This comment is submitted in response to the United States Federal Trade Commission (“FTC”) hearing on Concentration and Competitiveness in the U.S. Economy as part of the Hearings on Competition and Consumer Protection in the 21st Century. We submit this comment based upon our extensive experience and expertise in antitrust law and economics. As an organization committed to promoting sound economic analysis as the foundation of antitrust enforcement and competition policy, the Global Antitrust Institute commends the FTC for holding these hearings and for inviting discussion concerning a range of important topics.

Businesses today have greater access to data than ever before. Firms now have access to data at high volume, high velocity, and high variety—a phenomenon known
as “big data.” The increasing prevalence of big data creates new questions for antitrust enforcers. This Comment will discuss how big data should be considered in the context of antitrust analyses.

**Introduction**

Big data is often associated with online platforms that collect and use vast amounts of user data to improve their services and to target advertising, but the phenomenon is not limited to the digital space. Brick-and-mortar retailers use vast quantities of data as well. For example, Target once discovered that a teenager was pregnant before her parents did, based solely on her purchase history. Target uses such information to personalize advertisements and coupons.

Walmart is even more active. As one of the heaviest users of big data, Walmart analyzes over 2.5 petabytes of data every hour in its Data Café in Bentonville. In addition to helping it personalize advertisements, Walmart’s data analysis helps it run pharmacies more efficiently, speeding checkout, managing its supply chain, and optimizing placement on store shelves.

---


Background

While the volume and speed of data available to retailers today is dramatic, businesses have been using customer data to increase efficiency and create value for a long time. Mid-twentieth century grocery giant A&P used customer data to discern regional preferences, and to forecast demand in order to reduce food waste.\(^6\) An article from Harper’s Magazine in the late 1980s used the term “big data” to refer to the variety of customer data lists used to increase the effectiveness of direct mail advertisements.\(^7\) The author found that direct mail services at the time had access to data from credit-reporting agencies, magazine subscriptions, catalog purchases, customer questionnaires, real estate records, and other sources that could sketch a picture of the likely preferences of millions of individuals.\(^8\)

The sense that advances in data science are leading to an “overwhelming deluge of information” is not new either.\(^9\) That feeling traces back at least to the “avalanche of numbers” that occurred in the 1820s and 1830s, when national governments began accumulating more data than ever before to classify and tabulate their subjects in an


\(^7\) See Erik Larson, What Sort of Car-Rt-Sort Am I? Junk Mail and the Search for Self, HARPER’S MAG. (July 1989).

\(^8\) See id.

\(^9\) David Beer, How Should We Do a History of Big Data?, 3 BIG DATA & SOCIETY 1, 2-4 (2016) (“[T]he sense that we are being faced with a deluge of data about people is not new, in fact it has a long history.”).
attempt to improve governance. Sociologist David Beer points out that while “features of the current data movement are in some ways novel,” data has been scaling up for hundreds of years.

Today firms have both more data than ever and increasingly sophisticated methods for analyzing it. Firms collect and store data from online traffic, credit card purchases, smart phones and wearables, and numerous other data streams. These advances benefit consumers by, among other things, allowing firms to improve products and lower costs, and enabling the growth of the free-to-consumers model widely employed in the online space. The growth in the importance of data has also led to concerns that incumbent firms’ access to data can create barriers to entry in many industry.

Is Big Data an Entry Barrier?

The prior discussion leads naturally one to a question whether big data is itself a barrier to market entry. Before addressing this question, it is important to consider how economists define “barriers to entry,” and how this concept relates to the analysis of competitive outcomes.

---

10 See id.
11 See id.
14 See id. at 1147-49 (discussing these concerns and suggesting they “may be grossly overstated”).
Bain (1956) defined barriers to entry as structural factors that allow incumbents to persistently price above the competitive level without incurring the threat of entry.\textsuperscript{15} Examples include economies of scale that require large capital expenditures, product differentiation, and absolute cost advantages. Stigler (1968) considered barriers to entry as costs that an entrant must incur but incumbents do not.\textsuperscript{16} Examples include patents and grandfathered government regulations but not economies of scale to the extent that an entrant has access to the same cost function. The appeal of Stigler’s definition is its recognition that incumbents can earn supra-normal profits over the long-term only if they have some persistent advantage over potential rivals. What is missing from both Bain and Stigler, however, is an assessment of welfare. Both Fisher (1979) and von Weizsäcker (1980) filled this void with a normative definition that incorporates social welfare.\textsuperscript{17} Fisher found “a barrier to entry exists when entry would be socially beneficial but is somehow prevented.”\textsuperscript{18} von Weizsäcker stated, “a barrier to entry is a cost of producing which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry and which implies a distortion in the

\textsuperscript{15} See Joseph Bain, Barriers to New Competition 3 (1956).
\textsuperscript{16} See George J. Stigler, The Organization of Industry 67 (1983) (“A barrier to entry may be defined as a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry.”).
\textsuperscript{17} Franklin M. Fisher, Diagnosing Monopoly, 19 Q. Rev. Econ. & Bus. 7 (1979); C.C. von Weizsäcker, A Welfare Analysis of Barriers to Entry, 11 The Bell J. Econ. 399 (1980).
\textsuperscript{18} Fisher, supra note 12, at 23.
allocation of resources from the social point of view.”¹⁹ If economies of scale can increase overall welfare and we associate entry barriers with inefficiencies, then, von Weizsäcker asks, “in which sense can we speak of a barrier to entry?”²⁰ According to Fisher, “the right issue is not whether there are barriers to entry into the production of a particular mousetrap, but whether there are barriers to entry into innovation in mousetraps.”²¹ In other words, it is not the input itself that is the barrier to entry but, if they exist, hindrances to obtaining and utilizing the inputs, which incumbents do not face.

