible antitrust scrutiny of mergers and acquisitions.[16] Because it squares market share, the HHI weights large firms more highly than smaller ones, thereby taking inequality into account. However, as noted by Scherer [1990, *op. cit.*, fn. 46, p. 73], the choice of squared terms appears arbitrary, asserting, "There is no *a priori* reason why the weighting scheme need be quadratic, as with the HHI index."

Although the sum of squared shares formulation of the HHI may appear to be arbitrary, as noted in the Introduction, under certain assumptions it can be shown that in an oligopoly, the HHI is related to firms' price-cost margins. Specifically, assume an oligopoly consisting of N firms producing a homogeneous product having a common unit price P for all firms. If each firm i chooses its profit-maximizing output q_i assuming that all other firms j $(j \neq i)$ will not change their choice of profit-maximizing output levels q_i -- what the economic literature terms a Cournot model-- then it can be shown that in equilibrium where no profit-maximizing firm faces incentives to change its output quantity, industry-wide (variable) profits Π can be written as

---

[16] In public policy applications, typically the S_i shares ($0 \leq S_i \leq 1$) shares are multiplied by 100 so that they are expressed in terms of percentage points (between 0 and 100), implying that the "augmented" HHI metric ranges between 0 and 10,000. Legend has it that this multiplicative transformation of the HHI occurred after Department of Justice antitrust officials expressed frustration when attempting to square fractions. According to the 1982 US Department of Justice horizontal merger guidelines (https://www.justice.gov/atr/herfindahl-hirschman-index), mergers in markets with pre-merger HHIs above 1800 and involving an HHI increase of greater than 50 will likely involve close scrutiny and possibly a merger challenge by the Department of Justice or the Federal Trade Commission. Note that with an HHI of 1800 or 1850, the numbers-equivalent is between 5.4 and 5.6 equal-sized firms. For further discussion, see Dennis W. Carlton and Jeffrey M. Perloff, [2005], *op. cit.*, p. 645; and Michael Whinston, "Antitrust Policy Toward Horizontal Mergers", ch. 36 in Mark Armstrong and Robert Porter, eds., *Handbook of Industrial Organization,* Vol. 3, Amsterdam: North Holland Elsevier, 2007, 2369-2440. These 1982 guidelines were updated in 2010, so that mergers that increased HHI by more than 200 points and resulted in a post-merger HHI of > 2500 would be "presumed to be likely to enhance market power". See https://www.justice.gov/atr/file/810276/download. Note that with an HHI of 2500, the numbers-equivalent is 4.0.

$$\Pi = \sum_{i=1}^{N} \Pi_i = \sum_{i}^{N} (P - c_i) q_i = \left(\frac{PQ}{\varepsilon}\right)\left(\sum_{i=1}^{N} S_i^2\right) = \left(\frac{PQ}{\varepsilon}\right) HHI$$

implying that

$$HHI = \left(\frac{\Pi}{PQ}\right)\varepsilon, \tag{Eqn. 5}$$

where c_i is the marginal cost for firm i, ε is the absolute value of the market price elasticity of demand, and in the derivation the S_i share terms in HHI are quantity shares, where $S_i \equiv q_i/Q$ and $Q \equiv \sum_{i}^{N} q_i$.[17] Thus, under certain assumptions within a Cournot oligopoly framework, in equilibrium the HHI equals the product of the (variable) profit share in industry revenues times the absolute value of the market price elasticity of demand, thereby linking industry (variable) profitability shares directly and positively to the HHI.[18] Intuitively, when the variable profit share increases, firms are able to extract more profits from their revenues, thereby manifesting greater market power.

As an historical aside, we note the HHI was apparently first introduced by Albert O. Hirschman [1945] to analyze the composition of the United States' balance in foreign trade accounts; his analysis actually examined the square root of HHI.[19] When writing his Columbia University Ph.D. dissertation several years later and unaware of Hirschman's original contribution, Orris C. Herfindahl [1950] independently created the HHI, using it to analyze empirically the U.S. steel industry.[20] Gideon

---

[17] As noted above, the proof of this relationship linking industry profitability to the HHI is essentially anticipated in George J. Stigler, "A Theory of Oligopoly", *Journal of Political Economy* [1964], 72(1), 44-61, specifically on page 55. For related proofs and discussions, see *inter alia*, Jean M. Tirole [1988], *op. cit.*, pp. 221-223; F. M. Scherer and David Ross [1990], *op. cit.*, pp.70-73; and Dennis W. Carlton and Jeffrey M. Perloff [2005], *op. cit.* Appendix 8A, p. 283. Also note that the derivation of Eqn. 5 assumes a common output price P for all firms, so that quantity shares are identical to revenue shares.

[18] Note that variable profits depend on marginal rather than average costs, and that the HHI-profitability relationship is uninformative with respect to fixed and sunk costs. For discussion, see Sutton [1991], *op. cit*.

[19] Albert O. Hirschman, *National Power and the Structure of Foreign Trade,* Berkeley and Los Angeles: University of California Press, 1945, ch. 7 and pp. 157-162.

[20] Orris C. Herfindahl, *Concentration in the Steel industry,* unpublished Ph.D. dissertation, Department of Economics, Columbia University, 1950.

Rosenbluth [1955], a fellow Columbia economics Ph.D. student with Herfindahl, erroneously attributed

the original HHI development to Herfindahl [1950].[21] Both Rosenbluth [1961][22] and Hirschman [1964][23]

attempted repeatedly but unsuccessfully to rectify the historical record, and today the HHI is typically

called the "Herfindahl-Hirschman Concentration Index" (in alphabetical rather than chronological

order)[24], although Adelman [1969] respectfully and neutrally designated it as the "H" concentration

measure. In a bit of a pique, Hirschman has summarized the paternity of his "hybrid" inequality and

fewness index as follows: "The net result is that my index is named after Gini who did not invent it at all

or after Herfindahl who reinvented it. Well, it's a cruel world."[25]

Although under certain assumptions one can derive a relationship between industry profitability

and the industry HHI as in Eqn. 5, the seemingly arbitrary functional form of the HHI where shares are

squared and summed raises the issue of whether more intuitive microeconomic conceptual foundations

can be identified that yield the HHI formulation.[26] As we now document, this intuition comes from the

usage of HHI outside economics and instead in ecology/biology.

---

[21] Gideon Rosenbluth, "Measures of Concentration", in *Business Concentration and Price Policy,* Princeton, NJ: Princeton University Press, 1955, 57-99.

[22] Gideon Rosenbluth, "Round Table Gesprach uber Messung der industriellen Konzentration". *Die Konzentration in der Wirtschaft; Verhandlungen auf der Tagung des Vereins fur Socialpolitik, Gesellschaft fur Wirtschafts- und Sozialwissenschaften, in Bad Kissingen.* Berlin: Duncher & Humblot, 1960, 391-395.

[23] Albert O. Hirschman, "The Paternity of an Index", *American Economic Review,* 1964, 54(5).

[24] Herfindahl-Hirschman (HHI) indexes for U.S. manufacturing industries based on U.S. Census of Manufacturers data for 1947, 1954, and 1957 were published in Ralph L. Nelson, *Concentration in the Manufacturing Industries of the United States,* New Haven, CT: Yale University Press, 1963, pp. 17-19 and have been published in regular U.S. Census Bureau publications since then. For Canada, early HHI calculations were published in Department of Consumer and Corporate Affairs, *Concentration in the Manufacturing Industries of Canada*, Ottawa: Information Canada, 1971, Table A-1.

[25] Hirschman [1964], *op. cit.*, p. 761.

[26] Adelman [1969], *op. cit.,* p. 100, interpreted the HHI as the weighted average slope of the cumulative concentration curve – an alternative description having no behavioral foundations. Hall and Tideman [1967] have shown that the HHI is consistent with axioms they believe concentration metrics should possess. But since other indexes are also consistent with their axioms, they argue there is no best measure of concentration, and in any case they provide no behavioral foundations that generate a concentration index. See Marshall Hall and Nicolaus Tideman, "Measures of Concentration", *Journal of the American Statistical Association*, March 1967, 62(317):162-168.

III.   FROM BIOLOGICAL DIVERSITY TO ECONOMIC CONCENTRATION: SIMPSON'S DIVERSITY INDEX

There is a long tradition of measuring variety, diversity and fewness in biology. The two main characteristics typically considered when measuring fewness and diversity are *richness* and *evenness*. Richness is a measure of the number of different kinds of organisms or species present in a particular sample geographical habitat, i.e. fewness. However, diversity depends not only on richness, but also on evenness; evenness compares the similarity of the population size of each of the species present, i.e., equality in relative shares.

As a measure on its own, species richness is unaffected by the number of individuals of each species present. For example, one daisy has as much influence on the richness of a geographic sample as 1000 buttercups.[27] Evenness is a measure of the relative abundance of the different species making up the richness of the geographic sample habitat, the equality of relative shares.

For example, consider two sample geographic habitats each containing 1000 organisms but only three distinct species, say, daisies, dandelions, and buttercups. Sample habitat 1 consists of 300 daisies, 335 dandelions, and 365 buttercups. Sample habitat 2 has 20 daisies, 49 dandelions, and 931 buttercups. Both habitat samples have the same richness (3 species) and the same total number of organisms (1000). Sample habitat 1 has more evenness than sample habitat 2, because the total number of organisms in sample habitat 1 is quite evenly distributed between the three species, i.e., relative shares are roughly equal. In habitat 2, most of the organisms are buttercups, with only a few daisies and dandelions present. Sample habitat 2 is therefore considered to be less diverse than sample habitat 1. More generally, in ecology a habitat dominated by one or two species is considered to be less diverse (more concentrated) than one in which several different species have a similar abundance.

---

[27] The numerical example that follows is taken from "Simpson's Diversity Index", available online at http://www.countrysideinfo.co.uk/simpsons.htm, last accessed May 4, 2018.

