



# CPI Antitrust Chronicle

May 2015 (2)

## Big Data as a Barrier to Entry

Robert P. Mahnke  
eBay Inc.

## Big Data as a Barrier to Entry

Robert P. Mahnke<sup>1</sup>

So-called “big data” has been a technological development for some time, and lately it has matured into a phenomenon that competition lawyers have noticed. Mindful of the vast quantities of data that some companies are amassing, some argue that access to data can be a barrier to entry in online markets, protecting incumbents from competition. However, Geoffrey Manne & Ben Sperry recently argued “the notion of data as an antitrust-relevant barrier to entry is simply a myth.”<sup>2</sup> Manne & Sperry have it wrong: The fact that “big data” can be an antitrust-relevant barrier to entry has been well established.

At the outset, one problem is that the term “big data” seems to mean different things to different people. Darren Tucker & Hill Welford emphasize the extent to which once-sophisticated tools are increasingly ubiquitous: “Big data is everywhere. ... Big data is used by organizations of all sizes. Small businesses, entrepreneurs, and government agencies, in addition to large companies, are avid users of big data.”<sup>3</sup> Used in this way, “big data” is “bigger” than yesterday’s data, but what gets lost are the profound differences among some of these different users.

Those who are concerned about big data as a competition problem have something different in mind. Some firms are amassing data at a completely different scale. To take one public example, Google recently opened a new server farm in Oregon, “a massive, 164,000-square-foot building” that cost \$600 million.<sup>4</sup> According to the *Oregonian*, “Google now has data centers all over the country and around the world.” Small businesses and investors cannot afford such investments, nor can many large companies. When big data gets that big, most market participants will be excluded, and competition problems will be more likely.

Manne & Sperry suggest that data matters in offline markets as well. As more of us carry and use smartphones and tablets at all hours, the distinction between online and offline business will seem increasingly archaic, but, still, Manne & Sperry are not wrong. So, they say, the notion

---

<sup>1</sup> Global Competition Counsel at eBay Inc.; formerly Trial Attorney at the U.S. Department of Justice, Antitrust Division. The views here are those of the author alone and should not be attributed to eBay

<sup>2</sup> Geoffrey Manne & Ben Sperry, *Debunking the Myth of a Data Barrier to Entry for Online Services*, TRUTH ON THE MARKET (March, 2015), available at <http://truthonthemarket.com/2015/03/26/debunking-the-myth-of-a-data-barrier-to-entry-for-online-services/>. See also Darren S. Tucker & Hill B. Welford, *Big Mistakes Regarding Big Data*, ANTITRUST SOURCE (December 2014) (“Online markets are notable for their low entry barriers and typically do not require big data for entry.”); Daniel O’Connor, *Is Big Data an Entry Barrier? What Tinder Can Tell Us*, (April 2, 2015), available at <http://www.project-disco.org/competition/040215-big-data-entry-barrier-tinder-can-tell-us/> (“However plausible this argument sounds, a review of the short history of the Internet economy ... seems to cast doubt on the soundness of the theory.”); Andres V. Lerner, *The Role of ‘Big Data’ in Online Platform Competition*, available at <http://ssrn.com/abstract=2482780> (August 26, 2014).

<sup>3</sup> Tucker & Welford, *Id.*

<sup>4</sup> Mike Rogoway, *Google opens new, \$600 million Oregon data center*, (April 9, 2015), available at [http://www.oregonlive.com/silicon-forest/index.ssf/2015/04/google\\_opens\\_new\\_oregon\\_data\\_c.html](http://www.oregonlive.com/silicon-forest/index.ssf/2015/04/google_opens_new_oregon_data_c.html).

that companies like Google or Facebook could have “a monopoly on data is silly.” Well, this is patently true, but not because offline firms have data too. Data are not a commodity. What Google knows about where your phone went last weekend is not like what Facebook knows about your circle of friends from high school.

Axiomatically, product markets are comprised of goods which are reasonable substitutes for each other. Data is no such thing, and so it makes no sense to talk about a product market in “data.” There are so many, many, kinds of data. It very well may be the case that Tinder, to take an example discussed by Daniel O’Connor, did not face data-related barriers to entry when it started to compete with other dating services, but that hardly dictates that the same will be true in other product markets. Competitive markets may all be alike, at least in theory, but every uncompetitive market is uncompetitive in its own way.