The tension in defining barriers to entry is that there are really two ways in which the term is discussed. As Carlton stated, “Trying to use ‘barriers to entry’ to refer to both the factors that influence the time it takes to reach a new equilibrium and to whether there are excess long-run profits is confusing.”²² Therefore, for the purpose of competition policy, Carlton recommends that “rather than focusing on whether an entry barrier exists according to some definition, analysts should explain how the industry will behave over the next several years ... [which] will force them to pay attention to uncertainty and adjustment costs.”²³

---

¹⁹ C.C. von Weizsäcker, supra note 12, at 400.
²⁰ Id. at 401.
²¹ Fisher, supra note 12, at 27.
²² Dennis W. Carlton, Barriers to Entry, 1 ISSUES IN COMPETITION L. & POL’Y 601, 606 (2008).
²³ Id. at 615.
Consequently, we find that it is best to avoid suggesting that big data is or is not a barrier to entry. Rather, it is one potential factor when examining “the timeliness, likelihood, and sufficiency of entry efforts an entrant might practically employ.”

There are clearly impediments that an entrant must overcome in order to compete effectively. Common examples include regulatory compliance costs, expenditures on specialized equipment, and hiring skilled labor. Some obstacles are nominal. Some obstacles are substantial. Attempting to classify these impediments as entry barriers or not creates the conflation mentioned by Carlton. If this classification exercise is undertaken, however, it is important to clarify whether the discussion regards either the time required to reach a new equilibrium or a cost that is preventing socially desirable entry.

It is also relevant to note that big data is not an exogenous factor that dictates the number of firms in a market, which in turn determines the degree of competition and the rate of return. The contrary view is a vestige of the discarded structure-conduct-performance paradigm. Rather big data is endogenous, as are other dimensions of non-price competition. For instance, if a firm invests heavily in research and development, which allows it to introduce a new product or substantially to improve an existing

---

25 See Carlton supra note 16, at 604 (“Models that focus on only price competition may fail miserably to correctly predict industry concentration and consumer welfare when there are other product dimensions along which competition occurs. This is likely to be particularly true in industries requiring investment and creation of new products.”).
product, we would not normally view this as anticompetitive conduct or even conduct that ultimately leads to anticompetitive results. Rather, we would consider investment in innovation to be procompetitive. Similarly, investments in big data can create competitive distance between a firm and its rivals, including potential entrants, but this distance is the result of an activity that antitrust laws should encourage, not penalize.\textsuperscript{26} Moreover, the observation that a firm is making large margins gives no indication whether this reflects supra-competitive pricing if we properly consider the rate of return required over the whole process, including its investment in big data.\textsuperscript{27}

Even if big data represents an important component in the success of an incumbent, entrants can differentiate their products along other dimensions important to consumers. For instance, a grocery store entrant might focus more on carrying locally made produce or products that cater to specific diets. An online firm might focus more on building network effects or greater integration with complementary products rather than the use of big data. As Tucker and Wellford state, “The fact that some established online firms collect a large volume of data from their customers or other sources does not mean that new entrants must have the same quantity or type of

\textsuperscript{26}Michael L. Katz & Carl Shapiro, \textit{Systems Competition and Network Effects}, 8 J. ECON. PERSPECTIVES 93, 107 (1994) (“[M]erely observing a firm with a position of market dominance does not imply that the firm is earning super-normal profits: the firm’s quasi-rents may merely reflect costs incurred earlier to obtain the position of market leadership.”).
data in order to enter and compete effectively … [L]ack of asset equivalence should not be a sufficient basis to define a barrier to entry.”  

**Big Data in Perspective**

As the previous section indicates, big data does not represent a barrier to entry but rather a factor of production much like any other input. It should therefore be treated as such. While big data may be an important factor of production, it is not an end in and of itself. Just as labor, innovation, capital and entrepreneurial skill differentiate firms, so too does the ability to turn “big data” into something useful.

Hal Varian (2014) discusses how, in order to make good use of the data they have collected, companies must understand causality rather than just correlation. To understand causality, between, say, changes in price and in sales, firms must conduct experiments with control groups. Simply having the data is not enough to understand and make use of the information. For example, he reports that Google conducts roughly 10,000 experiments a year in search and advertisements, which help Google continually to improve its system. Varian also emphasizes that the need for continuous monitoring and experiments is nothing new: Henry Ford and his colleagues

---

30 Id.
31 Id.
were on the factory floor every day in order to perfect the assembly line, the key factor of production in Ford’s business.\textsuperscript{32}

These examples show that information collected by firms is not a new phenomenon, nor does it give them an inherent advantage simply by obtaining the information. It is how firms use this information, and the additional steps taken, that ultimately makes “big data” a key factor of production. Policy makers should be cognizant of this when making decisions that will treat companies with large amounts of information differently.

**Conclusion**

Data has attributes that differentiate it from other types of inputs, such as specialized equipment and skilled labor. It is non-rivalrous, which means that data can be used again and again without it losing its value and the collection of data by one firm does not inhibit or raise the cost of collection by other firms. Further, big data can be part of a positive feedback loop where increased quality begets an increase in users which begets further increases in quality. This process is undoubtedly procompetitive and welfare enhancing. Could, on the other hand, big data be used to insulate from entry an incumbent with antitrust market power? Perhaps, but again the issue is not whether it is a barrier to entry but rather what role data plays in that particular industry. If big data alters the competitive process, then we should focus on the testable

\textsuperscript{32} Id.
implications that emerge. What we should avoid, however, are shortcuts and regulation, or unwarranted antitrust intervention, based on the “bigness” of big data, and inappropriate labels such as “barriers to entry.”