UNIVERSIDADE DE BRASÍLIA

Faculdade de Direito

Curso de Graduação em Direito

João Felipe Aranha Lacerda

ANTITRUST IN THE DATA-DRIVEN ECONOMY:

IMPACTS OF BIG DATA AND ALGORITHMS ON TRADITIONAL ANTITRUST ANALYSIS IN BRAZIL

Brasília

2017

UNIVERSIDADE DE BRASÍLIA

Faculdade de Direito

Curso de Graduação em Direito

João Felipe Aranha Lacerda

ANTITRUST IN THE DATA-DRIVEN ECONOMY:

IMPACTS OF BIG DATA AND ALGORITHMS ON TRADITIONAL ANTITRUST ANALYSIS IN BRAZIL

Monografia apresentada à Banca Examinadora da Faculdade de Direito da Universidade de Brasília como requisito parcial para a obtenção do grau de Bacharel em Direito, elaborada sob a orientação do Prof. Paulo Burnier da Silveira.

Brasília

2017

UNIVERSIDADE DE BRASÍLIA

Faculdade de Direito

Curso de Graduação em Direito

Monografia apresentada à Banca Examinadora da Faculdade de Direito da Universidade de Brasília como requisito parcial para a obtenção do grau de Bacharel em Direito

João Felipe Aranha Lacerda

Banca Examinadora

____

Professor Paulo Burnier da Silveira

Universidade de Brasília

____

Professora Ana Frazão

Universidade de Brasília

____

Professora Eleanor Fox

New York University

Brasília, 8 de novembro de 2017

AKNOWLEDGEMENT

I would like to use the event of my final law degree dissertation to thank some people without whom everything would not have been possible.

First and foremost, I would like to thank my family, in particular my parents, my brother Guilherme and my grandmother Marly. Thank you for your unconditional support and the encouragement given to me. My utmost motivation is to make you proud.

I would like to thank my girlfriend Marina for all of her support and love.

Furthermore, I would like to thank my friends for making my life much more pleasant and enjoyable.

I would also like to thank Professor Paulo Burnier, who is already included in the previous paragraph but deserves a special mention as my dissertation adviser and mentor. Thank you for all the support.

Finally, I want to thank to all the great professors I had in the University of Brasília. Thank you for the knowledge and experience shared.

ABSTRACT

We are witnesses of a new economic reality. Big Data and algorithms are increasingly important in various sectors and are present in our daily lives. The economic reality has significantly changed since competition policy tools were designed. Therefore, some antitrust analysis tools should be revisited in order to adapt to this new reality. In this sense, this dissertation examines the impacts of the new economic reality – in special of Big Data and machine learning algorithms – to antitrust analysis. First, it describes the new economic reality, defining Big Data, algorithms and machine learning and showing their economic significance. Then, it analyzes the impacts of this new reality on antitrust, including merger review, abuse of dominant position and collusion.

Key-words: Competition policy; Big Data; Algorithms

RESUMO

Somos testemunhas de uma nova realidade econômica. Big Data e algoritmos são cada vez mais importantes em vários setores da economia e cada vez mais presentes em nosso cotidiano. A realidade econômica mudou significativamente desde que os instrumentos de defesa da concorrência foram pensados. Dessa forma, algumas ferramentas do direito antitruste devem ser revisitadas para se adaptarem a essa nova realidade. Nesse sentido, a presente monografia examina os impactos da nova realidade econômica – em especial, do surgimento do Big Data e de algoritmos de aprendizado de máquina (*machine learning*) – na análise antitruste. Inicialmente, a dissertação descreve a nova realidade econômica, definindo Big Data, algoritmos e aprendizagem de máquina e demonstrando sua importância econômica. Em seguida, analisamos os impactos dessa nova realidade na política de defesa da concorrência, incluindo em controle de atos de concentração, abuso de posição dominante e colusões.

Palavras-chave: Defesa da concorrência; Big Data; Algoritmos

TABLE OF CONTENTS

I – INTRODUCTION

An angry man enters into a Target store – a supermarket chain in the United States – demanding to see the manager. The angry man says: “My daughter got this in the mail! She’s still in high school and you’re sending her coupons for baby clothes and cribs? Are you trying to encourage her to get pregnant?”. The coupons contained advertisement for maternity clothing, cribs, and picture of smiling babies. The manager apologized and, a few days later, his superior called to apologize one more time. However, the father sounded embarrassed: “I had a talk with my daughter and it turns out there’s been some activities in my house I haven’t been completely aware of. She´s due in August. I owe you an apology”. Target knew the girl was pregnant before her own father.

This frequently told story[1] illustrates how companies – even in traditional businesses as supermarkets – are using Big Data and data-mining processes to collect information about its clients and to enhance their marketing strategies. In a discussion held in 2010, named “How Target Gets the Most out of Its Guest Data to Improve Marketing ROI”, Target statistician and senior manager Andrew Pole explains how the company extracts and uses all the data collected from its customers.[2] In the example above, Target identifies the pattern of purchases of pregnant women and send those coupons to the ones who fit that pattern.

Pole explains that each customer has a “guest ID”, associated with all the information regarding that person, such as name, address, usual method of payment, history of purchases (both online and in the store), mobile phone ID, actions in response to advertisement e-mails, internet browsing activity when the person uses a link provided by one of those e-mails, how people react to coupons, etc.

