Nexa Center for Internet & Society
Politecnico di Torino

Contestable markets and price discrimination in data-driven businesses

Massimiliano Nuccio, Università di Torino

Working paper nr 2/2017

Via Pier Carlo Boggio 65/A, 10129 Torino, Italia
(http://nexa.polito.it/contacts-en)
+39 011 090 7217 (Telephone)
+39 011 090 7216 (Fax)
info@nexa.polito.it

Mailing address:
Nexa Center for Internet & Society
Politecnico di Torino - DAUIN
Corso Duca degli Abruzzi, 24
10129 TORINO
ITALY

The Nexa Center for Internet & Society is a research center affiliated to the Department of Control and Computer Engineering of Politecnico di Torino (http://dauin.polito.it).

Table of contents

Contestable markets and price discrimination in data-driven businesses

Massimiliano Nuccio[1]
massimiliano.nuccio@unito.it

Abstract

"Data is the new oil" claimed the marketing guru Clive Humby back in 2006[2] and since then this mantra has been obsessively relaunched by academics, practitioners and policy makers. Although most of them are quite confident about the disruptive effect of data on our economies and societies, it seems we are just at the beginning of a paradigmatic change, yet uncertain of its direction and consequences.

From an economic perspective, information has been considered a major asset for a long time and a necessary assumption to allow markets to function properly. To what extent the exponential growth in data production will improve economic performances and, above all, generate benefits to firms and consumers? Are predicted benefits equally distributed or some players may take advantage of data and exploit new forms of market power?

The emergence of the Internet of Things (IoT) has raised compelling issues on the control of data flows in the human-machine relationship. We claim that the combination of a focused regulation to create a real market for data and the high innovative potential in business models can avoid monopolistic control and entry barriers to many of the rising data-driven markets.

Keywords: Big Data, IoT, entry barriers, price discrimination, data marketplace

[1] Department of Economics and Statistics, Univesity of Turin (Italy)
[2] http://ana.blogs.com/maestros/2006/11/data_is_the_new.html

1 Introduction

Many scholars and analysts had considered the diffusion of the Internet on a global scale as tremendous opportunity to change, if not even subvert, some traditional power relations in capitalistic economy. In particular, the possibility for consumers to access directly the production markets and to choose goods and services allegedly without intermediation had been celebrated as the end of long value chains of distributors which typically apply a relevant mark up on products, enlarging gap between production cost and market price. Lower transaction costs has been predicted for both producers who could eventually target their audience without paying a fee to any form of gatekeeping and final consumers who can access more information and broader markets.

From the perspective of the economics of information , the Internet, and more in general the digital revolution, has been supposed to reduce informative asymmetries and facilitate the exchange in many consumer and B2B markets, reducing the need to signal the quality of a product or the reliability of the seller. In fact, this has been seldom the case. It is probably true that many traditional oligopolistic markets have profoundly changed over the past two decades, but as a matter of fact new *digital intermediaries* have gained market power redefining the rules of their business (Smith and Telang, 2016).

The example of the cultural and creative industries can help showing the disruptive effect of data-driven innovation since most of cultural contents (books, news, videos, music, advertising, etc.) have been transformed into digital products reducing the need of any material support and dropping dramatically the cost of their production and distribution. Has this process diminished or increased competition in these markets? In the specific literature two opposite views seem to confront. Following Frank et al. (1996) and Elberse (2013) digitalization has brought more concentration where “winner takes it all” and incumbent firms can exploit non-temporary control on monopolistic or oligopolistic markets by building on entry barriers. Above all, big players can take advantage of network externalities generated by a large customer base and develop scale economies on high fixed cost production and almost zero marginal cost. These business models can thrive also because of difficulty in vertical differentiation of cultural contents which can hardly be ordered according to their technical quality. On the demand side, cultural consumption have always suffered from the “snowball effect” in taste and the subsequent rise of blockbusters (Adler, 2005).

