



# Big Data: What does it really mean for competition policy?

A look into the emergence of Big Data, its fundamental importance to businesses and the wider economy, and the critical role of competition authorities in ensuring Big Data is not exploited

---

## Context

A recurrent theme in competition policy recently has been the implications of big data. While Big Data may be a key element of a firm's business strategy to compete more effectively in the market, it is also increasingly an important consideration for regulators and competition authorities. The opportunity it affords firms to develop a competitive edge are also the very same features through which competition concerns may arise. Big Data therefore brings both opportunities and risks.

On the one hand, Big Data might increase operative efficiency, mitigate information asymmetries and contribute to more efficient sales transactions. At the same time, these benefits may raise competition concerns by changing the dynamics in the market. In particular, they can contribute to existing or create new barriers to entry as well as give the dominant players sufficient advantage to hinder competition through anticompetitive conduct. Equally, while consumers and firms may benefit from greater transparency created by Big Data, such transparency could also facilitate collusive behaviour.

## What is Big Data?

Big Data is more of a buzzword than a well-defined concept. It is a term gaining great popularity, with the seemingly exponential growth in data volumes being generated and the ever increasing power of modern computers to meaningfully and quickly analyse such data. Today, every single click on a website we view, every purchase we make, and even every single step we take can, and typically is, being turned into a data point and analysed.

As the name suggests, Big Data is often comprised of large volumes of data, but the term can sometimes be used only in reference to the techniques employed to

analyse the data. Literature often refers to the key characteristics of Big Data as the three, or sometimes four, Vs, namely:<sup>1</sup>

- **high-volume** – the sheer amount of data available;
- **high-velocity** – the rate at which new data are generated and analysed;
- **high-variety** – the differences in types of data used and the increasing complexity of data analysis; and
- **high-variability** – the different interpretations of data analysis and the extent to which data is consolidated, cleaned and consistent.

With the myriad of ways in which it can be generated and collected, typologies have emerged to rationalise the Big Data universe. A common classification of data is by whether they are: **internal** (i.e. user generated), **structured** (i.e. generated by a third-party but in a machine-readable format) or **unstructured** (i.e. not currently in a machine-readable format). Another popular delineation is between 'transactional' data which is typically structured, e.g. sales, and non-transactional data, with the latter further broken down into machine data, e.g. GPS results, and social data, e.g. data generated from Facebook or Twitter activity.

## The good, the bad, and the ugly of Big Data

Big Data is widely seen as of great benefit to the modern economy, and to both the businesses and consumers who operate in that economy. The Centre for Economics and Business Research (Cebr), for example, estimated that Big Data would, on average, contribute £36billion annually between 2012 and 2017.<sup>2</sup>

---

<sup>1</sup> See: Kobiulus (2013), "Measuring the business value of Big Data", IBM Big Data Hub Blog Article.

<sup>2</sup> See: Cebr (2012), "Data equity: Unlocking the value of big data". Report for SAS, April 2012.

It is an extremely valuable tool for modern businesses, with those failing to recognise and capture its benefits risking being left behind. Big Data can add value to almost any type of business and at almost any stage of the value chain. Businesses can use Big Data in a number of ways:

- **Social analytics** – e.g. tracking success of a firm's advertising campaign by social media exposure.
- **Decision science** – e.g. analysing consumer reviews to support future product development.
- **Performance management** – e.g. assessing most profitable customer segments and geographies.
- **Data exploration** – e.g. working out where to target advertising or up-selling efforts.

Such data-driven insights are an increasingly critical competitive differentiator, allowing firms to optimise business decisions at all levels of the value chain. It is clear that Big Data presents businesses with a whole host of welfare-enhancing opportunities – increased efficiency of production processes, enhanced responsiveness to customer sentiment and improved risk assessment and management to name but a few.

However, aside from important questions about personal data protection and privacy, Big Data raises questions of how it may be harnessed to the detriment of market competition.

### What does Big Data mean for competition?

It is, at first, important to note that Big Data may in many ways be conducive to greater competition in the market place. Specifically, by making information more readily available it can: lower search costs and increase propensity to switch; diminish barriers to prospective entrants; and even create new channels for entry and expansion.