With this information, Target is able to determine the customer potential to spend. They can assess if someone lives in a rich neighborhood, if she has a good job or if she buys good

[1] See Forbes. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. 2012. Available at https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#57186b4f6668 and Business Insider. The Incredible Story Of How Target Exposed A Teen Girl's Pregnancy. 2012. Available at http://www.businessinsider.com/the-incredible-story-of-how-target-exposed-a-teen-girls-pregnancy-2012-2.

[2] The audio of the discussion is available at http://www.predictiveanalyticsworld.com/patimes/how-target-gets-the-most-out-of-its-guest-data-to-improve-marketing-roi/6815/

quality products. Then, if Target detects that this customer could spend more than she has been spending, Target would direct more advertisements. Target can also analyze one's sensitivity to discounts. Some people will not be affected by a 5% discount, while others will. This determines the amount of the discount coupon sent to each customer. There are many other ways of extracting value of all the information collected. The fact is that, even in traditional business, data collection and processing has become extremely important.

Nowadays we are experiencing the beginning of a new economic reality. We can watch the rise of Big Data, Big Analytics, multi-sided internet platforms, concepts such as the sharing economy and the Internet of Things, as well as the rise of the so-called Internet Giants and of algorithm-driven business, which use artificial intelligence and machine-learning.

All these new features have innumerous potential applications that can benefit humanity in various sectors, such as health, transportation, retail, education and security, just to mention a few. Nonetheless, this new dynamic also poses new challenges for governments. Some aspects of this new reality can amplify market failures, such as market power and information asymmetry, which, in turn, calls for governmental intervention. One form of intervention that has become prominent in tech markets is competition policy.

Characteristics of some markets of this new environment, such as network effects, may lead to concentration of market power and to higher barriers to entry. To give an example of the size of the so-called Internet Giants, Apple's turnover is bigger than the GDP of some countries, such as Hungary or Equator.[3] Google's market value is higher than the combined value of all companies listed in the Brazilian stock market.[4] One may argue that this concentration is a result of market failures present in data-driven sectors.

Competition authorities worldwide have been discussing the impacts of Big Data and algorithms on competition analysis. Brazil should following this trend. For instance, Google's business practices in the search engine market are under scrutiny in three Administrative

---

[3] Apple Reports. Available at http://www.apple.com/pr/library/2015/10/27Apple-Reports- Record-Fourth-Quarter- Results.html. World Bank Database. Available at http://data.worldbank.org/indicator/NY.GDP.MKTP.CD

[4] Available at http://corporate.canaltech.com.br/noticia/google/google-vale- mais-do- que-todas- as-empresas-brasileiras-na-bovespa- 55453/

Proceedings before the Brazilian Competition Authority (CADE).[5] Those are search bias cases similar to the European Commission case that was recently decided against Google.[6]

In this sense, the present dissertation will address the following question: should traditional antitrust analysis tools be revisited in order to deal with the challenges posed by the new economic reality? The answers seems to be yes.

The new challenges apply to merger review, to abuse of dominant position, to exclusionary practices and to collusion.

In this dissertation, we will analyze the impact of those new challenges in each of these areas. In short, the analysis will concern itself with the impacts of the new economic reality on traditional antitrust analysis.

In chapter II, the paper describes the new economic reality, defining concepts like Big Data, algorithms and machine learning and showing their importance and applications in today's economy.

Chapter III describes possible challenges posed by this new economic reality on traditional antitrust analysis. First, we will address the impacts on merger review and on the analysis of abuse of dominant position. Secondly, we analyze the impacts on the analysis of collusion by antitrust authorities.

Finally, we conclude that we are all witnesses of a significant transformation in the market dynamics, with impacts on competition policy. In this sense, some traditional antitrust tools should be revisited to properly address data-driven markets.

II – A NEW ECONOMIC REALITY

Data is now everywhere[7]. People and objects now constantly generate a vast amount of data. Companies collect enormous amount of data from transactions and consumer behavior.

---

[5] Administrative Proceeding nº 08012.010483/2011-94; Administrative Proceeding nº 08700.009082/2013-03; and Administrative Proceeding nº 08700.005694/2013-19.

[6] Press release available at http://europa.eu/rapid/press-release_IP-17-1784_en.htm

[7] See The Economist. "Data, data everywhere,". February 25, 2010. Available at http://www.economist.com/node/15557443

According to a McKinsey estimate[8], in 2009, almost every sector of the US economy had at least an average of 200 terabytes of stored data per company with more than one thousand employees.

Not only people and companies generate data, but also objects. Everyday objects are increasingly interconnected. The rise of the Internet of Things is a major change regarding the collection of data. We now see devices such as smartphones, energy meters, vehicles and industrial equipment all interconnected, generating and communicating data. According to Xia, Yang, Wang and Vinel:

> We are witnessing the dawn of a new era of Internet of Things (IoT; also known as Internet of Objects). Generally speaking, IoT refers to the networked interconnection of everyday objects, which are often equipped with ubiquitous intelligence. IoT will increase the ubiquity of the Internet by integrating every object for interaction via embedded systems, which leads to a highly distributed network of devices communicating with human beings as well as other devices.[9]

The production of data is becoming a usual by-product of almost all human activities. When we buy something, communicate with each other, watch a movie, travel, play games, research, etc., we create data as a by-product.

