

**DIRECTORATE FOR FINANCIAL AND ENTERPRISE AFFAIRS  
COMPETITION COMMITTEE**

**BIG DATA: BRINGING COMPETITION POLICY TO THE DIGITAL ERA**

-- Background note by the Secretariat --

**29-30 November 2016**

*This document was prepared by the OECD Secretariat to serve as a background note for Item 3 at the 126th Meeting of the Competition Committee on 29-30 November 2016.*

*The opinion expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.*

*More documentation related to this discussion can be found at [www.oecd.org/daf/competition/big-data-bringing-competition-policy-to-the-digital-era.htm](http://www.oecd.org/daf/competition/big-data-bringing-competition-policy-to-the-digital-era.htm)*

Please contact Mr. Antonio Capobianco if you have any questions regarding this document [E-mail: [Antonio.Capobianco@oecd.org](mailto:Antonio.Capobianco@oecd.org)].

**JT03401771**

Complete document available on OLIS in its original format

*This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.*



## **BIG DATA: BRINGING COMPETITION POLICY TO THE DIGITAL ERA**

*Background note by the Secretariat\**

### **Abstract**

Business models based on the vast collection and process of user data in nearly real-time in recent years have enabled companies to offer a wide range of innovative and customised services, often at zero prices, with substantial gains for consumers. At the same time, data-driven network effects reinforced by user feedback loops, and high economies of scale associated with information technology infrastructures, may provide companies that own the data with market power and create a tendency for markets to tip. Concern is rising that the increasing reliance and use of personal data is harmful to consumers. While some practitioners have proposed adapting competition tools and antitrust policy to tackle such issues, others believe that these can be better addressed by data and/or consumer protection agencies. This issues paper attempts to define Big Data and its role within a competition context, and then identifies some of the potential implications for the enforcement of competition law in the areas of merger review, abusive of dominance and cartels. It also discusses how regulations on data ownership, access and portability may affect consumer protection and competitive neutrality.

---

\* This background paper was written by Ania Thiemann and Pedro Gonzaga of the OECD Competition Division, with the kind support of Professor Maurice E. Stucke, Professor of Law, University of Tennessee, College of Law.

## TABLE OF CONTENTS

1.	Introduction.....	5
2.	Big Data: what is it and why should we care .....	5
2.1.	Definition of Big Data .....	5
2.2.	The potential gains from data-driven innovation.....	7
2.3.	Competition challenges posed by Big Data.....	10
2.4.	The Big Data ecosystem.....	12
2.4.1.	Platform technology .....	13
2.4.2.	Content providers .....	13
2.4.3.	Sellers .....	14
2.4.4.	Infrastructure, including cloud computing and storage .....	14
2.4.5.	The public sector .....	14
3.	Implications of Big Data for competition law enforcement .....	15
3.1.	Competition tools .....	15
3.1.1.	Identifying the relevant market for antitrust purposes.....	15
3.1.2.	Assessing market power .....	17
3.2.	Merger review .....	18
3.2.1.	Bringing a privacy dimension to merger review .....	18
3.2.2.	Notification thresholds in Big Data mergers .....	20
3.3.	Abuse of dominance .....	21
3.3.1.	Exclusionary conduct .....	21
3.3.3.	Data as an essential input and the essential facilities doctrine .....	22
3.4.	Collusive practices.....	22
3.4.1.	The advent of digital cartels .....	23
3.4.2.	Big data antitrust.....	24
4.	Pro-competitive impact of Big Data regulation .....	24
4.1.	Consumer protection.....	25
4.1.1.	Inappropriate use of the word ‘free’ .....	25
4.1.2.	Ownership rights over data and privacy standards .....	26
4.1.3.	Rights on data portability .....	27
4.2.	Competitive neutrality .....	28
4.2.1.	Access to public sector information .....	28
4.2.2.	Leveraging private data for public goals .....	28
5.	Conclusion .....	29
	Endnotes .....	30
	Bibliography .....	33

**Boxes**

Box 1. Now-casting or contemporaneous forecasting .....8  
Box 2. Walmart.....9  
Box 3. Google .....9  
Box 4. Market definition in multi-sided platforms .....16  
Box 5. Assessing the market power of platforms .....17  
Box 6. The Facebook/WhatsApp merger.....19  
Box 7. The US Airline Case .....23

**Figure**

Figure 1. Feedback Loops ..... 10  
Figure 2. The Business Learning Curve..... 11  
Figure 3. Big Data Ecosystem..... 12  
Figure 4. Example of Privacy Standards ..... 27

## 1. Introduction

1. The exponential growth of computing power and the expansion of Internet access globally in recent years have spurred the advent of the digital economy and enabled the rise of business models based on the collection and processing of “Big Data”. This concept, originally used by computer scientists and increasingly popularised among academics, regulators and politicians, is now widespread across multiple disciplines.

2. Collecting, processing and exploiting personal data for commercial use is seen by many observers as a question of consumer protection rather than one of competition law enforcement. However, recent high-profile mergers and acquisitions in digital or Internet markets have raised the question of a possible competition impact of bringing together and gaining control over large data sets, as well as a desire to better understand the possible implications for consumers and markets.

3. This paper on “Big Data” represents the first step in a broader work stream of the OECD Competition Committee on Competition, Digital Economy and Innovation.<sup>1</sup> In November 2016, the OECD will be holding a hearing discussion on Big Data to explore the implications on competition authorities' work and whether competition law is the appropriate tool for dealing with issues arising from the use Big Data. As such this background paper aims to take a broad view of the topic and to present as many angles as possible in order to identify and discuss the most important channels through which Big Data may affect competition policy and competition law enforcement.

4. The paper is structured as follows. In section 2 we start by defining Big Data, identifying the possible competition issues related to the use of Big Data and describing the main actors and the market topology of the ‘Big Data ecosystem’. In section 3, we discuss how Big Data may affect the enforcement of competition laws by looking in turn at the main tools available to competition authorities, and the issues thrown up by Big Data in addressing merger review, abusive conduct and collusion. Section 4 addresses regulatory concerns and consumer protection issues. It also discusses the role of the public sector and competitive neutrality. Section 5 concludes.

## 2. Big Data: what is it and why should we care

### 2.1. Definition of Big Data

5. While the use of the term ‘Big Data’ is often vague and lacks precision (De Mauro et al, 2016), the most frequently used definitions of Big Data<sup>2</sup> usually refer to (1) the large dimension of datasets; and (2) the need to use large scale computing power and non-standard software and methods to extract *value* from the data in a reasonable amount of time. According to De Mauro et al (2016), “*Big Data is the information asset characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value.*”

6. To distinguish ‘Big Data’ from data in general,<sup>3</sup> we propose to follow Stucke and Grunes (2016). To the “3 Vs” definition originally introduced by Laney (2001), i.e. the *volume* of data; the *velocity* at which data is collected, used and disseminated; the *variety* of information aggregated; Stucke and Grunes add a fourth V: the *value* of the data. As the authors point out for personal data, each ‘V’ has increased enormously in magnitude over the past decade, and indeed continues to expand. We will look at each of these in turn.

7. The *volume* of data processed globally is forecast to continue to expand almost exponentially. Cisco, a US tech firm, forecasts that annual global data centre IP traffic will reach 10.4 zettabytes<sup>4</sup> by the end of 2019, up from 3.4 zettabytes (ZB) per year in 2014, a compound annual growth rate (CAGR) of 25% from 2014 to 2019.<sup>5</sup> The volume is a consequence of the ubiquity of online and Internet activity.

OECD (2015a) highlights the fact that as nearly all media as well as social and economic activities migrate to the Internet (including e-commerce and electronic government services); this generates petabytes of data every second. The increase in the *volume* of data has been facilitated by the realisation of Moore's Law<sup>6</sup> which has brought about ever more powerful, smaller, smarter and cheaper devices that are accessible to almost every individual across the globe. This in turn has led to the decrease in the cost of collecting, storing, processing and analysing data, while the access to data has been enormously facilitated by the rise of Internet platforms, e-commerce and the spread of smartphones.

8. Stucke and Grunes point out that the *velocity* at which some firms access, process and analyse data now approaches real time. Rising out of the ability to use data in real time is the phenomenon known as 'now-casting' (Box 1). This concept consists of taking an event occurring now and using it to forecast things as they happen, such as an outbreak of a 'flu' epidemic being notified owing to a sharp uptake in online searches for 'flu' remedies. However, this can also be used to spot a potential competitor, for instance by noticing the number of downloads of an application from an app-store, and cross-reference it with online usage or search preferences. Potentially, the ability to "now-cast" can give an incumbent leverage over new entrants.

9. This leads to another difference between Big Data and traditional data: the time-value. Being able to process vast amounts of data in real time has intrinsic value, and is more valuable in some cases than having data with a lag, for instance in accessing traffic information in a road-map application.

10. The *variety* of data has also increased, thanks to the ability to collect and process it, enabling companies to know not just customers' address (whether physical or IP), birthday and gender, but also multiple other bits of information, such as household composition, dietary habits, purchasing history, frequency and duration of visits to physical and online stores, as well as information from other databases to flesh out the customer's profile.<sup>7</sup> Not only does this enable the retailer to price discriminate,<sup>8</sup> but also to target customers with marketing and behavioural advertising.

11. The *Autorité de la Concurrence* and the *Bundeskartellamt* (2016) emphasise that the change in consumer habits towards an ever-greater use of the Internet for everything from shopping to reading the news, watching films and posting videos of themselves enables companies to "*record actions [of a large part of the population] in such a precise way that detailed and individualised conclusions on their receptiveness to specific sales messages can be drawn*" (p.11). The flashing ads from Steven Spielberg's 2002 blockbuster, "*Minority Report*", where iris scans in public spaces identify an individual and then flashes a personalised advertising at them, no longer belong purely to science fiction.

12. What this example demonstrates is the importance of data-fusion – when large sets of data are fused and then mined, they bring together new information that may enable a seller, or a competitor, to better understand and exploit the market. Sometimes, the potential of data fusion can be further exploited by combining personal data with many other types of data, such as weather conditions, public events, inventories or even data collected on machine components to detect wear and tear.

13. The final 'V', the *value* of Big Data, is both a cause and a consequence of the increase in *volume*, *variety* and *velocity*. While data in itself may be considered to be 'free' – depending on how it is collected – the process whereby information is extracted from the data generates the *value*. L'Autorité de la Concurrence and the *Bundeskartellamt* (2016) concur and point to the "*development of new methods capable of extracting valuable information from extremely large accumulations of (often unstructured) data*" (p.8). OECD (2015a) defines data analytics as "*the technical means to extract insights and the empowering tools to better understand, influence or control the data objects of these insights (e.g. natural phenomena, social systems, individuals)*".

14. Stucke and Grunes (2016) point out that Big Data is closely related to what may be called ‘Big Analytics’ and to the phenomenon known as ‘deep learning’, whereby computers teach themselves to improve in solving complex problems by crunching large datasets using sophisticated algorithms and neural networks that increasingly resemble the human brain. One such example is the Rubicon Project, a vast online platform that automates the buying and selling of advertising:

*“Relentless in its efforts for innovation, Rubicon Project has engineered one of the largest real-time cloud and Big Data computing systems, processing trillions of transactions within milliseconds each month.”<sup>9</sup> As the company added in a statement, “as we process more volume on our automated platform, we accumulate more data, such as pricing, geographic and preference information, data on how best to optimize yield for sellers and more. This additional data helps make our machine-learning algorithms more intelligent and this leads to more effective matching between buyers and sellers. As a result, more buyers and sellers are attracted to our platform, from which we get more data, which further reinforces the network effect...”<sup>10</sup>*

15. This quote highlights not only how the *value* of Big Data is derived (using the other three ‘V’s), but also the importance of online platforms and network effects in the ‘Big Data ecosystem’ which we discuss in section 2.4. Deep or machine-learning are key elements of this process. Stucke and Grunes (2016) correctly point out that the *volume* and *variety* of data can enable firms at times to uncover correlations from large, unstructured data bases, and these can outperform the findings from cleaner, but smaller datasets. It is not just a question of having a performing algorithm, but also the fact that seemingly disparate databases can be fused and mined to extract information that would otherwise not be tangible or useful; such as an insurance company uncovering the (risk-taking) proclivities of a client.

## **2.2. *The potential gains from data-driven innovation***

16. Being able to harness Big Data can lead to important and positive gains for a business, which in turn may benefit consumers, employees, and society in general. Indeed, the use of Big Data for innovative and creative purposes, in a process known as *data-driven innovation* (DDI), allows companies to improve the quality of their products and develop entirely new services, by better understanding and targeting individual consumer needs. Using machine learning algorithms, for instance, clearly improves the usefulness of web search engines: more searches, together with the ability to observe and record which choices the user clicks on, will help to *“improve and refine the search engine, as well as the implementation of its supporting algorithm”* (Autorité de la Concurrence and Bundeskartellamt, 2016).

### Box 1. Now-casting or contemporaneous forecasting

Now-casting is defined by Banbura et al (2013) as “*the prediction of the present, the very near future and the very recent past*”. It consists in the use of new, up-to-date and high-frequency data to produce early estimates, usually with great degree of accuracy, about events that are taking place very close to the present. Now-casting is particularly useful to obtain close to real-time information about relevant variables that are normally collected at low-frequency and published with a great lag. For instance, Banbura et al (2010) illustrate how a now-casting statistical model can make use of industrial production, which is published monthly with short delay, to construct a precise early-estimation of the euro area GDP, which is reported at a quarterly frequency and usually with a six-week lag.

The concept now-casting has been popularised for long among meteorologists, who use a combination of the latest radar, satellite and observation data to describe with good precision the present state of the weather and the expectable evolution for the following hours. The development of now-casting methods in the context of meteorology has contributed to reduce the number of fatalities and property damage caused by weather hazards, as well as to improve safety and efficiency in some industries, including aviation, water and power management, and construction.

More recently, the increasing number of businesses leveraging Big Data has allowed the expansion of now-casting to multiple economic applications. As an example, the online real estate marketplace Auction.com has launched a real estate now-cast that reports real-time information about home sales in the US and which will be used in the future to calculate ongoing pricing trends and other events. For that, Auction.com is implementing data models developed by Google’s chief economist Hal Varian, using a combination of private industrial datasets and publicly available Google search data. Google data has also been used to predict real-time market trends in other sectors, such as auto sales, retail and travel. As Hal Varian states in an overview video about Auction.com Real Estate now-cast:

*“Now-casts are a major shift in how predictive market analysis is done today. Most government and industry reports on things like housing, employment and consumer sentiment are published weeks or even months after those sales or actions are realised. For market participants trying to determine where the market is today and where is going, this is a bit like driving forward while looking in the rear-view mirror. (...) I believe real estate investors, banks, financial institutions, government agencies and others will look very closely to now-cast like this for more timely and accurate forecast.”<sup>11</sup>*

17. Using Big Data is also useful for businesses to generally improve the efficiency of production processes, forecast market trends, improve decision-making and enhance consumer segmentation, through target advertising and personalised recommendations. Although the efficiency gains from data driven-innovation are inherently hard to measure, some studies suggest that DDI users benefit, on average, from a 5% to 10% faster productivity growth than similar companies that do not use DDI.<sup>12</sup> Buchholtz et al (2014) also estimate that the multiple applications of Big Data will allow the EU economy to grow by an additional 1.9% by 2020.