In 1949, Edward H. Simpson, a Christ College, Cambridge-educated mathematical statistician published a brief one-page note in *Nature* entitled "The Measurement of Diversity".[28] Like the HHI, Simpson's proposed construct combined aspects of richness and evenness into a single metric. It was also appealing owing to its being elegant and straightforward in its underlying probability intuition. Today, the measure is known as Simpson's Diversity Index.

Remarkably, Simpson's Diversity Index is also analytically identical to the HHI index of economic concentration.[29] Formally, let N be the number of distinct species in the habitat, and let P^* be the total number of organisms over all the N species in the particular habitat. Here note, defining the number of species in a habitat in ecology raises issues analogous to defining product markets in economics.[30] Suppose a sample of P individuals is chosen at random from a population P^i of this kind, and we observe p_i organisms of each species i. What is the probability that two individual organisms chosen at random and independently from the population will be found to belong to the same species? From combinatorics, we know that the number of all pairs in the sample is $1/2\ P(P-1)$, and the number of pairs drawn from the same species is $\frac{1}{2}\sum_{i=1}^{N} p_i(p_i-1)$. Thus, it follows that the probability that two individual organisms chosen at random and independently from the population will be found to belong to the same species is

$$D = \sum_{i=1}^{N} \frac{p_i(p_i-1)}{P(P-1)}. \qquad\qquad \text{(Eqn. 6)}$$

---

[28] Edward H. Simpson, "The Measurement of Diversity", *Nature*, Vol. 163, April 30, 1949, p. 688. Simpson later gained fame as a crypt-analyst of Alan Turing's Banburismus code-breaking process developed during World War II.
[29] The subsequent ecology literature has expanded the decomposition of diversity into variety, balance and disparity components. See, for example, Mark O. Hill, "Diversity and Evenness: A Unifying Notation and Its Consequences", *Ecology,* March 1973, 54(2):427-432, https://esajournals.onlinelibrary.wiley.com/doi/10.23071934352; Lou Jost, "Entropy and Diversity", *Oikos*, May 2006;, 113(2):363-37j5, https://onlinelibrary..wiley.com/doi/10.1111/j.2006.0030-1299.14714.x; and Alje van Dam, "Diversity and its decomposition into variety, balance and disparity", *Royal Society Open Science,* July 2019, 6(7), https://royalsocietypublishing.org/doi/full/10.1098/rsos.190452.
[30] See, for example, Hill [1973], *op. cit.*, Jost [2006], *op. cit.*, and van Dam [2009], *op. cit*.

When p and P are large numbers,

$$D \approx \sum_{i=1}^{N} \left(\frac{p_i}{P}\right)^2 = \sum_{i}^{N} S_i^2 \ . \qquad \text{(Eqn. 7)}$$

The probability foundations of Simpson's Diversity Index provide attractive intuition for the HHI. Paradoxically, however, while Simpson's Diversity Index increases within its range from 0 to 1, over the same range the economic concentration metric increases as it approaches 1. This is because the economic interpretation of D in (Eqn. 7), analogous to its probability foundation in ecology, is what is the probability of two products randomly and independently selected in a given product market being sold/produced by the same manufacturer? Ecology texts therefore define Simpson's Diversity Index as $1 - D$, and interpret it as "the probability that two individuals randomly selected from a sample will belong to different species."[31]

Regarding the seemingly arbitrary squared or quadratic share term in the HHI, based on the probability interpretation provided by Simpson's Diversity Index, we observe that it derives from the following thought experiment explained by Simpson: What is the probability of a *pair* of randomly and independently chosen individuals/products belonging to the same species/manufacturer? Presumably, if instead of the probability of a pair the thought experiment involved the probability of a triplet of randomly and independently chosen individuals/products belonging to the same species/manufacturer, the HHI/Simpson expression would involve a third order term.[32]

IV. *GENERALIZATIONS OF AND ALTERNATIVES TO THE HERFINDAHL-HIRSCHMAN AND*

SIMPSON'S DIVERSITY INDICES

---

[31] "Simpson's Diversity Index", available online at http://www.countrysideinfo.co.uk/simpsons.htm, last accessed May 4, 2018. Others define the diversity measures as $1/D$, which has, however, a much different numerical range.
[32] More complex thought experiments for HHI/Simpson diversity expressions involving n-tuples are of course possible as well.

Numerous other measures of industry concentration have been proposed and implemented in the economics and related literatures. The Hannah-Kay index (HK) is a generalized form of the HHI discussed above, and allows the analyst to weight the share components of the HHI to be more sensitive to firms with larger or smaller shares.[33] HK is calculated in terms of a parameter α where $\alpha > 0$ but $\neq 1$. Specifically,

$$HK(\alpha) \equiv \left(\sum_{i=1}^{N} S_i^{\alpha} \right)^{\frac{1}{1-\alpha}}.$$ (Eqn. 8)

When $\alpha = 2$, the HK takes on $1/HHI$ as a special case. As $\alpha \to 0$, HK tends toward N, while as $\alpha \to \infty$, HK tends toward $1/S_1$ – the reciprocal of the largest firm's share. As α decreases, smaller firms are weighted more highly, and as α increases, larger firms are weighted more heavily. This allows the analyst to apply differential weights to firms of different sizes, which may be particularly relevant in markets where small firms are thought to discipline large firms. For HK, the numbers-equivalent is simply the index, i.e., $N_e(HK) = HK$.

The HK index has additional links in economics and ecology through Shannon's Diversity Index,[34] and specifically through the concept of *entropy*.[35] The concept of information entropy has its roots in physics. However, as a measure of concentration, entropy has its foundations in information theory, where it describes the frequency of characters in a data stream. [36] The underlying intuition is that if one observes a rarer character in a data stream, one is more surprised to observe it precisely because it is

[33] Leslie Hannah and J.A. Kay, *Concentration in Modern Industry: Theory, Measurement, and the U.K. Experienc*e, London, Macmillan, 1977.

[34] For discussion, see the review by H. G. Washington, "Diversity, biotic and similarity indexes: A review with special relevance to aquatic ecosystems", *Water Research* 1984, 18(6):653-694.

[35] Henri Theil, *Principles of Econometrics*, New York: John Wiley and Sons, 1971, pp. 639-641, 644, 662. For an early application in economics, see Michael O. Finkelstein and Richard M. Friedberg, "The Application of an Entropy Theory of Concentration to the Clayton Act", *Yale Law Journal*, 1967, 76(2):677-717. Also see P. E. Hart, "Entropy and Other Measures of Concentration", *Journal of the Royal Statistical Society, Series A (General)*, 1971, 134(1):73-85.

[36] Wikipedia, "Entropy (information theory)", last edited April 2, 2019. Available online at https://en.wikipedia.org/wiki/Entropy_(information_theory), last accessed May 4, 2019.

rare. "Information content" and "surprisal" are technical terms used throughout the information theory literature to measure rareness, which is quantified as $-\ln(S_i)$, where S_i is the relative frequency of a character. Entropy is simply calculated as the weighted average of information contents across all observations, i.e.,

$$E \equiv -\sum_{i=1}^{N} S_i \ln(S_i).$$ (Eqn. 9)

Applied to economics, entropy measures how much information content about a market is gained by observing the firm that sells a random product in the marketplace.[37] Interestingly, the entropy measure in Eqn. 9 can be derived by taking the limit of the HK index as α approaches 1, since $\lim_{\alpha \to 1} HK(\alpha) = E$.[38] Therefore, entropy as a measure of concentration need not rely on foundations from information theory. Specifically, by setting $\alpha = 1$, entropy weights all species (or, in economics, firms within a product market) by their frequency, without favoring either common or rare species.[39] Given this equivalence, entropy ranges from 0 to $\ln N$, and the numbers-equivalent of entropy is $N_e(E) = e^E$.

V.   AN EXTENSION: USING THE INDEXES TO DISCRIMINATE AMONG CONCENTRATION, DIVERSITY AND INEQUALITY

An implication of the above discussion is that the HHI concentration measure can describe a very large number of possibly different situations involving varying combinations of inequality and population size. This insight raises the issue of whether concentration, diversity, variability and inequality are fundamentally distinct concepts, or could they be related to each other within an integrated conceptual framework. We begin to address this issue by focusing on inequality and the Gini

---

[37] Stefan Baumgartner, "Measuring the Diversity of What? And For What Purpose? A Conceptual Comparison of Ecological and Economic Biodiversity Indexes", 2006. Available at SSRN: https://ssrn.com/abstract=894782 or http://dx.doi.org/10.2139/ssrn.894782.
[38] Mark O. Hill [1973], *op. cit*.
[39] Lou Jost [2006], *op. cit*.

coefficient[40], which in turn derives from the Lorenz curve[41] that graphically portrays the distribution of income, output or wealth. As we shall see, this metric focuses only on relative shares, not absolute levels, nor on fewness or population size.

A. *Inequality Measurement Using Lorenz Curves and Gini Coefficients*

Consider a two-dimensional graph known as the Lorenz curve, displayed in Figure 1 below.[42] On the horizontal axis is the cumulative portion of the total population (of firms, or households), from zero percent at the origin to 1 (100 percent) at the furthest point to the right. For each firm or household, there is some non-negative measure of output, income or wealth, and with firms ordered from left to right in order of increasing wealth. The vertical axis represents the corresponding cumulative portion of total output, income or wealth. Thus, the Lorenz curve plots cumulative portion of total output, income or wealth against the percentage of total output, income or wealth achieved by the bottom x percent of the population; neither axis reflects absolute amounts.

If the distribution of output, income or wealth were perfectly equal, everyone would have the same amount of output, income or wealth. In such a case the bottom b percent of the population would always have b percent of the cumulative output, income or wealth, and the Lorenz curve would be a 45-degree diagonal straight line on which $y = x$, called a "line of perfect equality".

Now consider a perfectly unequal distribution in which one person or firm had all the output, income or wealth, and everyone else had none. In that case, the Lorenz curve would consist of a straight

---

[40] Corrado Gini, "Variability and Mutability" (Italian: "Variabilita e mutabilita"), Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche. C. Cuppini, Bologna, 1912.

[41] Max O. Lorenz, "Method of measuring the concentration of wealth", *Journal of the American Statistical Association,* June1905, 9(70):209-219.

