Big data hell or heaven? Digital platforms and market power in the data-driven economy

Massimiliano Nuccio and Marco Guerzoni

Despina -Big Data Lab- Univeristy of Turin

Email: marco.guerzoni@unito.it

1 Introduction

In the 1990s, many scholars and analysts had considered the global diffusion of the internet as a tremendous opportunity to change –if not subvert– some traditional power relations in the capitalistic economy (Rosenzweig, 1998). In particular, the possibility for consumers to deal directly with suppliers and to choose goods and services allegedly without intermediation has often been celebrated as the end of the long value chain of distributors, which typically apply a significant mark-up on products and enlarge the gap between production cost and market price (Tapscott, 1996). Many predicted lower transaction costs both for producers, who could eventually target their customers without paying a fee for any form of gatekeeping (Bakos, 1998), and for final consumers, who could access more information and broader markets (Van den Poel and Leunis, 1999). However, in the digital economy, the market structure is characterized by the presence of large incumbents and this concentration has stimulated a debate on the risk associated with market power. Influential opinion makers on the world press are also calling for an antitrust intervention[1], while a few do not believe this should be the case[2] .

Our paper focuses on the business of so-called *digital platforms*, which are the leading

---

[1] See for example: Kahn L., "Amazon Bites Off Even More Monopoly Power", The New York Times (June 2017) https://www.nytimes.com/2017/06/21/opinion/amazon-whole-foods-jeff-bezos.html; Elliot L., "Is it time to break up the tech giants such as Facebook?", The Guardian (March 2018) https://www.theguardian.com/business/2018/mar/25/is-it-time-to-break-up-the-tech-giants-such-as-facebook; Schiller B., "This Is How We Take Power Back From Facebook (And Every Other Monopoly)", Fast Company (March 2018) https://www.fastcompany.com/40508850/this-is-how-we-take-power-back-from-facebook-and-every-other-monopoly

[2] See: Worstall T., "Google And Facebook Are Dominant But Not Monopolies", Forbes, (May 2017) https://www.forbes.com/sites/timworstall/2017/05/10/google-and-facebook-are-dominant-but-not-monopolies/#8b7cb6136d4e

players in digital markets and raise concern due to their market power (Brousseau and Pènard, 2007; Smith and Telang, 2016). The competitive advantage of these companies relies on the data they can gather from their users and customers in order to fuel predictive models capable of probabilistically determining preferences and purchase behaviors. The massive and heterogeneous flows of data generated by any kind of internet transaction or digital footprint are the pivotal asset to improve customer-centered business intelligence and anticipate what consumers want.

The aim of this paper is to contribute to the ongoing debate on the risk associated with the market power of a few oligopolistic players in the data-driven economy by using theoretical arguments and empirical evidence to assess any abuse of a dominant position. Specifically, we claim that the present market structure is the predetermined result of the emergence of big data and, as such, the existence of the risk of abuse of a dominant position should be assessed by analyzing the impact of using a large amount of data on firms' behaviors.

The paper is organized as follows. In section 2, we discuss why, instead of steering the economy towards a form a perfect competition, the large exchange of information leads to market concentration. Then, we question whether market concentration in the digital economy has brought more or less societal value. We assess whether incumbent firms can exploit market power to the detriment not only of competitors, but also of consumers and the society as a whole by looking at three major aspects of economic theory:

- Entry barriers (section 3). According to Rubinfeld and Gal (2017), technological, legal, or behavioural barriers can arise in every part of the data value chain. Moreover, Mahnke (2015) is persuaded that big data can be an antitrust-relevant barrier to entry, in particular when mergers of data-intensive companies raise the risk of anticompetitive behaviour (De Peyer, 2018)

- Level of price discrimination (section 4). Has the growing availability of digital information around consumers made more room for price discrimination? If this is the case, what type of discrimination and to the detriment of which consumers? According to Newman (2014), most digital companies adopt zero nominal price exchange, exploiting a non-transparent use of data for the purposes of price discrimination, and this also harms consumers who are ignorant of pricing options.

- Intensity of innovation (section 5). To what extent do these oligopolies reduce the incentive to innovate so as not to cannibalize their products? Or, conversely, does the

presence of large incumbents generate a positive effect on product quality and stimulate continuous innovation?

Finally, in section 6 we summarize the paper's main arguments and draw a baseline for policy makes.


2   Internet, Big Data, and Platforms: how competitive are digital markets?

The idea of the internet as a global market place where consumers can access an infinite number of competitive producers with virtually no transaction costs has recently turned out to be a utopian paradigm.

From a technological perspective, Arpanet, which became the technical foundation of the Internet, was created so as to have a decentralized network with a swarm-like structure and no central authority (Rosenzweig, 1998). Levy (1984) points out the role played by students at Duke University, and even more so at Berkley with their Community Memory Project, in influencing Arpanet with the Hacker Ethic, *"in that among its values was the belief that systems should be decentralized, encourage exploration, and urge a free flow of information"* (ibid, p. 137). This melting pot of ideas converged in the "techno-libertarian approach", probably the central ideology on the Internet (Rosenzweig, 1998), and has been translated into economics by free-markets prophets such as Gilder (1990) and Toffler (1980). However, early on, Rosenzweig (1998) stated that despite *"[…] free marketers today celebrate the Internet as the home of people's capitalism, it also seems headed down the road to oligopoly"* (ibid. p. 1551). Rosenzweig focuses on the role of infrastructure and surmises that this path was not inevitable. In the remainder of this section, we argue that there exists a natural and preordained tendency towards monopolization in the internet-based economy, and that this lies in the nature of the production and exchange of information.