The information extracted from the vast amount of data collected has great economic importance. Information has always been a relevant source of competitive advantage between firms. The use of techniques of extracting valuable information about the market and about customers from large and complex datasets is a significant competitive advantage in today's economy. As noticed by the McKinsey report:

> The use of big data is becoming a key way for leading companies to outperform their peers. For example, we estimate that a retailer embracing big data has the potential to increase its operating margin by more than 60 percent. We have seen leading retailers such as the United Kingdom's Tesco use big data to capture market share from its local competitors, and many other examples abound in industries such as financial services and insurance. Across sectors, we expect to see value accruing to leading users of big data at the expense of laggards, a trend for which the emerging evidence is growing stronger. Forward-thinking leaders can begin to aggressively build their organizations' big data capabilities. This effort will take time, but the impact of developing a superior capacity to take advantage of big data will confer enhanced competitive advantage over the long term and is therefore well worth the investment to create this capability. But the converse is also true. In

---

[8] MCKINSEY GLOBAL INSTITUTE. Big Data: The Next Frontier for Innovation, Competition, and Productivity, June 2011, available at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

[9] XIA, F.; YANG, L. T.; WANG, L.; VINEL, A. *Internet of things*. International Journal of Communication Systems, 25(9), 1101. 2012.

> a big data world, a competitor that fails to sufficiently develop its capabilities will be left behind.[10]

With the increased economic relevance and actual capacity of data collection and processing, we shall acknowledge that we live in a new economic reality. It is a different competitive environment for market players.

In this chapter, we will describe the concept of Big Data, examine some examples of its economic importance and briefly analyze the environment in which data-intensive enterprises flourish – multi-sided platforms. We will also analyze the use of algorithms and machine learning to extract value from Big Data.

A) BIG DATA AND MULTI-SIDED MARKETS

a.1. Defining Big Data

There is no consensus regarding the definition of Big Data. The McKinsey report presented the following definition:

> "Big data" refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don't define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes).[11]

However, we believe the most popular definition of Big Data refers to the volume of data, the velocity at which the data is collected and processed, the variety of information aggregated and its value[12]. This definition is called the "4Vs definition" (volume, velocity, variety and value).

[10] Idem.

[11] Idem.

[12] OECD. Big Data: Bringing Competition Policy to the Digital Era. Background Note by the Secretariat. DAF/COMP(2016)14. Available at https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf

(i) Volume

According to a report made by the American tech firm Cisco, the annual global data center IP traffic will reach 10.4 zettabytes by the end of 2019, from 3.1 zettabytes in 2013.[13] Stucke and Grunes put this amount in perspective stating that it would require more than 73 billion iPhones 6, with the largest storage, in order to store 8.6 zettabytes (the forecast for 2018)[14].

As the OECD stated, an IBM estimate suggest that nowadays more than 2.5 exabytes (a billion gigabytes) of data are generated every day, which is comparable to 167.000 times the information stored in all the books of the Library of the Congress in the US[15].

The migration of various social and economic activities to the internet is a reason why the amount of data produced has increased and will continue to increase exponentially. Moreover, this trend is likely to continue due to the fall in the cost to store and process large sets of data.

To illustrate this point, the OECD stated that in 2014, Facebook had over 900 million active users worldwide, who generate more than 1500 status updates per second, on average.

With the rise of the Internet of Things, those figures tend to increase exponentially, since almost every object will be transformed in a connected data generator. Our cars, homes, phones and even clothes will be constantly connected and providing data. As the Federal Trade Commission Chairwoman Edith Ramirez noted in 2014:

> We are at a pivotal stage in the information age. Thanks to smartphones and smart meters, wearable fitness devices, social media, connected cars, and retail loyalty cards, each of us is generating data at an unprecedented rate. In fact, in 2013 it was reported that an astonishing 90 percent of the world's data was generated in the two preceding years. Today, the output of data is doubling every two years[16]

---

[13] CISCO. Cisco Global Cloud Index: Forecast and Methodology, 2013–2018 (2014). Available at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_ White_Paper.pdf.

[14] STUCKE, Maurice; GRUNES, Allen. *Big Data and Competition Policy*. Oxford University Press. 2016. p 17.

[15] OECD. *Big Data for growth and well-being*. Interim Synthesis Report. 2014.

[16] Opening Remarks FTC Chairwoman Edith Ramirez Big Data: A Tool for Inclusion or Exclusion? Washington, DC September 15, 2014. Available at https://www.ftc.gov/system/files/documents/public_statements/582421/140915bigdataworkshop.pdf

(ii) Velocity

Together with the increase in the volume of data collected, the velocity at which it is processed and analyzed is approaching real time. The phenomenon of using real time data analysis to predict future events is known as "nowcasting". The term is defined by Banbura et al (2013) as "*the prediction of the present, the very near future and the very recent past*"[17] and by Google's Chief Economist as "*contemporaneous forecasting*"[18].

According to the OECD:

> it consists in the use of new, up-to-date and high-frequency data to produce early estimates, usually with great degree of accuracy, about events that are taking place very close to the present. Now-casting is particularly useful to obtain close to real-time information about relevant variables that are normally collected at low-frequency and published with a great lag.[19]

One frequently used example of nowcasting is the prediction of flu epidemics based on data from Google searches of related terms such as medicine names. This is a suitable case for nowcasting, since the official reports about infection normally are only published with a considerable lag.

Other application of nowcasting is the prediction of traffic jam based on data provided by GPS users. Waze, for instance, has a business model based on the instant analysis of data provided by the drivers. In the real estate market, Google invested in an online real estate auction company, who will analyze Google's data in order to predict the direction and trends in the market[20].