Conversely, Anderson (2006) welcomes the positive effects of digitalization of cultural contents, namely an opportunity for producers to access a broader market and for consumers to satisfy their specific taste. Monopolistic players are only temporary because the market is extremely fragmented in successful niches, which draw a “long tail” of products and micro-producers.

This paper analyses the market of *data-driven businesses* from the perspective of industrial economics trying to address mayor concerns for competition and consumer welfare. From a legislative perspective it is important to stress that such an complex issue is regulated by at least three major laws: while the first problem is supposed to affect the antitrust law, issues on consumers should be addressed in consumer law and, eventually in privacy law, since the use of personal data allows different form of price discrimination and non-price competition. Although different data governance regimes have been developed in EU and US, both these major sources of regulation try to address similar questions: Who owns the data? Who can use and access the data?

The paper is organised around 4 paragraphs. In the first we offer some definitions of the topics discussed in order to draw the perimeter of the analysis around competition in data-driven markets. In particular, we try to understand whether the Internet of Thing (IoT) has qualitatively changed the competition in data-driven markets or it represents another side of the same Big Data coin. In the second and third paragraphs we deal respectively with the issues of entry barriers and price discrimination to show the alleged and real risks of market power in the access to sources of large behavioural data. The last paragraph discusses the interaction between economics and law to draw critical aspects that regulation should look after and possible research directions to explore the creation of a real marketplace for data.

2 Platforms, Internet of Things and Big Data

One of the major problems of *data-driven businesses*, but more in general of online competition, concerns with the definition of the perimeter of what firms do and sell or, in other words, with a clear classification of their actual industry. Apple, Alphabet (Google) and Amazon, just to mention three major players in the digital markets, are among the most capitalized and revenue-making companies in the world, although rarely at the top of rankings when we consider the number of employees. They are the champions of the network economy (Shapiro and Varian, 2013) and thrive on a daily use of huge amount of information in the form of data. Nevertheless, being data-driven organizations is not enough to claim that they compete on the same market. They may contend in some connected industries, but most of them developed different business models in which data is only one of the assets relevant for their market success.

The literature on the economics of information have broadly investigated the possible interaction between two or more groups of agents via physical or digital intermediaries, also defined *platforms*, which usually exploit cross-groups externalities. For example, Caillaud and Jullien (2003) studied dating agencies, real estate agents, and internet "business-to-business" websites; Rochet and Tirole (2003) focused on the profitable credit cards industry; Anderson (2006) analysed the determinants of equilibrium prices in two-sided markets. They all agree that digital technologies have facilitated the diffusion of these intermediaries which typically do not generate their own contents, but provide access to products and services acting alternatively as online retailers or marketplace. Nevertheless, it would be extremely misleading to conceive these new intermediaries as the mere digital transformation of traditional brick-and-mortar distribution. Although some scale and "localization" economies are still at work even in the digital markets, many fundamental aspects have changed (Belleflamme, 2016): i) many physical products have dematerialized into digital contents; ii) for many products both digital and physical we do not require the permanent ownership anymore, but we tend to enjoy their benefits as a service (e.g. movies, songs but also cars and holiday apartments) iii) the opportunities of interaction and communication have exponentially grown and affected the nature of consumption and distribution.

Haucap and Heimeshoff (2014) focused on three relevant online markets, namely search engines, auction and trading platforms and social networks, which show a relatively high degree of industrial concentration. According to the authors, the distinction into different markets is not exclusively based on the characteristics of products or customers or business models, which in fact may often overlap or look alike, but on the effects of specific drivers of industry concentration. In particular, Evans and Schmalensee (2008) found that while scale economies and indirect network effects can favour industrial concentration in these markets, capacity limits, product differentiation and the potential for multi-homing stimulate a more competitive environment.

A major impulse to the "datafication" of our economies and societies originated with the diffusion of *sensors* embedded in devices and able to capture and transmit data from humans to machines and from machines to machines. It is worth to notice that IoT does not actually define any specific industry, but labels interconnected networks of objects, which can typically be described by three layers:

i) objects/devices, for example sensors, smartphones, cars, etc.;
ii) communication networks connecting them, for example broadband, 4G, Wi-Fi, Bluetooth, etc.;
iii) computing systems that make use of the data, including storage, analytics and applications.