The Internet has been a key driver in the generation and collection of vast amounts of information, both offline and online. To paraphrase Herbert Simon, a wealth of information means a dearth of attention on the part of its recipients, since attention should be allocated wisely given the overabundance of information stimuli (Simon, 1971).

Figure 1 A framework for the evaluation of welfare loss in data-driven digital markets
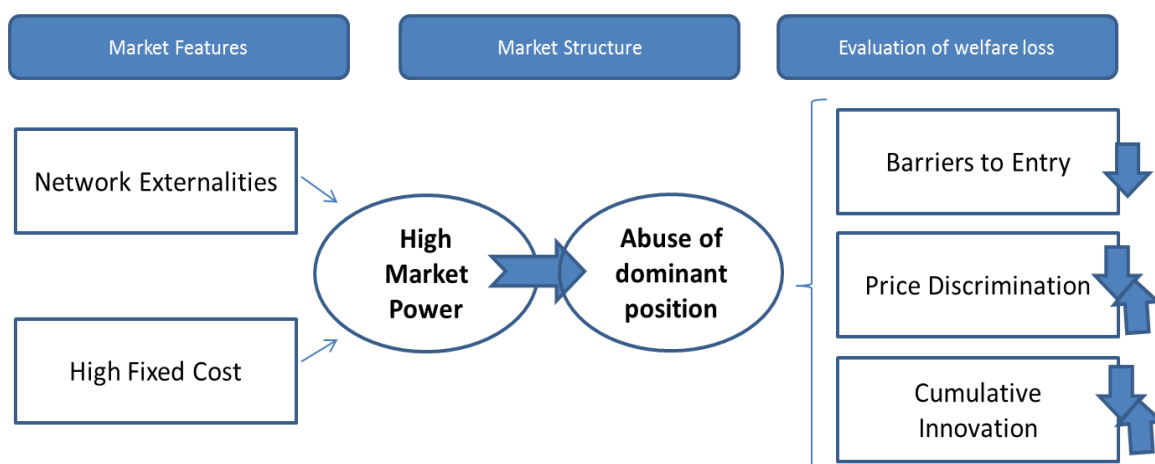
The rise of data-driven digital markets (see Figure 1) contains two seeds, both of which hinder the utopian libertarian economy foresighted by the internet pioneers. Firstly, as widely discussed in economic literature, the digital economy is characterized by two-sided markets, which benefit *network externalities* both from the demand and from the supply side (among others, see Rochet and Tirole, 2003; Evans and Schmalensee, 2013). On the demand side, consumers are overwhelmed with information and have a shortage of attention for a full screening of all available opportunities. On the supply side, firms willing to profit from growing markets need to become visible to their potential customers. These two mechanisms generate the need for new digital intermediaries, which on the one hand collect information from firms operating in the market and, on the other, organize and filter said information, supporting consumer searches. The more consumers use a specific platform, the more they make it appealing for firms; and, conversely, the more firms are visible on a platform, the more consumers are prone to choose it. Clearly, the positive feedback generated in two-sided markets abets a form of industry concentration.

Literature on the economics of information has broadly investigated the possible interaction between two or more groups of agents via physical or digital intermediaries, 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; Armstrong (2006) analyzed the determinants of equilibrium prices in two-sided markets. They all agree that digital technologies have facilitated the diffusion of these intermediaries,

who typically do not generate their own contents but provide access to products and services, acting alternatively as online retailers or as a marketplace.

A second aspect to consider is that platforms also compete in terms of the ability to leverage on big data and successfully filter same for consumers: a dating site should have an effective matching mechanism; a search engine should correctly predict what consumers are looking for; an e-commerce portal should suggest similar items to consumers. Storing a large amount of information is not enough since firms also needs to develop the ability to process information efficiently. This requires *high fixed costs* and specific investments in capital assets such as software and servers, as well as in generating analytical competencies and attracting skilled data scientists. For instance, in 2010, Google, which deals daily with a computational power equal to 30 million personal computers, developed the software MapReduce, which became one of the most widespread standards for parallel computing. In 2016, Google's capital expenditure in production equipment, facilities and datacenters peaked at 10.9 billion dollars.[3] Facebook is also setting a new standard with the 2-billion-dollar "Open Compute Project" redesigning its software and its physical network infrastructure.[4] Finally, Amazon is diversifying into cloud services, which account for a large share of the group's profits.[5]

The increased need for large investments in new knowledge in order to deal with large volumes of information generates heavy economies of scale. As clearly pointed out in traditional literature on industrial organization, the larger the economies of scale generated by fixed costs, the larger the efficient size of firms, and the smaller the number of firms in the market (Cowling 1982, Pitelis & Tomlinson, 2017). Moreover, the return on investment, either in advertising to launch a platform or in R&D, increases with market size due to the positive feedback generated by network externalities, making the sunk cost endogenous. The presence of endogenous sunk costs generates a further tendency towards market power (Sutton, 1989).