Finally, nowcasting can also be used for regulatory purposes. The Eurostat, in a paper named "Big Data and Macroeconomic Nowcasting", conducted a study regarding the application of big data analysis to allow policy-makers and market players to forecast macroeconomic variables.

[17] BANBURA, Marta; GIANNONE, Domenico; MODUGNO, Michele; REICHLIN, Lucrezia. *Now-Casting and the Real-time Data Flow*. ECB Working Paper Nº 1564. 2013. p. 2.
[18] STUCKE; GRUNES, Op. cit. p. 19
[19] OECD. *Big Data: Bringing Competition Policy to the Digital Era*. Background note by the Secretariat. DAF/COMP(2016)14. 2016.
[20] GITTELSOHN, John. *Google Data to Help Auction.com Predict Homebuying Trends*. Bloomberg Business, 30/10/2014. Available at http://www.bloomberg.com/news/2014-10-30/google-data-to-helpauction-com-predict-homebuying-trends.html;

(iii) Variety

The variety of the data collected by companies has also increased. Nowadays, companies collect more than just basic information about its customers. Due to increased volume of data collected and to augmented processing power, companies can extract a broad myriad of information about a single customer, sometimes after the combined analysis of the data provided. With data about age, name, address and purchase history, a firm is able to infer the customer's financial capacity, personal preferences, household composition, dietary habits, etc.

The commercial strategies of famous retailers increasingly focus on acquiring and analyzing consumer data in order to better understand its client's behavior. Target[21], Whole Foods[22] and Amazon[23] already use these strategies in order to obtain competitive advantage in relation to old-fashioned brick-and-mortar retailers.

(iv) Value

The value of Big Data derives from the increased volume and velocity at which the data is collected and processed and from the variety of the data aggregated. More importantly, however, is the capacity of algorithms to access, analyze and extract organized and valuable information from large and complex datasets – which is known as Big Analytics. According to Federal Trade Commission Chairwoman Edith Ramirez:

> Today, the output of data is doubling every two years. Advances in computational and statistical methods mean that this mass of information can be examined to identify correlations, make predictions, draw inferences, and glean new insights.[24]

---

[21] See Target´s Senior Manager Andrew Pole´s speach, available at http://www.predictiveanalyticsworld.com/patimes/how-target-gets-the-most-out-of-its-guest-data-to-improve-marketing-roi/6815

[22] NORTON, Steven. Big Companies Rein in Data Sprawl: Whole Foods, GE and Others Merge Disparate Data Sets to Cut Costs, Learn About Customers. Wall Street Journal, 21 October 2015. Available at: http://www.wsj.com/articles/big-companies-rein-in-data-sprawl-1445482009

[23] PACKER, George. Cheap Words: Amazon is good for customers. But is it good for books?. The New Yorker. February 2014. Available at http://www.newyorker.com/magazine/2014/02/17/ cheap-words

[24] Opening Remarks FTC Chairwoman Edith Ramirez Big Data: A Tool for Inclusion or Exclusion? Washington, DC September 15, 2014. Available at https://www.ftc.gov/system/files/documents/public_statements/582421/140915bigdataworkshop.pdf

As explained by Ezrachi and Strucke, Big Data and Big Analytics have a "*mutual reinforcing relationship*".[25] The ability to rapidly analyze and extract organized information from the data is what enhances its value. An immense dataset would be useless if no one could access and extract valuable information from it.

On the other hand, Big Analytics algorithms can only perform well because they have at their disposal a great volume of data. Self-learning algorithms, in special, need a vast amount of data in order to function well and to increase its capacity to better analyze data. Therefore, the more data we have, the better we can analyze it and, as a result, the more valuable the data.

Self-learning algorithms, which will be further explained in section B of this chapter, are increasingly important to extract value from Big Data. As noticed by Ezrachi and Strucke, investment in Artificial Intelligence start-ups increased more than 300% on a year over year basis.[26] Self-learning algorithms learn based on experience and, hence, need large data sets to analyze in order to learn.

According to the European Data Protection Supervisor: "*deep learning computers teach themselves tasks by crunching large data sets using (among other things) neural networks that appear to emulate the brain*".[27]

In conclusion, Big Data, combined with Big Analytics, is increasingly valuable. The information extracted from large sets of data has potential application in various fields. Moreover, the control over strategic data can mean significant competitive advantage to firms. These topics will be addressed in the next section.

a.2. The economic importance of Big Data

With the decrease in the costs to obtain, store and analyze data, more social and economic activities will migrate to the data-driven environment. In the words of Stucke and

[25] EZRACHI, Ariel; STUCKE, Maurice. Virtual Competition. *The promise and perils of the algorithm-driven economy*. Cambridge, Massachusetts. Harvard University Press. 2016. p 16.
[26] Idem. p. 258, note 38. See also "Artificial Intelligence Startups See 302% Funding Jump in 2014", CBInsights. Available at https://www.cbinsights.com/research/artificial-intelligence-venture-capital-2014/
[27] THE EUROPEAN DATA PROTECTION SUPERVISOR. *Towards a new digital ethics Data, dignity and technology*. Opinion 4/2015. Available at https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf

Grunes, "*we are entering the age of datafication, where we take all aspects of life and turning them into data*".[28]

For instance, a report made by McKinsey[29] showed that personal geo-location data is increasingly more abundant with the spread of mobile phones. This fact creates relevant opportunity for many fields of applications. McKinsey identified three categories of useful employment of personal location data: (i) location-based applications and services for individuals; (ii) organizational use of individual personal location data; and (iii) macro-level use of aggregate location data.