18. In addition to the overall productivity gains, the exploitation of Big Data can generate other substantial social benefits that are usually not accounted for by standard measures. OECD (2013b) estimates that, in the transport sector, the tracking of mobile devices to reduce traffic congestion could provide time and fuel savings of up to USD500 billion worldwide by 2020; in the electricity sector, the adoption of smart grid applications to control the operation of household appliances, send feedback to consumers about energy consumption and adjust production capacity to demand forecasts, could reduce the cost of CO2 emissions by EUR79 billion by 2020; and, in the US health-care sector, the creation of electronic health records could reduce medical errors, improve diagnosis, increase efficiency in management and pricing, foster R&D and achieve other goals that would allow savings of about USD300 billion by 2020.

19. At the micro level, the impact of Big Data can more easily be illustrated by the business success of disruptive companies that collect vast amounts of data from consumers to offer data-driven services, such as the US chain-store Walmart (see Box 2), the high technological company Google (Box 3), the UK-owned supermarket chain Tesco or the US transportation network company Uber, to list just a few. At the same time that these companies become more efficient and profitable, consumers benefit from a variety of innovative services that provide greater convenience, customisation and, sometimes, significantly lower prices.



### Box 2. Walmart

Walmart, a US retail company with a network of hypermarkets, discount stores and groceries stores, is currently the largest corporation in the world by consolidated revenue. The business strategy of Walmart involves the practice of low profit-margins to obtain high scalability. It also uses Big Data to improve operational efficiency. Walmart collects around 2.5 petabytes of data per hour and is estimated to have increased online sales by 10% to 15% as a result of data analytics (Dezyre, 2015).

Walmart collects consumer data about historical purchases, living location, clickable actions / keywords entered in the website, as well as information from social networks. Then, using data mining, it analyses the pattern of consumer data and crosses it with information about other events (such as sports, weather...), in order to improve predictive analysis, launch new products and provide personalised recommendations.

There are several creative ways through which Walmart leverages Big Data. To name just a few, the company uses demand estimation to improve inventory management and shipping policy; it launches entrepreneurship contests in social media in order to place the most popular products in the shelves; and sends recommendations to consumers of gifts for friends based on their Facebook profiles.

*“A familiar example of effective data mining through association rule learning technique at Walmart is finding that Strawberry pop-tarts sales increased by 7 times before a Hurricane. After Walmart identified this association between Hurricane and Strawberry pop-tarts through data mining, it places all the Strawberry pop-tarts at the checkouts before a hurricane.” (Dezyre, 2015).*

### Box 3. Google

Google, as one of the biggest companies in the world by market capitalisation (with its parent company Alphabet Inc.), has attracted a great deal of attention from antitrust agencies and is at the centre of the discussion on Big Data regulation. In the search engine market, Google is currently the company with the highest share of search queries, having also a substantial share in the search advertising market, in the EU and US.<sup>13</sup>

The market position of Google does not necessarily imply competitive harm, as it is relatively consensual that the enormous success of the high-technology company has been largely achieved through means of innovation (for instance with the revolutionary search algorithms introduced) and by hiring highly talented individuals. Google has also brought extensive benefits to the consumer, by providing a wide variety of high quality services for which it does not charge any monetary price. This includes not only its multiple searching tools, but also e-mail services, online translators, detailed maps for navigation and the Android operating system, among many others. Indeed, Bork and Sidak (2012), as well as Manne and Wright (2011) strongly defend Google’s business strategy based on continuous investment in efficiency and quality as pro-competitive, arguing further that antitrust actions against Google would only punish a successful competitor.

Still, academic researchers such as Newman (2013), propose a theory of harm according to which Google is aggressively expanding its business to new product sectors in order to improve its user data and foreclose the market to other companies, thus reinforcing its position in search-based advertising. While during the acquisition of DoubleClick by Google the FTC<sup>14</sup> claimed that user data is not an essential input to compete in the search advertising market, some commentators question, particularly given market developments, whether the claim is compelling.<sup>15</sup>

The fact that Google has been estimated to charge a higher cost-per-click (CPC) than Bing, its main competitor,<sup>16</sup> suggests that advertisers attribute a higher probability of converting a viewer of Google’s ads into a customer. Since advertisers can publish exactly the same content on both platforms, Google’s premium may partially come from the associated benefits of the network effects and its resulting ability to use data to better target potential consumers with behavioural ads.

Google’s vast and ongoing investments to continuously develop new products that are offered to users at a zero price also reflect the perceived value of data. By combining all the data collected through Android and other products, and using its own algorithms as well as machine-learning programmes, Google is able to enhance its detailed user profiles with information that no other competitor has and which should be valuable enough to recover the money invested.<sup>17</sup> Moreover, the importance of controlling user data to compete in this market has become apparent in one of Google’s acquisitions:

*“Indeed, as the Office of Fair Trading (“OFT”) later found in a merger between Google and Waze, it was Waze’s inability to achieve sufficient scale of data that hindered its competitive significance in mapping services in the United Kingdom. The OFT agreed that the more users supplied Waze with data on traffic*

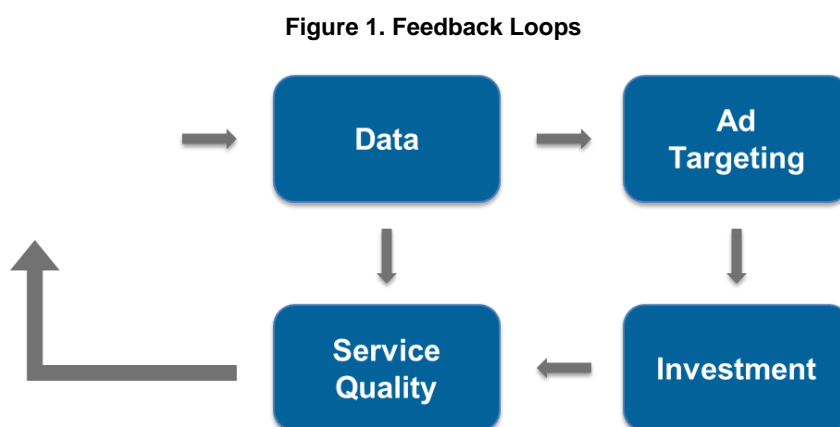
*conditions, the better Waze’s turn-by-turn application became, and the more likely Waze would attract additional users. But this presented a chicken-and-egg dilemma. Users would not be attracted to mapping sites unless the quality was good, and the quality won’t be good absent a sufficient amount of data from users.”<sup>18</sup>*

**2.3. Competition challenges posed by Big Data**

20. As the acquisition and use of Big Data becomes a key parameter of competition, companies will increasingly undertake strategies to obtain and sustain a data advantage. As argued by Stucke and Ezrachi (2016, p. 30), “Companies are increasingly adopting business models that rely on personal data as a key input. (...) companies offer individuals free services with the aim of acquiring valuable personal data to assist advertisers to better target them with behavioral advertising.” While the competitive rivalry and drive to maintain a data advantage can be pro-competitive, yielding innovations that benefit consumers and the company, some competition authorities emphasise that network effects and economies of scale driven by Big Data can also confer market power and a durable competitive advantage.<sup>19</sup>

21. One question is whether the use of ‘Big Data’ poses a different problem from the use of plain or traditional data. Indeed, the corner store has thrived by knowing its customers well. Traditional salesmen have always engaged in close relations with their clients to learn about their preferences and offer a customised product. Likewise, manufacturers use historical data to estimate demand and improve their products in highly competitive industries. Can we therefore assert that Big Data raises a new competition challenge that was not observed before?

22. Unlike the brick-and-mortar retail economy, modern business models are frequently characterised by data-driven network effects that can improve the quality of the product or service. These data-driven network effects are the result of the two user feedback loops depicted in Figure 1. On the one hand, a company with a large base of users is able to collect more data to improve the quality of the service (for instance, by creating better algorithms) and, this way, to acquire new users – ‘user feedback loop’. On the other hand, companies are able to explore user data to improve ad targeting and monetise their services, obtaining additional funds to invest in the quality of the service and attracting again more users – ‘monetisation feedback loop’. These interminable loops can make it very difficult for any entrant to compete against an incumbent with a large base of customers.



23. As an illustration, if a search engine only has one thousand daily queries, its algorithms have less data to learn responsive search results (other than more straightforward inquiries) and fewer related searches that it can suggest to users. With poorer quality search results, it will be unlikely to attract many users from the larger search engines; with fewer users, the search engine will attract fewer advertisers,

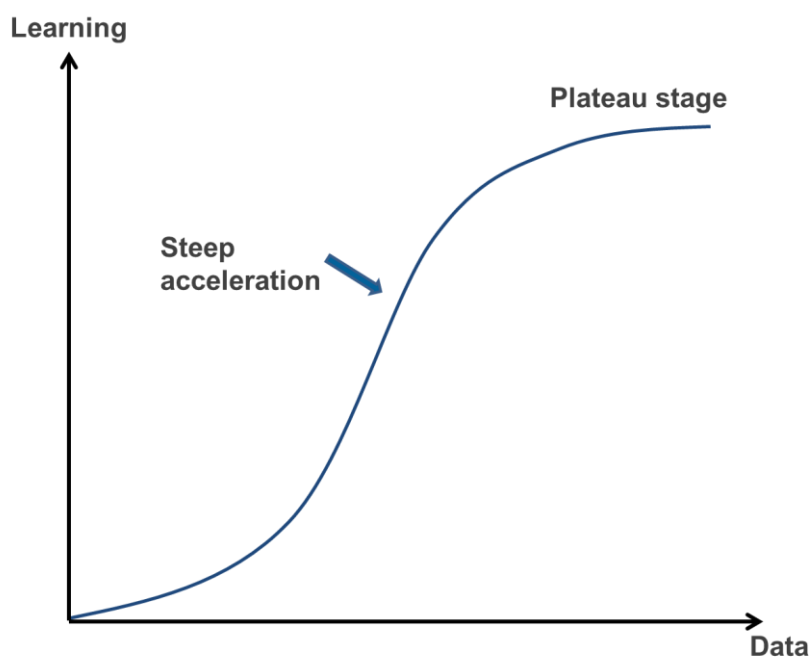
which means fewer occasions for users to click on paid search results and less advertising revenue to expand the platform to other services.

24. With each user a company acquires relative to its competitors, a quality gap may emerge. If the quality differences become apparent to users, the feedback loop can accelerate – attracting both new users and users of the competitors’ products. In markets with data-driven network effects, such as search engines, social networks, and community-sourced navigation apps, the winner not only gains potential revenue, for example, when the user clicks on sponsored ads; that user’s data also helps improve the quality of the product itself, which affects the product’s attractiveness to future users and advertisers. Some of these data-driven network effects may eventually taper off. But the data-driven network effects in these online markets can amplify the processes of gaining and losing users.

25. As a result of such data-driven network effects, users may become reliant on the dominant platform even though they prefer a different platform model. For instance, while online users may prefer the privacy options promised by some search engines, the larger search engines provide better targeted results. Another example is a turn-by-turn navigation app, where a smaller app might offer better features, but one reluctantly uses the dominant app which has better traffic information provided by its many users. The dominant platform may not do anything that can be properly qualified as anticompetitive, and yet the feedback loop can reinforce dominance and prevent rival platforms from gaining customers.

26. Another difference between modern applications of Big Data and traditional business models is the lack of physical bounds to the quantity and variety of data that can be collected in a digital world and the unlimited knowledge that can be obtained by running data mining algorithms on a variety of datasets, or using data-fusion. As a result, Big Data has shifted the slope of the business learning curve (Figure 2), allowing the steep acceleration phase of the Big Data incumbent to last longer and making the increasing returns on data harder to exhaust. When a Big Data player finally reaches the plateau stage, its dimension is so big that may be very difficult for any small player to effectively exert competitive pressure, creating a potential for market ‘tipping’ and winner-takes-all outcomes.

**Figure 2. The Business Learning Curve**



27. Competition concerns may also result from the fact that the cost structure of treating and using information is rather unusual, involving high up-front sunk costs and close-to-zero marginal costs.<sup>20</sup> This is particularly true in the case of Big Data, where the information technologies required to store and process the data can be very costly, involving vast data centres, servers, data-analytical software, internet connections with advanced firewalls and expensive human resources, such as computer scientists and programmers. Once the system is fully operational, the incremental data can ‘train’ and improve the algorithms at a low cost (thereby also the product or service quality). This cost structure is characterised by high economies of scale and scope and can therefore facilitate market concentration of Big Data in the hands of a few players.