In these networks it is possible to distinguish between two levels: the edge, which is composed by an infinite number of endpoints (devices) and the core, which is based on cloud computing and where most of data are stored and analysed. Even from this basic definition it is evident that many industries are in fact involved and data of different nature can be extracted from different sources. Things may seem smarter because they intensively use the information which is produced by the interaction between humans and machines: what keeps the systems operating is the continuous stream of information and the possibility of adapting actions on the basis of more informed decisions. In this scenario, as suggested by Surblyté (2016), overcoming the distinction between personal and non-personal data can help conceiving *Big Data* as the pivotal resource not only in the digital economy, but also in the old economy performed by smart objects. Although some digital companies have shown a competitive advantage in gathering and processing information, it still has to be proved that they may use it to conquer dominant positions in markets where the adoption of the IoT has succeeded like energy, mobility and transportation, wearables, etc. The core of value and focus of the economic analysis should address the capabilities of transforming information to support choice making of both humans and machines.

The definition of Big Data is still in evolution and controversial since it is the result of different disciplinary contributions. This term started circulating freely in the early 2000s (Chen, et al., 2012), making its first appearance in articles and publications in 2008, though the very first signs of its usage can be traced back to the 1970s (Ularu et al., 2012). At the beginning of e-commerce Laney (2001) had proposed three challenges for data management, which have since been co-opted as a key definition based on the celebrated “3 Vs”: Volume, Velocity, and Variety.

Volume refers to the quantity of data a company or institution manages to gather (Ularu et al., 2012): in particular we talk about great agglomerations of data, too big and far too complex to be used in their raw form. It’s clear that, as the standard methods are not enough, it becomes necessary to make further improvements in the management and the deciphering process of these massive sets of data, even in terms of software innovation and capacity building (Chen et al., 2012).

Variety relates to the multiple types of data that are stored in the archives and that make up Big Data in general (Ularu et al., 2012). This concept can be read from two different perspectives: on the one hand, it seems incredibly difficult, if not impossible, to predict the infinitely different potential interactions between such big dataset, as well as all the possible combinations of data within those same dataset.

Velocity is about how fast data can be collected and then processed (Ularu et al., 2012): the speed at which data are gathered is continuously increasing, often too much for analysts to be able to keep up. Eventually, it may happen that the major part of the information collected will be wasted for lack of processing time or capabilities (LaValle al., 2011). However it has to be noted that the process of technology and software innovation (Schmarzo B., 2013), as well as personnel formation, is also picking up speed, while costs for accessing the IT support needed to analyse data are steadily dropping; this means that the whole decoding process is gaining both speed and accuracy, and the risk of throwing away valuable resources is consistently getting lower.

In some views a fourth key characteristic of Big Data can be added, that is veracity (Ularu et al., 2012). It essentially refers to the "trustworthiness" of data, the level of accuracy and truthfulness with which the data reflect reality. Regardless of it being a key characteristic or not, it is extremely important that the data gathered are as complete and as close to the truth as possible, for it will make it easier for business analysts and managers to use the information to implement successful policies.

Given the intensive use of technology, software vendors play a relevant role also in shaping the paradigm. Oracle (2013), for example, avoids employing any Vs in offering a definition, and, instead, contends that Big Data is the derivation of value from traditional relational database driven business decision-making, augmented with new sources of unstructured data. Such new sources include blogs, social media, sensor networks, image data and other forms of data that vary in size, structure, format and other factors.

The existence of multiple, ambiguous and often contradictory descriptions of the term led Ward and Barker (2013) to search for a concrete definition in order to eliminate ambiguity and to further research goals. They found three points of similarity among the various definitions of Big Data: (a) *size*: the volume of the datasets is a critical factor; (b) *complexity*: the structure, behaviour and permutations of the datasets is a critical factor and; (c) *technologies*: the tools and techniques which are used to process a sizable or complex dataset is a critical factor.

