Deconstructing the Antitrust Implications of Big Data
Keynote Remarks of FTC Chairwoman Edith Ramirez[1]
43rd Annual Conference on International Antitrust Law and Policy
Fordham Competition Law Institute
New York, NY
September 22, 2016

I am delighted to be back in New York for this annual conference on international antitrust law and policy. My thanks to James Keyte and Fordham for inviting me.

As antitrust enforcers, we have to pay attention to technological developments in the markets we are analyzing because they can affect how we think about the nature and sources of competition. This morning I would like to talk about a feature of the digital economy that has been a topic of much discussion in the antitrust community of late, namely, big data.

In describing big data, people often refer to the "three Vs" – volume, velocity, and variety.[2] The volume and variety of data stem in part from the ubiquitous collection and compilation of smaller data – a tap of a smartphone, a website login, or a movement collected by a shopping center sensor, to cite a few examples. The exponential increase in the volume of data also stems from the plunging cost of data storage, which can extend the lifespan of data indefinitely. The variety of big data reflects more than just the ever-increasing collection and storage of raw data. It is also the product of an unprecedented power to analyze data, in order to draw inferences about the world and make predictions about events to come.

Big data has the potential to save lives, enhance government services, increase marketplace efficiency, and boost economic productivity. It can be used to deliver higher quality

---

[1] As prepared for delivery.

[2] *See, e.g.*, Doug Laney, *3D Data Management: Controlling Data Volume, Velocity, and Variety* (Feb. 6, 2001), *available at* http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf. Some experts identify a fourth "V" – veracity – whether the data is credible and reliable.

health care at lower cost; enable forecasters to better predict the weather and spikes in energy consumption; and improve industrial efficiencies in order to deliver better and lower-cost products and services to consumers. Some say it will revolutionize how we live, work, and think.

But what are the implications for competition? As antitrust enforcers, how should we think about big data? And does the equation change in situations where the big data at issue consists mainly of consumer information?

I will focus on the issue of the aggregation of data, and, using the experience of the Federal Trade Commission, share some thoughts about how we can begin to answer these questions.

I. The Competitive Significance of Data

Like big data itself, the idea that data may have competitive significance or strategic value to a firm is not entirely new. For at least as long as we have had computers and databases, firms have had some ability to make use of the data they possess in their business or operations.

In assessing its potential significance or value, we generally view data as we would any other asset – either as a product or as an input to a product or service. I will illustrate the FTC's analytical approach with a few examples from our merger work.

An example of a merger case where we viewed data as a key input is the FTC's 2008 challenge of a proposed acquisition by Reed Elsevier of its market rival, ChoicePoint.[3] The Commission alleged a relevant market for "electronic public record services for law enforcement customers," in which Reed Elsevier's subsidiary, LexisNexis, and ChoicePoint were, by far, the two largest providers.

[3] *Reed Elsevier NV, et al.*, File No. 081-0133 (filed Sept. 16, 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/09/080916reedelseviercpcmpt.pdf.

The electronic records at issue encompassed a wide array of public and non-public records about individuals and businesses, including certain credit data, criminal records, motor vehicle records, property records, and employment records. LexisNexis and ChoicePoint each combined their respective databases of such records with robust analytics to offer a suite of services that allows law enforcement customers to uncover previously unknown information about persons of interest and to generate leads in their investigations.

The Commission alleged that the proposed merger likely would have resulted in a loss of the intense head-to-head competition between LexisNexis and ChoicePoint to provide information services tailored to the needs of their law enforcement customers. Although the electronic records themselves were readily available from the same government and private sources to competitors, the FTC alleged that entry would be difficult because of the time and cost associated with developing comparable services and gaining customer acceptance. In particular, an entrant would need analytical tools that could turn the records into services that customers would find valuable. Ultimately, the parties chose to resolve the Commission's competitive concerns with a divestiture of ChoicePoint's electronic public records services to Thomson Reuters.

An example of a merger case where data was itself the product or service is the FTC's 2010 challenge of Dun & Bradstreet's already consummated acquisition of Quality Education Data, or QED.[4] Our complaint charged that the acquisition substantially lessened competition in the market for kindergarten through twelfth grade educational marketing data.

This K−12 data included contact, demographic, and other information relating to teachers, administrators, schools, and individual school districts. Customers of QED and D&B

[4] *Dun & Bradstreet Corp.*, Dkt. No. 9342 (filed May 7, 2010), https://www.ftc.gov/sites/default/files/documents/cases/2010/05/100507dunbradstreetcmpt.pdf.

used this data to market their products and services to teachers, administrators, and other school personnel through direct mail and e-mail. The FTC determined that the parties' customers did not regard other sources of marketing data as close substitutes. The data, by virtue of its unique characteristics, had greater utility and value to customers than alternative datasets. Consequently, we viewed the combination of the parties' respective K−12 data products as enhancing D&B's market power under a straightforward unilateral effects theory.