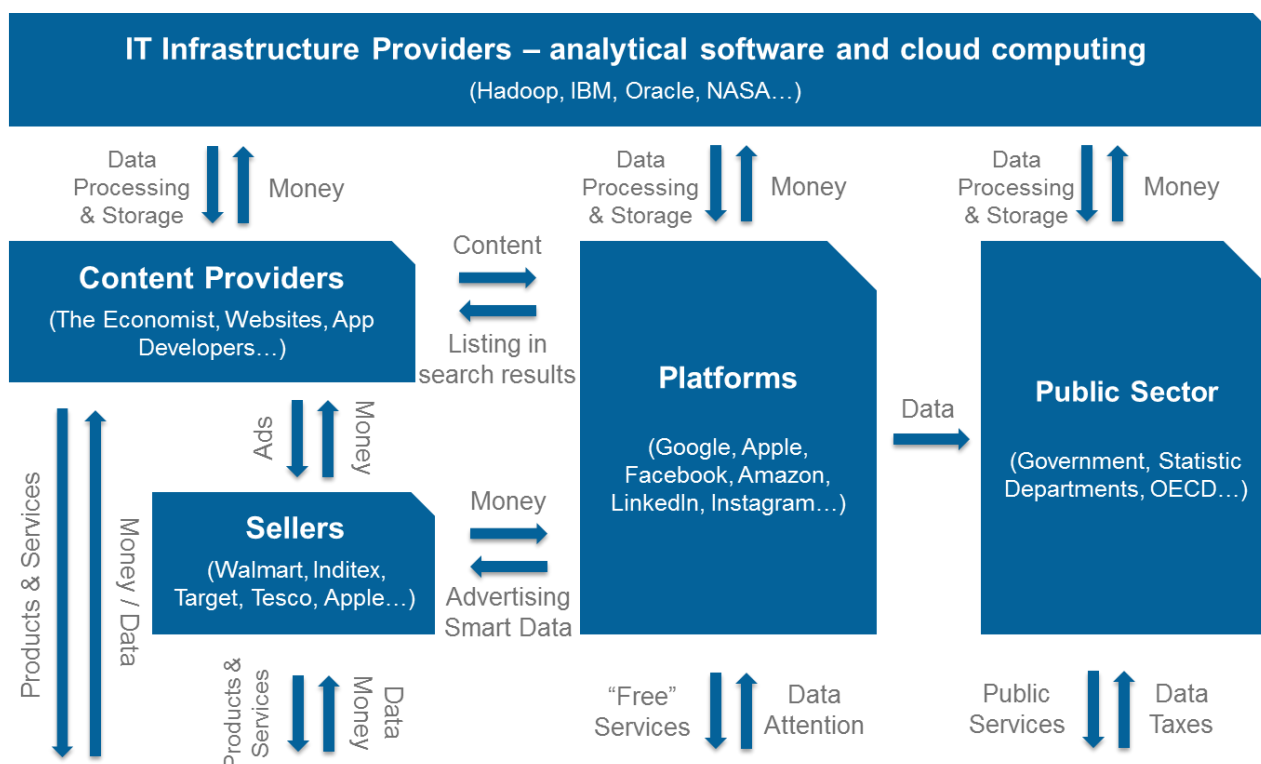
28. Moreover, unlike “small data”, whose units of information provide meaningful and valuable insights comprehensible for a human being, the value of an individual Big Data observation is rather small. For instance, data about a single website click is useless unless it is collated with billion other similar actions, which must be then related to actual purchasing decisions. As a result, Big Data-sets need a certain scale to be profitable and are more frequently collected by large players.

29. Finally, other competition problems may arise due to the particular structure of the markets where Big Data is typically transacted. To better understand the operating environment, the next section takes a close look at the ‘Big Data Ecosystem’.

### 2.4. The Big Data ecosystem

30. Big Data is collected, transacted and converted to money value in a complex ecosystem<sup>21</sup> composed of multiple interconnected markets, many of which are multi-sided. This section gives a brief descriptive overview of the main types of business and agents involved, as represented in Figure 3.

Figure 3. Big Data Ecosystem



#### 2.4.1. Platform technology

31. At the epicentre of the Big Data ecosystem, where many of the competition concerns discussed above are observed, platforms operate as the main interface between consumers and other market players. Two main categories can be distinguished: attention and matching platforms.<sup>22</sup>

32. Attention platforms, such as search engines or social networks, typically provide a set of ‘free’ services that are subsidised by advertising sold on a ‘per-click’ basis. This way, instead of paying a monetary price for the service, consumers pay with their attention, by having paid results, organic results interspaced with publicity or by being required to watch an advertisement before gaining access to a content video. Arguably, consumers also pay by submitting their data, either indirectly – through the website recording clicks for online searches or shopping – or directly – through entering personal data into an online form. The attention platform then uses the consumer’s private data to improve the quality of the services and to better target advertisements, allowing the platform to attract new consumers and to charge a higher cost-per-click to advertisers.

33. Matching platforms provide a marketplace where different types of players can interact, such as buyers and sellers, employers and employees, or even individuals in online dating sites. Matching platforms earn money by charging fixed fees to access the platform and variable fees per transaction. Frequently, the group of users with higher elasticity of demand is subsidised by the other group (for example, clients do not pay to use shopping sites; job seekers do not pay to use employment sites, and so on). Yet, all groups have their private data collected, which is used to improve the quality of the platform and of the matching algorithms, ultimately leading to a greater number of transactions.

34. The multisided features of platforms tend to lead, as a result of direct and indirect network externalities, to the concentration of users and their respective data in the hands of a few players. In turn, the use of Big Data provides online platforms with substantial market power in the supply of essential information services, upon which all companies and consumers rely. Indeed, these business models have proven highly profitable, allowing some attention platforms and matching platforms to be among the top 10 most valuable companies in the world by market capitalisation. High profitability *per se* does not imply competitive harm, as long as business success is achieved through means of data-driven innovation, and not through exploitation of Big Data to discriminate against some players, impose switching costs, enforce exclusive contracts, or conduct other forms of abuse.

#### 2.4.2. Content providers

35. Another set of players in the Big Data ecosystem are the content providers such as journals, websites and application developers that create the informative content available in many platforms, in exchange for a position in listing results. The contents created are not only displayed by search engines as part of their core business, but also by other platforms like social networks, which need creative content to attract and retain consumers’ attention and keep high levels of traffic. Unfortunately, because content providers are many and platforms are few, the informative content that is actually passed to consumers may not be only the result of a competitive process, but rather the outcome of platforms’ strategic decisions.

36. Content providers make money either by selling their product directly to consumers, or by selling advertising space to sellers. However, as they lack the Big Data required to conduct proper target advertising, it is increasingly common for websites to run, for instance, ads through the platform (such as Google’s) and to receive a share of the advertising revenues.

### 2.4.3. *Sellers*

37. The main subsidisers of the players discussed so far are the sellers or service providers, who offer products and services to the final consumer in exchange of money. This includes manufacturers, wholesalers, professionals, real estate agencies, consultants, finance institutions and any other types of businesses that may use platforms' marketing channels to persuade the consumers to purchase their products. The vast majority of the players in this group face vigorous competition.

38. There are, however, a number of large sellers who can grow enough in scale to use Big Data themselves, just as in the case of Amazon, Tesco or Target (the second-largest discount retailer in the US). By collecting data from online transactions, fidelity cards and forms submitted by consumers, sometimes in exchange for price discounts and free products, these companies are able to further increase their size and to acquire an irretrievable distance from smaller competitors, who lack the scale and expensive infrastructures to process Big Data and to become viable competitors. Again, this does not necessarily imply competitive harm, as the efficiency gains and innovations achieved by the large players may bring benefits for society.

### 2.4.4. *Infrastructure, including cloud computing and storage*

39. The operations conducted by Big Data users are crucially supported by information technology (IT) infrastructure providers, such as Hadoop, IBM and Oracle. Indeed, data-driven innovation companies are quickly faced with petabytes of data that are expensive to store and even harder to process, for which they lack resources. IT infrastructure providers not only develop the adequate software to handle Big Data, but most importantly they provide cloud computing and storage, that is, they act as third-party data centres where companies can store and process their data on-demand. These data centres usually consist of big clusters of computers connected by fast local area networks, constantly operating and largely benefiting from economies of scale.

40. The creation of cloud computing has partially reduced the problem of scale associated with IT infrastructures, by converting these fixed costs into variable costs and allowing small companies to operate without owning the physical infrastructures. With companies such as Amazon, Google and Microsoft providing machine learning algorithms as part of their cloud computing services, small companies find it increasingly more convenient to have their data processed and mined using external IT infrastructures. Indeed, Cisco forecasts that, by 2019, 86% of all business workload processing will be processed by cloud computing.<sup>23</sup> But, as a greater number of companies become dependent on the infrastructures of a few providers, the latter get access to significant volumes and variety of data that allows them to improve further their own data analysis algorithms. If the trend continues, a competition problem may arise in the future, as new entrants may not be able to build sufficiently powerful IT infrastructures whose analytical software can compete with those of incumbents.

### 2.4.5. *The public sector*

41. Finally, on the opposite side of the ecosystem, the public sector, including central and local government, as well as public hospitals, clinics, social security and other public services, collects Big Data from citizens and, occasionally, from platforms and sellers, when the latter are required to provide information to comply with the law. The public sector is, indeed, one of the most data-intensive sectors of the economy, using national databases for scientific research and to support the provision of public services. Still, there is a potential to exploit further the data in hands of governments for public purposes, by implementing the new data mining and machine learning techniques that have been developed by the private sector. At the same time, the use of Big Data for the provision of public services may pose a

problem of competitive neutrality, by making it hard or even impossible for private firms to compete in some areas, at least without access to the public data.

### **3. Implications of Big Data for competition law enforcement**

42. There is currently no agreement in the literature about the implications of Big Data for competition law enforcement. There seems to be an increasing realisation that competition authorities have a clear role to play in preventing the accumulation of market power through acquisitions, for instance, where privacy protection in itself becomes an intrinsic product quality. Others however assert that current privacy and/or consumer protection legislation is sufficient to take care of these concerns. In this section we will review this debate and discuss potential implications of Big Data for the effectiveness of existing competition tools and for the main activities of competition authorities: merger review, assessment of abuse of dominance and cartel enforcement.

#### **3.1. Competition tools**

43. Many of the current instruments of competition analysis, such as market definition, may be insufficient to fully account for the features of digital markets, for instance in the presence of ‘zero’ prices. In those cases, tools such as the SSNIP test, as well as the most consensual measures of market concentration, fall short of capturing the specific features of these markets.

44. Sometimes a simple adaptation of the current competition tools in light of the existing literature may suffice; in other cases competition law enforcers may find it necessary to complement their assessment with a number of new criteria, in a case-by-case analysis. The next section attempts to identify not only the limitations of current tools, but also some solutions that have been proposed from recent experience.

##### **3.1.1. Identifying the relevant market for antitrust purposes**

45. Identifying the relevant markets inside the Big Data ecosystem can be a particularly daunting task, as a result of the many different players involved that may take multiple roles, as well as the complex relations that link them. For example, a company like Apple is simultaneously a platform (through the operating system iOS, Apple Store and iTunes); a seller of multiple technological products, such as computers, tablets, phones and watches; and an IT infrastructure provider, through the provision of the iCloud service. At the same time, Apple interacts with many types of players, by transacting products and services with consumers, charging content providers (the ‘app’ developers) for the use of Apple’s platforms, selling advertising space, and even cooperating with platforms, such as Facebook or LinkedIn.

46. The multi-sided platform structure may require competition authorities to adapt the traditional SSNIP and hypothetical monopoly tests (Box 4). While the theory of the economics of multi-sided platforms is not new, it may be particularly hard to identify the multiple sides of a market when digital platforms engage in non-monetary transactions in exchange for data. For instance, a traditional platform such as a newspaper clearly operates simultaneously in the news and advertising market, charging a price both to readers and advertisers, but it is less evident in which markets a company such as Google participates, as it provides multiple money-free services such as searches, translation, GPS navigation, video uploading and a social network, among others. Therefore, in order to identify a multi-side market, it is not enough to look for monetary transactions; it is equally important to search for any data flows that may be observed in the market.

47. The fact that the collection of data for commercial purposes allows companies to increasingly offer a wide range of products at zero prices has implications for measuring the size of the relevant markets, since the SSNIP and hypothetical monopolist tests crucially rely on price mechanisms. As a

result, when products and services are money free, one of the few solutions available to define the market might be through means of a quantitative assessment of quality, for instance by using a SSNDQ test to measure the effect of a ‘small but significant non-transitory decrease in quality’ (OECD, 2013). This test, while sometimes applied in industries where quality measures are well-accepted and quantifiable (e.g., health sector), is used sparingly in other industries, where appropriate measures of quality have still to be developed.

#### **Box 4. Market definition in multi-sided platforms**

Digital markets are often characterised by multi-sided features and cross externalities that make market definition considerably complex. Several authors have written about the issue, starting with Rochet and Tirole’s (2000) classical paper on competition in multi-sided markets. More recently, Evans and Noel (2008) and Filistrucchi et al (2014) have provided useful insights on how to adapt the usual market definition tools to attention and matching platforms. In particular, they propose a modified SSNIP test to account for the cross externalities of a price increase on the various sides of the market.

With respect to attention platforms, the literature broadly agrees that, as long as consumers, advertisers, content providers and any other agents involved do not directly transact with each other, law enforcers should define a different market for each side of the platform (see Filistrucchi et al, 2014 and Wright, 2004). The underlying reasoning for identifying multiple markets is that products may be perceived with different degrees of substitutability in the different sides of the platform – for instant, social networks and search engines may be regarded as substitutes for advertisers, but not for consumers.

Then, as suggested by Evans and Noel (2008) and Filistrucchi et al (2014), when defining each market one should consider all externalities imposed on the remaining sides. For that, competition authorities could implement a modified SSNIP test that evaluates the impact of a price increase in one market on the overall profitability of the platform, by incorporating into the analysis the cross demand elasticities between the multiple sides.

Market definition tends to be less complex in matching platforms, where most authors agree that it is usually enough to define a single market, as long as all transactions take place simultaneously at the different sides of the platform. In this case, Filistrucchi et al. (2014) recommend competition authorities to apply a single modified SSNIP test, by measuring the overall profitability of a small increase in the total price set by the platform. A fundamental difference of this approach is that the price variation may include simultaneous variations in fixed and variable transaction fees charged to all sides of the platform.

Nonetheless, several issues remain that are not fully accounted for across the literature. What happens for instance if one side of the platform benefits while the other side experiences (consumer) harm? Would a competition agency try to add up and net out the effects? Or should the agency give more weight to the consumer side, or treat the two sides separately? The latter is hard in the case where one side subsidises the other side. Moreover, because the case involves data, would an agency also consider the effect that the data-driven merger might have in helping the firm attain or maintain market power in other markets involved in the platform? There are no easy answers to these questions, and little precedence as yet from case law.

In the case of a data-driven merger, for instance, it may be the case that efficiencies from the merger would reduce advertising costs and benefit advertisers with better targeted behavioural ads, while consumers would still receive free services from the other side of the platform. This would be positive in light of our discussion above. However, suppose then that quality, in the form of privacy protection for the data, would be reduced following the merger. Then there would be consumer harm, because the privacy protections would be reduced by a small, but significant, non-transitory degree. In that case, how does a competition agency go about balancing the gain to the advertisers compared to the loss in quality to consumers?

These are issues that are likely to come increasingly to the fore in data-driven markets in the near term.



### 3.1.2. Assessing market power

48. Market power is particularly difficult to assess when companies provide zero-price services to consumers in exchange for data, in which case enforcers may underestimate the degree of market power, or even assume that the market presents no competition problem. However, a zero-price offer may be part of a profit-maximising strategy to attract price-sensitive consumers and, then, to exert market power over other groups of participants, for example by selling information in other sides of the market (this is, for instance, the model of some types of dating platforms where access may be free for women, but men have to pay). Also, market power may be exerted through non-price dimensions of competition, allowing companies to supply products or services of reduced quality, to impose large amounts of advertising or even to collect, analyse or sell excessive data from consumers.