The presence of network externalities in two sided-markets and large fixed costs leads to

---

[3] Amazon Annual Report, 2016,
https://www.sec.gov/Archives/edgar/data/1018724/000101872417000011/amzn-20161231x10k.htm#sDD564D1D97D95265902D853FC83D523D

[4] Bort, J. "How Facebook is eating the $140 billion hardware market" Business Insider (June 2015)
https://www.businessinsider.com/facebook-open-compute-project-history-2015-6?IR=T

[5] Amazon has turned AWS into a business that accounts for 10% of its net revenue and generated $17.4 billion in 2017 (+43% on 2016). Amazon Annual Report, 2017,
https://www.sec.gov/Archives/edgar/data/1018724/000101872418000005/amzn-20171231x10k.htm. Sep 20th, 2018

the emergence of large companies that quickly gain market power and possibly a dominant position. The scalability of internet services is also an additional advantage, which speeds up the concentration process by attaching a high premium to being the first mover. Focusing on three relevant online markets, namely search engines, auction and trading platforms and social networks, Haucap and Heimeshoff (2014) 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 (2013) find that while scale economies and indirect network effects can favor industrial concentration in these markets, capacity limits, product differentiation and the potential for multi-homing stimulate a more competitive environment.

To sum up, platforms share similar competitive advantages. Firstly, by coordinating agents, they reduce the quantity of attention needed by consumers to navigate the vast amount of information offered by the digital economy. For this reason, a distinctive feature of platforms is to *search* for something in the infinite array of information provided by the internet, be it news, videos, cars, apartments or hotels. Indeed, it is worth noting that the home page of any internet company is designed around a well-highlighted search field.

Secondly, the same firms also share specific competences and a physical structure for analyzing customer-generated big data in order to optimize matching, marketing, and logistics, depending on the type of service offered.

The resulting concentration is striking. The world's five big tech public companies, namely Apple, Alphabet (Google), Microsoft, Amazon and Facebook, are collectively worth nearly 4.2 trillion US dollars[6] and typically compete in monopolistic markets. In 2008, Google launched Chrome, which ten years later controls about 70-80% of the worldwide search engine market, despite having a very limited presence in China, where Baidu owns 80% of the market. In 2017, Chrome (Google) dominated 59% of global browsers, followed by Firefox and Internet Explorer (Microsoft) with a 13% market share each[7]. Google and

---

[6] According to the NASDAQ, in September 2018 both Apple and Amazon reached the record in market capitalization of $1 trillion dollar, followed by Windows and Google (over 800 bn each) and Facebook (about 500 bn).

[7] Net MarketShare, 2018, https://www.netmarketshare.com; Marketing to China, 2018, The search engine market in China https://www.marketingtochina.com/search-engine-market-share-china-2018/. Last accessed: Sep 20th, 2018.

Facebook account for 73% of all U.S. digital advertising[8]. Operating Systems in the mobile segment are 99% either Android (Google) or iOS (Apple), while in the desktop market 90% of PCs run on Windows[9]. In other markets, these same companies constitute either duopolies or oligopolies, with one leader and one or more global followers. Amazon accounts for almost half of US e-commerce sales, eBay following from a distance at 6.8%[10]. Much the same happens in the cloud market, where Amazon Web Services has a 47% market share compared to Microsoft Azure's 10% or thereabouts[11].

Also, if we look at specific markets, internet-based services are taking market shares from equivalent traditional services, and tend to be very concentrated. For example, the online travel agency (OTA) market is basically a duopoly of Expedia and Priceline (booking.com), which hold two-thirds of the global online travel-booking industry and 90% of the U.S.[12] In the accommodation industry, as of November 2016, Airbnb reported having just over 3 million listings worldwide. By raw number of listings, this was nearly three times the size of the newly combined entity made up of Marriott International and Starwood Hotels & Resorts Worldwide[13]. Although the statistics are often hidden, pornography sites get more visitors each month than Netflix, Amazon and Twitter combined: Pornhub alone collects about 3 bn visits per month, which is far from the 23 bn registered by YouTube, and still slightly more than the 2.6 bn recorded by Amazon[14]. In broadcasting, millions of U.S. households are cutting the cord and switching to over-the-top (OTT) devices, including ad-led video on demand, pay-per-view and subscription video on demand, a market led by Netflix, Amazon (Prime) and Google (YouTube)[15]. Uber and Lyft are gaining market share over taxis and

---

[8] D'Onfro J., "Google and Facebook extend their lead in online ads, and that's reason for investors to be cautious" CNBC (December 2017), https://www.cnbc.com/2017/12/20/google-facebook-digital-ad-marketshare-growth-pivotal.html.

[9] Net MarketShare, 2018, https://www.netmarketshare.com Last accessed: Sep 20th, 2018.

[10] Molla R. "Amazon could be responsible for nearly half of U.S. e-commerce sales in 2017" Recode (October 2017) , https://www.recode.net/2017/10/24/16534100/amazon-market-share-ebay-walmart-apple-ecommerce-sales-2017

[11] Coles C., "Cloud Market in 2018 and Predictions for 2021", McAfee (July 2018), https://www.skyhighnetworks.com/cloud-security-blog/microsoft-azure-closes-iaas-adoption-gap-with-amazon-aws/.

[12] Quinby R., "Hotels vs. the (OTA) World", Phocuswright (May 2017), https://www.phocuswright.com/Travel-Research/Research-Updates/2017/Hotels-vs-the-OTA-World

[13] STR, 2017, "Airbnb & Hotel Performance. An analysis of proprietary data in 13 global markets". Available at http://www.str.com/Media/Default/Research/STR_AirbnbHotelPerformance.pdf

[14] Website performance rankings are provided by https://www.similarweb.com Last accessed: Sep 20th, 2018

[15] Perez S., "Netflix reaches 75% of US streaming service viewers, but YouTube is catching up" TechCrunch (2017) https://techcrunch.com/2017/04/10/netflix-reaches-75-of-u-s-streaming-service-viewers-but-youtube-is-catching-up/

rentals, their combined share accounting for 72% of the ground transportation segments for business expense reimbursement[16].