In the first category, the report mentions examples of smart routing applications, automotive telematics and mobile phone location based services.

Smart routing application, based on real-time traffic data, is an important application of Big Data for consumers. The navigation mobile applications use the information instantly extracted from the data to nowcast traffic jam, congestions, accidents, etc. With this information and with the mapping data, the navigation system can suggest alternative routes to save driver's time and fuel.

With the growth in the use of smartphones and free navigation applications, the estimate of value that will be generated by those applications is significant. McKinsey report estimates that the potential value of smart routing will be around US$ 500 billion by 2020 in fuel and time savings, which is equivalent of 20 billion hours saved in traffic and about US$ 150 billion in fuel consumption. This would have associated environmental benefits, with an estimated reduction in carbon dioxide emission of 380 million tons.[30]

Furthermore, location data can be applied to automotive telematics, with impacts on personal and property safety. For example, the possibility to monitor the location of a vehicle can significantly inhibit car theft.

In the organization use of personal location data, an example would be geo-targeted advertising. In the example given by the McKinsey report:

> Consumers who choose to receive geo-targeted ads might have a personalized advertisement for a favorite store pop up on their smartphone when they are close to that store. Or a smartphone user meeting with friends at a bar or restaurant might receive a coupon offer for drinks or food from that establishment. This technology could direct users to the nearest ATM, provide

[28] STUCKE; GRUNES. Op. Cit. p. 24
[29] MCKINSEY. Op. Cit. p 85
[30] Idem. p 89

> location and time-based restaurant reviews, and offer a range of special offers for stores based on the smartphone user's location or destination[31]

Personal location data can also be applied to insurance pricing, in order to better discriminate the behavior of individual policyholders. Even public authorities' response to emergencies can be a field of application.[32]

Moreover, aggregate location data can be used for macro-level purposes, such as urban planning. Information about traffic can indicate the best place to build a new road, for instance.

Personal location data is just an example of the myriad of potential efficiency gains from the use of Big Data. The OECD produced an extensive book – Data-Driven Innovation, Big Data for Growth and Well-Being – about real and potential applications of Big Data in various sectors, such as infrastructure, healthcare, science, public sector, etc.[33]

According to the OECD:

> Using Big Data is also useful for businesses to generally improve the efficiency of production processes, forecast market trends, improve decision-making and enhance consumer segmentation, through target advertising and personalised recommendations. Although the efficiency gains from data driven innovation [DDI] are inherently hard to measure, some studies suggest that DDI users benefit, on average, from a 5% to 10% faster productivity growth than similar companies that do not use DDI.[34]

Considering the important value of Big Data in the economy and the innumerous possibilities of application in various fields, Big Data is increasingly a competitive significant factor. Companies can acquire and sustain relevant competitive advantage in the market due to the control of consumer data. As noted by European Commission officials:

> In the digital economy, large sets of data (so-called 'big data') are becoming increasingly valuable as they reveal patterns of information that enable companies to understand user behaviour and preferences and improve (or target) their products and services accordingly. This makes the availability of 'big data' a significant competitive advantage for companies active in, for instance, targeted online advertising, online search, social networking services and software products.[35]

Ezrachi and Stucke illustrate this point by mentioning the Walmart-Amazon example[36]. Back in the early 2000s, Walmart was the prominent example of market power. Its purchase

---

[31] Idem. p 90
[32] Idem. p. 91
[33] OECD. Data-driven innovation. Op. Cit.
[34] OECD. Big Data. Op. Cit. p. 8
[35] OCELLO, Eleonora; SJÖDIN, Cristina; SUBOČS ,Anatoly. *What's Up with Merger Control in the Digital Sector? Lessons from the Facebook/WhatsApp EU Merger Case*. 1 Competition Merger Brief. February 2015. p 6.
[36] EZRACHI; STUCKE. Op. Cit. p. 11-15.

power could “make or break” suppliers. Few years later, however, Walmart announced it would close 269 stores globally.

What happened is that a significant part of its customers was migrating from brick-and-mortar stores to online shopping. The company who was attracting many of those customers was Amazon. In the article “Walmart plays catch-up with Amazon”, published in The New York Times in 2015[37], Jim Stewart states that “*Walmart was once a disrupter in its own right but now it finds itself falling behind in the race for online customers*”. The article also quotes a financial analyst, who affirms that “*with every passing year, it becomes harder and harder for Walmart to compete with Amazon*”.

In 2017, Amazon market value was almost the double of Walmart’s.[38] Ezrachi and Strucke stated that the “*sentiment is that Walmart´s distributional efficiencies from its brick and mortar store model do not translate to the data-driven analytics and dynamic pricing of the online world*”.[39] In order to illustrate the competitive significance of Big Data, the authors compare Amazon’s and traditional retailers’ business models.