[42] The left panel in Figure 1 is similar to that in Wikipedia, "Gini Coefficient", available online at https://en.wikipedia.org/wiki/Gini_coefficient, page 2 of 11, last accessed March 1, 2019. Also see Wikipedia, "Lorenz curve", available online at https://en.wikipedia.org/wiki/Lorenz_curve#Explanation, last accessed May 3, 2019.

line at $y = 0\%$ for all $x < 100\%$, and when $x = 100\%$ it would suddenly exhibit a right angle perfectly

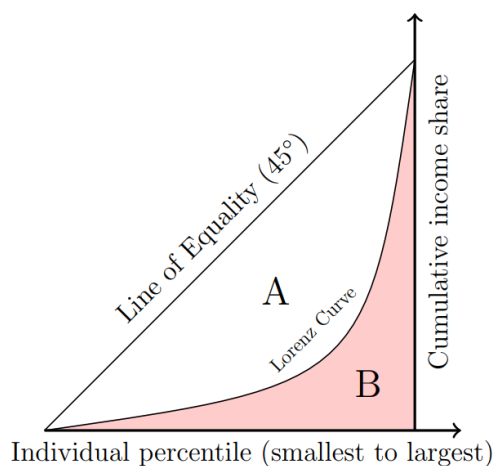vertical straight line to $y = 100\%$. Call this the "line of perfect inequality".

Figure 1:

Gini Coefficient = Area between Line of Perfect Equality and Lorenz Curve divided by area of the entire triangle, which equals the area between Line of Equality and the Line of Perfect Inequality. The area of the entire triangle is $A + B = 0.5$. Thus the Gini Coefficient = $A/(A + B) = 2A = (1 - 2B)$.

As seen in Figure 1, the Gini coefficient is the ratio of the area between the line of perfect

equality and the observed Lorenz curve (area A) to the area between the line of perfect equality and the

line of perfect inequality (areas $A + B$), i.e. Gini = $A/(A + B)$.[43] The greater the Gini coefficient, the

more unequal is the distribution of output, income, or wealth. Since both the horizontal and vertical

axes scale from zero to one, the area of the triangle $(A + B) = 0.5$ (base times height) = 0.5 (1 times

1) = 0.5, implying that the Gini coefficient is also equal to $2A$ and to $(1 - 2B)$.

There are many studies estimating income inequality across the globe. For example, the Gini

coefficient income distribution averages for the total population in OECD countries in 2008-2009 have

been estimated to range between 0.34 and 0.53 on a pre-tax basis, and between 0.25 and 0.48 on a

disposable income basis (post-taxes and transfers), implying that taxes and transfers can mitigate

---

[43] The left panel in Figure 1 is similar to that in Wikipedia, "Gini Coefficient", available online at https://en.wikipedia.org/wiki/Gini_coefficient, page 2 of 11, last accessed March 1, 2019. Also see Wikipedia, "Lorenz curve", available online at https://en.wikipedia.org/wiki/Lorenz_curve#Explanation, last accessed May 3, 2019.

inequality. For the U.S., the pre-tax and disposable income distribution Gini coefficients in 2008-2009 were estimated at 0.49 and 0.38.[44]

Although economists have long used Lorenz curves and Gini coefficients to measure inequality of income, wealth or consumption, and in consumer finance to measure, say, the actual y percent of delinquencies that are attributable to the x percent of people with the worst risk scores, the Lorenz-Gini concepts are also useful in very different contexts whenever one seeks to study inequality in a distribution. Examples include inequality in health related quality of life, measuring the inequality of universities, the selectivity of protein kinase inhibitors against a panel of kinases, evaluating the fairness achieved by Internet routers in scheduling packet transmissions from different flows of traffic, describing inequality among the size of individuals in ecology, and in studies of biodiversity where the cumulative proportion of species is plotted against the cumulative share of organisms.[45] One could even imagine using Lorenz curve and Gini coefficient analogs to describe the inequality of firms in a market, plotting firms' cumulative market share of outputs or revenues on a vertical axis and the cumulative percentile of firms, ordered from largest to smallest, on the horizontal axis.

However, it is worth noting that the Lorenz curve and the Gini coefficient as measures of concentration suffer from the defect they do not depend on the number of firms. The Gini coefficient of the Lorenz curve would not distinguish between a single firm industry and one with a thousand equal sized firms.[46] If one believes that the absolute numbers of firms in an industry affect concentration – since that number reveals in part the nature of entry and exit conditions, with many firms suggesting easy entry, and few firms that entry is blockaded – then the Gini coefficient's neglect of fewness or

---

[44] Wikipedia, "Gini Coefficient", available online at https://en.wikipedia.org/wiki/Gini_coefficient, page 4 of 11, last accessed March 1, 2019.
[45] For discussion and references, see Wikipedia, "Gini Coefficient", available online at https://en.wikipedia.org/wiki/Gini_coefficient, page 8 of 11, last accessed March 1, 2019.
[46] Hall and Tideman [1967], p. 164.

population size may make it less attractive as a measure of economic concentration.[47] The next index we

consider does not have this drawback.

B. *LINKING CONCENTRATION MEASURES TO GINI COEFFICIENTS: ROSENBLUTH AND HALL-*

TIDEMAN

An alternative concentration measure that incorporates both number of firms and inequality

was developed initially by Rosenbluth [1960] as a graphical extension to concentration curves, but was

subsequently developed axiomatically by Hall and Tideman [1967].[48] The Rosenbluth/Hall-Tideman

(RHT) index depends on the ordinal rank of a firm, and therefore on the absolute number of firms in a

market. It is computed as

$$RHT \equiv \frac{1}{2\sum_{i=1}^{N} iS_i - 1}$$ (Eqn. 10)

where i is the ordinal ranking of the ith firm, and where the ith largest firm must be the ith term in the

summation. Note that the RHT index weights each firm by its rank, thus taking explicit account of the

absolute number of firms. The RHT index ranges between 0 and 1, being close to 0 for an industry with

an infinite number of equally sized firms, and reaching 1 in the case of a monopoly. Like the HHI, the

RHT measure equals $1/N$ for an industry with N equally sized firms, and the numbers equivalent of the

index equals $1/RHT$.[49]

Although not immediately obvious, the RHT metric is closely related to the Gini coefficient.

Rosenbluth [1960] motivated the RHT index using a figure analogous to the Lorenz curve, but where the

horizontal axis is the absolute number of firms, rather than firm percentile, ordered from largest to

---

[47] *Ibid.*, p. 165.

[48] Gideon Rosenbluth, "Round Table Gesprach uber Messung der industriellen Konzentration". *Die Konzentration in der Wirtschaft; Verhandlungen auf der Tagung des Vereins fur Socialpolitik, Gesellschaft fur Wirtschafts- und Sozialwissenschaften, in Bad Kissingen.* Berlin: Duncher & Humblot, 1960, 391-395; Marshall Hall and Nicolaus Tideman, "Measures of Concentration", *Journal of the American Statistical Association* 1967, 62(317):162-168.

[49] Jacob A. Bikker and Katharina Haaf, "Measures of Competition and Concentration in the Banking Industry: a Review of the Literature", *Economics & Financial Modelling*, Summer 2002, 1-45. See especially p. 10.

smallest, rather than largest to smallest (see Figure 2 below). Following Rosenbluth [1960], define the

area above the empirical concentration curve in Figure 2 as area K, and the area above the 45-degree

line but underneath the empirical concentration curve as area J. Note that since the horizontal axis in

Figure 2 is the number of firms N (rather than the percentile) and firms are ordered from largest to

smallest, Figure 2 is a horizontal scaling of Figure 1 after vertical and horizontal reflections. Thus, if we

charted a Lorenz curve using data for the same market, we would have that $B = K/N$. Rosenbluth

[1960] then defined the RHT index of Eqn. 9 as the reciprocal of twice the area above the cumulative

concentration curve, so that in Figure 2, $RHT = 1/2K$, a geometric interpretation that is equivalent to

the computation defined in Eqn. 10. If we computed the Gini coefficient for this market, we would find
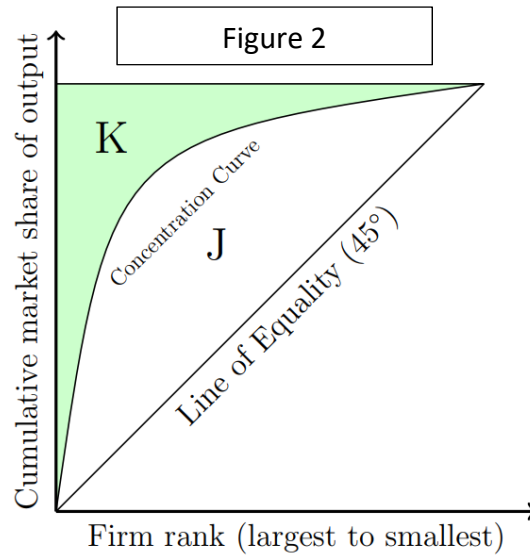
Figure 2

that $G = 2A = (1 - 2B)$ as in Figure 1. Then, if we substitute in $B = K/N$ into $G = 1 - 2B$ and

rearrange, we obtain $N(1 - G) = 2K$, so that

$$RHT = \frac{1}{2K} = \frac{1}{N(1 - G)} \text{, where } G = \text{Gini Index.} \qquad \text{(Eqn. 11)}$$

Eqn. 11 makes clear the dependence of the RHT index on the number of firms N and the

inequality in firm size G. Note that for a given value of relative inequality between firms G, RHT

decreases with increases in N, i.e., $\partial RHT/\partial N = -1/[N^2(1-G)] < 0$. Alternatively, for a given number of firms N, the RHT concentration index increases as the Gini index of inequality increases, i.e., $\partial RHT/\partial G = 1/[N(1-G)^2] > 0$.

Although the intuition underlying Eqns. 10 and 11 may be less than obvious, if for example sales revenues are an appropriate measure of firm size in a given market, then S_i could be interpreted as a sample estimate of the probability that a firm of rank i will obtain a random dollar of sales, and in that case $\sum_{i=1}^n iS_i$ is the expected value of that rank.