Following the acquisition, only one other competitor offered K−12 data but its dataset did not compare in size, breadth, or scope of coverage to that of either merging party, pre-acquisition. After four months of administrative litigation, the matter was resolved with a consent agreement that called for a divestiture of an updated and augmented K−12 database, as well as the QED brand and related intellectual property, to Mailings Clearing House, a fringe competitor.

Let me make one other observation about the *Reed Elsevier* and *Dun & Bradstreet* cases. In both matters, one of the alleged anticompetitive effects of the merger was harm to innovation. In *Reed Elsevier*, the intense rivalry between LexisNexis and ChoicePoint had incentivized ChoicePoint to launch a new and advanced electronic public records service designed specifically for law enforcement customers.[5] The Commission thus expressed concern that the proposed acquisition would reduce LexisNexis's incentives to innovate and develop new services. Similarly, in *Dun & Bradstreet*, competition from QED had spurred D&B to improve its product quality, which included the development of new product features.[6]

---

[5] Analysis to Aid Public Comment at 3, *Reed Elsevier NV, et al.*, File No. 081-0133 (filed Sept. 16, 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/09/080916reedelseviercpanal.pdf.

[6] Analysis to Aid Public Comment at 1, *Dun & Bradstreet Corp.*, Dkt. No. 9342 (filed Sept. 10, 2010), https://www.ftc.gov/sites/default/files/documents/cases/2010/09/100910dunbradstreetanal.pdf.

A key take-away from these cases is that we need to be especially vigilant of the potential effects on innovation where mergers involve big data as an input or a product. As I mentioned at the outset, what makes the world of big data look so promising are the advances in computational and statistical analysis that allow firms to make better sense of, and derive greater utility from, an ever-growing mass of information. These advances have the potential to make data more competitively significant and strategically valuable to a firm, and thus to augment the intensity of competition between that firm and its market rivals. In such a setting, the harm posed by a merger may well be the dampening or elimination of this important dimension of competition. As antitrust enforcers, we need to ensure that a merger does not break or impede a virtuous cycle of competition and innovation.

II. Implications of Consumer Data

Big data encompasses all types of data from every conceivable source. But much of the exponential growth in the amount of data that is being generated today is driven by the ubiquitous collection of consumers' personal information.

To understand the big data economy when it comes to consumer data, it is important to recognize the ever-growing sources of "little data," which can quickly become big data. Online tracking is one method of gathering little bits of data about who consumers are, what they do, and where they go. With every search, every click, every tweet, and every post, data is collected. Traditionally, companies have engaged in this type of tracking through cookies, but they also use other methods like Flash cookies, "history sniffing," and "device fingerprinting" to track consumers across websites. On mobile devices, cross-app tracking is routine. In fact, we have gone from cross-site tracking to cross-app tracking, and now to cross-device tracking, where companies can track the same consumer across her desktop, laptop, tablet, and smartphone.

But data collection about consumers is not limited to computers and mobile devices. Companies also follow consumers across Internet-connected devices referred to as the Internet of Things ("IoT"). Three and one-half billion sensors are already in the marketplace. Today, connected devices are in our homes, our cars, and on our bodies, in the form of connected smoke detectors and light bulbs, connected cars, and wearable computers, among others.

This is only the beginning. Some experts estimate there are 18 billion connected devices around the world today and that there will be more than 30 billion by 2020.[7] Others view those estimates as conservative, predicting the number of sensors will increase from 15 billion as of 2015 to a trillion within the next decade.[8] All of these connected devices mean much more data about consumers – their activities, their habits and preferences, their family and friends.

And, finally, although many big data discussions focus on these high-tech methods of data collection, one should not forget more traditional brick-and-mortar data collection. Through loyalty programs, warranty cards, surveys, sweepstakes entries, and credit card purchases, we leave yet another trail of "little data" bread crumbs. The trail has only gotten longer and wider as retail has expanded into electronic commerce and digital platforms.

The volume and variety of consumer information being collected on a daily basis raises the question about how combinations involving firms with sizable collections of personal data should be treated.

The FTC's 2007 investigation of the *Google/DoubleClick* merger is instructive. At the time of the investigation, Google was – and still is – a dominant firm in search and search advertising. It was also an ad intermediary, aggregating and selling online advertising space to

---

[7] SAM LUCERO, IHS TECHNOLOGY, IOT PLATFORMS: ENABLING THE INTERNET OF THINGS 5 (Mar. 2016), https://cdn.ihs.com/www/pdf/enabling-IOT.pdf.

[8] *See, e.g.*, R. Colin Johnson, *Roadmap to Trillion Sensors Forks*, E.E. TIMES, Dec. 10, 2015, 7:25 PM, http://www.eetimes.com/document.asp?doc_id=1328466.

advertisers and ad agencies. DoubleClick was a leading firm in the so-called "third-party ad serving market." It managed the selection, delivery, and placement of advertisements for both publishers selling online advertising space and advertisers buying that space, either directly or indirectly through intermediaries like Google.

The FTC asked whether the combination of Google's database of user search information and DoubleClick's consumer data collected on behalf of its publisher customers would enhance Google's market power in the ad intermediation market. After close investigation, the Commission concluded that it would not.