That said, Big Data is also presenting new challenges to competitive markets. While it is difficult to cover all these issues in one article, we present below some of the key concerns that have so far arisen in this debate.

**Data network effects** – Big Data can drive network effects, raising barriers to entry and enabling the large incumbent firms to consolidate their position in the market. Data network effects explain the phenomenon whereby your product (generally through some form of machine-learning) becomes smarter, and thus more attractive to consumers, the more data it gets from consumers. The more consumers use a given product,

the more data they contribute and the 'smarter' the product becomes; this in turn attracts more consumers, who contribute new data, further increasing the products' performance, and so the cycle continues. These data network effects can combine with more traditional network effects, increasingly concentrating data and thus market power in the hands of fewer firms.

**Platforms and market foreclosure** – Big Data can also diminish competitive pressures, by denying potential competitor firms access to this data: either through an outright refusal to supply, or a constructive refusal to supply (e.g. by selling data to competitors significantly above the competitive rate). This is of particular concern where: (a) few substitutes of the data are available; and (b) the data are of significant value to the production process, such that absence of the data results in an inferior product or service offering. This effect may be exacerbated by the aforementioned data network effects, which sees Big Data increasingly concentrated in a small number of 'super-platforms' – think Google, Amazon and Facebook.

**Collusion** – by increasing the speed at which price changes are observed, and thus the ability to detect and punish deviations, Big Data may also facilitate greater collusive practices. This is of particular concern as collusion in a Big Data context is likely to take more tacit forms. Ezrachi and Stucke (2015)<sup>3</sup> consider how algorithmic models built around Big Data can promote greater collusion: firstly, by omitting human biases from the strategy, algorithmic modelling of Big Data should create the stability necessary for tacit collusion; and, secondly, by knowing that rival firms operating similar models can capture and respond to competitive price changes very quickly, this diminishes the incentive to undertake such strategies.

**Mergers** – we are starting to see the emergence of data-driven mergers, with firms merging (at least in part) to take advantage of complementarities in the data they collect. This may be in part reflect an attempt to capture the data network effects we described earlier. The interesting thing is that these mergers often do not fit into the traditional classification of horizontal or vertical mergers. Perhaps of more concern is the fact that the assessment of mergers traditionally looks at the impact on prices, but in data-

---

<sup>3</sup> Ezrachi, A. and M. E. Stucke (2015), "Artificial Intelligence & Collusion: When Computers Inhibit Competition". Oxford Legal Studies Research Paper No. 18/2015.

driven markets the products and services are often free and, as such, do not lend themselves to this analysis.

**Behavioural discrimination** – it has been suggested that the growth of Big Data (and Big Data analytics) has led to a move away from traditional price discrimination models to models of behavioural discrimination. Big Data is allowing firms to segment consumers into ever smaller groups to better identify their reservation prices and ultimately extract greater consumer surplus. The more data that is collected, the more personalised this experience is likely to become, perhaps making it increasingly difficult to assess a ‘general’ market price and thereby assess the attractiveness of outside options. Such behavioural discrimination by incumbent firms may also make it more difficult for potential entrants to compete, given that they lack sufficient scale or breadth of data.

### Recent cases involving Big Data

Big Data is an emerging field and, as such, existing precedent of Big Data issues in competition cases is expectedly light. That said, there is certainly growing regulatory interest in this field. To date, merger control cases are the key area in which Big Data issues have been investigated.

In the context of merger control, the European Commission (EC) investigated Google’s acquisition of DoubleClick in 2008<sup>4</sup> and, in particular, how important access to DoubleClick’s data was for firms to compete. No concerns were found in this respect and the merger was allowed to go ahead. Microsoft’s acquisition of Yahoo! search<sup>5</sup> was also cleared by the Commission, with the merger in fact seen as beneficial to competition, by increasing Microsoft’s scale and thus ability to compete with Google.

More recently, the EC investigated Facebook’s acquisition of WhatsApp<sup>6</sup> and, in particular, whether Facebook could use the additional user data from WhatsApp to the detriment of competition. The EC again concluded that this additional user data was not sufficient to hamper competition, as there a sufficient number of other companies in the market with their own databases of user data that can compete for the provision of targeted advertising.