49. In their joint report, the French *Autorité de la Concurrence* and the German *Bundeskartellamt* (2016) remark that, even when products are free, the possession of Big Data might be an important source of market power, particularly when the data can be used as a barrier to entry. This concern was behind the actions of the US Department of Justice (DoJ) to block the merger of Bazaarvoice and Power-Reviews which, if cleared, could have created serious barriers to entry in the market for ‘rating and reviews platforms’, through the potential monopolisation of data. This and other cases suggest that, in markets where zero-prices are observed, market power is better measured by shares of control over data than shares of sales or any other traditional measures.

50. Finally, the very specific features of the digital economy imply that, in many cases, firms compete *for* the market instead of competing *in* the market, leading to ‘winner takes all’ outcomes, as it was observed when Facebook was able to displace Myspace as the most popular social network. This form of competition is typical among digital platforms and may require new criteria for a proper assessment of market power (see Box 5). In these cases, focusing on promoting market contestability is crucial to guarantee that dominant companies still face competitive pressure to constantly improve their products and preserve low prices.

#### **Box 5. Assessing the market power of platforms**

The German *Bundeskartellamt* (2016) published a report with recommendations to better assess market power in the specific context of platforms and networks. Recognising that the presence of high returns to scale associated with Big Data, as well as direct and indirect network effects may result in highly concentrated markets, they propose the consideration of additional criteria to evaluate whether the market is contestable.

According to the report, market power may be restricted by the ability of consumers to multi-home and, more importantly, by their incentive to do so. In the search engine market, for instance, there are few if any restrictions to multi-home, but users (searchers) may be encouraged to systematically use the same search engine as a result of default options and network effects that are hard to change. Consumer inertia also means that there is a high propensity to remaining with the default option on a device. This makes it harder for new entrants to acquire a critical mass of searches in order to establish themselves. Other criteria suggested by the *Bundeskartellamt* include seeing whether platforms are sufficiently differentiated, which allows them to target different groups of consumers and reduces the risk of monopolisation. Another point is that any technical or physical limitations of platforms that could lead to congestion raise the incentives for new entry. Finally, if the market has a high innovation potential it will limit the ability of firms to exert market power, by improving dynamic competition and allowing new entrants to overthrow established incumbents.

### **3.2. Merger review**

51. The use of Big Data analysis in the competition policy framework has been motivated, at least in part, by a series of cross-border transactions, such as the Google/DoubleClick<sup>24</sup> and the Facebook/WhatsApp<sup>25</sup> mergers, which have attracted a high degree of attention from the general public, as well as competition practitioners. These transactions challenged traditional categorisation, being hard to classify as horizontal or conglomerate mergers and involving a difficult evaluation process by competition authorities.

52. When competition authorities focus solely on the price effects of a transaction, some anti-competitive mergers may end up being cleared unconditionally, with potentially a significant future cost for consumers. However, if the risk of monopolisation of data or the privacy costs imposed on consumers are taken into consideration, the decisions may dramatically change to account for other important dimensions of competition policy. In this section, we discuss how privacy considerations may affect merger reviews and whether the current notification thresholds are able to capture transactions motivated by Big Data.

#### *3.2.1. Bringing a privacy dimension to merger review*

53. The accumulation of vast amounts of data about consumer behaviour and the expansion of targeted advertising has imposed costs in form of the loss of privacy on consumers. In fact, the price effectively paid by consumers for Internet services now extends far beyond punctual advertising breaks (such as when using the music-streaming service, Spotify) or banner ads flashing next to a search entry. Consumers' data and search entries are also analysed by data mining software, involving sometimes a more serious degree of intrusiveness. This can be illustrated by the now-famous anecdotal case of Target, the second largest discount retailer in US, which used historical purchasing data to estimate, among other things, a pregnancy probability score for female clients. According to a press release (Hill, 2012), a US retail company, Target, used the results from its own probability calculations to send multiple coupons for baby products to a teen-age girl, eventually alerting her father to the fact that she was pregnant.

54. This and a number of similar cases have contributed to a growing concern with the protection of consumer privacy in the context of the use of Big Data, raised not only by consumer protection and data protection offices, but also by antitrust authorities, which have already started bringing a privacy dimension to competition policy. The first competition law enforcement case involving privacy seems to be the Google/DoubleClick merger,<sup>26</sup> when Commissioner Pamela Jones Harbour of the US Federal Trade Commission (FTC) raised the concern that the merger would deprive consumers of meaningful privacy choices (ultimately, the merger was cleared by the FTC).<sup>27</sup> Similarly, when the joint venture between Microsoft and Yahoo was announced, the chairman Herb Kohl of the Senate Antitrust Subcommittee stated the importance of evaluating the impact of the transaction on internet users' privacy (Lande, 2008).

55. The introduction of a privacy dimension into competition policy is not a consensual practice. Some within the antitrust community believe that competition policy should have as sole objective the promotion of competition as a means to promote the efficient allocation of resources, while other public interests should be addressed by the respective public offices.<sup>28</sup> Cooper (2013) for instance argues that addressing privacy concerns through competition law would introduce an undesirable level of subjectivity into antitrust enforcement and could even conflict with fundamental rights of free speech, which are protected by the First Amendment of the US and in many other jurisdictions.

56. However, in circumstances where privacy violations by companies take place through the exercise of market power, it has been argued that there may be a legitimate justification for competition authorities to address privacy as an antitrust concern.<sup>29</sup> In the sense that data has been identified as the

‘new currency of the internet’, an increase in the collection of private data can be compared, to some extent, to a price increase. Or, equivalently, if consumers value privacy as a desirable characteristic, a reduction in privacy is analogue to a reduction in the quality of the service provided. For instance, in *Facebook/WhatsApp*<sup>30</sup> (see Box 6), the European Commission officials noted that if a website, post-merger, “would start requiring more personal data from users or supplying such data to third parties as a condition for delivering its ‘free’ product” then this “could be seen as either increasing its price or as degrading the quality of its product” (Ocello et al, 2015).

57. Competition authorities have generally recognised the importance of quality as a competitive feature, especially when the product or service is offered for free.<sup>31</sup> In fact, privacy considerations can fall within the ambit of non-price quality competition (Kimmel and Kestenbaum, 2014). As argued by Lande (2008):

*“Antitrust is actually about consumer choice, and price is only one type of choice. The ultimate purpose of the antitrust laws is to help ensure that the free market will bring to consumers everything they want from competition. This starts with competitive prices, of course, but consumers also want an optimal level of variety, innovation, quality, and other forms of non-price competition. Including privacy protection.”*

58. Some may claim that the collection of private data does not necessarily leave consumers worse off, as it allows companies to improve the quality of their products and improve consumer segmentation. Even so, privacy clearly confers a quality dimension that must be evaluated as a form of horizontal differentiation, since some consumers may prefer a higher degree of data protection, while other consumers may be willing to reveal their data to benefit from more personalised content and targeted ads (Cooper, 2013).

59. The consideration of privacy as a relevant parameter of non-price competition would have significant implications for merger review and ultimately affect the decision to clear or block a merger. In particular, in its evaluation of whether a potential merger might substantially reduce the welfare of consumers with high privacy preferences, competition authorities could decide to prevent the acquisition of the few companies in the market providing services with a greater extent of privacy protection. This is the case of the internet company DuckDuckGo, which provides search engine services without collecting or sharing any personal information, such as the IP address, search entries and search history (the company has other additional mechanisms to protect data, not allowing any type of browser cookies, appointing users to encrypted versions of major web sites and having an option to turn off ads).

#### **Box 6. The Facebook/WhatsApp merger**

Social network and text messaging are among the most popular services offered in the internet market, with a particularly high penetration rate among the adolescent population. According to a survey by Lenhart (2015), teenagers in the US sent and received, in average, 30 messages a day in 2015, and 71% of them had a Facebook account. With WhatsApp owning the leading messaging platform and Facebook offering the most widely used social network, as well as its own platforms to share messages, photos and videos, such as Facebook Messenger and Instagram, the merger between the two companies has been a focal point in the debate about Big Data, competition and privacy.

The Facebook/WhatsApp case illustrates a typical merger of two horizontally differentiated products that provide consumers with a different price/privacy trade-off. While WhatsApp was either free or, in some jurisdictions, charged users with a nominal fee in exchange for an ad-free service where no personal data was collected, Facebook messaging services have always been free, but involve the collection of data for target advertising purposes. As the Electronic Privacy Information Centre stated:<sup>32</sup>

*“Facebook messaging is notorious for its extensive data collection practices. When Facebook revamped its messaging system in November 2010, it automatically opted in all Facebook users and initially disabled users’ ability to delete individual messages. Without user consent, the new messaging system also pulled data from Facebook’s social*

*graph to prioritize messages from certain users. Currently, even when users delete a message, it continues to be stored on Facebook's servers. At the end of 2013, Slate<sup>33</sup> reported that even when a user chooses not to send a message, Facebook still tracks what the user wrote."*

Despite the privacy concerns, the merger was cleared both by the FTC and European Commission<sup>34</sup> on the condition that the WhatsApp service would continue to honour previous privacy policies and obtain users' consent before changing any policies. The merger review concluded that the higher concentration of data would still not allow the merged entity to obtain dominance in the advertising market, given the existence of other competitors with a considerable share of data collection across the web. However, Stucke and Grunes (2016) suggest that the review may have failed to account for potential future quality degradation impact on consumers, who may fail to perceive any future changes in privacy policies or not have the incentives to change to other messaging services with better privacy protection, due to lock-in resulting from network effects. Indeed, on 25 August, 2016, the New York Times reported that WhatsApp had announced that it would soon start to "share some member information with Facebook".<sup>35</sup>

### 3.2.2. Notification thresholds in Big Data mergers

60. Many jurisdictions use notification thresholds to screen mergers that are subject to notification to the competition authority. In most cases, these thresholds are based on the turnovers of the companies involved in the transactions.<sup>36</sup> In some cases however, a simple turnover threshold may leave out acquisitions with an important future competition impact where an established incumbent, motivated by the prospects of gaining access to a variety of additional data sources, buys a small entrant that it sees as a data-driven innovator or with access to valuable data.

61. In the Facebook/WhatsApp<sup>37</sup> transaction, the small value of the turnover of the latter company was not enough to trigger the notification threshold. Nonetheless, despite the comparatively small size of WhatsApp, Facebook paid USD19 billion for the company, hinting at the future value they expected to achieve from the acquisition. Ultimately, the European Commission reviewed the merger, as a result of Facebook's request to have a "one-stop-shop" review that would avoid notifying multiple jurisdictions with different thresholds and rules. The European Competition Commissioner, Margrethe Vestager (2016), stated in a public announcement, following the Facebook/WhatsApp case:

*"The issue seems to be that it's not always turnover that makes a company an attractive merger partner. Sometimes, what matters are its assets. That could be a customer base or even a set of data. (...) Or a company might be valuable simply because of its ability to innovate.*

*A merger that involves this sort of company could clearly affect competition, even though the company's turnover might not be high enough to meet our thresholds. So by looking only at turnover, we might be missing some important deals that we ought to review."<sup>38</sup>*

62. A possible solution to capture mergers motivated by the acquisition of a rival's data is the incorporation of an additional threshold based on the value of the transaction, reflecting the high price that buyers are usually willing to pay for the assets they are acquiring, such as data. Moreover, such transaction thresholds could help enable competition authorities to identify pre-emptive acquisitions intended to displace potential disruptive innovators (some of which may be data-driven innovators), as already discussed in OECD (2015b).

63. Transaction thresholds exist in the US and Mexican jurisdictions, and are under the consideration in others, such as Germany. Following a recommendation from the advisory body, the *Monopolkommission* (2015), the German Federal Ministry of Economic Affairs and Energy<sup>39</sup> published a draft amendment to the *Act against Restraints of Competition*, proposing the creation of a new threshold for the value of the transaction set at EUR 350 million, in addition to the existing turnover thresholds. The consideration of the value of the transaction has also been discussed by EU Commissioner Vestager (2016), who underlines the importance of setting the right threshold value to prevent harming innovative start-ups.

### 3.3. *Abuse of dominance*

64. Control over a high volume and variety of data can be an important source of productivity gains and product innovation. Accordingly, when Big Data is concentrated in the hands of a few large players, it may provide them with a substantial competitive advantage against which new entrants have difficulty competing. While the collection and control of even substantial amounts of data is not illegal, the misuse of Big Data to raise entry costs and gain or maintain market power might amount to a violation of competition law that requires the intervention of competition authorities.

65. In this section we identify some types of exclusionary conduct that may be enabled by controlling data, providing also illustrative examples. Then, we discuss whether data is an essential input in some industries and whether that could eventually justify the application of the essential facilities doctrine.

#### 3.3.1. *Exclusionary conduct*

66. The purpose of data-driven exclusionary and predatory behaviour may be to limit competitors' timely access to key data, prevent others from sharing data, inhibit data-portability or exclude rivals that threaten the data-related competitive advantage of an incumbent. These goals can be achieved, for instance, through the use of exclusive contracts with third-party data providers.

67. In their joint report, the *Autorité de la Concurrence* and *Bundeskartellamt* (2016) identify forms of abusive conduct that involve the exploitation of Big Data to foreclose the market. One example discussed in the report is the provision of discriminatory access to data, with the intention of providing one company with an unduly competitive advantage over other competitors. This can be observed, for instance, when a supplier, a platform or a marketplace operator is vertically integrated in the retail market and uses its access to data in the upstream market to obtain an unfair advantage over the other retailers. Even in the absence of vertical relations, a firm may discriminate against access to data to attempt to exclude a viable competitor. As an example, the French company Cegedim was challenged by the *Autorité de la Concurrence* (2014) for refusing to sell information from a medical database (over which they had exclusive control) to any customers that use software from one of their main competitors.

68. The UK Competition and Markets Authority (CMA, 2015) mentions the possibility that firms may leverage the data controlled in some markets to achieve enhanced power in other related markets, by using bundling or tying strategies. For example, firms may tie the purchase of their datasets to their own data analytics services. There are cases where bundling and tying may have efficiency gains. Hence each case must be assessed individually, in order to establish, *inter alia*, whether these strategies specifically aim at monopolising data and raising rivals' entry costs or preventing entry.

69. A less evident type of abusive conduct could take the form of a violation of consumers' privacy rights, as suggested by the German *Bundeskartellamt* during an investigation on Facebook. As argued by the President of the *Bundeskartellamt*, Andreas Mundt: "*it is essential to examine under the aspect of abuse of market power whether the consumers are sufficiently informed about the type and extent of data collected*".