The presence of large platforms as incumbents in many digital markets raises the question whether this development comes at the detriment of consumers due to the unfair exploitation of market power. Although this topic is discussed in antitrust law and has opened a debate with public regulators (AdC and BKamt, 2016), its consequence in terms of economics of competition and innovation are pivotal to understanding the effect of data-driven companies on social welfare.

It should be noted that market concentration *per se* does not necessary lead to a welfare loss, which can only occur when oligopolistic or monopolistic incumbents exploit market power to increase prices or engage in anticompetitive behavior such as price discrimination. Even in the case of a firm exploiting its market power, for instance by charging a very high mark-up, this would still not result in welfare loss for consumers if the extra-profit generated by market power is invested in innovative activities. Indeed, if innovation activities lead to higher product quality or new products, this can even enhance consumer welfare. These considerations are well received in antitrust laws, which do not punish market power *per se*, but its abuse. Moreover, they do not punish anticompetitive behavior when this favors technological advantages.

The next sections discuss this issue from three economic standpoints that we believe are the most important in terms of assessing the position of incumbents in digital markets: the existence of entry barriers, the presence of price discrimination, and the ability to generate welfare by increasing the capacity for innovation.

3   Entry barriers and digital platforms

According to economic theory, an incumbent cannot exploit its market power to the detriment of the consumer if it cannot build durable entry barriers, therefore facing the constant risk of potential entrants (Baumol, 1986). The aim of this section is to address five crucial potential sources of entry barriers related to the nature of big data.

1. *Big data as public goods reduce entry barriers.*

    By their nature, information goods are conceived as *public goods* that are neither rival

---

[16] Richter W., "Uber and Lyft are gaining even more market share over taxis and rentals" Business Insider (July 2018) https://www.businessinsider.com/uber-lyft-are-gaining-even-more-market-share-over-taxis-and-rentals-2018-7?IR=T

nor excludable, although exclusion can be artificially built. Certainly, the use of data by one company does not prevent another company from gathering and using the same data. Radinsky (2015) is particularly concerned with those businesses like search engines, based on past data and historical patterns that are not available to new competitors. Haucap and Heimeshoff (2014) have already called for increasing transparency in this specific market and have suggested not sharing the algorithm (which is the source of competitive advantage) but some historical data with possible competitors. However, big failures in big data have already shown that predictions based on historical data alone can be misleading, while data not even directly connected with the problem to be analyzed or solved, but occurring more recently, is much more valuable (Radinsky and Acriche, 2016). As for 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 and real-time value reduce entry barriers.*

The life of data is short and having a long-term relationship with costumers does not guarantee success in the future. Although time-series analyses of behavioral data can forecast customer preferences, their return on scale diminishes over time. In fact, in many businesses (Familiar and Barnes, 2017), predictive models have progressively been shifting towards nowcasting technologies, which analyze *real-time data.* For example, real-time bidding (RTB) has emerged as a promising business model for online advertising markets (Yuan et al., 2014) since it can identify in real-time the characteristics and interests of the target audience for each impression and automatically deliver best-matched advertising at optimized prices.

In this respect, big data does not just constitute a significant advantage for platforms; it has become even more strategic to non-digital companies (Lerner, 2014; Manne and Sperry, 2015). Traditional companies in insurance and banking, large and small retailers, logistics operators and public administrations have adopted various devices (for example apps, payment systems, loyalty or membership cards) to collect and analyze information on the behavior of clients, users and consumers (Lerner, 2014). For example, a smart grid system can measure how and when consumers use the most power (Amin, 2015), and help both users and utility companies to manage and plan energy use. In the case of industrial data generated by sensors, the value is in their flow (Surblyté, 2016), and therefore, the focus of regulation should partially move from data itself to its source, to determine under what circumstances companies and public agencies should provide compulsory access in

order to safeguard competition.

3. *Volume and variety: accumulation and ubiquity of data are not an entry barrier.*

Although volume is considered the main characteristic of big data, its value is attached to its heterogeneous contents (Rubinfeld and Gal, 2017), which include both structured and unstructured, purposely and randomly collected data.

Big data is ubiquitous by definition and is difficult to compress in proprietary data marts since it springs out of different sources (CRM, search engines, news, social media, sensors) and in different forms (numbers, text, images). Accumulation of data is not *per se* an entry barrier, as long as rivals are not prevented from doing it too. By drawing comparisons with 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 require the use 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 failures is still substantial.

4. *Feed-back loop: scale and network advantages are not necessary an entry barrier.*

Platforms are typically easy to build but difficult to scale, and they require continuous fine tuning to be reliable and efficient. A widely accepted advantage of a Big Data approach is the so-called "reinforcing effect", capable of generating 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 the more friends use 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 from the 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 customize marketing initiatives targeting specific groups.

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

The indirect network effect can be significant, even for search engines like Google: the

more users big digital platforms attract, the more information they analyze to improve their performance, thereby attracting even more users. To quote Arrow, Varian (2006), the chief economist at Google, argues that this is 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 only generate a *temporary advantage*, while evidence suggests that continuous innovation and product quality are the keys to succeed.