Building on this perspective, we claim two major dimensions which make a Big Data approach fruitful in applied research and particularly useful in business intelligence. First, it is deliberately conceived to deal with complex systems and describe different and conflicting aspects of human behaviour. Second, and strictly connected to this, Big Data is not merely a set of advanced techniques to analyse big dataset, but it has strong epistemological implication shaping the attitude by which we interpret social phenomena and comprehend their mutual relations. It must be clearly stressed that Big Data is not a framework only for nerds and computer scientists, but a truly multidisciplinary field of investigation. Although technological knowledge and skills are necessary to implement data management and computation, its explanatory power would be much limited without expertise both in statistical-mathematical sciences -to develop and implement algorithms and plan robust experiment- and in social sciences -to formulate relevant questions and to effectively interpret results.

3 Barriers to entry and market power in data-driven competition

According to the definition provided in the previous paragraph, it is quite clear that, if not combined in predictive models and analysed by a range of advanced computational techniques, "data, by itself, are often of low value" (Rubinfeld and Gal, 2017, 2). The aim of this paragraph is to address two crucial questions:

i) Is the access to and the exploitation of this information a real risk for competition in the respective multiple markets where these players are involved?
ii) In view of these considerations, to what extent data-driven business models thrive on informative entry barriers? Or, in other words, if we consider the supply-side, is Big Data a major obstacle to competition ?

Recalling four aspects on the nature of Big Data can help to answer the above questions.

1) Big Data as public good

Information goods by their nature are conceived as *public goods* (not rival and not excludable), although exclusion can be artificially built. For sure the use of data by one company does not prevent another company to gather and use the same data. Radinsky (2015) is particularly concerned with those businesses like search engines based on past data and historical patterns, which are not available to new competitors. Haucap and Heimeshoff (2014) have already called for increasing transparency in this specific market and suggested not to share the algorithm, which is the source of competitive advantage, but some historical data with possible competitors.

Big fails in Big Data have already shown that predictions based only on historical data can be misleading, while data even not directly connected to problem to be analysed or solved, but occurring more recently are much valuable (Radinsky and Acriche, 2016). As for the exclusive use the common practice of multi-homing (Sokol and Comerford, 2016) by consumers implies that incumbent providers do not ask for and, therefore, do not have explicit exclusivity on data.

2) Velocity: diminishing returns e real-time value.

The life of data is short. Although time-series have value in forecasting, their return on scale diminishes over time. Predictive models in consumer behaviour and risk assessment have progressively shifted to now-casting, thereby competition and technology have focused on analysing *real-time data* in many businesses. To this respect, Big Data is not a relevant advantage only for platforms and digital players, but it is becoming more strategic event to not-digital companies (Lerner, 2014; Manne and Sperry, 2015). Brick-and-mortar companies contending in insurance and banking, big and small retail, logistics, public administration etc., have adopted different devices (apps, payment systems, loyalty or membership cards, etc.) to collect and analyse information on clients, users and consumers' behaviour (Lerner, 2014). For example, the diffusion of IoT will spread quickly the adoption of iBeacons, sensors paired with mobile devices which allow the retail industry to track behavioural data even in closed and small environments and to set up commercial campaigns targeting customers in real-time mode. As widely shown in the video and movie industry, where the rise of streaming has basically replaced downloading, data should be considered as a flow more than an asset (Surblyté, 2016) and thereby, the focus of regulation should partially move from data itself to its source, to determine under which circumstances companies shall provide compulsory access to data.

3) Volume and variety: accumulation and ubiquity of data

Although volume is the major characteristic conferred to Big Data, its value is attached to its heterogeneous contents (Rubinfeld and Gal, 2017), which include structured and unstructured data, purposely and randomly collected.