70. Whether privacy violations should be evaluated under the scope of competition law or left to other Consumer Protection Offices is still an open question, whose answer might depend on the specific nature of the abuse. One example requiring the attention of the competition authority is where the privacy violation is reasonably capable of helping a company attain or maintain its monopoly power (especially in markets with strong data-driven network effects). A second question is where a privacy violation acts as an exclusionary abuse, by extracting private information that is not accessible to other rivals and using that data to foreclose rivals or raise entry barriers.

### 3.3.3. *Data as an essential input and the essential facilities doctrine*

71. Some practitioners are currently discussing whether data may be considered an essential input in some markets, without which companies cannot compete. It is clear that in some cases data and, more specifically, the knowledge extracted from the data are a source of a significant competitive advantage.<sup>40</sup> Subsequently, some commentators have debated whether the ‘essential facilities’ argument is worth consideration.

72. Recognising that the essential facility doctrine is not universally accepted by courts or competition law practitioners, the addition of fast moving and speculative claims to application of the doctrine is particularly challenging and has received strong opposition, not only from studies sponsored by current incumbents (Lerner, 2014), but also from some antitrust practitioners (Balto and Lane, 2016) and academics (Sokol and Comerford, 2016). These authors typically argue that data is not a crucial input for the success of any company, as innovative entrants have been able to establish themselves in spite of their initial small share of user data:

*“(...) the history of the digital economy offers many examples, like Slack, Facebook, Snapchat, and Tinder, where a simple insight into customer needs enabled entry and rapid success despite established network effects”.*<sup>41</sup>

73. While it is true that those entrants cited above were able to displace established incumbents in the past, the role of Big Data as an essential component of business strategies is relatively recent, and technological developments and the business models resulting from using deep learning are quite different now from when those companies entered the market. Hence it may be the case that it is becoming increasingly harder for new companies to come up with innovations that are sufficiently disruptive to exert competitive pressure over a current dominant, or well-established player.

74. Yet, in order to apply the so called essential facilities doctrine, it is not enough to show that Big Data is an indispensable input, being equally necessary to prove that it cannot be reasonably duplicated by competitors.<sup>42</sup> An argument frequently posed by the opponents of the essential facilities doctrine is that data cannot be easily monopolised: it is non-rival and, they argue, non-exclusive, since there are no contracts preventing users from sharing their personal information with multiple companies. Furthermore, they argue that there are few entry barriers to new platforms, as data is relative inexpensive to collect, shorted-lived and abundant.<sup>43</sup> However, as discussed previously, it may not be the collection of the data, as much as the ability to timely and swiftly extract useful information from a large volume and variety of data that leads to a competitive advantage being gained.

### 3.4. *Collusive practices*

75. There is little discussion in the current literature about the implications of Big Data for the detection and investigation of cartels, possibly as a result of the very few cases that have been investigated to date. Even so, the implications of Big Data for consumer welfare may be significant, as advanced methods of data analysis, programming tools and artificial intelligence, added to the greater transparency and ability to compare prices provided by the Internet, are likely to greatly facilitate market coordination.

76. This section describes methods that have been used to implement formal cartels or to facilitate tacit collusion, by improving market transparency and stability. It also makes some suggestions as to how competition authorities may adapt their tools to face new and improved forms of cartels, although additional work is still needed in this area.

### 3.4.1. *The advent of digital cartels*

77. Evidence suggests that digital cartels existed even before Big Data was ‘big’. In a well-known case investigated in the 1990s by the US Department of Justice (DoJ), major US airlines were accused of using a database with detailed airfare information to make repeated tariff announcements and fast price changes, in order to enable online collusion (Box 7). However, after three years of investigation the case was closed with a settlement agreement between the DoJ and the airline companies, not establishing a legal precedent.

78. In 2015, the DoJ prosecuted a cartel operating in the digital market for the first time, which involved several sellers fixing prices for posters sold in the Amazon marketplace. In particular, the executive who was charged by the DoJ had developed a pricing algorithm reactive to consumer preferences, which was shared with other sellers and implemented in parallel to allow for price coordination. As the then Assistant Attorney-General Bill Baer stated in a press release:<sup>44</sup>

*“We will not tolerate anticompetitive conduct, whether it occurs in a smoke-filled room or over the Internet using complex pricing algorithms.”*

79. Because the number of digital cartels that has raised the attention of competition enforcers is still scarce, the incentive for companies to search for new creative ways of using Big Data to collude is enormous, especially if the improved anti-competitive practices are hard to detect and prosecute in court of law. Stucke and Ezrachi (2015) identify four potential strategies to use Big Data to facilitate collusion, information that may be useful for antitrust authorities to consider.

80. Firstly, firms may use real-time data analysis to monitor compliance with an explicit agreement that, in all remaining aspects, resembles a traditional cartel. Secondly, firms may share identical pricing algorithms that allow them to simultaneously adjust prices based on the inflow of market data, just as in the poster price-fixing case. If competitors use a vertically integrated company to orchestrate the implementation of the algorithm, a classic hub-and-spoke cartel may ensue. Thirdly, at a more sophisticated level, firms may use Big Data to facilitate (tacit) collusion, either by improving market transparency or by making actions more interdependent – for instance, by programming immediate retaliations to price falls. Fourthly, companies may use artificial intelligence to create profit-maximising algorithms that, through machine learning, may achieve tacit collusion, even in cases where the programmer did not initially foresee such an outcome.

81. The two last strategies may pose serious challenges to competition authorities in the future, as it may be very difficult, if not impossible, to prove an intention to coordinate prices, at least using current antitrust tools. Particularly in the case of artificial intelligence, there is no legal basis to attribute liability to a computer engineer for having programmed a machine that eventually ‘self-learned’ to coordinate prices with other machines.<sup>45</sup>

#### **Box 7. The US Airline Case**

During the early 1990’s, the US Department of Justice investigated tariff fixing activities in the airline industry, where the cartel members were able to implicitly coordinate tariffs using a third party centre and sophisticated signalling mechanisms. The case is described in detail in Borenstein (1999).

In the US airline industry, airline companies send fare information on a daily basis to the Airline Tariff Publishing Company (ATPCO), a central clearinghouse that compiles all the data received and shares it in real time with travel agents, computer reservations systems, consumers and even the airline companies themselves. The database published by ATPCO includes, among other things, information about prices, travel dates, origin and destination airports, ticket restrictions, as well as first and last ticket dates, which indicate the time range when the tickets at a particular fare are for sale.

According to the case presented by the DoJ, airline companies were using first ticket dates to announce tariff raises many weeks in advance. If the announcements were matched by the rivals, when the first ticket date arrived all companies would simultaneously raise the tariff. Some of the coordination strategies were more complex, involving the use of fare code numbers and ticket date footnotes to send signals or negotiate multimarket coordination.

According to the DoJ's case it was the existence of a fast data exchange mechanism to monitor tariffs and react rapidly to price changes that enabled companies to collude without explicitly communicating. As tacit collusion is not forbidden by competition law and any explicit coordination was very hard to prove in a criminal case, eventually the DoJ reached a settlement agreement<sup>46</sup> with the airline companies, under which the latter agreed to stop announcing most price increases in advance, with the exception of a few circumstances where early announcements could enhance consumer welfare. All of the airline defendants' fares had to be concurrently available for sale to consumers.

### 3.4.2. *Big data antitrust*

82. It remains unclear how antitrust authorities can adjust their tools to fight digital cartels, but it is likely that any efficient response will require them to use game theory tools and play tit-for-tat. In other words, it may be necessary to introduce sophisticated methods of data analysis into competition law enforcement, in order to detect and prevent concerted practices in the digital markets. This may require the addition of new resources for antitrust authorities, such as computer scientists.

83. Although not yet common, the economic literature proposes methods to distinguish competitive and collusive regimes from observed data, usually known as screening methods.<sup>47</sup> Harrington (2008) reviews several existing empirical approaches to detect cartels, including models to identify bid rigging in public procurements (Bajari and Ye, 2003) and models to test for price collusion (Porter, 1983). New methods are continuously being developed, such as the framework proposed by Marmer et al (2016) to test for collusion in open ascending bid auctions, which are increasingly observed in the internet market. Screening methods based on data analysis have the advantage of being able to detect both formal and informal cartels at high speed, meaning that they can actually be effectively used to identify all forms of cartels operating online.

84. When companies use algorithms and machine learning to enable tacit collusion, screening is not enough to solve the problem, since firms can only be condemned if they engage in any form of explicit communication or, at least, reveal some intention to collude. Finding ways to prevent collusion between self-learning algorithms might be one of the biggest challenges that competition law enforcers have ever faced, and whose solution may involve artificially making market conditions more unstable and less prone to tacit collusion. Stucke and Ezechia (2016) propose some ways of doing so, such as sponsoring the entrance of a maverick company whose fast growth might disrupt the cartel; create a system of secret discounts, through which firms can reduce their prices without having their rivals noticed; or enforce a minimum time-lag for price changes, in order to incentivise companies to undercut prices. As Stucke and Ezechia note, however, each solution has its shortcomings. Thus any solutions proposed so far are still at an early stage and further research is needed on this topic.

## 4. **Pro-competitive impact of Big Data regulation**

85. Data-driven innovations are becoming an increasingly vital feature of our societies, leading to a growing dependence by consumers and firms alike upon the services they provide. For instance, firms depend on targeted advertising services and extensive databases collected by data brokers for their competitiveness; consumer welfare is tied in with being part of social networks and using smartphone applications for anything from shopping to travelling to booking medical appointments. Both firms and consumers depend on the supply of structured, organised and relevant information provided by search engines. With the development of the Internet of Things, from home appliances to driver-less transport, this dependence will increase exponentially in the next decade.



86. For the reasons discussed in this paper, several key data-driven markets tend to be highly concentrated. Based on observed examples, such as the Facebook/WhatsApp merger, it is not unreasonable to posit that the market structure, linked to the fact that the markets are as yet poorly understood by antitrust authorities and regulators alike, seems to have an intrinsic risk of moving towards even more concentration.<sup>48</sup> This is likely to happen through absorption of smaller competitors and new entrants through acquisitions by dominant incumbents, or through foreclosure by the biggest market players.

87. As a result, policy makers are increasingly focusing their attention on the digital economy and on new ways of regulating the use of Big Data, either to protect market competition, or to attain other public goals. In this section we will firstly discuss some consumer protection regulations that have been recently debated and which may have potential pro-competitive effects. Secondly, we will address regulations that aim at promoting an efficient use and exchange of Big Data between the private and public sectors, but which may raise concerns for competitive neutrality.

#### **4.1. Consumer protection**

88. Most data transactions between users and online service providers are characterised by asymmetry of information. Whenever users subscribe to an online service, they are rarely aware of the many types of data that may be collected from them (such as personal data, behavioural data, IP and location tracking data, business transactions data, and so on), nor are they entirely informed about how the data might be used or shared with third parties. Even though privacy policies and general terms and conditions are legally required in some jurisdictions to provide this information, the terms tend to be couched in complicated language, and are often so extensive that the time it takes to read them in full defies most users. A 2008 study by McDonald and Cranor estimated that “*reading privacy policies carries cost in time of approximately 201 hours a year, worth about \$3,534 annually per American Internet user*”. Stucke and Grunes (2016) put this figure at around 10 days a year. With such a heavy opportunity cost involved in verifying privacy policies, there is a strong argument in favour of having enhanced consumer protection and regulation of the digital markets.

89. Moreover, while consumer protection regulations may improve the security of online transactions, these will be insufficient if companies have all the bargaining power and consumers have no option but to accept the terms and conditions imposed on them. According to a US survey by the Pew Research Center (2014), “*91% of adults in the survey ‘agree’ or ‘strongly agree’ that consumers have lost control over how personal information is collected and used by companies*”. In order for the mechanisms of market competition to continue to operate, consumer protection rules should not only aim to keep consumers well informed, but quite possibly also restrict companies’ market power to freely collect massive amounts of data for any intended purposes.

##### *4.1.1. Inappropriate use of the word ‘free’*

90. One of the sources of information asymmetry between users and service providers is the fact that online companies often propose their products for ‘free’, when in fact these involve multiple non-pecuniary costs in the form of providing personal data, paying attention to ads, or the opportunity costs of reading privacy policies. According to Friedman’s (2008) work on behavioural economics, describing a product as ‘free’ is deceptive and may affect the consumer’s rational decision-making process, making them pay more for the product than they would if they were fully informed. In order to protect consumers, the FTC has applied its *Guide Concerning Use of the Word ‘Free’* to prevent companies to suggest that a product is costless if, for instance, it requires the consumer to engage in any other transaction or to buy additional units of product (as yet, the guide has not been applied to zero-price Internet products).

91. Moreover, some observers highlight that in a competitive market, in order to acquire data, which for some platforms is deemed intrinsically valuable, prices could actually be negative. In other words consumers would be paid to use the service. However, by keeping privacy policies deliberately vague, service providers make it difficult for consumers to evaluate the real value of their data. The user is given the immediate benefit of the zero-price service, but is unaware of the short or long-term costs in divulging information, as they do not know how the data will be used and by whom.

92. The designation of a product as ‘free’ can have even more serious consequences than behavioural effects, sometimes affecting the legal rights of consumers as well. In a private litigation case in California involving leakage of personal information,<sup>49</sup> Facebook users brought a claim against the company under the Unfair Competition Law and the Consumers Legal Remedies Act. However, because Facebook’s services do not require payment, the Court dismissed the claim, stating that non-paying users are not considered consumers under California law. Similarly, others have suggested that competition authorities lack jurisdiction over free products, under the claim that they do not represent economic activity. That claim was rejected by the OFT (2013, paragraphs 7 and 8) in the Google/Waze merger. In other words, it is not because a product or service is offered without money changing hands, that there is no economic activity.

93. Addressing the economic and legal concerns involving the use of zero-price online services, Hoofnagle and Whittington (2014) propose a *transaction cost economics* (TCE) approach to account for the actual costs involved in the transaction of online products, such as the track of users’ data, the cost of monitoring changes in privacy policy, switching and cancellation costs that lead to lock-in, and the lack of information security. Based on the TCE analysis, they discuss alternative policies to improve the efficiency of transactions, such as the legal recognition that the transfer of personal information is a non-free exchange of value; and forcing companies to notify subscribers that the service is provided in exchange for personal data and to detail the purpose for the data collected.