Hall and Tideman [1967] demonstrate that both the HHI and the RHT indexes satisfy six axioms they argue should be satisfied by measures of economic concentration, and while there may be many other measures that also satisfy their six axioms, the Gini coefficient does not. They report that if one rank orders the various CR$_k$ (for $k = N$ for each industry), HHI and RHT concentration measures using the 1958 U.S. data for 446 Standard Industrial Classification four-digit industries, the rank correlations between any pair are above 0.9, suggesting "that with respect to ordinal qualities there is scarcely any difference between the three measures", although in some cases the absolute ordinal ranking differed substantially: in 23 of the 446 industries the difference in rankings between RHT and HHI or RHT and CR was greater than 90.[50]

Regarding choice among the CR, HHI and RHT indexes, Hall and Tideman note first that with all three indexes, there is more than one distribution of firms that can generate a given value of the index, i.e., a particular value of the economic concentration index is not associated with a unique distribution of firms. Moreover, since their notion of concentration includes reference both to inequality and the number of firms in the industry, "under some situations the trade-off between numbers and inequality would be such as to allow industries with different size distributions to be viewed as equally

---

[50] Hall and Tideman 1967, *op. cit.*, p. 166.

concentrated."[51] Although they prefer the HHI and RHT measures to CR and Gini coefficients on *a priori*

grounds since the former satisfy their six axioms while the latter do not, they do not recommend a

choice between HHI and RHT. Rather, they view this as similar to the choice between various price index

formulae, since

> "…the root of the problem…is the selection of an appropriate set of weights with which
> to combine the [shares]. The weighting problem here is similar to the weighting
> problems with any index number and as with index numbers there are no a priori
> reasons free of crucial assumptions which will allow us to say that a particular weighting
> scheme is best. The properties set forth above [i.e., the six axioms] cannot uniquely
> determine the best measure of concentration but they do serve to delineate measures
> that are undesirable for theoretical reasons."[52, 53]

Hart [1971] compares the use of information-theory based entropy (and redundancy – another

metric based on observed entropy) as a measure of industry concentration to classical measures of

statistical dispersion (e.g., arithmetic means, variances, variances of logarithms, geometric means) and

what he calls *ad hoc* measures of concentration such as the concentration ratios or cumulative

concentration curve.[54] He concludes:

> "…when the number of firms is large enough to use statistical distribution theory, the
> classical statistical measures are superior to the entropy or the redundancy. When the
> number of firms is small, the entropy is superior to the redundancy, but both are
> inferior to the traditional measures of concentration derived from the cumulative
> concentration curve. Consequently, there is little point in using the information theory
> measures to measure business concentration."[55]

---

[51] *Ibid.*
[52] Hall and Tideman [1967], *op. cit.,* p. 164. Words in square brackets added.
[53] We note in passing that since the Hall-Tideman [1967] article was published, developments in index number
theory involving the interpretation of index numbers as exact Taylor series' approximations to underlying distinct
utility, indirect utility, cost and production functional forms imply that while the families of index number formulae
satisfying certain approximation properties may be unique to specific functional forms, choice among alternative
functional forms is typically not feasible based only on economic theory. On this, see, for example, W. Erwin
Diewert, "Exact and Superlative Index Numbers", *Journal of Econometrics* 1976, 4:115-145.
[54] P. E. Hart, "Entropy and Other Measures of Concentration", *Journal of the Royal Statistical Society, Series A
(General)*, 1971, 134(1):73-85.
[55] *Ibid.,* p. 73.

In summary, the existing economic literature favors use of the HHI and RHT metrics as measures of economic concentration. Hence, in the empirical analysis that follows, we limit our attention to these two concentration measures.

VI.   AN EMPIRICAL ILLUSTRATION INVOLVING THE PRESCRIPTION DRUG LIPITOR™

What happens to concentration measures as monopoly power gives way to competition and unfettered entry in a product market with high fixed/sunk costs but low marginal costs? How similar empirically are the various measures of concentration? Do they exhibit expected step patterns as sequential extensive entry occurs? Are there examples of product markets where concentration exceeds DOJ/FTC announced thresholds inviting scrutiny? We now illustrate these selected empirical relationships among various measures of industry concentration.

We choose to focus on a product market in which an event or series of events occurs that we expect materially affects market structure and measures of concentration. Prescription pharmaceutical markets are good candidates, since loss of exclusivity (LOE) through patent expiration or other market exclusivity is known to result in the brand's rapid loss of market share to off-patent generic entrants, although the speed and extent of generic entry may be affected by statutory or regulatory restrictions.[56] Since research has shown that the extent of generic entry depends in large part on the pre-LOE market size of the branded drug losing exclusivity,[57] it is desirable that the product we choose be a "blockbuster" drug having large sales revenues while patent-protected.

---

[56] Murray L. Aitken, Ernst R. Berndt, Barry Bosworth, Iain M. Cockburn, Richard G. Frank, Michael Kleinrock and Bradley T. Shapiro, "The Regulation of Prescription Drug Competition and Market Responses: Patterns in Prices and Sales Following Loss of Exclusivity," ch. 8 in Ana Aizcorbe, Colin Baker, Ernst R. Berndt and David M. Cutler, editors, *Measuring and Modeling Health Care Costs*, University of Chicago Press for the National Bureau of Economic Research, 2018, pp. 243-271.

[57] For a survey, see Fiona Scott Morton and Margaret Kyle, "Markets for Pharmaceutical Products", ch. 12 in Mark V. Pauly, Thomas G. McGuire and Pedro P. Barros, editors, *Handbook of Health Economics*, Vol. 2, Amsterdam: Elsevier, 2012, pp. 763-823, but especially Section 5.1, "Generic Entry", pp. 792-795.

The product market on which we focus is that for brand name Lipitor™ (chemical name

atorvastatin calcium), a 3-hydroxy-3methyl-glutaril-coenzyme-A (HMG-CoA) reductase inhibitor

antihyperlipidemic agent commonly known as a statin, approved by the U.S. regulator (the U.S. Food

and Drug Administration or FDA) as an adjunct to diet to reduce elevated total cholesterol, low density

lipoprotein cholesterol, triglycerides and to increase high density lipoprotein cholesterol. Four strengths

(10, 20, 40 and 80mg) of the once-daily dosed Lipitor™ tablet were approved for marketing by the FDA

on December 17, 1996. Lipitor™ rapidly became the U.S.'s and the world's largest selling drug. Since

being launched in 1997, Lipitor™ earned more than $100 billion for its manufacturer, Pfizer, and by

2011 nine million people in the U.S. were taking the Lipitor™ medicine.[58] Lipitor competed in a crowded

field with a number of other statins, such as lovastatin (original brand name Mevacor™, approved

3/28/91, but discontinued, replaced by Altoprev™, approved 6/26/02), pravastatin (Pravachol™,

approved 10/31/91), simvastatin (Zocor™, 12/23/91), fluvastatin (original Lescol™ brand discontinued,

replaced by Lescol XL™, 10/6/00), rosuvastatin (Crestor™, 8/12/03), and pitavastatin (Livalo™, 8/3/09).

Lipitor™ was also a component in a combination lipid-lowering and hypertension product branded as

Caduet™. Lipitor™ and its competitors all marketed once-a-day dosage tablets, although with different

mixes of dosage strengths.

In the analysis of atorvastatin calcium below, we ignore combination products in which

atorvastatin calcium is a component ingredient. Although we recognize that in antitrust matters market

definition can be a very contentious issue,[59] for our present purpose of merely illustrating empirically

---

[58] Veronique Dupont, "Pfizer's blockbuster drug Lipitor goes generic", *MedicalXpress*, November 30, 2011. Available online at https://medicalxpress.com/news/2011-11-pfizer-blockbuster-drug-lipitor.html, last accessed 5 August 2019.
[59] There is a large literature addressing issues of market definition issues in horizontal mergers and acquisitions. See, for example, F. Michael Scherer and David Ross [1990], *op. cit.*, pp. 174-190, and Dennis W. Carlton and Jeffrey M. Perloff [2005], *op. cit.,* pp. 642-648. As noted above, issues of market definition in economics are somewhat analogous to defining distinct species in ecology; see the Hill [1973], Jost [2006] and van Dam [2019] references cited earlier.

alternative concentration measures and their time trend breaks, here we simply assume that the various strengths of oral branded and generic atorvastatin calcium tablets constitute the same product market in the U.S.

The market dynamics involving Lipitor[TM] and its generic atorvastatin calcium competitors suggest that measures of concentration are likely to vary considerably in a step function pattern after Lipitor[TM] lost patent protection on November 30, 2011. With much generic entry anticipated following LOE, we expect that concentration measures will reflect considerable competition and greatly reduced concentration. Patents on all Lipitor[TM] formulations and strengths were successfully challenged in U.S. courts by Ranbaxy, who consequently was awarded 180-day exclusivity on its Abbreviated New Drug Application (ANDA) effective November 30, 2011. However, serious data integrity issues emerged with Ranbaxy's ANDA, creating uncertainty whether the FDA would permit Ranbaxy to launch on November 30, 2011.[60] Although Ranbaxy's 180-day exclusivity was scheduled to expire on May 28, 2012, any delay in the launch of Ranbaxy due to its data integrity issues with the FDA would likely also delay the timing of the 180-day exclusivity and the unfettered post-180-day exclusivity entry by numerous other ANDA applicants. These uncertainties may have affected the extent and timing of subsequent generic entrants.