Big Data is ubiquitous by definition and can be hardly compressed in proprietary data marts since it springs out of different sources (CRM, search engines, news, social media, sensors, etc.) and in different forms (numbers, text, images). Accumulation of data is not *per se* an entry barrier, if also rivals are not prevented to do the same. Drawing a comparison to the brick-and-mortar business, we learned that big stores of goods are not necessarily a condition for market power. There is no doubt that machine learning and artificial intelligence techniques assume the usage of a vast amount of information to train their algorithms, but without the selection of data dedicated to a specific problem the risk of spurious correlation, overfitting and other predictive fails is still substantial.

4) Feed-back loop: scale and network advantages

Platforms are typically easy to build but difficult in scaling and they require continuous fine tuning to be reliable and efficient. A widely accepted advantage of a Big Data approach is about the so called "reinforcing effect" able to generate direct and indirect network externalities. Direct network effects depend on the size of the network and have become very important for platforms like Facebook where users confer more value when many other friends are using the same social network. Indirect network effects are particularly effective in two-sided platforms like Ebay, since more users on one side (final customers) may attract more users on the other side (advertisers or merchants). Incumbents are supposed to benefit by cross-platform effect because, on the one hand, customers minimize their research cost when they can access a wider market for final goods and, on the other, advertisers can be connected to more customers and with the support of matching algorithms can customise marketing initiatives targeting specific groups.

As mentioned in the previous paragraph, carrying capacity and incentive mechanisms can counterbalance this effect. In fact, on the one side, customers are often disturbed by an overload of advertising and, on the other, advertisers risk a lot to put all their marketing budget into one player because they may generate negative externalities (congestion). Also, cost structure in online advertising is often based on pay-per-click strategy, which is not an incentive for advertisers to target only big platforms.

Even for search engines like Google, indirect network effect can be relevant: the more users big digital platforms attract, the more information they analyse to improve their performance, and therefore they keep on attracting even more users. Quoting Arrow, Varian (2006) argues that this not exactly a network effect, but is more akin to a "learning by doing" mechanism which rewards the most efficient and innovative firm.

To sum up, according to Lerner (2014) and Haucap and Heimeshoff (2014) network effects on competition are mixed and seem to generate only a *temporary advantage*, while evidence suggests that continuous innovation and product quality are the keys to succeed.

The nature of data and, above all, its use as competitive advantage of firms seem not to rise serious concerns on being a possible source of market power. Nevertheless, some cautions should be considered. An effective measure of monopolistic or oligopolistic behaviours in online competition based on Big Data is provided by *searching and switching costs*. Considering a value-chain perspective on Big Data markets, Rubinfeld and Gal (2017) claim that technological, legal and behavioural barriers do not exist only in the phase of collection, but may appear in the storage and usage (analysis) of information. Therefore, more than the accumulation of big quantity of data what regulators should evaluate and secure is the actual possibility for consumers to switch product or platform anytime and anywhere. Although switching costs depend also by the quality of products and, therefore, by the innovative capability of the system, this does not prevent regulators to monitor markets over some risky behaviours:

i) Platform lock-in. Adoption of standards fosters innovation only when it allows businesses and consumers to change operators and quickly access new markets. In multi-sided markets multi-homing is a not always guaranteed especially when is based on reputation mechanism which discourages sellers to change platform and lose an established market position.
ii) Data portability. Exclusive licence on personal data should be avoided and consumers should be free to grant their data to multiple parties.
iii) Third parties data. To foster competition customer data stored by some companies must be available to third parties, in particular, open data and even personal

micro-data should not be exclusive of specific players. The Payment Service Directive 2 adopted in 2015 by the EU[3] is a good example of how interoperability in banking data can foster financial service innovation and open up markets.

4 Data as a source of price discrimination

If we look at the demand-side and symmetrically to entry barriers, price discrimination is considered a second major concern attached to the rise of Big Data. Even in this case two opposite opinions confront with each other. Some believe that information on consumer behaviour harms consumers who could pay more than they would in absence of data sharing. Conversely, others believe that while keeping information private may not enhance social welfare and lower prices, in particular by sophisticated and forward-looking consumers (Kerber, 2016), more information would increase competition.