Manne & Sperry further argue that it doesn’t really matter how much data a company has; what matters is what they do with it when they design their products and services. Doubtless this can be true, and Manne & Sperry offer examples. But the online economy is incredibly diverse, featuring robust competition and terrific innovation in some corners, and dominant incumbents with remarkably durable market power in others. There is no single internet business plan. Many start-ups have been successful despite not (yet) having their own data because they have persuaded users to give it to them. Web 2.0 and social media provide many examples.

But not every industry can be won with a social model. If there is strong competition in some markets, that does not mean that there are no barriers to entry in others. And past performance is no guarantee of future results. Even if “Google achieved success over other search engines by conceiving of a better way of matching users queries to relevant websites,”<sup>5</sup> as Daniel O’Connor says, the information it has now compiled about what many of us have searched for and found useful cannot be easily replicated.

It is true that data are not like oil, in that Google and Facebook (for example) both can have the same data. (Though they also surely have many data the other does not, given the different ways that they interact with their users.) Indeed, in online businesses, this is a commercial reality. For example, shopping-comparison sites get their data about products and prices from merchants. The specifications of, say, a camera do not change when it is sold in different places. And merchants find it burdensome to create and deliver different data about their goods, so online shopping comparison sites tend to have the same product data and to compete in other ways—e.g., by generating their own unique content like product reviews.

But while two parties *could* have the same data, that hardly means that in any given industry they *will* have it. There are data, and then there are data. Some businesses are successful and valuable because they have access to data that others cannot easily obtain. For example, Craigslist. When other companies have attempted to scrape its listings, Craigslist has tried to stop

---

<sup>5</sup> Tom Hornby, *The Rise of Google, Beating Yahoo at Its Own Game*, LOWEND MAC (2013), hyperlink in the original, available at <http://lowendmac.com/2013/the-rise-of-google-beating-yahoo-at-its-own-game/>.

them, with success.<sup>6</sup> Is it impossible for anyone to compete with Craigslist in its core business? Maybe not, but anyone trying faces a substantial barrier to entry. Does every online or digital market work like this? Of course not, but some do.

Blanket assertions about the competitive dynamics of data do not illuminate the circumstances of specific markets. We have a concrete example of barriers to entry in data markets, a merger that led to precedents on both sides of the Atlantic. In 2007, Thomson announced it would acquire Reuters. Both companies offered, among other things, terminals to traders and other financial professionals. It was reported at the time that the merger was intended to enable the combined company to better compete with Bloomberg and its eponymous terminal. In essence, Thomson and Reuters offered bundles of financial data.

Both the U.S. Department of Justice (“DOJ”) and the European Commission’s Directorate General for Competition (“DG Comp”) investigated the acquisition, and eventually conditioned clearance on remedies. (As a DOJ trial attorney, I led the U.S. investigation, in cooperation with DG Comp staff.) Both agencies found that the combination raised competition concerns in the markets in specific sorts of data: “the distribution of aftermarket broker research reports, of earning estimates, of fundamental financial data of enterprises and of time series of economic data,”<sup>7</sup> to quote an EC press release. (The DOJ’s Complaint omitted time series of economic data.)<sup>8</sup> For these specific types of data, Thomson and Reuters were leading providers in concentrated markets and there were good reasons to believe that other companies would not be able to enter to compete with them.

For example, one of the products at issue was fundamentals data. Thomson and Reuters offered years of fundamentals data for publicly traded companies around the world. In some countries, this data is relatively easy to obtain. In the United States, one can scrape SEC filings. In other countries, however, this was not the case, and entrants faced a significant barrier to entry. DOJ alleged:

New entrants into the fundamentals data market, particularly with respect to international fundamentals data, must overcome significant barriers to entry. These include the difficulties of arranging for collection of data on tens of thousands of companies on a global basis, constructing a reliable historical database, the need to develop local expertise in each country’s accounting norms, and the ability to develop data normalization and standardization processes. Therefore, entry or expansion by any other firm will not be timely, likely, or sufficient to defeat an anticompetitive price increase.<sup>9</sup>

DG Comp explained further. They emphasized the time and expense required for an entrant to offer global fundamentals data. “The hurdles come, first, from the need to collect fundamental data with a global coverage and second, to collect fundamental data going back in

---

<sup>6</sup> Wikipedia, *Craigslist Inc. v. 3Taps Inc.*, available at [http://en.wikipedia.org/wiki/Craigslist\\_Inc.\\_v.\\_3Taps\\_Inc](http://en.wikipedia.org/wiki/Craigslist_Inc._v._3Taps_Inc).