5. *Switching costs can be a barrier.*

The nature of data and, above all, its use as the competitive advantage of firms seem not to raise serious concerns about being a possible source of market power. Nevertheless, some cautions should be considered. *Searching and switching costs* provide an effective measure of monopolistic or oligopolistic behavior in online competition based on big data. Considering a value-chain perspective on big data markets, Rubinfeld and Gal (2017) claim that technological, legal and behavioral barriers exist not only during the collection phase, but may also appear during 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 partly depend on the quality of products and, therefore, on the innovative capability of the system, this does not prevent regulators from monitoring markets for risky behaviors related respectively to platform lock-in, data portability and third-party data. Adopting standards only fosters innovation when it does not lock businesses and consumers into one platform, allowing them instead to change operators and quickly access new markets. In multi-sided markets, multi-homing is not always guaranteed especially when it is based on reputation mechanisms that discourage sellers from changing platform and losing an established market position. Also, data portability should be ensured. Exclusive licenses on personal data should be avoided and consumers should be free to grant their data to multiple parties. Eventually, in order to foster competition, the customer data stored by some companies must be made available to third parties; in particular, open data and even personal micro-data should not be exclusive to specific players. The Payment Service Directive 2 – PSD2 adopted in 2015 by the EU[17] is a

---

[17] The Directive (EU) 2015/2366 is available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015L2366

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 in parallel with entry barriers, price discrimination (PD) is considered a second major concern connected with the rise of big data in the digital economy. PD is the practice of charging a different price to different customers for the same product. Despite the negative implications inherent in the term, which sounds prejudicial, economic theory is not clear about its alleged negative or positive effects on total welfare, efficiency and equity (Kraemer and Kalka, 2017). Both second- and third-degree PD are in fact fairly accepted. Discounts based on purchased quantity are second-degree PD –e.g. buy-one-get-one-free– while in third-degree PD, a firm charges different prices for different groups –e.g. discounts for senior citizens– or consumers self-select different price segments by choosing slightly different versions of the good –e.g. color– or a different time or location of the service. Self-selection by consumers, based on versioning the same product or on their own socio-demographic characteristics, has often generated a net increase in the consumption of goods that would otherwise not have been purchased. In theory, first-degree PD, whereby a different price is set for every individual, would allow firms to extract full surplus as profits, but in fact it has always been considered fairly unfeasible, since it would be almost impossible to know the reservation price of each consumer for each good.

What scholars agreed upon is that big data analytics has favored a shift from second- and third-degree to first-degree PD, although very often data-driven firms combine these pricing strategies. By employing massive behavioral and multi-dimensional fine-grained data to segment consumers, digital platforms can get close to willingness-to-pay for very small groups or even individuals. Using data on Netflix subscribers, Shiller (2014) show that the increase in profits made feasible by personalized pricing is much higher when using web browsing behaviors rather that demographic data. Mikians et al. (2012) empirically show the effectiveness of both first-degree price- and search-discrimination by employing profiling data based on location and personal information based on internet transactions.

A report by CEA (2015) is more skeptical on the actual ability of data-mining firms to predict individual willingness-to-pay. Behavioral data can help to explore the demand curve and test price elasticity under different conditions. For specific products, steering practices reveal a different pricing for different target groups, but in general personalized prices are not at all common. Among the reasons attributed by CEA (2015), the reputation of incumbents

plays a major role, combined with an intense use of the 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 first-degree PD may be deeply censored by users and can certainly damage the image of any intermediary.

As for risk-based pricing, for example in insurance and health, the law should punish PD based on risk factors which fall outside the individual customer's control (CEA, 2015). In the European Union General Data Protection Regulation (GDPR), special categories of sensitive data are very tightly regulated to avoid harming fundamental civil and personal rights. On the other hand, when risk is mainly behavioral and could be avoided by legitimate and personal decisions, discrimination tends to predict and reward less-risky consumers.

In terms of competition policy, PD is negative in at least two situations (Armstrong, 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. Discussing the net-profit gain of dynamic pricing in data-driven competition Townley et al. (2017) conclude that "the effects on consumer welfare are ambiguous". When markets are competitive, discrimination can generate high benefits, serving price-sensitive customers with lower prices while raising prices for those willing to pay more – see for example ticket pricing in transportation or cinemas. On the contrary, one may argue that information on consumer behaviour harms consumers who might end up paying more than they would in the absence of data sharing. The strongest argument by opponents of discriminating practices concerns incumbent firms 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, thus reducing consumer surplus. The counter-argument stresses that, in a monopolistic market with particularly high fixed costs, it is only with price discrimination that the market is both profitable and can be accessed by low-value users (Armstrong, 2008). In other words, extra profit allows discriminating firms to serve all segments, including low-value markets. More generally, economists believe that social welfare, defined as the combination of consumers' and producers' surplus, might increase when discrimination actually expands the total output (Varian, 1985) and, *ceteris paribus*, consumers with a lower willingness-to-pay than the price equilibrium are offered a lower price. According to Kerber (2016), the results of theoretical models show the complexity of the effects of more information about customers. In particular, with sophisticated and forward-looking consumers (Kerber, 2016), while keeping information private may not enhance social welfare and lower prices, more information would increase competition and innovation.

In this context we should empirically observe both price dispersion, which can be a proxy for some discriminating practices, and a reduction in prices. Price dispersion in digital markets has been well documented (Gorodnichenko and Talavera, 2017), as has the overall reduction in prices due to e-commerce (Willis, 2004). These two pieces of evidence combined suggest that incumbent firms in digital markets pursue PD by lowering prices for consumers with a low willingness-to-pay, and this points to a likely positive effect of price discrimination. Thus, we can conclude that price discrimination has a positive effect on consumers. A final caveat should be made: while the effect on consumers is positive, the decrease in profit margins due to reduced prices goes to the detriment of sellers' profits. Indeed, the dominant position of an internet aggregator can impose a pricing strategy on the seller, as shown by the confrontation between Amazon and traditional publishers.[18]

5 Effects on innovation

While in a static framework, the exploitation of market power leads to welfare loss at the expense of the consumers, this is not necessarily true if we consider the dynamic efficiency gains due to the introduction of any innovation resulting from extra profits. Indeed, it could be argued that the over-profits generated by market power are a necessary condition for the R&D investments required by innovation (Acs and Audretsch, 1987).