However, in an effort to mitigate market share losses to Ranbaxy during the 180-day exclusivity period, Pfizer contracted with Watson (later acquired by Actavis) to market an authorized generic version of all dosage strengths of atorvastatin calcium for five years beginning November 30, 2011.[61] During the 180-day exclusivity period, therefore, there were three competitors on the U.S. market (a triopoly -- Pfizer's Lipitor[TM] brand, its authorized generic through Watson-Actavis, and Ranbaxy's

---

[60] For one account, see Katherine Eban, *Bottle of Lies: The Inside Story of the Generic Drug Boom*, HarperCollins Publishers, New York, 2019, especially chapters 6, 22 and 23.
[61] "Watson Launches Generic LIPITOR®: Largest generic product launch in US history", *Allergan Investor News Search*, 11.30.2011. Available online at https://www.allergan.com/news/thomson-reuters/watson-launches-generic-lipitor-r-aspx, last accessed 10 August 2019.

generic). On May 28, 2012 when 180-day exclusivity expired, three additional ANDA holders entered at all four dosage strengths (Apotex, Mylan and Sandoz).[62]

Given Lipitor's very large pre-LOE market size, industry observers noted it was somewhat surprising that the number of generic entrants at the time of unfettered generic entry following expiration of 180-day exclusivity was relatively small – only five as of late 2012. According to a Federal Trade Commission document, each of the known generic entrants as of May 2012, as well as branded Pfizer's Lipitor™, was known to have substantial manufacturing capacity, so conceivably Pfizer and the generics could flood the atorvastatin calcium market and greatly reduce generic prices, making potential entry by other generics less attractive. Moreover, Pfizer's highly publicized aggressive protection of its brand franchise (through rebating and widespread couponing that reduced considerably the customer copayment typically required for branded drugs, as well as its continued extensive direct-to-consumer marketing campaign) may have reduced the perceived payoff to entry by generic manufacturers, thereby constraining entry.[63] Recall also that these considerations involving Pfizer were in addition to the uncertainties created by Ranbaxy's data integrity dispute with the FDA, possibly jeopardizing Ranbaxy's 180-day exclusivity launch.

Pfizer also benefited from certain fortuitous events affecting its competitors. On November 9, 2012, Ranbaxy informed its customers of a voluntary recall of certain lots of the company's 10, 20 and 40 mg dosage strengths of atorvastatin calcium tablets, due to possible contamination with very small glass particles. On November 29, 2012, the FDA notified the public of Ranbaxy's voluntary recall.[64]

---

[62] For further discussion, see Murray L. Aitken, Ernst R. Berndt, Barry Bosworth, Iain M. Cockburn, Richard G. Frank, Michael Kleinrock and Bradley T. Shapiro [2018], *op. cit.*

[63] On the market's perception of the reputation of Pfizer and its generic subsidiary Greenstone, see Federal Trade Commission [2011], *Authorized Generic Drugs: Short-Term Effects and Long-Term Impact*, August, especially p. 82.. Available online at http://www.ftc.gov/opa/2011/08/genericdrugs.shtm.

[64] U.S. Food and Drug Administration, "Questions and Answers on the Ranbaxy Atorvastatin Recall", November 29, 2012. Available online at https://www.fda.gov/drugs/drug-safety-and-availability/questions-and-answers-ranbaxy-atorvastatin-recall, last accessed 6 August 2019. U.S. Food and Drug Administration, "FDA Statement on the Ranbaxy Atorvastatin Recall", November 30, 2012. Available online at https://www.fda.gov/drugs/drug-safety-and-availability/fda-statement-ranbaxy-atorvastatin-recall, last accessed 6 August 2019.

Ranbaxy resumed production of atorvastatin calcium intended for sale in the U.S. market in February 2013, but in March 2014 it issued another recall, this time for a dosage mix-up in which 20mg tablets were found in sealed bottles marked for 10 mg tablets. This recall followed by one month the FDA prohibiting Ranbaxy from shipping to the U.S. any pharmaceutical ingredient made at its Toansa plan in northern India after an inspection there revealed poor manufacturing practices.[65] As of 2014, the Ranbaxy-owned Ohm Laboratories Inc. in New Jersey was the company's only permitted manufacturer of drugs for U.S. sale.[66] Because of these various market developments we expect that, following Pfizer's LOE, realized market shares were affected, as were the time patterns of concentration measures. In particular, while we expect HHI/RHT to fall sharply after LOE, the speed with which they fall may be mitigated by the various uncertainties during the 180-day exclusivity period, and that in turn may have delayed subsequent entry.

Here we begin by describing time trends in market revenue shares, move on to various concentration metrics based on revenue shares, and then compare concentration metric results using alternatively revenue and quantity shares. Given our earlier discussion that documented a preference in the economic literature for the HHI and RHT metrics, we limit our empirical analysis to those two measures, although for each of them we define market shares in both volumes and in revenues (volumes * price).

A. DATA SOURCES

We analyze U.S. unit volume and revenue quarterly data for the various atorvastatin brand and generic oral tablet formulations (10, 20, 40 and 80 mg strengths) and construct volume and revenue-based shares utilizing data from the IQVIA (formerly IMS Health and QuintilesIMS) National Sales

---

[65] Zeba Siddiqui and Kanika Sikka, "Ranbaxy recalls some batches of generic Lipitor in latest quality blow", *Reuters Health News*, March 7, 2014. Available online at https://www.reuters.com/article/us-ranbaxylab-lipitor-usa-recall/ranbaxy-recalls-some-batches-of-generic-lipitor-in-latest-quality-blow-idUSBREA2703..., last accessed 6 August 2019.
[66] *Ibid.*

Perspective™ (NSP) database, 2004Q4-2016Q3. NSP data derive from a projected audit of purchases

from manufacturers or wholesalers to pharmacies, clinics, hospitals and other distribution outlets,

covering 100 percent of the national unit volume in all major classes of trade and distribution channels.

As a measure of sales volume, NSP reports "standard units" measuring the number of smallest

dosage form single items (such as tablets in bottles) contained in a unit or shipping package purchased

by pharmacies or other distribution outlets. Because the recommended dosage for each of the 10, 20,

40 and 80 mg strengths is once daily, the standard unit volumes can be considered also to measure the

number of daily doses.

The "dollar sales revenues" reported by NSP are the total amounts paid for purchases of

atorvastatin calcium standard units from a specific manufacturer or wholesaler by quarter. The invoice-

based dollar metrics reflect the amounts paid by retail pharmacies, hospitals and clinics, whether

purchased directly from a manufacturer or indirectly via a wholesaler or chain warehouse. Invoice line

item discounts are included, but prompt payment bottom-line invoice and 340B discounts are excluded.

Drug rebates paid by the manufacturer to an insurer or an intermediary are not reflected in these

invoiced revenues; such data are generally not publicly available. To the extent sales from wholesalers

include wholesaler margins and exclude off-invoice rebates paid by manufacturers to pharmacy benefit

managers, third party payers, pharmacies and insurers, the NSP data overstate net revenues received by
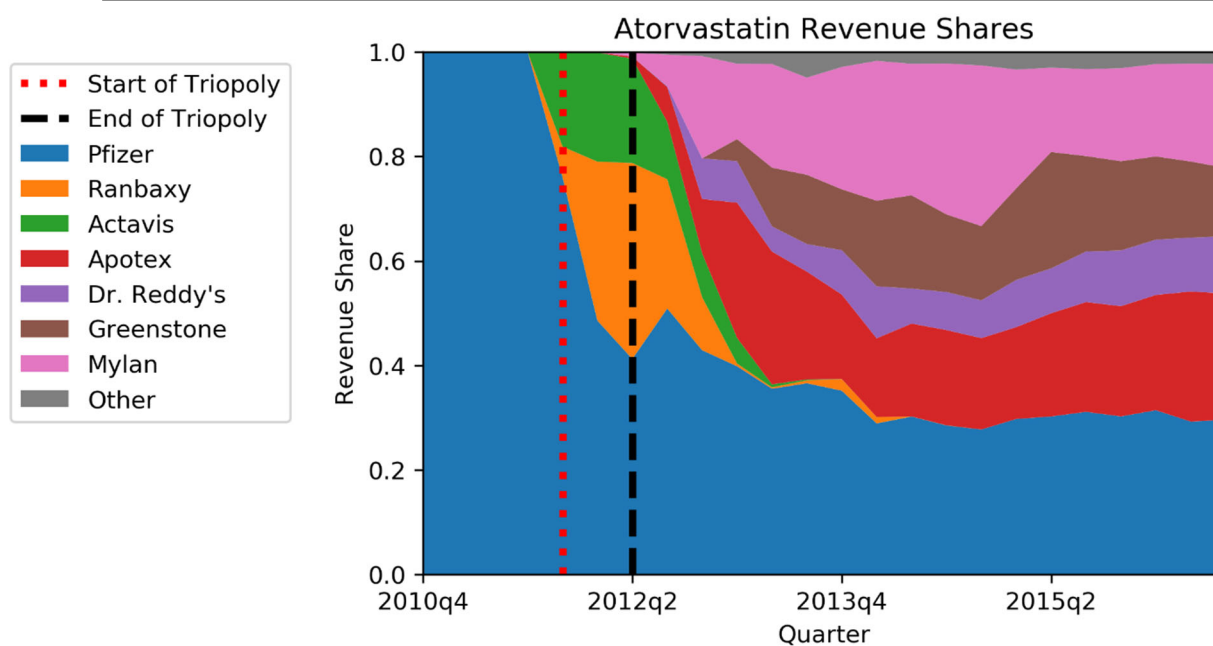
manufacturers.[67]

B.   *REVENUE AND QUANTITY SHARES AND HHI AND RHT CONCENTRATIONS*

Quarterly revenue shares for the various atorvastatin calcium manufacturers over the 2004Q4 –

2016Q3 time period are shown in Figure 3, portraying market shares during four eras: (i) the period of

Pfizer Lipitor™ exclusivity from 2004Q4 until its LOE on November 30, 2011 (2004Q4 – 2011Q4); (ii) the

---

180-day exclusivity period between November 30, 2011 and May 28, 2012 (2011Q4-201Q2), during

which time brand Lipitor™, generic Ranbaxy and authorized generic Watson/Actavis competed; (iii) a

transition period from the end of 180-day exclusivity on May 28, 2012 through the end of 2012 during

which unfettered generic entry and some exit occurred (2012Q2 – 2012Q4); and (iv) the post-2012 era

in which market shares among the atorvastatin calcium marketers stabilized, as did the identity of the

market leaders (after 2012Q4).