#### 4.1.2. *Ownership rights over data and privacy standards*

94. Hoofnagle and Whittington (2014) also propose the recognition of ownership rights for consumers over the data they produce, which could potentially increase their bargaining power to negotiate the conditions under which data is exchanged and even to receive a monetary compensation.<sup>50</sup> Following from the Coase theorem, establishing property rights and allowing users and companies to trade data in a free market could lead to a more efficient outcome that would account for heterogeneous preferences (for instance, consumers with a high preference for privacy could benefit from online services and have their data secured, while other consumers would be willing to provide detailed personal information for a sufficiently high price). As unlikely as it would appear to expect companies to pay for individual users’ data, one can increasingly observe new business models such as Handshake, which has created a virtual marketplace for consumers to negotiate and sell their personal data to other companies.

95. Another solution that has been proposed<sup>51</sup> is the creation of global standards for transparency of contractual terms. Standards can be used not only to make privacy policies easier to understand, but also to empower consumers with more control over their own data, by giving them a list of standardised options about the extent to which they authorise their personal data to be collected. While such options already exist, these can be made simpler, more explicit and uniformed. In Figure 4 we present a hypothetical example of what companies could be legally required to ask to any new subscriber of the online service.

Figure 4. Example of Privacy Standards

### Terms of Services

- I do not authorise the collection of my personal data.
- I authorise the collection of my personal data for internal purposes that are solely needed and used to provide the particular product or service in question.
- I authorise the collection of my personal data for the creation of aggregate databases that may be shared with third parties.
- I authorise my personal data to be collected and shared with third parties without any restrictions.

#### 4.1.3. Rights on data portability

96. Finally, rules on data portability play also an important role in restricting companies' market power, by reducing switching costs and allowing consumers to easily change to new and potentially better services. Taking the example of social networks, enabling consumers to transfer their profile's data and multimedia files across sites could promote market competition between incumbents and even encourage new entry. With this purpose, the European Commission has recently adopted a reform of the General Data Protection Regulation,<sup>52</sup> whose article 20 on *right of data portability* states:

*“The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance (...)”.*

97. In a nutshell, article 20 provides consumers with the right to download their personal data from any online service in an easy-to-use format and to have that data transferred to any other company.

98. While the goal of the *right of data portability* is certainly to prevent 'lock-in' and promote competition, Swire and Lagos (2013) argue that the specific provisions of article 20 may actually have perverse anti-competitive effects.<sup>53</sup> In particular, they criticise the fact that the new regulation, unlike traditional competition law, applies broadly to all types of firms and not only to those who hold a dominant position. Therefore, even small start-ups have the obligation of developing expensive software codes to enable the transference of data to other companies, a cost that may threaten their ability to compete.<sup>54</sup> Swire and Lagos (2013) equally criticise article 20 for not accounting for dynamic efficiency effects, since the legal obligation to share data collected from consumers may reduce expected profits and harm firms' incentives to innovate.

99. Whether the new rules on data portability will promote competition or reinforce the market position of dominant companies depends on how the courts will interpret Article 20 and how its provisions will be enforced. Still, this discussion clearly illustrates that any attempts to regulate the transactions and flows of Big Data must be carefully scrutinised, or there is a risk that new regulations will undermine the same goals they were intend to achieve, such as competitive and innovative markets.

## 4.2. *Competitive neutrality*

100. Governments, public agencies and state-owned enterprises have a privileged position to collect data that cannot be easily obtained by the private sector, not only because of their size, but also as a result of their role in law enforcement, provision of public services and collection of official statistics, making the public sector one of the most data-intensive sectors in the economy. The potential exploitation of public data in areas such as internal security, crime prevention, health, traffic and even macroeconomic policy can lead to massive savings of public funds, improvement of public services and ultimately economic growth. A McKinsey report (2011) estimates that if public sector data was exploited to its full extent, the governments of the European OECD member countries would be able to reduce operating expenditures by 15% to 20%, reduce fraud and errors by 30% to 40% and raise tax collection by 10% to 20%, generating an added value of between EUR 150 billion and EUR 300 billion.

101. Unfortunately, the potential exploitation of Big Data by the public sector may imply that, in some industries, private companies will not be able to compete against a government controlling such volume and variety of data across the whole economy, which could never be replicated by any other agent. In this context, initiatives to open public data and partnerships between the private and public sector may be an important measure to preserve competitive neutrality and improve efficiency in the use of existing data.

### 4.2.1. *Access to public sector information*

102. Given the enormous potential of Big Data to be re-used for multiple ends at small marginal costs, it may be desirable that governments allow for open access of public sector information to the private sector, which is well suited to commercially exploit it. Open data initiatives provide companies with the opportunity to innovate, prevent public agencies from forming governmental monopolies and promote more vigorous competition, by reducing the data gap between incumbents and small start-ups. Naturally, any open data initiative must be cautiously implemented, keeping all disclosed data at an aggregate level and eventually securing more sensitive data on taxes, health or social transfers.

103. In this matter, the OECD in 2008 adopted a recommendation to enhance access and effective use of public sector information, with the aim to “*promote more efficient distribution of information and content as well as the development of new information products and services particularly through market-based competition*”. Likewise, the CMA (2015) has recently evaluated the benefits of their former recommendations on the commercial use of public information, which aimed to foster business competition in the supply of value-added products and services.

104. In general, the OECD’s and CMA’s recommendations involve making public sector information available as a default rule (with some exceptions such as national security); creating simple, fast and less restrictive licensing systems; enhancing data quality; and decreasing the price charged per user, if possible to match the marginal cost of maintenance and distribution. The last recommendation may be hard to implement as it affects the financial position of the public agency, case in which the government may be required to inject additional public funds.

### 4.2.2. *Leveraging private data for public goals*

105. In the same way that public sector data can be of high value for private companies, in some cases the data gathered by the private sector can also be of great utility for the public counterpart, allowing the Government to complement official statistics with less traditional sources. As an example, the company CitiVox currently sells data about unreported crimes collected through text messages and social media to the governments of several Latin American countries,<sup>55</sup> helping the authorities fight crime. In a more impressive case, Ginsberg et al (2009) have created an empirical model to detect influenza epidemics using

Google's search data, which is implemented to monitor the spread of influenza and similar illnesses in the US. By analysing search queries about influenza submitted by users in real time, the model allows for faster detection and preventive response, as compared to traditional surveillance systems such as the number of visits to physicians, which has a reporting lag of one-to-two weeks.

106. Occasionally, the private and public interest may conflict. This is the case of the transport sector,<sup>56</sup> where private transportation companies hold Big Data that could be leveraged to improve road safety, traffic management and urban planning, but they have concerns that sharing their data with the government would allow other competitors to access the same information and eliminate their competitive advantage, and hence the public ends up losing out. One solution may be to create the right economic incentives for private companies to share such data, such as pecuniary compensations, fiscal deductions, confidential treatment of data or even data-sharing partnerships between the private and public sectors.

## **5. Conclusion**

107. The rise of new businesses models based on the collection and processing of Big Data is currently shaping the world. With the development of data mining and machine learning, businesses are able to offer innovative, high-quality and customised products and services at low or even zero prices, with great gains for consumers. Nowadays, it is possible in many parts of the world to obtain detailed navigation instructions with real-time traffic conditions; epidemics can be forecast and fought in a useful timespan, potentially saving millions of lives; efficiency is boosted in digital marketplaces, where prices are adjusted to supply and demand in near-real time, while consumers benefit from customised suggestions and users' feedback; social and professional networks allow individuals to keep in contact with existing friends and colleagues, or even to be matched with potential partners; advanced self-learning algorithms enable online searchers to find the exact information they need in a minimum amount of time.

108. Yet, the endless achievements of Big Data do not come without a cost. Consumers may be increasingly facing a loss of control over their data, and their privacy; they are confronted with intrusive advertising and behavioural discrimination, and are ever more locked-in to the services upon which they rely. This dependency will increase exponentially in the next decade when the Internet of Things becomes a reality, running home appliances, transport and even business processes. Companies can use the most advanced computing technologies to coordinate practices, impose abusive conditions on consumers, use their enhanced market power to charge higher prices and even foreclose the market to potential competitors. Data-driven network effects tend to become self-sustaining, favouring incumbents and enabling them to entrench their positions once they reach the tipping point of a critical mass of users.

109. Whether Big Data gains will overcome the potential costs for society, it will partly depend on whether and how competition authorities and regulators will be able to react to the new challenges of the digital economy. Depending on how challenges are addressed, we may either face increasingly competitive, contestable and dynamics markets in the future, where efficiency and continuous innovation prevail; or a sharp rise of market concentration, resulting in abuse of power and stagnation.

110. This paper has identified and discussed some of the most recent competition concerns thrown up by Big Data. Our discussion does not claim to be exhaustive, rather it aims to highlight some of the issues that competition practitioners will likely need to deal with in the very near future. Given the very fast evolution of the digital markets and the continuous development of new applications for Big Data, antitrust authorities and regulators need not only to be aware of the current market reality, but also anticipate new forms of strategic interactions and potentially anti-competitive conducts by market players, in order to bring competition policy fully into the digital era.

## ENDNOTES

- 
- 1 See OECD (2016a) scoping paper on the Digital Economy.
- 2 Other definitions can be found in McKinsey Global Institute (2011), Mayer-Schönberger and Cukier (2013) and Laney (2001), who introduced the ‘3Vs’ definition.
- 3 A recent joint paper by the French *Autorité de la Concurrence* and the German *Bundeskartellamt* (2016) contains a highly useful discussion of the types of data, and their categorisation. They propose to distinguish data by the type (personal versus other types); the structure of the data (whether in structured data bases, e.g. consumer data listed by alphabetical order, or unstructured); or the way that the data is gathered (whether by asking consumers to enter data into a form; or by observation, using cookies or “crawling” whereby webpages are systematically collected). This definition is useful in better understanding how data may be collected.
- 4 10.4 zettabytes correspond to 10.4 trillion gigabytes. These are staggering numbers. As an illustration, in order to store 10.4 zetabytes of data, each and every individual in the world (including infants) would have to own eleven iPhones of 128 gigabytes.
- 5 For more information and forecasts see Cisco Global Cloud Index, Forecast and Methodology, 2014-2019, available at [http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud\\_Index\\_White\\_Paper.html](http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html). One zettabyte equals 1,099,511,627,776 gigabytes.
- 6 Moore’s law describes the velocity at which the number of transistor components in an integrated circuit increases, roughly doubling every two years, therefore reducing the cost of electronics over time.
- 7 Stucke and Grunes use the example of the UK retailer, Tesco. See also Brad Howarth, ‘How Tesco’s loyalty card transformed customer data tracking’, CMO, 21 May 2015.
- 8 This issue is examined in detail in the Background Paper on Price Discrimination, which will also be presented at the Committee meeting in November 2016.
- 9 <http://rubiconproject.com/whoweare/>.
- 10 The Rubicon Project, Amendment No 3 to Form S-1 Registration Statement, 30 April 2014.
- 11 Available at <https://www.auction.com/blog/auction-com-launches-real-estates-first-nowcast/>.
- 12 See OECD (2015a), pages 28 and 29.
- 13 See comScore (2016), eMarketer (2016) and Newman (2013).
- 14 Statement of FTC concerning Google/DoubleClick, FTC File No. 071-0170, available at [https://www.ftc.gov/system/files/documents/public\\_statements/418081/071220googledc-commstmt.pdf](https://www.ftc.gov/system/files/documents/public_statements/418081/071220googledc-commstmt.pdf).
- 15 See Newman (2013) and Stucke and Grunes (2015).
- 16 Klishina (2011).

- 
- 17       Vise and Malseed (2005).
- 18       Stucke and Grunes (2015).
- 19       See the recent report by the Autorité de la Concurrence and Bundeskartellamt (2016).
- 20       See Shapiro and Varian (1999).
- 21       For a more detailed description of the Big Data ecosystem see OECD (2015a).
- 22       Sometimes the literature also distinguishes transaction from non-transaction platforms, depending on whether the players from the different sides of the market directly transact with each other. While attention platforms are usually of the non-transaction type and matching platforms of the transaction type, this is not always the case. For instance, dating platforms match couples who do not engage in any market transaction.
- 23       See reference op.cit.
- 24       Case No COMP/M.4731.
- 25       Case No COMP/M.7217.
- 26       Case No COMP/M.4731.
- 27       Harbour and Koslov (2010).
- 28       For instance, Sokol and Comerford (2016) claim that antitrust is ill-equipped to solve consumer law problems.
- 29       See, for instance, the EDPS (2012), Lande (2008), Stucke and Grunes (2015) and Newman (2013).
- 30       Case No COMP/M.7217.
- 31       *Microsoft/Skype* (Case Comp/M.6281), Commission Decision C(2011)7279, 7 October 2011, para 81 (noting that “*Since consumer communications services are mainly provided for free, consumers pay more attention to other features*” and “*quality is therefore a significant parameter of competition*”); *Microsoft/Yahoo! Search Business* (Case Comp/M.5727), Commission Decision C(2010) 1077, 18 February 2010.
- 32       Electronic Privacy Information Center, in Re: WhatsApp, <http://216.92.162.125/privacy/internet/ftc/whatsapp/> (quoting WhatsApp’s privacy policy from 2012).
- 33       [www.slate.com/articles/technology/future\\_tense/2013/12/facebook\\_self\\_censorship\\_what\\_happens\\_to\\_the\\_posts\\_you\\_don\\_t\\_publish.html](http://www.slate.com/articles/technology/future_tense/2013/12/facebook_self_censorship_what_happens_to_the_posts_you_don_t_publish.html).
- 34       Case No COMP/M.7217.
- 35       [http://www.nytimes.com/2016/08/26/technology/relaxing-privacy-vow-whatsapp-to-share-some-data-with-facebook.html?\\_r=1](http://www.nytimes.com/2016/08/26/technology/relaxing-privacy-vow-whatsapp-to-share-some-data-with-facebook.html?_r=1)
- 36       See OECD (2016c) on local nexus and notification thresholds in mergers.
- 37       Case No COMP/M.7217.