Addressing virtues and shortcomings of discriminative practices in prices inevitably embraces an evaluation of both freemium and subscription models, which are among the most diffused forms of competition in digital industries. In particular, freemium can be assimilated to a form of *versioning* where one pays only for an upgraded version of the product, while subscription is a form of *bundling* which allows consumers to choose a combination of products. The former model is an evolution of media models based on advertising: differently from the past, nowdays advertising can be targeted almost at the individual level because of the information released by users. As suggested by Manne and Sperry (2015) non-price competition relays on the fact that most consumers prefer "free and useful" to "more private". It is often claimed that free users services based on Big Data are pro-competitive benefits because they expand the market and show highly innovative potential.

While the effect of bundling (subscription model) is still controversial, the effect of serving loyal consumers within the framework of the freemium base is considered an incentive to new entries in the market.

What kind of advantage would data-driven organizations take over consumers, both

[3] http://ec.europa.eu/finance/payments/framework/

directly by *price discrimination* and indirectly by rising advertising costs, which would turn out in higher prices?

Price discrimination is the practice of charging a different price to different customers for the same product. It falls off the scope of the present paper to debate around the acceptability of such behaviour. Despite the bad implication of the term which sounds unfair, economic theory is not definitive about its alleged negative or positive effects on total welfare, efficiency and equity.

In terms of competition policy price discrimination is negative at least in two situations (Amstrong, 2008): first, when a dominant firm exploits final consumers by diminishing total or consumer welfare, and, second, when it excludes or weakens actual and potential rivals. Regarding to the former case, when markets are competitive, discrimination can generate high benefits serving price-sensitive customer with lower prices while raising prices for those willing to pay more – see for example ticket pricing in transports or cinemas.

As for the risk-based pricing, for example in insurance and health, price discrimination should be limited by law when it harms fundamental civil and person rights, or in other words, when it punishes risk factors which fall outside individual customer's control (CEA, 2015). On the other hand, when risk is mainly behavioural and could be avoided by legitimate and personal decisions, discrimination tends to predict and reward less-risky consumers.

In less competitive markets, buyers can undermine differential pricing by becoming sellers and exploiting arbitrage opportunities of buying in one market at a low price and selling in the high-priced market.

What scholars agreed upon is that Big Data analytics has favoured a shift from II- and III-degree to I-degree price discrimination (PD), although very often these strategies are not mutually excluding but combined by data-driven firms. Self-selection by consumers based on versioning of the same product or on their own socio-anagraphic characteristics (III degree PD) has often allowed a net increase in consumption of goods that otherwise would have not been purchased. By employing massive behavioural and multidimensional fine-grained data (Big Data) in segmenting consumers firms can improve prediction models (Shiller, 2014) and get close to willingness-to-pay of each group or even individuals. As suggested by Shiller (2014, 2), new equilibrium models should consider the "near-universal" I-degree PD. A report by CEA (2015) is in fact more sceptical on the actual ability of data-mining firms to predict individual willingness-to-pay. Behavioural data can help to explore the demand curve to test price elasticity under different conditions. For specific products steering practices reveal different pricing for different target groups, but in general personalized prices are not common at all. Among the reasons found by CEA (2015), reputation of incumbents plays a major role in combination with the intense use of internet to compare prices: consumers have been improving their ability to switch platforms, browse price aggregators and meta-engines and even explore secondary markets, a false positive application of I degree PD may be deeply censored by the users and can absolutely damage the image of any intermediary.

The strongest argument by opponents of discriminating practices concerns the incumbent firm with appropriating most or all of consumer surplus by getting closer to reservation prices. It is true that in a monopolistic market price discrimination can lead to efficient prices reducing consumer surplus. The counter argument stresses that in a monopolistic market with particularly high fixed cost only with price discrimination both the market is profitable and can be accessed by low-value users. In other words, extra profit by discriminating firms allows all markets not being shut down and low-value market being served too (Armstrong, 2008). More in general, economists believe that social welfare increases when discrimination actually expands the total output. From this perspective the evaluation of Big Data competition should not focus simply on the positive or negative effect of price discrimination but should include the understanding the economic mechanisms of the markets it affects.