Figure 3: Atorvastatin Calcium Revenue Shares by Manufacturer

Please note: Pfizer's atorvastatin calcium product was marketed under the brand name Lipitor™.
The remaining manufacturers listed are selling generic atorvastatin calcium products.

Watson/Actavis' authorized generic each gained market share rapidly during their triopoly 180-

day exclusivity ending in 2012Q2.[68] However, they then each quickly saw their market shares erode as

Ranbaxy experienced manufacturing problems in 2012Q4 and encountered post-180-day exclusivity

entry by Apotex, Mylan and Dr Reddy's Laboratories that took market share from Pfizer. Pfizer's

---

[68] Recall that during the 180-day exclusivity period between November 2011 and May 2012, the three atorvastatin competitors on the U.S. comprising the triopoly were Pfizer's branded Lipitor, Ranbaxy – the successful patent challenger, and Watson – the authorized generic.

Lipitor[TM] brand share loss was largely offset, however, by the 2013Q1 entry of Greenstone, Pfizer's wholly owned generic subsidiary.[69] At the end of 2013 – eight quarters after initial generic entry – Pfizer's Lipitor[TM] revenue market share was 35% (much greater than its 2% volume share, reflecting Lipitor's relatively high price), Mylan captured 23% of market revenues (20% of volume), Apotex 16% of revenues and 27% of volume, Pfizer's Greenstone (12% of revenues and 27% of volume), and Dr. Reddy's Laboratories (8% of revenues and 14% of volume), so together the largest five marketers garnered 94% of atorvastatin US market revenues (90% of volume), up from 81% of revenues in 2012Q4, but down from 92% of 2012Q4 volumes. Watson/Actavis, the authorized generic, had only a 0.02% market share in 2013Q4, falling sharply from a high share of 21% in 2012Q1. Notably, with its Lipitor[TM] brand and Greenstone generic subsidiary, Pfizer's combined brand/generic franchise maintained a 47% market revenue (29% volume) share after eight quarters following LOE. Meanwhile, Ranbaxy's initial successful generic entry was also very short-lived; by 2013Q4 Ranbaxy's revenue share was but 2.2% (volume share 6.1%), dramatically less than its peak revenue share of 38% (41% volume share) in 2012Q2 when it enjoyed 180-day exclusivity. Although Sandoz was one of the first three generic entrants after 180-day exclusivity expired, Sandoz was not ever able to penetrate the atorvastatin calcium U.S. market – its peak revenue share never exceeded 1.6% and at the end of 2013 it had only a 0.5% market revenue (1.8% volume) share.

From the end of 2013 through the end of our study sample – 2016Q3 – additional generic entry occurred as did some exit; altogether, 17 distinct entities marketed atorvastatin calcium between 2011Q4 and 2016Q3 (including a repackager[70]), although by 2016Q3 only 12 had positive sales, with five

---

[69] "Greenstone LLC—your partner for generic medications", undated. Available online at https://www.greenstonellc.com/about-us, last accessed August 10, 2019.
[70] In the U.S. prescription pharmaceutical market, a repackager is not a manufacturer but instead simply repackages and markets a product (perhaps into bottles having different numbers of tablets) made by an FDA-approved manufacturer, often under contract.

others having exited, either temporarily or permanently.[71] The five largest manufacturers at the end of 2013 were still the five largest in 2016Q3, able to keep most of their market share, together garnering 98% of atorvastatin calcium U.S. market revenues and volume: Pfizer Lipitor™ – 30% revenues, 0.6% volume, Apotex 24% revenues, 39% volume, Mylan 20% revenues, 30% volume, Greenstone 13% revenues, 16% volme, and Dr. Reddy's Laboratories 11% revenues, 14% volume. The only change in the ranking identity of the five largest revenue suppliers between end 2013 and 2016Q3 involved switching between second and third place by Mylan and Apotex. Notably, the extent of inequality/variability in revenue market shares among the five largest marketers at the end of 2013 (ranging from 8% to 35%) gradually fell over time, so that by the end of 2016 the inequality/variability in market revenue shares among the five largest marketers declined, ranging from 11% to 30%. These post-LOE revenue share movements among the atorvastatin calcium manufacturers resulted in step function declines in market concentration, as measured by the HHI and RHT revenue metrics, displayed in Figure 4. Although the Pfizer atorvastatin calcium franchise included both the brand and its wholly owned subsidiary, Greenstone, here we initially treat Pfizer's brand Lipitor™ and its Greenstone generic subsidiary as distinct manufacturers (but we modify this later). Prior to November 30, 2011, Lipitor™ market share was 100%, so the uniform 1.000 HHI and RHT concentration values before 2011Q4 are not displayed in Figure 4.

There are two important takeaways from Figure 4. First, as expected, a rapid decline in the HHI and RHT concentration measures occurs in the four quarters between LOE in 2011Q4 and the end of 2012 during which time the triopoly's 180-day exclusivity expired (from 0.614 to 0.256 for HHI, and 0.626 to 0.252 for RHT), but after that both concentration metrics exhibit relative stability. Second, the HHI-Revenue and RHT-Revenue concentration measures are remarkably similar, with the RHT value

---

[71] In addition, there may have been potential entrants with approved Abbreviated New Drug Applications who chose not to enter for commercial or other reasons.

Figure 4

Atorvastatin, Revenue, After Entry
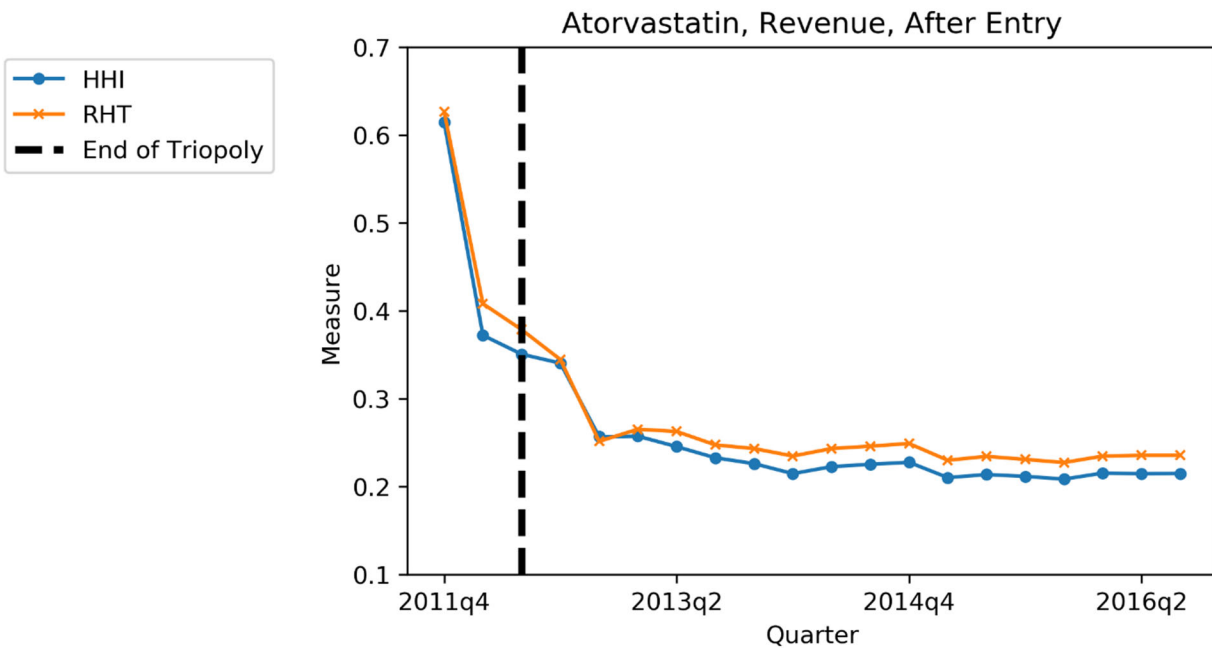
typically being very slightly greater or almost equal to the HHI value. In 2016, for example, the mean

quarterly RHT revenue metric was 0.235, while that for HHI was 0.215. The simple Pearson correlation

between HHI-Revenues and RHT-Revenues over the 2011Q4 - 2016Q3 time period is 0.996.

While Figure 4 displays RHT and HHI metrics based on revenue shares, in Figure 5 we report RHT

and HHI metrics based on quantity shares. Like HHI-Revenues and RHT-Revenues in Figure 4, in Figure 5

we observe that HHI-Quantity and RHT-Quantity concentration measures fall sharply as expected in the

four quarters after LOE in 2011Q4 to the end of 2012 (from 0.555 to 0.192 for HHI-Quantity, and from

0.575 to 0.211 for RHT-Quantity). Levels and trends in the RHT and HHI quantity metrics are very similar,

with the simple Pearson correlation between them over the 2011Q4-2016Q3 time period being 0.997.

Moreover, when comparing Figures 4 and 5, we observe that levels and trends in RHT measures based

on revenue shares are very similar to those based on quantity shares up through the end of 2013.