This view has been well received in antitrust laws, which allow a firm with market power to engage in anticompetitive behaviors or to apply a high mark-up on marginal costs if this might generate more consumer value, as for instance in the case of product improvements.

At the European level, for instance, Art 101 of the TFEU clearly states that the article's provisions are inapplicable when anticompetitive behavior "*contributes to improving the production or distribution of goods or to promoting technical or economic progress, while allowing consumers a fair share of the resulting benefits*". The introduction of such a suspensive clause has sparked a decade-long debate, in the realm of economics of innovation, on the relation between market power and innovation.

The debate is rooted in the Structure-Conduct-Performance (S-C-P) paradigm, which states that firms' decisions (Conduct) are based on the structure of the industry and have a direct impact on economic and innovative performance. This approach dates back to the

---

[18] Albanese A., "How German Publishers Deal With Amazon", Publishers Weekly (October 2015) https://www.publishersweekly.com/pw/by-topic/international/Frankfurt-Book-Fair/article/68477-how-german-publishers-deal-with-amazon.html

seminal work by Arrow (1962), who showed that, in theory, a monopolist has less incentive to introduce new process innovations than firms in a competitive environment, since this might cannibalize its own position. Conversely, Dasgupta and Stiglitz (1980) adopt a game theory approach to show that, in a competitive environment, the race to innovate generates more R&D investments and, therefore, more innovation than a more oligopolistic market characterized by a few large incumbents.

Hypotheses whereby large firms with market power innovate more than small firms in a competitive market, are known as neo-Schumpeterian hypotheses, since the early Schumpeter (1912) highlighted the quasi Nietzschean role of a single entrepreneur capable of introducing innovative solutions in a competitive environment. On the other hand, the late Schumpeter (1942), influenced by the observation of the trustified American capitalism, suggested that the engine of innovation lies in large monopolistic corporations that can rely on large profits to invest in R&D laboratories (Nelson and Winter, 1982; Kamien and Schwartz, 1982).

However, empirical evidence on the neo-Schumpeterian hypotheses has been inconclusive, and a firm's sector of belonging seems to be a better predictor than market structure or the firm's size, both in terms of propensity to innovate and type of innovation (product vs. process). Consequently, many scholars (Pavitt, 1984) have suggested abandoning the S-C-P paradigm and considering market structure as endogenous to firms' choices, rather than other way around. The most interesting approach along this line is the "technological regimes" approach (Malerba and Orsenigo, 1996 and 1997; Breschi, et al., 2000), which considers the underlying technology of a sector as the main determinant of both market structure and innovative performance.

In those industries where technology opportunity levels are low, appropriability conditions are high, and the cumulativeness of knowledge is important; here, we tend to observe a few large, stable innovators ("Schumpeter Mark II"), whereas in the opposite case we observe a turbulent change in market leaders ("Schumpeter Mark I"). Malerba and Orsenigo (1997) showed that most industries can be classified in one of these two cases.

The industry life-cycle perspective takes a dynamic view on the same issue (Cohen and Klepper, 1996; Klepper, 1996). Usually, an infant industry is characterized by high opportunity, by low appropriability due to low codification of the knowledge involved in the process and, naturally, by a low degree of cumulativeness of the competencies involved. Conversely, a mature industry reveals the opposite situation. For this reason, infant industries are characterized by a high rate of entry of new innovative firms and of exit of former incumbents, while mature industries remain stable.

In the 1990s, at the time of the dot-com bubble, the internet economy was clearly a Schumpeter Mark I, in which innovation was introduced by a constant influx of new companies. However, it seems that now the low-hanging fruit of innovation has been harvested and further exploitation of technological opportunities requires large R&D investments as proven by the recent intensification of research by Facebook, Alphabet, and Amazon. Alphabet and Amazon entered the top 20 chart of largest R&D spenders in 2013, quickly reaching positions 1 and 2, respectively, in 2017[19]. According to the US Patent and Trademark Office in 2017 Google scored almost 2500 patent applications and was exceeded only by IT producers[20].

Moreover, the knowledge involved in big data analytics and the business model are rather stable, making the process of knowledge accumulation rather cumulative. According to this interpretation, we argue that market concentration in the internet economy is the result of the underlying technology, but it has allowed a fast pace of technological change that would otherwise have been impossible. This evidence makes the decisions of antitrust authorities rather difficult, as they need to carefully balance the negative and direct impact of market power exploitation against the resulting dynamic efficiency.

6   Conclusions

The technology underlying digital transformation has triggered a process of concentration in several markets in terms of information goods, and a few global players (*platforms*) have typically risen by leveraging on network externalities and economies of scale. When information is both the input and output of an industry, network externalities make a single platform an efficient tool to coordinate the market. Moreover, the exploitation of big data requires a large R&D investment in knowledge both for the creation of algorithms and for the development of the data processing infrastructure. In this paper, we discuss whether or not technology-driven dominant positions can favor the anticompetitive behavior of large incumbents.