- 
- 38 Vestager, M. (2016), “Refining the EU Merger Control System”, Speech at the *Studienvereinigung Kartellrecht*, Brussels, 10 March 2016, [https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system\\_en](https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system_en).
- 39 Bundesministeriums für Wirtschaft und Energie (2016).
- 40 See Stucke and Grunes (2015), Stucke and Ezrachi (2015) and Newman (2013).
- 41 Sokol and Comerford (2016), page 5.
- 42 Lipsky and Sidak (1999).
- 43 In this matter, Balto and Lane (2016) compare an attempt to monopolise data to the exercise of monopolising water in a tsunami or in a rainstorm.
- 44 DOJ (2015).
- 45 These two collusive strategies and corresponding competitive concerns are elaborated in Stucke and Ezrachi (2016).
- 46 DOJ and American Airline, INC., Settlement Agreement, September 23 2004, available at <https://www.justice.gov/atr/case-document/settlement-agreement-and-order>.
- 47 See OECD’s (2013) roundtable on the use of screens to detect cartels.
- 48 For a deeper discussion, see also the Special Report in *The Economist*, September 17-23, 2016.
- 49 Facebook Privacy Litig., 791 F. Supp. 2d at 708-709.
- 50 A similar solution is proposed by Jaron Lanier (2013) in his seminal book “Who Owns the Future”, which involves an economy of micropayments where online users are compensated for any information they post online.
- 51 EDPS (2014), page 35, paragraph 80.
- 52 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, available at [http://ec.europa.eu/justice/data-protection/reform/files/regulation\\_oj\\_en.pdf](http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf).
- 53 Swire and Lagos’s (2013) critique applies to the draft regulation proposed by the European Commission in 2012. In the meanwhile, the regulation was approved in 2016 with modifications, which may partially address some of the authors’ concerns.
- 54 In this respect, a study by Shah and Kesan (2012) provides evidence of the complexity and high costs involved of achieving interoperability between different data formats, even for large companies.
- 55 OECD (2013).
- 56 International Transport Forum (2015).



## BIBLIOGRAPHY

- Acquisti, A., C. R. Taylor and L. Wagman (2016), “The Economics of Privacy”, *Journal of Economic Literature*, Vol. 52, No. 2, <http://dx.doi.org/10.2139/ssrn.2580411>.
- Autorité de la Concurrence (2014), “Décision No. 14-D-06 du 8 juillet 2014 relative à des pratiques mises en oeuvre par la société Cegedim dans le secteur des bases de données d’informations médicales”, <http://www.autoritedelaconcurrence.fr/pdf/avis/14d06.pdf>.
- Autorité de la Concurrence and Bundeskartellamt (2016), “Competition Law and Data”, <http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.html>.
- Bajari, P. and L. Ye (2003), “Deciding Between Competition and Collusion”, *Review of Economics and Statistics*, Vol.85, pp. 971-989, <http://www.econ.ohio-state.edu/lixinye/Research/restat1-3-03.pdf>.
- Bakhshi, H., A. Bravo-Biosca and J. Mateos-Garcia (2014), “Inside the Datavores: Estimating the Effect of Data and Online Analytics on Firm Performance”, *Nesta*, London, [https://www.nesta.org.uk/sites/default/files/inside\\_the\\_datavores\\_briefing.pdf](https://www.nesta.org.uk/sites/default/files/inside_the_datavores_briefing.pdf).
- Balto, D. A. and M. C. Lane (2016), “Monopolizing Water in a Tsunami: Finding Sensible Antitrust Rules for Big Data”, <http://ssrn.com/abstract=2753249>.
- Bauer, M. (2016), “Big Data: New Frontier for Competition Law”, IBC Competition EU Competition Law Conference, Brussels, <http://oecdshare.oecd.org/daf/competition/Knowledge%20Database/Non-OECD%20events/IBC%20AdvancedCompLaw%20Brussels%20Feb2016/Michael%20Bauer.pdf>.
- Borenstein, S. (1999), “Rapid Price Communication and Coordination: The Airline Tariff Publishing Case”, in J. E. Kwoka Jr. and L. J. White (Eds.), *The Antitrust Revolution: Economics, Competition, and Policy*, <http://faculty.haas.berkeley.edu/borenste/download/atpcase1.pdf>.
- Bork, R. H. and J. G. Sidak (2012), “What Does the Chicago School Teach About Internet Search and the Antitrust Treatment of Google?”, *Journal of Competition Law & Economics*, Vol. 8, No. 4, pp. 663-700, <http://jcle.oxfordjournals.org/content/8/4/663>.
- Brynjolfsson, E., A. McAfee, M. Sorell and F. Zhu (2008), “Scale Without Mass: Business Process Replication and Industry Dynamics”, *Harvard Business School Technology & Operations Mgt. Unit Research Paper*, No. 07-016, <http://dx.doi.org/10.2139/ssrn.980568>.
- Buchholtz, S., M. Bukowski and A. Sniegocki (2014), “Big and Open Data in Europe - A Growth Engine or a Missed Opportunity?”, Warsaw Institute for Economic Studies, <https://www.microsoft.com/global/eu/RenderingAssets/pdf/2014%20Jan%2028%20EMEA%20Big%20and%20Open%20Data%20Report%20-%20Final%20Report.pdf>.
- Bundesministeriums für Wirtschaft und Energie (2016), “Entwurf eines Neunten Gesetzes zur Änderung des Gesetzes gegen Wettbewerbsbeschränkungen” (Draft Ninth Act amending the Act against Restraints of Competition), <http://www.bmwi.de/BMWi/Redaktion/PDF/G/neunte-gwb-novelle,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf>.

- CEBR (2012), “Data Equity: Unlocking the Value of Big Data”, Centre for Economic and Business Research Ltd, Report for SAS, <http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf>.
- Chui, M., M. Löffler and R. Roberts (2010), “The Internet of Things”, McKinsey Quarterly, <http://www.mckinsey.com/industries/high-tech/our-insights/the-internet-of-things?cid=other-eml-clc-mip-mck-oth-1603>.
- Clark, D. (2012), “Control Point Analysis”, 2012 TRPC Conference, <http://dx.doi.org/10.2139/ssrn.2032124>.
- CMA (2015), “The Commercial Use of Consumer Data”, Report on the CMA's Call for Information, [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/435817/The\\_commercial\\_use\\_of\\_consumer\\_data.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435817/The_commercial_use_of_consumer_data.pdf).
- comScore (2016), “comScore Releases January 2016 U.S. Desktop Search Engine Rankings”, <https://www.comscore.com/Insights/Rankings/comScore-Releases-January-2016-US-Desktop-Search-Engine-Rankings>.
- Cooper, J. C. (2013), "Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity" *George Mason Law Review*, forthcoming, *George Mason Law & Economics Research Paper*, No. 13-39, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2283390](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2283390).
- Covington (2016), “Germany’s Proposal to Introduce ‘size-of-transaction’ Merger Review Thresholds Steps up Wider European Debate over Acquisitions of Start-ups”, *Antitrust Transactions*, [https://www.cov.com/-/media/files/corporate/publications/2016/07/germany\\_merger\\_review\\_thresholds.pdf](https://www.cov.com/-/media/files/corporate/publications/2016/07/germany_merger_review_thresholds.pdf).
- Crofts, L. (2016), “Antitrust watchdogs are realizing power of ‘Big Data’, EU data chief says”, *Mlex Global Antitrust*, <http://www.mlex.com/GlobalAntitrust/DetailView.aspx?cid=783261&siteid=190&rdir=1>.
- Curtis, M. (2014), “New Data Sources: A Conversation with Google’s Hal Varian”, Macroblog of the Federal Reserve Bank of Atlanta, <http://macroblog.typepad.com/macroblog/2014/04/new-data-sources-a-conversation-with-googles-hal-varian.html>.
- De Mauro, A., M. Greco and M. Grimaldi (2016), “A Formal Definition of Big Data Based on its Essential Features”, *Library Review*, Vol. 65., No. 3, pp. 122-135, <http://www.emeraldinsight.com/doi/pdfplus/10.1108/LR-06-2015-0061>.
- Dezyre (2015), “How Big Data Analysis Helped Increase Walmart’s Sales Turnover”, accessed in June 2016, [https://www.capgemini-consulting.com/resource-file-access/resource/pdf/walmart\\_pov\\_15\\_7\\_2015.pdf](https://www.capgemini-consulting.com/resource-file-access/resource/pdf/walmart_pov_15_7_2015.pdf).
- DOJ (2015), “Former E-Commerce Executive Charged with Price Fixing in the Antitrust Division’s First Online Marketplace Prosecution”, Press Release by the Department of Justice on Monday, April 6, 2015, [http://www.justice.gov/atr/public/press\\_releases/2015/313011.docx](http://www.justice.gov/atr/public/press_releases/2015/313011.docx).
- EDPS (2014), “Privacy and Competitiveness in the Age of Big Data: The Interplay Between Data Protection, Competition Law and Consumer Protection in the Digital Economy”, Preliminary opinion of the European Data Protection Supervisor, [https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2014/14-03-26\\_competition\\_law\\_big\\_data\\_EN.pdf](https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2014/14-03-26_competition_law_big_data_EN.pdf).

- Engels, B. (2016), “Data portability among online platforms”, *Internet Policy Review*, Vol. 5, No. 2, <http://policyreview.info/articles/analysis/data-portability-among-online-platforms>.
- eMarketer (2016), “Google Still Dominates the World Search Ad Market”, <http://www.emarketer.com/Article/Google-Still-Dominates-World-Search-Ad-Market/1014258>.
- EU Parliament (2015), “Challenges for Competition Policy in a Digitalised Economy”, Directorate General for Internal Policies, Policy Department A: Economic and Scientific Policy, <http://www.europarl.europa.eu/studies>.
- European Commission (2010), “Google”, press release, case 38740, [http://europa.eu/rapid/press-release\\_IP-10-1624\\_en.htm?locale=en](http://europa.eu/rapid/press-release_IP-10-1624_en.htm?locale=en).
- Evans, D. S. (2011), “Antitrust Economics of Free”, *Competition Policy International*, Vol. 7, No. 1, <http://dx.doi.org/10.2139/ssrn.1813193>.
- Evans, D. S. (2009), “The Online Advertising Industry: Economics, Evolution, and Privacy”, *Journal of Economic Perspectives*, Vol. 23, No. 3, pp. 37-60, <http://ssrn.com/abstract=1376607>.
- Federal Trade Commission (2014), “Data Brokers: A Call for Transparency and Accountability”, <https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>.
- Federal Trade Commission (2016), “Big Data: A Tool for Inclusion or Exclusion?”, FTC Report, <https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf>.
- Federal Trade Commission (2012), “Protecting Consumer Privacy in an Era of Rapid Change: Recommendation for Business and Policy Makers”, Federal Trade Commission Report, <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.
- Federal Trade Commission (2013), “What Information do Data Brokers Have on Consumers and How do They Use It”, Prepared Statement of the Federal Trade Commission before the Committee on Commerce, Science & Transportation, [https://www.ftc.gov/sites/default/files/documents/public\\_statements/prepared-statement-federal-trade-commission-entitled-what-information-do-data-brokers-have-consumers/131218databrokerstestimony.pdf](https://www.ftc.gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-entitled-what-information-do-data-brokers-have-consumers/131218databrokerstestimony.pdf).
- Filistrucchi, L., D. Geradin, E. V. Damme and P. Affeldt (2014), “Market Definition in Two-Sided Markets: Theory and Practice”, *Journal of Competition, Law & Economics*, Vol. 10, No. 2, pp. 293-339, <http://jcle.oxfordjournals.org/content/10/2/293.abstract>.
- Franklin, M. and L. Crofts (2016), “Privacy Should Be Weighted in Mergers, EU Watchdogs Suggests”, *MLex*.
- Friedman, D. A. (2008), “Free Offers: A New Look”, *New Mexico Law Review*, Vol. 38, No. 49, pp. 68-69, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1010238](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1010238).
- Frontier Economics (2016), “Big Data: A New Frontier for EU Competition Law”, Presentation by Frontier Economics Europe Ltd., <http://oecdshare.oecd.org/daf/competition/Knowledge%20Database/Non-OECD%20events/IBC%20AdvancedCompLaw%20Brussels%20Feb2016/David%20Foster.pdf>.

- Gal, M. S. and D. L. Rubinfeld (2016), “The Hidden Costs of Free Goods: Implications for Antitrust Enforcement”, *Antitrust Law Journal*, Forthcoming, [http://awards.concurrences.com/IMG/pdf/gal-rubinfeld\\_final2.pdf](http://awards.concurrences.com/IMG/pdf/gal-rubinfeld_final2.pdf).
- Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski and L. Brilliant (2009), “Detecting Influenza Epidemics Using Search Engine Query Data”, *Nature*, Vol. 457, pp. 1012-1014, <http://www.nature.com/nature/journal/v457/n7232/full/nature07634.html>.
- Guniganti, P. (2016), “Laitenberg: DG Comp may enforce in e-commerce”, *Global Competition Review*, <http://globalcompetitionreview.com/news/article/41886/laitenberg-dg-comp-may-enforce-e-commerce/>.
- Graef, I. (2015), “Market Definition and Market Power in Data: The Case of Online Platforms”, *World Competition*, Vol. 38, No. 4, pp. 473–505, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2657732](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2657732).
- Harbour, P. J. and T. I. Koslov (2010), “Section 2 in a Web 2.0 World: An Expanded Vision of Relevant Product Markets”, *Antitrust Law Journal*, Vol. 76, pp. 769-794, <http://www.nortonrosefulbright.com/files/us/images/publications/20100816Section2InWebWorld.pdf>.
- Harrington, J. (2008), “Detecting Cartels”, in P. Buccirossi (Ed.), *Handbook in Antitrust Economics*, MIT Press, <http://assets.wharton.upenn.edu/~harrij/pdf/DetectingCartels-10.8.05.pdf>.
- Hill, L. (2012), “How Target Figured out a Teen Girl Was Pregnant before Her Father Did”, *Forbes*, 16 February, <http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#4e4def1434c6>.
- Hildebrandt, M., and B.-J. Koops (2010), “The Challenges of Ambient Law and Legal Protection in the Profiling Era”, *Modern Law Review*, Vol. 73, Issue 3, pp. 428-460, <http://dx.doi.org/10.1111/j.1468-2230.2010.00806.x>.
- Hoofnagle, C. J. and Whittington, J. (2014), “Free: Accounting for the Costs of the Internet’s Most Popular Price”, *UCLA Law Review*, Vol. 61, pp. 606-670, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2235962](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2235962).
- International Transport Forum (2015), “Big Data and Transport: Understanding and Assessing Options”, *Corporate Partnership Board Report*, [http://www.internationaltransportforum.org/Pub/pdf/15CPB\\_BigData.pdf](http://www.internationaltransportforum.org/Pub/pdf/15CPB_BigData.pdf).
- Kimmel, L. and J. Kestenbaum, (2014), “What’s Up with WhatsApp?: A Transatlantic View on Privacy and Merger Enforcement in Digital Markets”, *Antitrust*, Vol. 29, No. 1, <http://connection.ebscohost.com/c/articles/102841624/whats-up-whatsapp-transatlantic-view-privacy-merger-enforcement-digital-markets>.
- Klishina, N. (2011), “PPC Search: Google AdWords vs. Microsoft AdCenter (Bing)”, CallFire, <http://www.callfire.com/blog/2011/01/07/ppc-search-google-adwords-vs-microsoft-adcenter-bing/>.
- Lahart, J. (2016), “How Wal-Mart’s Store Closings Paint Wider Retail Picture: Shift to Online Sales Shows Difference between Retailing’s Haves and Have-Nots”, *Wall Street Journal*, January 15, 2016, <http://www.wsj.com/articles/how-wal-marts-store-closings-paint-wider-retail-picture-1452871692>.
- Lambrecht, A. and C. E. Tucker (2015), “Can Big Data Protect a Firm from Competition?”, <http://dx.doi.org/10.2139/ssrn.2705530>.