5 Discussion

In this paper we have not deliberately mentioned the implication in privacy policy rising with Big Data because, following Sokol and Comerford (2016), we believe that the antitrust regulation should start from the economics of competition, while the economics of privacy concerns indeed with consumer and privacy laws. In paragraph 1 we show that as long as switching costs are low and technology lock-in is avoided, it is difficult to claim that Big Data can be a source of market power. At the moment for major digital (and non-digital) markets, data is an input: the online providers do not have market power in a dedicated market for data-as-a-product and thereby there is no ground to observe a monopoly. Data-driven companies and online platforms (or aggregators) have established themselves as new intermediaries in the markets of information goods which were *already* oligopolistic. The real change we can observe concerns a progressive shift with market power from product and content producers to service providers and distributors.

Scholars stressing the risk of market control by Big Data companies tend to overestimate two aspects. First, the business value of data is built around the capabilities of extracting knowledge, not its mere acquisition and storage. These capabilities are never technological or analytical only, but rely on heterogeneous skills and a new approach to continuous innovation. It is often the case that companies already generate a lot of information on their customers without need of Big Data and could possibly process it with advanced data mining and machine learning techniques. The major effort to become a data-driven firm is the pursue of a cultural and organizational change able to focus on the economic value of information.

Second, some hints suggest that the establishment of monopolistic positions is not easy to gain nor to maintain. Evidence shows that (i) the rate of innovation in digital markets is particularly high; (ii) most of major players compete in multiple markets; (iii) being a first-mover is not necessarily an advantage (e.g. Google, Amazon, Facebook); (iv) the average birth of the major players is recent (excluding Microsoft and Apple).

In conclusion, we claim that a good business model (easy, convenient and reliable) is facilitated by a data-driven approach, but *underlying solutions* are more important. WhatsApp, Tinder, Airbnb, just to quote some well-known cases, have implemented effective ideas that found the approval of consumers before using Big Data.

A different point of view on regulation (Kerber, 2016) considers antirust and privacy strictly connected in the case of Big Data and calls for an holistic approach to the topic. We can agree that given the ambiguous nature of price discrimination and the obvious concerns on privacy, consume protection and privacy law are more effective than antitrust and better fitting the issues around Big Data.

In particular, a research area that should be more investigated is the *economics of privacy*. We all know a basic difference in EU and US approach to consider this aspect. Data protection presents a trade-off in benefits and costs: while in Europe personal data and privacy tend to be considered a fundamental individual right, the US legislation is more keen to conceive their allocation as a problem of economic efficiency. In the present privacy protection the virtual price for data leads to a non-price competition (Sokol and Comerford, 2016; Manne and Sperry, 2015) where services are free in exchange for personal information. A possible solution to this controversial topic would be a regulation which accelerates the creation of a real market for personal data (Kerber, 2016). Such a regulation would leave with individuals the choice on which data to be released also on the basis of a monetary value attached to the personal information. On the one hand, the creation of a data marketplace would allow a full portability of personal data without undermining the risk of privacy connected with a competitors' access to private data given out to big platforms. On the other hand, monetization could limit the use of particular sensitive data which allows real discriminative practice based on ethnicity, gender, health conditions, etc.

According to McKinsey and Co. (2016), a key for a successful IoT is exactly the creation of a marketplace for data open to third parties and regulated by a system of licencing which can ensure the neutrality and safety of transactions. In other words, we can figure out a sort "commodification" of the information by itself: data providers could offer a platform their data which can be purchased by data users according to different possible business contracts. We believe that prospective research should follow this direction also to explore possible risks of data monetization, for example those concerning with monopsony to avoid that only a few big players may afford to purchase relevant information and, therefore, control the relative market.

6 References

Adler, M. (1985). Stardom and talent. The American economic review, 75(1), 208-212.

Anderson, C. (2006). The long tail: Why the future of business is selling less of more. Hachette Books.

Armstrong, M. (2008) "Price Discrimination" in Handbook of Antitrust Economics, ed. P. Buccirossi, The MIT Press.