We evaluate the risk of abuse of a dominant position by looking at three economic aspects. First, we claim that firms cannot easily abuse their dominant position because digital markets show a high degree of contestability. The nature of the data itself and rapidly changing

---

[19]   Statista (2017) https://www.statista.com/statistics/265645/ranking-of-the-20-companies-with-the-highest-spending-on-research-and-development/ Last accessed: 20th Sep 2018

[20] USPTO https://aztecip.com/patent-news/top25-2017.html Last accessed: 20th Sep 2018

technologies do not grant long-term entry barriers. Secondly, the coordinating role of incumbents and the large amount of information place them in the position of discriminating among consumers. However, we show that consumers tend to benefit from discriminating practices in online markets, since firms pursue these practices by lowering prices for the low-end markets, thus expanding the output, while leaving welfare unaltered for old consumers.

However, on looking at the huge profits generated by some of these incumbents, it would be naïve not to recognize some sort of rent exploitation. Therefore, thirdly, we highlight that antitrust laws tolerate market power exploitation when it aims to enhance consumer welfare, for instance by investing in improving existent products or developing new ones. Despite big tech companies generating billions of dollars in net incomes, they rarely –if ever– distribute dividends; so far, they sometimes employ profits for share buybacks, but mostly for R&D investments and for the acquisition of high-tech companies. Amazon, Apple and Microsoft spend more in R&D than any other company in the S&P500 index, including those in the pharmaceutical and automotive industries.[21]

We explain that, in this context, it is precisely the nature of the technological progress that generates market power, and that, conversely, only extraordinary efforts in R&D allow incumbents to stay ahead of competitors and potential entrants. Furthermore, the need for huge R&D investments is increasingly exacerbated by emerging technologies which can disrupt the market.

Data-driven companies and digital platforms have established themselves as new intermediaries in information goods markets that were already oligopolistic. The change we have observed in the digital economy concerns a progressive shift in market power along the value chain, from product and content producers to service providers and distributors.

Policy makers and scholars stressing the risk of market control by big data companies tend to overestimate the technological aspect over the human factor. The business value of data is built around the capability for extracting knowledge, and not its mere acquisition and storage. What is not stressed enough in the big data paradigm is that *"data, by itself, is often of low value"* (Rubinfeld and Gal, 2017); what is required is not just technological or analytical capabilities, but heterogeneous skills and, above all, a new approach to continuous innovation. Many companies already generate a lot of information on their customers without the need for big data and may even be able to process it with advanced data mining and

---

[21] Molla R., "Tech companies spend more on R&D than any other companies in the U.S."Recode (2017) https://www.recode.net/2017/9/1/16236506/tech-amazon-apple-gdp-spending-productivity

machine learning techniques. The major effort in becoming a data-driven firm is the achievement of cultural and organizational change, focusing on the economic value of information. Digital transformation shows that, for any (digital) business model, *underlying solutions* and *innovative culture* are more important than data. WhatsApp, Tinder, Airbnb, to quote some well-known cases, have implemented effective ideas that earned the approval of consumers before using big data.

The baseline for any data-related competition policy should consider the nature of big data: on the one hand, it has triggered market concentration, but on the other, it limits the possibility of its abuse. This claim is not a-priori apologia of large incumbents in digital markets, which should in any case be kept under control by antitrust authorities as in any other industry, but rather an attempt to argue that market concentration is a matter of fact and not necessarily detrimental. Nonetheless, the concentration of power in a few large global players should raise other concerns linked with the supranational nature of these firms, which can easily cherry-pick locations to exploit tax competition among countries or more favorable privacy legislation and the fair use of data. The recent Facebook-Cambridge Analytics case has not only shown the potential misuse of personal data for propaganda purposes; it has also stressed that trivial control over privacy issues can severely erode the trust people have placed in many big tech companies. The adoption of the GDPR in Europe in May 2018 has taken back control of personal data, allowing individuals to choose what kind of data use to allow companies, and above all it has introduced an effective set of enforcement measures which will affect companies and citizens even outside Europe.

References

Acs, Z. J., & Audretsch, D. B. (1987). Innovation, market structure, and firm size. *Review of Economics and Statistics*, 69(4) 567-74.

Adler, M. (1985). Stardom and talent. *American Economic Review,* 75(1), 208-212.

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

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

Arrow, K. (1962). "Economic welfare and the allocation of resources for invention" in The rate and direction of inventive activity: Economic and social factors, 609-26. New York: Princeton University Press.

AdC Autorité de la Concurrence and BKamt Bundeskartellamtes zu Daten und Auswirkungen auf das Wettbewerbsrecht (2016). Competition Law and Data. Available at: https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.html

Amin, M. (2015). Smart Grid. *Public Utilities Fortnightly*, March 2015 Available at: http://www.ourenergypolicy.org/wp-content/uploads/2015/06/20150604091846-Amin-Maaterials-PUF-1503.pdf

Bakos, Y. (1998). The emerging role of electronic marketplaces on the Internet. Communications of the ACM, 41(8), 35-42.

Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological regimes and Schumpeterian patterns of innovation. *The Economic Journal*, 110(463), 388-410.

Brousseau, E., & Pénard, T. (2007). The economics of digital business models: A framework for analyzing the economics of platforms. *Review of Network Economics*, 6(2).

Caillaud, B. and Jullien, B. (2003) Chicken and Egg: Competition Among Intermediation Service Providers. *RAND Journal of Economics*, Vol. 34, 309–28

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

Cohen, W. M., & Klepper, S. (1996). Firm size and the nature of innovation within industries: the case of process and product R&D. The review of Economics and Statistics, 232-243.

Cowling, K. (1982). Monopoly capitalism. London.