- Lanier, J. (2013), *Who Owns the Future*, Penguin Books Limited, [https://books.google.ie/books?id=w\\_LobtmRYmQC](https://books.google.ie/books?id=w_LobtmRYmQC)
- Laney, D. (2001), “3D Data Management: Controlling Data Volume, Velocity and Variety”, Meta Group (Gartners Blog post), posted on 6 February 2001, available at <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.
- Lerner, A. V. (2014), “The Role of 'Big Data' in Online Platform Competition”, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2482780](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2482780).
- Lenhart, A. (2015), “Teens, Social Media & Technology Overview 2015”, Pew Research Center, <http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/>.
- Lipsky Jr., A. B. and J. G. Sidak (1999), “Essential Facilities”, *Stanford Law Review*, Vol. 51, No. 5, pp. 1187–1249, [https://www.criterioneconomics.com/docs/essential\\_facilities1.pdf](https://www.criterioneconomics.com/docs/essential_facilities1.pdf).
- McDonald, A. M. and L. F. Cranor (2008), “The Cost of Reading Privacy Policies”, *A Journal of Law and Policy for the Information Society*, Privacy Year in Review, pp. 540-565. [http://moritzlaw.osu.edu/students/groups/is/files/2012/02/Cranor\\_Formatted\\_Final.pdf](http://moritzlaw.osu.edu/students/groups/is/files/2012/02/Cranor_Formatted_Final.pdf).
- McLennan, M. (2016), “Netherlands Starts Big Data Probe”, *Global Competition Review*, <http://globalcompetitionreview.com/news/article/41893/netherlands-starts-big-data-probe>.
- Mandel, M. (2012), “Beyond Goods and Services: The (Unmeasured) Rise of the Data-Driven Economy”, *Progressive Policy*, [http://www.progressivepolicy.org/wp-content/uploads/2012/10/10.2012-Mandel\\_Beyond-Goods-and-Services\\_The-Unmeasured-Rise-of-the-Data-Driven-Economy.pdf](http://www.progressivepolicy.org/wp-content/uploads/2012/10/10.2012-Mandel_Beyond-Goods-and-Services_The-Unmeasured-Rise-of-the-Data-Driven-Economy.pdf).
- Manne, G. and B. Sperry (2015), “The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework”, *CPI Antitrust Chronicle*, <http://ssrn.com/abstract=2617685>.
- Manne, G. and B. Sperry (2014), “The Law and Economics of Data and Privacy in Antitrust Analysis”, *2014 TPRC Conference Paper*, <http://dx.doi.org/10.2139/ssrn.2418779>.
- Manne, G. and J. Wright (2011), “Google and the Limits Of Antitrust: The Case Against the Case Against Google”, *Harvard Journal of Law and Public Policy*, Vol. 34, No. 1, pp. 171 -244, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1577556](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1577556).
- Marel, E., M. Bauer and H. Lee-Makiyama (2014), “A Friendly Fire on Economic Recovery: A Methodology to Estimate the Costs of Data Regulation”, *ECIPE Working Paper*, No. 02/2014, <http://ecipe.org/app/uploads/2014/12/WP22014.pdf>.
- Marmer, V., A. Shneyerov and U. Kaplan (2016), “Identifying Collusion in English Auctions”, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2738789](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2738789).
- Mayer-Schönberger, V. and K. Cukier, (2013), *Big Data: A Revolution That Will Transform How We Live, Work and Think*, John Murray Publishers, Great Britain.
- McKinsey Global Institute (2011), “Big Data: The Next Frontier for Innovation, Competition and Productivity”, McKinsey & Company, <http://www.mckinsey.com/business-functions/business-technology/our-insights/big-data-the-next-frontier-for-innovation>.

Monopolkommission (2015), “Competition Policy: The Challenge of Digital Markets”, Special Report No. 68 by the Monopolies Commission pursuant to the section 44(1)4 of the Act against Restraints on Competition, [http://www.monopolkommission.de/images/PDF/SG/s68\\_fulltext\\_eng.pdf](http://www.monopolkommission.de/images/PDF/SG/s68_fulltext_eng.pdf).

Newman, N. (2013), “Search, Antitrust and the Economics of the Control of User Data”, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2309547](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309547).

Ocello, E., C. Sjödin and A. Subočs (2015), “What’s Up with Merger Control in the Digital Sector? Lessons from the Facebook/WhatsApp EU Merger Case”, *Competition Merger Brief* No. 1/2015, [http://ec.europa.eu/competition/publications/cmb/2015/cmb2015\\_001\\_en.pdf](http://ec.europa.eu/competition/publications/cmb/2015/cmb2015_001_en.pdf).

OECD (2016a), “Digital Economy”, Scoping Note prepared by the Secretariat for the Competition Committee Meetings, DAF/COMP(2016)3.

OECD (2016b), “Price Discrimination”, Background Paper prepared by the Secretariat for the Competition Committee, DAF/COMP(2016)15.

OECD (2016c), “Local Nexus and Jurisdictional Thresholds in Merger Control”, Background Paper prepared by the Secretariat for the Working Party No. 3 on Co-operation and Enforcement, DAF/COMP/WP3(2016)4.

OECD (2015a), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264229358-en>.

OECD (2015b), “Disruptive Innovation and Competition Policy Enforcement”, Background note prepared by Alexandre de Streel and Pierre Larouche for the Global Forum on Competition, DAF/COMP/GF(2015)7, [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/GF\(2015\)7&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/GF(2015)7&docLanguage=En).

OECD (2014a), *Cloud Computing: The Concept, Impacts and the Role of Government Policy*, OECD Digital Economy Papers, No. 240, OECD Publishing, <http://dx.doi.org/10.1787/5jxzf41cc7f5-en>.

OECD (2014b), “International Cables, Gateways, Backhaul and International Exchange Points”, OECD Digital Economy Papers, No. 232, OECD Publishing, <http://dx.doi.org/10.1787/5jz8m9jf3wkl-en>.

OECD (2013a), “Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value”, *OECD Digital Economy Papers*, No. 220, OECD Publishing, <http://dx.doi.org/10.1787/5k486qtxldmq-en>.

OECD (2013b), “Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by 'Big Data'”, *OECD Digital Economic Papers*, No. 22, OECD Publishing, <http://dx.doi.org/10.1787/5k47zw3fcp43-en>.

OECD (2013c), “The Role and Measurement of Quality in Competition Analysis”, *OECD Policy Roundtables*, <http://www.oecd.org/daf/competition/Quality-in-competition-analysis-2013.pdf>.

OECD (2013d). “Roundtable on Ex Officio Cartel Investigations and the Use of Screens to Detect Cartels”, Background Note prepared by the Secretariat for the Competition Committee, DAF/COMP(2013)14, [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP\(2013\)14&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP(2013)14&docLanguage=En).

- OECD (2008), *OECD Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, C(2008)36, [www.oecd.org/internet/ieconomy/40826024.pdf](http://www.oecd.org/internet/ieconomy/40826024.pdf).
- OECD (1985), “OECD Declaration on Transborder Data Flows”, *OECD Digital Economy Papers*, No. 1, OECD Publishing, <http://dx.doi.org/10.1787/20716826>.
- OFT (2013), “Completed Acquisition by Motorola Mobility Holding (Google, Inc.) of Waze Mobile Limited”, ME/6167/13, [http://webarchive.nationalarchives.gov.uk/20140402142426/http://www.oft.gov.uk/shared\\_of/mergers\\_ea02/2013/motorola.pdf](http://webarchive.nationalarchives.gov.uk/20140402142426/http://www.oft.gov.uk/shared_of/mergers_ea02/2013/motorola.pdf).
- Ohlhausen, M. K. and A. Okuliar (2015), “Competition, Consumer Protection, and the Right (Approach) to Privacy”, *Antitrust Law Journal*, Vol. 80, No. 1, pp. 121-156, <http://ssrn.com/abstract=2561563>.
- Pasquale, F. A. (2013), “Privacy, Antitrust, and Power”, *George Mason Law Review*, Vol. 20, No. 4, pp. 1009-1024, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2309965](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2309965).
- Pew Research Center (2014), “Public Perceptions of Privacy and Security in the Post-Snowden Era”, [http://www.pewinternet.org/files/2014/11/PI\\_PublicPerceptionsofPrivacy\\_111214.pdf](http://www.pewinternet.org/files/2014/11/PI_PublicPerceptionsofPrivacy_111214.pdf).
- Poeter, D. (2012), “Android a Loss Leader for Google in 2010”, *PC Magazine*, <http://www.pcmag.com/article2/0,2817,2403972,00.asp>.
- Porter, R. H. (1983), “Optimal Cartel Trigger Price Strategies”, *Journal of Economic Theory*, Vol. 29, pp. 313-338, <http://www.sciencedirect.com/science/article/pii/0022053183900509>.
- Redmond, W. (2011), “Microsoft Expands Data Platform With SQL Server 2012, New Investments for Managing Any Data, Any Size, Anywhere”, Microsoft News Center, 12 October, <http://news.microsoft.com/2011/10/12/microsoft-expands-data-platform-with-sql-server-2012-new-investments-for-managing-any-data-any-size-anywhere/#sm.001455vie5k4era114u1dxjpf0f6>.
- Schepp, N. P. and A. Wambach (2016), “On Big Data and its Relevance for Market Power Assessment”, *Journal of European Competition Law & Practice*, Vol. 7, No. 2, pp. 120-124, <http://jeclap.oxfordjournals.org/content/7/2/120.abstract>.
- Shah, R. and Kesan, J. P. (2012), “Lost in Translation: Interoperability Issues for Open Standards”, *Journal of Law and Policy for the Information Society*, Vol. 8, No. 1, pp. 113-141, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1201708](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1201708).
- Shapiro, C. and H. R. Varian (1999), *Information Rules: A Strategic Guide to the Network Economy*, Harvard Business Review Press, Boston, Massachusetts, <http://www.uib.cat/depart/deeweb/pdi/acm/arxiu/premsa/information-rules%20VARIAN%20SHAPIRO.pdf>.
- Sokol, D. and R. Comerford (Forthcoming - 2016), “Does Antitrust Have a Role to Play in Regulating Big Data?”, in *Cambridge Handbook of Antitrust, Intellectual Property and High Tech*, Cambridge University Press, [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2723693](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2723693).
- Stewart, J. B. (2015), “Walmart Plays Catch-Up with Amazon”, *New York Times*, [http://www.nytimes.com/2015/10/23/business/walmart-plays-catch-up-with-amazon.html?\\_r=0](http://www.nytimes.com/2015/10/23/business/walmart-plays-catch-up-with-amazon.html?_r=0).
- Stucke, M. E. and A. Ezrachi (2017, forthcoming), “Artificial Intelligence & Collusion: When Computers Inhibit Competition”, *University of Illinois Law Review*, <https://awards.concurrences.com/IMG/pdf/ssrn-id2591874.pdf>.

- Stucke, M.E. and A. Ezrachi (2016, forthcoming), *Virtual Competition*, Harvard University Press, United States, <http://www.hup.harvard.edu/catalog.php?isbn=9780674545472>.
- Stucke, M.E. and A.P. Grunes (2016), *Big Data and Competition Policy*, Oxford University Press, United Kingdom, <https://global.oup.com/academic/product/big-data-and-competition-policy-9780198788133?cc=fr&lang=en&>.
- Stucke, M. E. and A. P. Grunes (2015), “Debunking the Myths Over Big Data and Antitrust”, *CPI Antitrust Chronicle, University of Tennessee Legal Studies Research Paper*, No. 276, <http://ssrn.com/abstract=2612562>.
- Swire, P. and Y. Lagos (2013), “Why the Right to Data Portability Likely Reduces Consumer Welfare: Antitrust and Privacy Critique”, *Maryland Law Review*, Vol. 72, No. 2, <http://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=3550&context=mlr>.
- Tambe, P. (2014), “Big Data Investment, Skills, and Firm Value”, *Management Science*, Vol. 60, No.6, pp.1452-1469, <http://dx.doi.org/10.1287/mnsc.2014.1899>.
- Tucker, D. S. and H. B. Wellford, “Big Mistakes Regarding Big Data”, *The Antitrust Source*, [http://www.americanbar.org/content/dam/aba/publishing/antitrust\\_source/dec14\\_tucker\\_12\\_16f.auth\\_checkdam.pdf](http://www.americanbar.org/content/dam/aba/publishing/antitrust_source/dec14_tucker_12_16f.auth_checkdam.pdf).
- Varian, H. (2016), “Use and Abuse of Network Effects”.
- Varian, H. (2013), “Beyond Big Data”, presented at the NABE Annual Meeting, September 10, 2013, San Francisco, CA, <http://people.ischool.berkeley.edu/~hal/Papers/2013/BeyondBigDataPaperFINAL.pdf>.
- Varian, H. (2006), “The Economics of Internet Search”, *Rivista di Politica Economica*, Vol. 96, No. 6, pp. 9-23, <http://www.di.ens.fr/~lelarge/soc/varian2.pdf>.
- Varian, H., J. Farrell and C. Shapiro (2005), *The Economics of Information Technology: An Introduction*, Raffaele Mattioli Lectures, Cambridge University Press, <http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521605212>.
- Vestager, M. (2016), “Refining the EU Merger Control System”, Speech at the *Studienvereinigung Kartellrecht*, Brussels, 10 March 2016, [https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system\\_en](https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system_en).
- Vise, D. A. and M. Malseed (2005), *The Google Story*, Random House, Inc, New York, New York, <http://www.penguinrandomhouse.com/books/184072/the-google-story-by-david-vise-with-mark-malseed/9780385342735/>.
- Wright, J. (2004), “One-Sided Logic in Two-Sided Markets”, *Review of Network Economics*, Vol. 3, No. 1, pp. 44-64. [http://ap4.fas.nus.edu.sg/fass/ecsjkdw/wright\\_mar04.pdf](http://ap4.fas.nus.edu.sg/fass/ecsjkdw/wright_mar04.pdf).