Besides other differences, Amazon uses dynamic pricing algorithms that constantly collects and analyses customers and market data in order to set prices in the most efficient level.[40]

Walmart, nowadays, is also investing in Big Data to compete in an increasingly data-driven market. Since Walmart is still one of the biggest retailers, it has the ability to collect a vast amount of data from its customers. As mentioned by the OECD:

> It also uses Big Data to improve operational efficiency. Walmart collects around 2.5 petabytes of data per hour and is estimated to have increased online sales by 10% to 15% as a result of data analytics (Dezyre, 2015).
>
> Walmart collects consumer data about historical purchases, living location, clickable actions / keywords entered in the website, as well as information from social networks. Then, using data mining, it analyses the pattern of consumer data and crosses it with information about other events (such as sports, weather…), in order to improve predictive analysis, launch new products and provide personalised recommendations. There are several creative ways through which Walmart leverages Big Data. To name just a few, the company uses demand estimation to improve inventory management and shipping policy; it launches entrepreneurship contests in social media in order

[37] STEWART, Jim. *Walmart plays catch-up with Amazon*. The New York Times. 22.10.2015. Available at https://www.nytimes.com/2015/10/23/business/walmart-plays-catch-up-with-amazon.html?mcubz=0
[38] BUKHARI, Jeff. *Amazon Is Worth More Than Walmart, Costco, and Target Combined*. Fortune. 05.04.2017. Available at http://fortune.com/2017/04/05/amazon-walmart-costco-target-market-cap/
[39] EZRACHI; STUCKE. Op. Cit. p. 12.
[40] Idem. p. 13.

> to place the most popular products in the shelves; and sends recommendations to consumers of gifts for friends based on their Facebook profiles.
>
> A familiar example of effective data mining through association rule learning technique at Walmart is finding that Strawberry pop-tarts sales increased by 7 times before a Hurricane. After Walmart identified this association between Hurricane and Strawberry pop-tarts through data mining, it places all the Strawberry pop-tarts at the checkouts before a hurricane.[41]

In conclusion, there is a growing consensus about the competitive importance of Big Data in today's economy. Information extracted from Big Data is becoming a crucial asset in various sector.

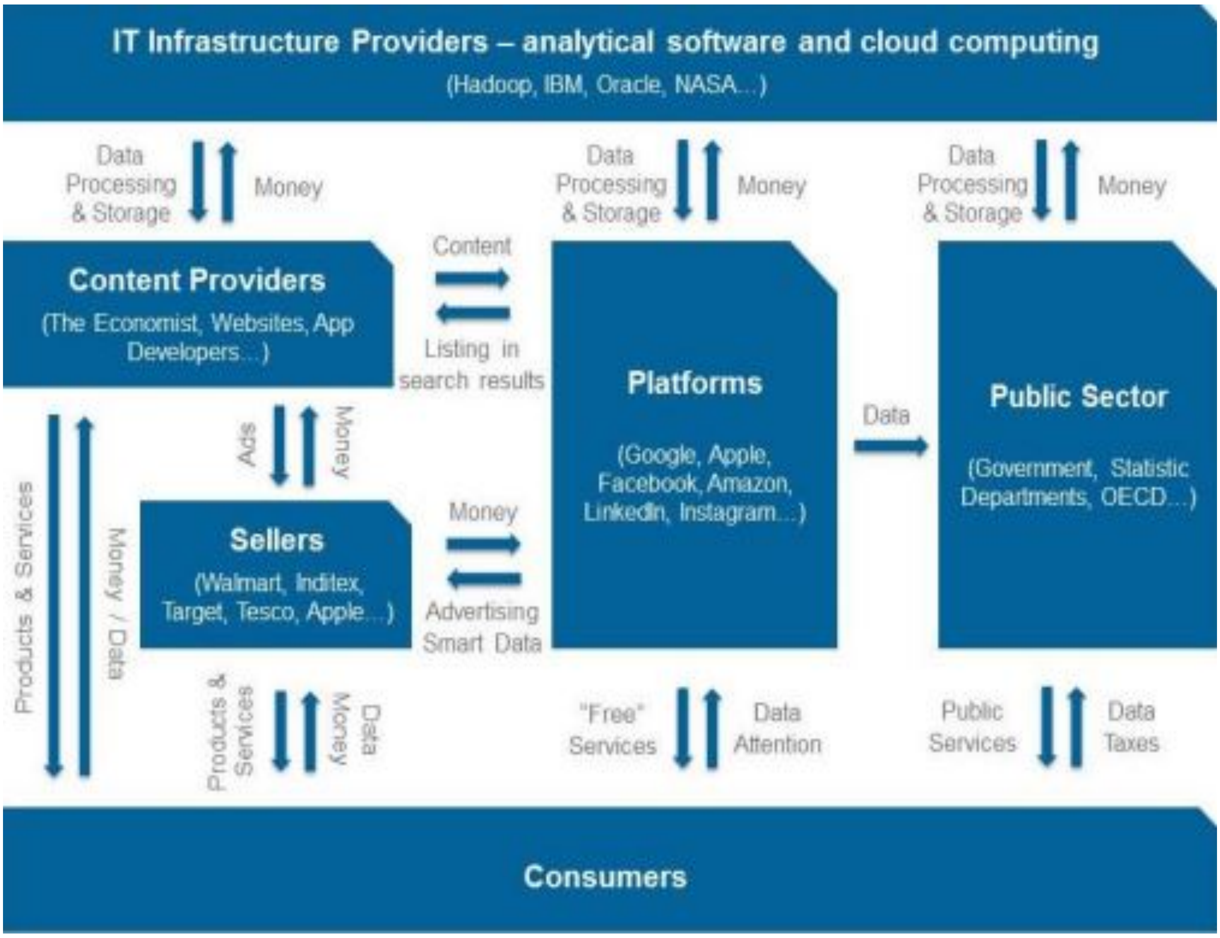
In the next section, the ecosystem in which Big Data is collected and used will be examined. It is made of interconnected, often multi-sided, markets. Understanding the agents involved in the Big Data ecosystem is important in order to have a clear view of the big picture and of the dynamic of those markets.

a.3 The Big Data ecosystem: multi-sided markets

The OECD made a didactic graphic representing the Big Data ecosystem and the interactions that take place.[42]

[41] OECD. Big Data. Op. Cit. p. 8. Citing DEZYRE. *How Big Data Analysis Helped Increase Walmart's Sales Turnover*. Available at https://www.capgemini-consulting.com/resource-fileaccess/resource/pdf/walmart_pov_15_7_2015.pdf.