The current AT&T and Time Warner merger in the USA looks set to be another big test of how competition authorities analyse and treat the risks posed by big data in merger analysis.

Outside the area of merger control, there only appear to be have been two abuse of dominance cases related to Big Data (while there is no evidence of any collective dominance cases to date):

- In 2012, the EC imposed binding commitments on Thomson Reuters to create a new licence allowing customers to use Reuters Instrument Codes (RICs) for data sourced from Thomson Reuters’ competitors.<sup>7</sup> This was based on the EC’s concern that Thomson Reuters could be abusing its dominant position in the market for consolidated real-time data feeds through its licensing practices. In effect, the existing licensing agreements were seen as an attempt to restrict access to an essential facility and thus foreclose competition in the market for real-time data feeds.
- In 2016, the German Federal Cartel Office (FCO) opened proceedings against Facebook for an alleged abuse of dominance. The concern is that Facebook is abusing its dominant position in social networking by imposing privacy terms and conditions that would otherwise not be accepted by its users. This investigation is ongoing.

### Where do the authorities go from here?

Despite a limited number of Big Data cases, we are nevertheless witnessing a large increase in new initiatives by European competition bodies in understanding how to apply competition law rules to users of Big Data. National antitrust enforcers are also testing the applicability of existing competition law tools to Big Data issues, particularly with regard to potential abuse of dominance cases.

The EC is also currently consulting on changes to referral thresholds for merger cases, in order to close the perceived enforcement gap. Specifically, they are considering the introduction of a ‘deal-size’ threshold to capture significant M&A deals which do not meet the current turnover based thresholds. The Facebook WhatsApp merger referred to above did itself fall below the current turnover-based threshold. The deal-

<sup>4</sup> EC Case No COMP/M.4731 – Google/ DoubleClick.

<sup>5</sup> EC Case No COMP/M.5727 – Microsoft/Yahoo! Search.

<sup>6</sup> EC Case No COMP/M.7217 – Facebook/WhatsApp.

<sup>7</sup> EC Case No COMP/D2/39.654 – Reuters Instrument Codes (RICs).

size threshold, therefore, could be triggered by a merger with data volumes exceeding a certain level.

At a national level, the UK's Competition and Markets Authority (CMA) published a report in 2015 in which they highlighted the anticompetitive effects arising from Big Data.<sup>8</sup> Furthermore, the UK's Financial Conduct Authority (FCA) is looking at the potential implications of Big Data for competition in retail insurance products and the possible implications for end consumers. However, in September of last year, the FCA confirmed that insurers will not face a market study into their use of Big Data.

Elsewhere, the German and French competition authorities collaboratively published a report on Competition Law and Data (with the French Competition Authority also launching a general inquiry into data-related markets and strategies). The joint report looks at the implications and challenges presented by data collection in the digital economy and wider industries, and in particular at firms looking to protect their already established data advantage. In particular, the report proposes that future cases could be based on the logic that abuse of dominance can arise from a firm's ability to derive market power from Big Data that a competitor is unable to match. They raise two specific questions to consider in this regard:

- Is there a scarcity of data or is it possible for competitors to easily obtain or replicate this data?
- Does the scale and scope of the relevant data matter for the assessment of market power?

The report calls for a case-by-case assessment of competition law risks arising from companies with a significant data advantage.

So far it seems that Big Data does not challenge the fundamentals of our existing competition frameworks. It simply represents an additional factor to consider, for example when defining the relevant markets or assessing the market power of a firm. As with any new development in how businesses operate and what businesses can do, test cases (such as the AT&T Time Warner merger) will help all of us in the competition community find our way.

If you have any questions about this topic please contact:



Iona McCall  
Managing Consultant  
[iona.mccall@europe-economics.com](mailto:iona.mccall@europe-economics.com)  
Direct dial: +44 207 269 2639



Sam Winward  
Consultant  
[sam.winward@europe-economics.com](mailto:sam.winward@europe-economics.com)  
Direct dial: +44 207 269 2654

---

<sup>8</sup> See: CMA (2015), "The commercial use of consumer data: Report on the CMA's call for information". June 2015.