Armstrong, M. (2006). Competition in two-sided markets. The RAND Journal of Economics, 37(3), 668-691.

Belleflamme, P. (2016). The economics of digital goods: A progress report. Review of Economic Research on Copyright Issues, 2016, vol. 13(2), pp. 1-24

Caillaud, B. and Jullien, B. (2003) "Chicken and Egg: Competition Among Intermediation Service Providers."RAND Journal of Economics, Vol. 34, pp. 309–328

CEA (2015) Big Data and Differential Pricing, The White House, URL: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf

Chen, H., Chiang, R.H. & Storey, V.C., (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS quarterly, 36(4), 1165-1188

Elberse, A. (2013). Blockbusters: Why Big Hits–and Big Risks–are the Future of the Entertainment Business. Faber & Faber.

Evans D. and Schmalensee R. (2008) Markets with two-sided platforms. In: ABA section of antitrust law (ed) Issues in competition law and policy. 667–693

Frank, R. H., Cook, P. J., & Rosen, S. (1996). The winner-take-all society. Journal of Economic Literature, 34(1), 133-134.

Haucap, J., and Heimeshoff, U. (2014). Google, Facebook, Amazon, eBay: Is the Internet driving competition or market monopolization?. International Economics and Economic Policy, 11(1-2), 49-61.

Kerber, W. (2016). Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection. MAGKS, Joint Discussion Paper Series in Economics, (14-2016).

Laney D. (2001) 3d data management: Controlling data volume, velocity and variety. Gartner. URL: https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N., 2011. Big data, analytics and the path from insights to value. MIT Sloan Management Review, 52(2), p.21.

Lerner, A. V. (2014). The Role of 'Big Data' in Online Platform Competition. URL: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2482780

Manne, G. A., and Sperry, B. (2015). The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework. *CPI Antitrust Chronicle, May*.

McKinsey and Co. (2016) Creating a successful Internet of Things data marketplace URL: http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/creating-a-successful-internet-of-things-data-marketplace

Oracle (2013) Big Data for enterprises. URL: http://www.oracle.com/us/products/database/big-data-for-enterprise-519135.pdf

Radinsky, K. (2015) "Data Monopolists Like Google Are Threatening the Economy." Harvard Business review. 2 march 2015. URL: https://hbr.org/2015/03/data-monopolists-like-google-are-threatening-the-economy

Radinsky K. and Acriche Y. (2016) How to make better predictions when you don't have enough data Harvard Business review. 29 december 2016 URL: https://hbr.org/2016/12/how-to-make-better-predictions-when-you-dont-have-enough-data

Rochet, J.-C. and Tirole, J. (2003) "Platform Competition in Two-Sided Markets." Journal of the European Economic Association, Vol. 1, pp. 990–1029.

Rubinfeld D. L. and Gal M.S. (2017) Access barriers to Big Data, Arizona Law Review

Schmarzo, B., 2013. Big Data: Understanding how data powers big business. John Wiley & Sons.

Shapiro, C., & Varian, H. R. (2013). Information rules: a strategic guide to the network economy. Harvard Business Press.

Shiller, B. R. (2014). First degree price discrimination using big data. Presented at The Federal Trade Commission. Working Paper 58, Brandeis University

Smith M.D. and Telang R. (2016) Streaming Sharing Stealing The MIT press

Sokol, D. D., and Comerford, R. E. (2016). Does Antitrust Have a Role to Play in Regulating Big Data? URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2723693

Surblyte, G. (2016). Data as a Digital Resource. *Max Planck Institute for Innovation & Competition Research Paper No. 16-12*

Ularu, E.G., Puican, F.C., Apostu, A. & Velicanu, M., 2012. Perspectives on big data and big data analytics. Database Systems Journal, 3(4), 3-14

Varian, H. (2006). The Economics of Internet Search. *Rivista di Politica Economica*, 96(6), 9-23.

Ward, J.S. and Barker, A., 2013. Undefined by data: a survey of big data definitions. URL: arXiv preprint arXiv:1309.5821