[42] Idem. p 12.



We will start by the platforms, which is the category that interacts with all agents in the Big Data ecosystem. Platforms are tools that connect or provide services to different groups of consumers, in multi-sided markets. They can be divided in two groups: attention and matching platforms[43].

Attention platforms usually provide non- paid services for users and, on the other side, sell targeted advertising to companies. Examples are social networks, like Facebook, search engines, such as Google, and video platforms, such as YouTube. Users do not pay a monetary price. Instead, they pay with attention and personal data, which the platform sells to advertisers.

Matching platforms provide a common place in which interested users can interact and transact. Examples are e-commerce platforms, job offers websites and online dating platforms. The company can charge a fixed fee to allow access to the platform or charge a fee per transaction. The users also provide the platform with their personal data.

[43] Idem. p. 12.

It is worth noting that those two types of platforms can be mixed in some cases. For instance, an online dating website that has advertisement would have aspects of both attention and matching platforms.

In both cases, there is a so-called feedback loop, associated with network effects. In other words, an increase in the amount of users of a platform results in more data to be collected and, consequently, in improved service quality, which, by its turn, attracts new users. Moreover, the higher the number of customers in one side of the market, the more valuable the platform for the other side. For example, the more people Facebook has as users, the higher the benefit for a new user and for an advertiser and the higher the amount of user data collected. This, by its turn, leads to a better targeting of the advertisement, increasing, once more, the benefit for the advertisers.

According to the OECD, "*the multisided features of platforms tend to lead, as a result of direct and indirect network externalities, to the concentration of users and their respective data in the hands of a few players*".[44]

Together with platforms, content providers are another category of agents in the Big Data ecosystem. They are the ones who produce informative and entertainment content available in the platforms, such as journals, news agencies, websites and application developers. An example would be The New York Times. They provide content to platforms such as search engines and social networks. They normally make money by direct sales or by advertising in their website. Nonetheless, as noted by the OECD, since they do not have the Big Data necessary to provide proper targeted advertising, it is common for them to run advertisements through the platform (such as Google's)[45].

In addition, there are the sellers, who are the ones who advertise their products and services through the platforms in order to persuade consumers to buy their products and services. This category includes companies from almost every field of economic activity. They are the main subsidizers of the other agents in the ecosystem.

There are, as noticed by the OECD[46], sellers that are large enough to collect Big Data themselves and extract value from it. As already mentioned, Walmart and Target, for instance, have Big Data strategies for pricing and marketing.

[44] Idem. p. 13.
[45] Idem. p. 13.
[46] Idem. p. 13

Furthermore, the Big Data environment heavily depends on information technology (IT) infrastructure providers, including cloud computing and storage. They are necessary because data-driven companies have to collect and analyze vast amount of data, so they outsource those activities to technology companies like IBM and Oracle.

The development of this market for IT infrastructure, as explained by the OECD[47], mitigated the problem of scale regarding Big Data, by converting high fixed costs into variable costs, which allowed small companies to operate without the need of owning an expensive IT infrastructure.

Finally, the public sector is a player who can collect great amount of Big Data from its citizens and companies, normally in the context of public services or due to legal obligations. In the words of the OECD, "*the public sector is, indeed, one of the most data-intensive sectors of the economy, using national databases for scientific research and to support the provision of public services*".[48] There is a significant potential to extract value from the Big Data held by governments.

A good illustration of the potential application of Big Data by governmental bodies is the Project Brain (Projeto Cérebro) in the Brazilian Competition Authority (CADE). The project uses Big Data and Big Analytics in order to identify patterns in public procurements that indicate a possible collusion between the participants. The Project has attracted the attention of authorities from other jurisdiction, such as Portugal and Switzerland.

This year, the information extracted from the Big Data within Project Brain led to the opening of a cartel investigation on the markets of orthosis, prosthesis and special medical materials.[49] According to an interview of the former acting General Superintendent Diogo Thompson, the intention is to develop intelligence methods in order to rely less on leniency agreements to fight cartels.

This project is an example of the enormous potential for the use of Big Data by governments. Normally, governmental data is disorganized and underused.

[47] Idem. p. 14.
[48] Idem. p 14.
[49] Administrative Proceeding 08700.003699/2017-31; Administrative Proceeding 08700.003709/2017-38. Press release available at http://www.cade.gov.br/noticias/superintendencia-instaura-processos-para-apurar-cartel-no-mercado-de-orteses-proteses-e-materiais-medicos-especiais

B) PRICING ALGORITHMS AND MACHINE LEARNING

As already mentioned, Big Data is only valuable because there are advanced algorithms that can process and extract information from it. The development of Big Data is interconnected with the development of fast and smart algorithms – known as Big Analytics.

Algorithms that learn from experience can only develop if they have a large data set to use as input. The value of Big Data and of self-learning algorithms are self-reinforcing. The more data is available, the better the algorithms will become and the more value they will be able to extract.