However, after that the quantity-based HHI and RHT metrics increase and remain above their respective

revenue-based HHI and RHTs. The intuition underlying this difference in measured concentration

depending on whether it is based on revenues or on quantities is as follows. As Pfizer's Lipitor™ brand

price relative to that of generics increases and as generics' volume market shares increase, the gap

between Pfizer's and the generics' revenue and volume shares increases, resulting in the numerous

generics' volume shares being weighted more heavily in the HHI-quantity metric, increasing it overall. [72]

At the end of the sample period in 2016Q3, the quantity-based HHI and RHT values are 0.280 and 0.301,

while the revenue-based numbers are considerably smaller at 0.215 and 0.237. The simple Pearson

correlation over the entire 2011Q4-2016Q3 time period between RHT-Revenues and RHT-Quantities is

Figure 5

---
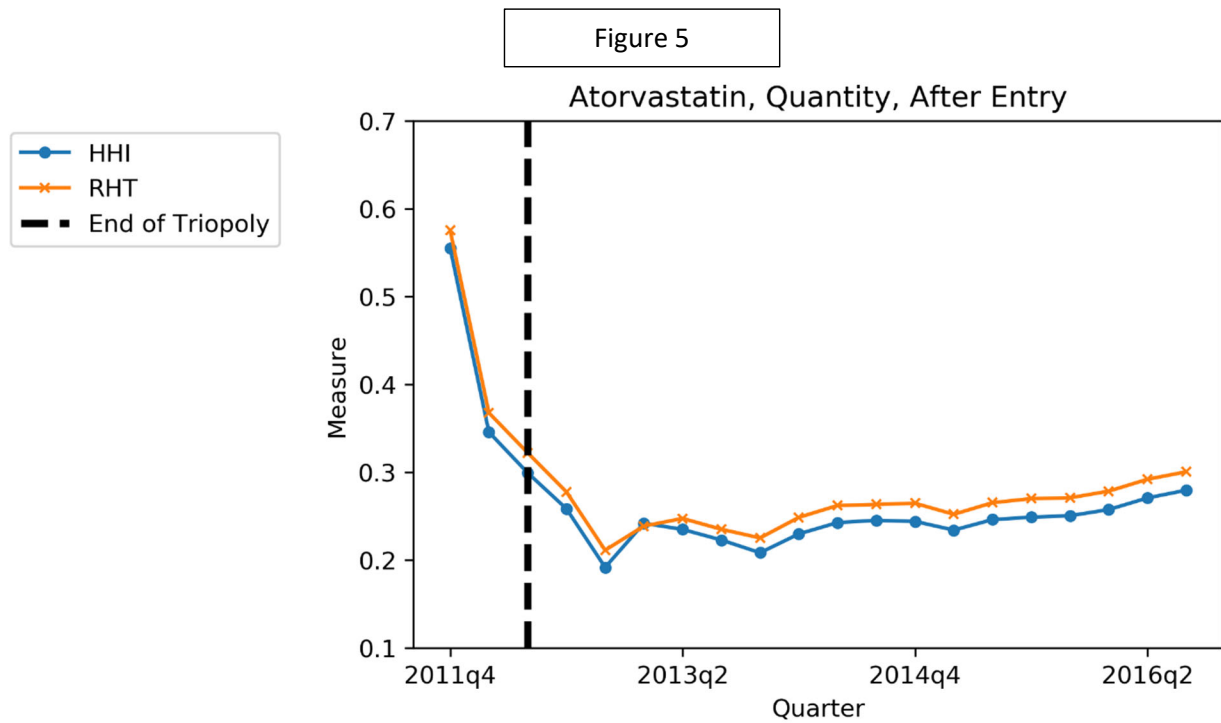
[72] During the 180-day triopoly period (2011Q4-2012Q2), the mean (standard deviation) of Pfizer's Lipitor™ brand price was $4.54 (0.10), while that for the Actavis authorized generic entrant was $2.94, and that of the Paragraph IV filer Ranbaxy was $2.80 (0.42). Following the 180-day exclusivity, Pfizer's brand Lipitor™ price increased further to $5.47 (0.84), while prices for the various generic products (including the non-Pfizer triopolists Actavis and Ranbaxy) all were much lower than the branded Lipitor™, converging over the 2012Q2-2016Q3 interval to a mean of $0.32 (less than 6% of the Lipitor™ brand price), and a standard deviation of 0.40. Notably, even though they were components of the 2011Q4-201Q2 triopoly, the authorized generic Actavis and Paragraph IV filer Ranbaxy competed aggressively with other generics after the triopoly expired; prices of Actavis and Ranbaxy between 2012Q2 and 2016Q3 had a mean of $0.30 and a standard deviation of 0.49. Note that even if each of the dosage strengths of generic atorvastatin were priced identically across manufacturers, differences in average price could emerge because of variations across manufacturers in the mix of dosage strengths.

Similar patterns hold for HHI measures based on revenue shares and quantity shares (the Pearson simple correlation between HHI Revenues and HHI Quantities is 0.908). While RHT quantity measures are typically greater than HHI quantity metrics, as was the case for their revenue-based counterparts, there does not seem to be a consistent inequality between the RHT revenue and RHT quantity values (while pre-2014Q1 RHT revenue > RHT quantity, thereafter the inequality is reversed), nor is there a consistent inequality between the HHI revenue and the HHI quantity metrics (although a similar inequality pattern holds pre- and post-2014Q1). Finally, the trends in concentration as measured by HHI - Revenues and RHT-Quantities are also very similar - the Pearson simple correlation is 0.899, while that between RHT Revenues and HHI quantities is 0.924. Thus in the empirical analysis reported below, we focus only on HHI revenues.

It is worth noting that the high correlations between the various RHT and HHI concentration measures based on either revenue or quantity shares could reflect the fact that the various atorvastatin calcium manufacturers' products are each dosed once daily, all are rated to be bioequivalent to each other by the FDA, and the FDA's bioequivalence certification could mitigate price differentials across manufacturers.[73] It is worth emphasizing, however, what we observe in the U.S. atorvastatin calcium market may not be generalizable to other pharmaceutical, and especially non-pharmaceutical, markets.

C. ACCOUNTING FOR THE EFFECTS OF MANUFACTURER'S SUBSIDIARIES WHEN MEASURING HHI CONCENTRATION

In the HHI and RHT concentration measure calculations reported above, we treated each of the firms marketing atorvastatin calcium as independent, implicitly assuming they were vigorous competitors. If the various manufacturers coordinate their pricing and marketing actions, however, treating them as independent would understate their combined market power and ability to affect

---

[73] https://www.fda.gov/media/89135/download.

prices and quantities sold. To the extent one believes that the ability of a firm to control pricing, marketing, and product quality is affected by whether its span of control encompasses wholly owned subsidiaries or other consolidated entities, [74] one may believe it informative and necessary that the industry's concentration measures take into account mergers and acquisitions, and aggregate the contributions from all the firms' subsidiaries. Moreover, whether the branded firm has control over pricing by its authorized generic during the 180-day exclusivity has not been studied extensively.[75] Thus, it is instructive to examine the sensitivity of the HHI concentration measures to whether one treats subsidiaries as independent of vs. being combined with the parent company.

Suppose one inquired, how competitive was the atorvastatin calcium market at the end of our study period in 2016Q3? Using the HHI Revenue concentration measure and treating Pfizer and its wholly owned subsidiary Greenstone as independent suppliers, we obtain a concentration measure of 0.21498, and when employing the FTC-DOJ convention of multiplying the 0-1 range by 10,000, we observe an HHI of 2149.8. This number is modestly under the DOJ's 2500 post-merger threshold for scrutinizing any merger that might enhance market power, and suggests the atorvastatin calcium market was quite competitive. The numbers-equivalent of this HHI Revenue metric is 4.6516.

However, if one recognizes that Greenstone who markets generic atorvastatin calcium is a wholly owned subsidiary of brand manufacturer Pfizer, and that it is the Upjohn Division of Pfizer that markets brand Lipitor^{TM}, then to assess competitiveness one might want to combine the shares of Greenstone's generic atorvastatin calcium with the shares of Pfizer-Upjohn's brand Lipitor^{TM} division. If one does that, the combined HHI Revenue metric in 2016Q3 increases by 0.07746 to 0.29244, which,

---

[74] In particular, until 2014Q1 market shares based on quantities are quite similar to market shares based on revenues. After 2014Q1, however, the high price of brand Lipitor^{TM} (about $6.00) relative to generics' average price (of about $0.30) increases the gap between HHI revenue and HHI quantity metrics.

[75] In the initial quarters Actavis-Watson was the authorized generic, it captured about a 20% market revenue share. But this fell to 11% in 2012Q3, 5% in 2013Q1, and never exceeded 1% thereafter. Thus, in terms of market revenue share, the impact of the authorized generic was primarily manifested during the 180-day exclusivity that ended in 2012Q2, and sharply declined thereafter.

when multiplied by 10,000, becomes 2924.4, well above the DOJ's 2500 threshold, and possibly calling into question just how competitive was the U.S. atorvastatin calcium market in late 2016.[76] With Pfizer's Lipitor[TM] and Greenstone's generic atorvastatin calcium combined, the numbers-equivalent of this HHI Revenue metric falls to 3.4195.

This is not the only consolidation involving Pfizer. On July 29, 2019, Pfizer announced that its Upjohn Division that manages brand Lipitor[TM] not only was being combined with its wholly owned Greenstone subsidiary that marketed generic atorvastatin calcium, but that it was also acquiring and merging with Mylan pharmaceuticals – another major marketer of generic atorvastatin calcium in the U.S.[77] Pfizer contemplated the announced deal would close in mid-2020, but the merger needs to survive scrutiny by U.S. antitrust authorities such as the Department of Justice (DOJ) or the Federal Trade Commission (FTC) before it can be consummated.[78] The Pfizer announcement raises the issue: Had this Pfizer-Greenstone-Mylan consolidation occurred earlier at the end of our 2004Q4-2016Q3 study period instead of being announced in July 2019, what would have been the consequences for atorvastatin calcium concentration measures? Would the likely increased HHI concentration warrant U.S. antitrust authority scrutiny of the proposed merger?

---

[76] For the HHI metric, if one combines the shares of Firm j and Firm k into the share of a single aggregated (Firm j + Firm k), thereby incorporating the combined firm into the HHI calculation, the HHI with the combined firm will increase by $2*(S_j + S_k)$. In 2016Q3, the revenue shares of Greenstone and Pfizer's brand Lipitor[TM] were 0.1308 and 0.2960, respectively. We ignore the market share of the authorized generic Actavis/Watson, since after 2014Q4 its market revenue share never even reached 0.01%.