<sup>7</sup> European Commission, *Mergers: Commission clears acquisition of Reuters Subject to Conditions*, available at [http://europa.eu/rapid/press-release\\_IP-08-260\\_en.htm](http://europa.eu/rapid/press-release_IP-08-260_en.htm).

<sup>8</sup> U.S. Department of Justice, Antitrust Division Decision of February 19, 2008, available at <http://www.justice.gov/atr/cases/f230200/230281.htm>.

<sup>9</sup> *Id.* at ¶37.

time several years.”<sup>10</sup> In addition, significant work had to be done once data were collected. Vendors standardized data “by making accounts of companies reporting under different accounting rules comparable and responding to codification systems relevant to users.”<sup>11</sup> “This also involves significant and particularly skilled manpower.”<sup>12</sup> To carry out these processes “requires several years and significant investment and hence represents a huge barrier to entry.”<sup>13</sup> But that may be the optimistic view, as DG Comp further explains that, according to market participants, “the raw materials needed to create these databases are simply unavailable at any price.”<sup>14</sup>

The DOJ and DG Comp both undertook this sort of analysis separately for each of the data products as to which they brought a case. They did this because each of the products—while data—was different. The competitive dynamics at play in each market, while perhaps similar, were distinct. Also, of the scores of data products offered by Thomson and Reuters, most raised no competition concerns. As noted, DG Comp obtained commitments concerning only four products, and the DOJ’s consent decree addressed three.

Maybe the Thomson/Reuters matter was unusual, particularly in that the data products at issue were sold to users as such. Would a different result apply where data are a key input to another product? Tucker & Wellford suggest so, arguing, “the notion of a relevant market consisting of internally used data is inconsistent with longstanding precedent that recognizes a market only where a product or service is sold to consumers.”<sup>15</sup> But if internally used data are a key input rather than the product itself, surely a barrier to obtaining it can be a barrier to entry just the same. Suppose that Thomson and Reuters had not sold fundamentals data separately (as was Bloomberg’s practice), but that it was an essential part of a terminal offering: While the relevant markets alleged by DOJ and DG Comp would have differed, the ultimate result would have been no different.

Manne & Sperry also question how market power fortified by data could be remedied. The Thomson and Reuters matter provides one answer, though surely not one which necessarily will fit other markets. In an agreement reached jointly with DOJ and DG Comp, Thomson and Reuters agreed to a form of remedy which would not have been possible with a non-digital good:

[T]he parties committed to divest copies of the databases containing the content sets of such financial information products, together with relevant assets, personnel and customer base as appropriate to allow purchasers of the databases and assets to quickly establish themselves as a credible competitive force in the marketplace in competition with the merged entity, re-establishing the pre-merger rivalry in the respective fields.<sup>16</sup>

---

<sup>10</sup> DG Comp, *supra* note 7 at ¶361.

<sup>11</sup> *Id.* at ¶362.

<sup>12</sup> *Id.* at ¶363

<sup>13</sup> *Id.* at ¶364.

<sup>14</sup> *Id.* at ¶365

<sup>15</sup> Tucker & Wellford, *supra* note 2.

<sup>16</sup> DG Comp, *supra* note 7.

The combined Thomson Reuters continued to offer its predecessors' data products, but an entrant was enabled to go to market with a competing product based on a duplicate of the data. It is exactly the fact that data bases are not rivalrous that made this remedy possible.

Big data certainly can pose new and interesting problems for competition lawyers, but the notion that we categorically cannot identify barriers to entry or potential remedies in such product markets is not one of them.