Dasgupta, P., & Stiglitz, J. (1980). Industrial structure and the nature of innovative activity. *The Economic Journal*, 90(358), 266-293.

De Peyer, B. H. (2018). EU merger control and big data. *Journal of Competition Law & Economics*, Available at: https://doi.org/10.1093/joclec/nhx026

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

Evans, D. S., & Schmalensee, R. (2013). The antitrust analysis of multi-sided platform businesses *NBER Working Paper* 18783.

Familiar B., Barnes J. (2017) Business in Real-Time. In: Business in Real-Time Using Azure IoT and Cortana Intelligence Suite. Apress, Berkeley, CA

Frank, R. H., Cook, P. J., & Rosen, S. (1996). The winner-take-all society. *Journal of

Economic Literature, 34(1), 133-134.

Gilder, G. (1990). Microcosm: the quantum revolution in economics and technology. Simon and Schuster.

Gorodnichenko Y. and Talavera O. (2017). Price setting in online markets: Basic facts, international comparisons, and cross-border integration. *American Economic Review*: 1;107(1):249-82.

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.

Kamien, M. I., & Schwartz, N. L. (1982). Market structure and innovation. Cambridge University Press.

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

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 562-583.

Kraemer A. and Kalka R. (2017) How Digital Disruption Changes Pricing Strategies and Price Models. In: Khare A., Stewart B., Schatz R. (eds) Phantom Ex Machina. Springer, Cham

Laney D. (2001) 3D data management: Controlling data volume, velocity and variety. Gartner. Available at: https://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. Available at: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2482780

Levy, S. (1984). Hackers: Heroes of the computer revolution (Vol. 14). Garden City, NY: Anchor Press/Doubleday.

Mahnke, R. (2015) Big Data as a Barrier to Entry, *CPI Antitrust Chronicle,* (May) 2, 11

Malerba F. and Orsenigo, (1997). Technological Regimes and sectoral patterns of innovative activities. *Industrial and Corporate Change* 6(1):83-118

Malerba and Orsenigo, (1996), The Dynamics and evolution of industries, *Industrial and Corporate Change*, 5(1):51-87.

Manne, G. A., and Sperry, B. (2015). The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework. *CPI Antitrust Chronicle,* (May) 2, 6

Mikians, J., Gyarmati, L., Erramilli, V., & Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In Proceedings of the 11th ACM Workshop on Hot Topics

in Networks (pp. 79-84). acm.

Nelson, R. R., & Winter, S. G. (1982). The Schumpeterian tradeoff revisited. *American Economic Review*, 72(1), 114-132.

Newman, N. (2014) Search, antitrust, and the economics of the control of user data. *Yale J. on Reg.*, 31, p.401.

Pavitt, K. (1984). Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy*, 13(6), 343-373.

Pitelis, C.N. & Tomlinson, P.R. (2017) Industrial organization, the degree of monopoly and macroeconomic performance – a perspective on the contribution of Keith Cowling (1936-2016), *International Journal of Industrial Organisation*, 55, 182-189

Radinsky, K. (2015) Data Monopolists Like Google Are Threatening the Economy. *Harvard Business Review*. 2 March 2015. Available at: 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 Dec 2016 Available at: 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.

Rosenzweig, R. (1998). Wizards, bureaucrats, warriors, and hackers: Writing the history of the Internet. *The American Historical Review*, 103(5), 1530-1552.

Rubinfeld D. L. and Gal M.S. (2017) Access barriers to Big Data, *Arizona Law Review*, 59(2)

Schumpeter, J.A. (1912). Theorie der wirtschaftlichen Entwicklung. Dunker & Humblot, Leipzig. The Theory of Economic Development translated by R. Opie [1934]. Cambridge: Harvard University Press

Schumpeter, J. A. (1942). Socialism, capitalism and democracy. New York: Harper and Brothers.

Shiller, B. R. (2014). First degree price discrimination using big data. Available at: http://benjaminshiller.com/images/First_Degree_PD_Using_Big_Data_Jan_18,_2014.pdf.

Simon H., (1971) Designing Organizations for an Information-Rich World. In Greenberger M. (ed.) Computers, communications, and the public interest, 37-72. Baltimore: Johns Hopkins University Press

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

Toffler, A. (1980). The third wave. New York: Bantam books.

Sokol, D. D., and Comerford, R. E. (2016). Does Antitrust Have a Role to Play in Regulating Big Data? Available at: 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

Sutton J. (1989) Endogenous sunk costs and the structure of advertising intensive industries. *European Economic Review*: ;33(2-3):335-44.

Townley, C., Morrison, E., & Yeung, K. (2017). Big Data and Personalized Price Discrimination in EU Competition Law. Yearbook of European Law, 36, 683-748.

Tapscott, D. (1996). The digital economy: Promise and peril in the age of networked intelligence (Vol. 1). New York: McGraw-Hill.

Van den Poel, D., & Leunis, J. (1999). Consumer acceptance of the Internet as a channel of distribution. *Journal of Business Research*, 45(3), 249-256.

Varian, H.R. 1985. Price Discrimination and Social Welfare. *American Economic Review*, 75(4): 870-875.

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

Yuan, Y., Wang, F., Li, J., & Qin, R. (2014). A survey on real time bidding advertising. In Service Operations and Logistics, and Informatics (SOLI), 2014 IEEE International Conference 418-423.

Willis, J.L., (2004). What impact will e-commerce have on the US economy?. Economic Review-Federal Reserve Bank of Kansas City, 89(2), 53.