[77] CBS News, "Pfizer to merge off-patent drug unit with Mylan, creating new drug giant", July 29, 2019. Available online at https://www.cbsnews.com/news/pfizer-mylan-merger-pfizer-to-merge-off-patent-drg-unit-with-mylan-creating-new-drug giant/, last accessed 10 August 2019. Also see Frank Vinluan, "Upjohn & Mylan to Merge, Freeing Pfizer to Focus on Biopharma R&D", *Xconomy*, July 29, 2019. Available online at https://xconomy.com/new-york/2019/07/29/upjohn-mylan-to-merge-freeing-pfizer-to-focus-on-biopharma-rd/, last accessed 10 August 2019; and Upjohn, "Our Products: Portfolio of Brands", undated. Available online at https://www.pfizer.com/products/upjohn/portfolio, last accessed 10 August 2019.

[78] Subsequent to the July 29, 2019 announcement of the Upjohn-Mylan consolidation, in November 2019 Pfizer and Mylan announced that it would rename its combined Upjohn-Mylan division firm Viatris. See Patricia Sabatini, "Viatris picked as new name for a merged Mylan/Upjohn", Pittsburgh Post-Gazette, November 12, 2019. Available online at https://www.post-gazette.com/business/healthcare-business/2019/11/12/Viatris-Mylan-Upjohn...

Recall that in 2016Q3, Mylan's revenue market share at 0.1992 was second largest to Pfizer's

0.2960. If one first combined Pfizer's Lipitor™ brand share of 0.2960 with Greenstone's generic

atorvastatin calcium share of 0.1308, and then added the 0.1992 market share of the proposed Mylan

acquisition, one would obtain a combined Pfizer atorvastatin calcium revenue market share of 0.6260 in

2016Q3. Based on Eqn. 3 and footnote 82, the change in the HHI Revenue concentration from the two-

division (Pfizer-Upjohn and Greenstone) to the three-division firm (Pfizer-Upjohn, Greenstone and

Mylan) in 2016Q3 would be from 0.29241 to 0.46248, which when multiplied by 10000 is a 1700.4

change in the HHI-Revenue metric from 2924.1 to 4624.5. This increase in the HHI Revenues as a result

of the Mylan acquisition implies a decrease in the numbers-equivalent metric from 3.4195 to 2.1622.

Under the 2010 revised DOJ horizontal merger guidelines, mergers that increased HHI by more than 200

points and resulted in a post-merger HHI of >2500 would be "presumed to be likely to enhance market

power" and invite close scrutiny from U.S. antitrust authorities.[79] Therefore, the very substantial

increase in the HHI to its post-merger high level above 4500 suggests the proposed merger might in fact

receive close scrutiny. However, as Carl Shapiro, former Chief Economist at the Antitrust Division of the

DOJ has recently noted, quoting from a District of Columbia 1990 District Court decision, "[e]vidence of

market concentration simply provides a convenient starting point for a broader inquiry into future

competitiveness"[80], any scrutiny of this proposed merger is likely to involve many considerations other

than merely levels or changes in HHI.

Nonetheless, these calculations demonstrate the sensitivity of the HHI/RHT concentration

measures to assumptions concerning the market power potentially exercised by Pfizer's atorvastatin

calcium franchise, and highlight their equivocal implications. Whether the proposed Pfizer-Mylan

---

[79] See https://www.justice.gov/atr/file/810276/download for further discussion on the DOJ's horizontal merger guidelines and their implementation.
[80] Carl Shapiro, "Protecting Competition in the American Economy: Merger Control, Tech Titans, Labor Markets", *Journal of Economic Perspectives*, Summer 2019, 33(3):69-93.

merger would significantly harm or possibly enhance competition in the U.S. market for atorvastatin

calcium, whether the merger should be approved or challenged by antitrust authorities, and/or whether

divestment of a portion of Pfizer's Greenstone subsidiary should be a requirement for the merger to

proceed, are issues beyond the scope of this research.[81] In addition, the U.S. atorvastatin calcium market

in 2019 or 2020 may differ greatly from that in 2016Q3. As implied by the above 2019 quote from Carl

Shapiro, when contemplating whether to challenge mergers and acquisitions, contemporary U.S.

antitrust analyses consider issues beyond the historic "structure, conduct and performance" paradigm

that focused particularly on market concentration measures such as the HHI. Among other issues,

antitrust authorities now also take into account buyers' market power, barriers to entry and exit, pricing

actions following previous mergers and likely ones should the proposed merger occur, non-price

competition, implications for R&D, global trade matters, and cost implications from scale and scope

economies facilitated by the merger. Consideration of HHI is but one component of a thorough antitrust

competition analysis.[82]

VII.   SUMMARY AND CONCLUDING REMARKS

The characterization of industry structure and industry concentration has long been a task facing

empirical economic researchers, for it is widely believed that market structure, market behavior and

various market performance outcomes are important interrelated phenomena. Although a number of

---

[81] There is a recent precedent for the FTC requiring divestment in order for a proposed merger to be consummated. See "FTC Requires Teva to Divest Over 75 Generic Drugs to Settle Competition Concerns Related to its Acquisition of Allergan's Generic Business: Settlement Preserves Competition and Marks Largest Divestiture Order in a Pharmaceutical Merger Case", Federal Trade Commission, Press Release, July 27, 2016. Available online at https://www.ftc.gov/news-events/press-release/2016/07/ftc-requires-teva-divest-over-75-g..., last accessed 24 August 2019.

[82] Although they acknowledge the multiplicity of factors that are taken into account when evaluating competitive effects of mergers and acquisitions on potential consumer harm, Volker Nocke and Michael D. Whinston [2019] have recently concluded , "…we see it as likely that for evaluating unilateral effects the form of current concentration schemes should be modified to focus more on the change in the Herfindahl index and less on its post-merger level, probable that safe harbors are allowing mergers to proceed that lead to consumer harm, and possible that current concentration schemes are more generally too lenient." Volker Nocke and Michael D. Whinston, "Concentration Screens for Horizontal Mergers" (First draft), unpublished manuscript, MIT Sloan School of Management, September 6, 2019.

alternative measures of market concentration are commonly used, such as the k-firm concentration measure and the Herfindahl-Hirschman index (HHI), their foundations in economic theory and statistics are limited and have not been developed extensively, leaving their unqualified use as measures of market power potentially vulnerable to the criticism of "measurement without theory".

Notably, the characterization of stochastic size distributions is not unique to economics, but is a common practice in many other disciplines and fields of study. In this paper, we have traced the pedigree of the much-used Herfindahl-Hirschman (HHI) economic concentration index to the Simpson Index of diversity originally developed in ecology, where an identical calculation to the HHI is interpreted as the probability of two organisms randomly selected from a sample habitat belonging to the same species (analogous in economics to the probability a pair of randomly and independently selected products are being marketed by the same manufacturer). This probabilistic foundation of the HHI to some extent shields it from the allegation that the sum of squared shares calculation is arbitrary and unscientific, even as its links to market power and antitrust competition analysis remain ambiguous.

We have also considered alternative proposed measures of concentrations, some of them mathematical generalizations of the HHI, others such as entropy originating from information theory in engineering and physics, another set that is developed axiomatically, and still others incorporating related concepts such as inequality and absolute population size. We have considered computational and interpretability aspects of the various concentration measures, and noted the extent to which they incorporate considerations not only of relative inequality such as the Gini coefficient and Lorenz curve, but also of absolute population size. Other things equal, markets with a large number of competitors suggest barriers to entry are limited, and therefore such markets could plausibly be expected to be competitive, other things equal. Therefore, to economists concentration metrics incorporating both variability/relative inequality and absolute population size considerations are preferable, for if one believes that economic performance outcomes depend not only on relative sizes but also on the

absolute number of competitors in a market, then one prefers a concentration measure that incorporates both features.

The existing economic literature comparing the various concentration metrics on *a priori* statistical and axiomatic criteria appears to view the HHI and the closely related Rosenbluth/Hall-Tideman (RHT) metrics most favorably. Choice between these two measures on *a priori* grounds is indeterminate, since the choice involves selection of weights and is therefore similar to choice among alternative index number formula in economic index number theory.

Finally, we have illustrated empirically the HHI and RHT measures based on revenue shares and on quantity shares, using data from the U.S. market for Lipitor^{TM} -- a large-selling branded statin drug and its generic atorvastatin calcium equivalents, 2004Q4 to 2016Q3, whose market status changed from patent-protected monopolist to one of facing many competitors in the generic commodity market. The very large size of the branded Lipitor^{TM} market would suggest likely considerable generic entry post-LOE, and indeed in the five years between Lipitor^{TM} LOE in 2011 and the 2016Q3 end of our sample a total of 17 distinct entities marketed atorvastatin calcium. Therefore, there is reason to expect dramatic declines in the concentration indexes following LOE to levels consistent with competition. We observe that HHI measures based on revenue shares and on quantity shares show similar sharp declines following Lipitor^{TM}'s LOE, although towards the end of our sample concentration metrics based on quantity shares are greater than those based on revenue shares reflecting the fact that at that time the price of the branded drug was more than sixteen times the average price of various generic drugs. Over the entire 2011Q4 – 2016Q3 post-LOE time period, the quantity and revenue-share based HHIs are highly correlated, as are RHT measures based on revenue shares and on quantity shares. Moreover, levels and trends of HHI measures are highly correlated with corresponding RHT metrics.

However, we have also demonstrated the limited and equivocal role played by concentration measures in evaluating the extent of competition by documenting the dramatic numerical sensitivity of

the HHI revenue metrics to how one defines the identity and market shares of competitors when mergers and acquisitions occur. Specifically, HHI measures rise substantially above U.S. antitrust authority scrutiny threshold levels when Pfizer's wholly owned Greenstone generic subsidiary is combined with Pfizer's branded Lipitor^{TM} division, and especially when in addition Pfizer's proposed acquisition of generic manufacturer Mylan is included as part of the Pfizer atorvastatin calcium franchise.

Finally, based on *a priori* reasoning, there is little on which to base a choice between the HHI and RHT measures (both meet requirements satisfying axiomatic conditions, while the HHI has an attractive probabilistic foundation). Moreover, we find very similar empirical trends and high correlations among them. Together, these considerations offer support and encouragement for the continued use of the HHI metric in measuring economic concentration, provided that users recognize that market structure as measured by concentration is only one of the various factors affecting market performance outcomes, and that market behavior and strategic considerations also play important roles.