For one thing, the data collected by DoubleClick belonged exclusively to its publisher customers, and Google had agreed to abide by those contractual restrictions. Furthermore, the Commission determined that, even if Google were to access DoubleClick's data, there was not enough competitively sensitive information to give Google an anticompetitive advantage over its advertising rivals. In short, neither the data available to Google, nor the data available to DoubleClick, was, "an essential input to a successful online advertising product."[9]

The Commission also did not see how Google and DoubleClick's consumer data had particular strategic value in preserving Google's position in the ad intermediation market. Google's then-most significant competitors in that market – Microsoft, Yahoo!, and Time Warner – each had access to their own unique stores of consumer data, including data drawn from their own rival search engines, ad servers, and ad intermediation services. As long as Google had no access to its competitors' data stores, which it did not, it was difficult to see how Google held any strategic advantage over them in the ad intermediation market.[10]

---

[9] Statement of the Federal Trade Commission at 12, *Google/DoubleClick*, File No. 071-0170 (issued Dec. 20, 2007), https://www.ftc.gov/system/files/documents/public_statements/418081/071220googledc-commstmt.pdf.

[10] *See id.* at 12–13.

Google/DoubleClick thus raised the issue whether a firm's seemingly greater ability to accumulate consumer data through its dominant platform, standing alone, created competitive concerns. The Commission concluded that, at least in the ad intermediation market, Google's database of consumer data from its popular search engine, alone or in combination with DoubleClick's customer data, did not give Google additional market power.

This matter underscores the point that one firm's stores of consumer data may not be more valuable than another firm's, even though they are collected from different platforms. On the facts there, the collections of consumer data were functional substitutes for one another. As long as Google could not prevent other firms from accumulating their own consumer data, there was not any competitive concern.

Of course, a different set of facts could lead to a different result. Whether there is a competitive advantage associated with access to a large volume of data will depend on the particular set of facts, including the specific market at issue. We therefore have to analyze big data issues on a case-by-case basis and be on the lookout for transactions and conduct that seem inconsistent with competition on the merits.

The *Google/DoubleClick* merger also raised a separate issue at the intersection of competition and privacy. Privacy advocates urged the FTC to block the merger on the premise that the combination of Google and DoubleClick's respective personal datasets could be exploited in a way that threatened consumers' privacy. The Commission responded by emphasizing that the sole objective in merger review is to identify and challenge transactions that harm competition. However, with that in mind, and although the investigation focused mainly on the merger's impact in online advertising markets, the Commission also considered the possibility that the transaction might have an adverse effect on consumer privacy as a non-price

dimension of competition. Ultimately, the Commission did not find any evidence to support this theory of possible harm.[11]

The take-away is that, as antitrust enforcers, we must be prepared to consider privacy protection as an aspect of competition. In the same way that firms may compete with one another by offering warranty protection with their products, they may also offer different privacy options and features to entice consumers to purchase their services. Of course, as with any theory of competitive harm, it needs to find support in the evidence. Otherwise, it is just a theory.

I will note, however, that because of the FTC's dual consumer protection and competition mission, even if we find no harm to competition on privacy grounds, we can take other steps to protect the privacy of consumers. For example, at the same time that the Commission closed its investigation of the *Google/DoubleClick* merger, the Commission announced a set of proposed principles for online behavioral advertising.[12] The FTC recognized the need for greater transparency and consumer control in light of the privacy issues raised by behavioral advertising and saw the set of principles as a way to promote industry self-regulation. The principles took into account the potential economic benefits of online advertising to consumers, including free online content, and balanced those benefits against the risks that might flow from consumers not fully appreciating the way behavioral advertising works.

III. Conclusion

There is no question that the aggregation of data may have important implications for competition. But, as I have noted, I do not believe big data requires new competition rules or

---

[11] *Google/DoubleClick* Statement at 2–3.

[12] Press Release, Fed. Trade Comm'n, FTC Staff Proposes Online Behavioral Advertising Principles (Dec. 20, 2007), https://www.ftc.gov/news-events/press-releases/2007/12/ftc-staff-proposes-online-behavioral-advertising-privacy.

modes of analysis. Instead, what is called for is a careful untangling of the issues so that there is a clear understanding of the relationship, if any, between big data and the exercise of market power.

Of course, given that the analysis, manipulation, and use of big data are still evolving as new tools are created to mine that data, we have to be open to new fact patterns and new theories of harm. Moreover, when the data at issue is consumer information, privacy concerns may be a relevant consideration, but only if there is evidence that the merging parties compete with each other on privacy. Otherwise, such concerns are more appropriately addressed outside the context of antitrust through a separate inquiry of the parties' privacy and data security practices.

The other point I will leave you with is that by enforcing the antitrust laws solely to protect the competitive process and promote consumer welfare, we ensure that our actions do not inadvertently deprive consumers of the potential benefits of big data. In some circumstances, big data may play a competitively benign role, and in other circumstances, it may provide a fertile source of innovation competition. We should allow such scenarios to follow their natural course, provided no harm is done to competition or the competitive process.

Thank you.