In the following section, some basic concepts regarding algorithms and machine learning will be explained. Furthermore, we will examine some applications of self-learning algorithms in the economy.

b.1 Defining algorithms and machine learning

According to the simple definition of the OECD, "*an algorithm is a sequence of rules that should be performed in an exact order to carry out a certain task. Thus, an algorithm is an instance of logic that generates an output from a given input* [...]"[50].

With the development of computers, algorithms were written into codes that computers can read and perform the steps in order to execute tasks in high speed, compared with humans.

That is why nowadays, algorithms are generally understood as related to computer programs. According to the Merriam-Webster dictionary, an algorithm is "*a step-by-step procedure for solving a problem or accomplishing some end especially by a computer*".[51]

With increased computer power, complex tasks that humans were not able to complete, are easily and rapidly executed by algorithms. This is possible due to Artificial Intelligence.

Artificial Intelligence (AI) is the science of creating intelligent machines. According to Swarup, this field of study is based on the idea that human intelligence can be so precisely

[50] OECD. *Algorithms and Collusion - Background Note by the Secretariat*. DAF/COMP(2017)4. 2017. p 6. Available at file:///Users/joao.alacerda/Downloads/Algoritms%20and%20collusion%20OECD.pdf. p 6
[51] Merriam-Webster Dictionary. Available at https://www.merriam-webster.com/dictionary/algorithm

described that machines can simulate it[52]. This fact raises numerous philosophical questions. Alan Turing, in its famous 1950 paper, proposed the question "can machines think?", stating that to address this question, one should first define "machine" and "think", which is not a simple task[53].

According to the OECD, Artificial Intelligence developed significantly after the creation of algorithms that teach machines to learn.[54]

Machine learning is a subfield of AI, which develops machines that are able to learn by applying algorithms to data and experience[55]. In the precise words of Anitha, Krithka and Choudhry:

> Machine learning is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs and using that to make predictions or decisions, rather than following only explicitly programmed instructions[56]

Depending on the nature of the signal or feedback available to learning system, machine learning tasks can be separated in three categories[57]: (i) supervised learning, in which the algorithm is presented with a sample of inputs and their outputs in order to extract a general rule that maps inputs to outputs; (ii) unsupervised learning, in which the algorithm has to find hidden patterns from unlabeled data; and (iii) reinforcement learning, in which the algorithm interacts with a dynamic environment and has to perform a task, learning through trial and error.

One frequently told example of machine learning is the program Libratus, developed by Carnegie Mellon University to play the complex version of poker (No-Limit Texas Hold´em). The program used machine learning technology to improve its game by trial and error. In January 2017, Libratus competed in a tournament against top poker players. It played during the day and at night it analyzed the data collected to improve its strategies. The program learned how to bluff and, at the end, it beat the world's top poker players. According to one of the creators of Libratus:

> The computer can't win at poker if it can't bluff. Developing an AI that can do that successfully is a tremendous step forward scientifically and has

[52] PRAKHAR, Swarup. *Artificial Intelligence*. International Journal of Computing and Corporate Research, Vol. 2, No. 4. 2012. Available at http://www.ijccr.com/july2012/4.pdf.
[53] TURING, Alan. Computing Machinery and Intelligence. Mind 49. 1950. p. 433-460.
[54] OECD, algorithms, p. 7
[55] Idem
[56] ANITHA, P.; KRITHKA, G.; CHOUDHRY, Mani Deepak. *Machine Learning Techniques for learning features of any kind of data: A Case Study*. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3, Issue 12. December 2014.
[57] Idem.

> numerous applications. Imagine that your smartphone will someday be able to negotiate the best price on a new car for you. That's just the beginning.[58]

Another concept mentioned by computer science is deep learning. Deep learning is a subfield of machine learning that uses artificial neural networks that try to replicate the activity of human neurons in order to learn faster and more accurately than traditional machine learning.[59]

For the purpose of this paper, the important characteristic of deep learning is that it is not possible to know what features and information the algorithm used to process inputs and to provide outputs.[60] In the words of the OECD: *"regardless of the quality of the results produced, deep learning algorithms do not provide programmers with information about the decision-making process leading to such results".*[61]

b.2 Applications of algorithms and machine learning in the economy

The rise of algorithms by firms from various sectors results from the synergy of mathematics, computer power and the internet.[62] This combination makes possible the collection and analysis of a significant amount of data almost instantly, allowing empirically-driven decision-making.[63]

The OECD mentions two important employments of algorithms and machine learning by companies: (i) predictive analytics and (ii) optimization of business processes.[64]

Predictive analysis is the use of algorithms to predict the probability of future events based on the analysis of past data. Predictive models can estimate market variables, such as demand and customer behavior, price variations due to some external event, the likely changes in the market with the entry of a new competitor.[65]

---

[58] CARNEGIE MELLON UNIVERSITY. *Carnegie Mellon Artificial Intelligence Beats Top Poker Pros*. Published on January 31, 2017. Available at https://www.cmu.edu/news/stories/archives/2017/january/AI-beats-poker-pros.html
[59] OECD. Algorithms. Op. Cit. p 8.
[60] Idem.
[61] Idem.
[62] MEHRA, Salil K. *De-Humanizing Antitrust: The Rise of the Machines and the Regulation of Competition*. Legal Studies Research Paper Series. Research Paper Nº 2014-43. 11/06.2014.
[63] Idem.
[64] OECD. Algorithms. Op. Cit. p. 9.
[65] Idem.

Algorithms can also be used to improve internal business processes, reducing costs, enhancing production efficiency and, more importantly, setting prices in optimal levels.

There are examples of algorithms and machine learning employment in different areas. We can mention fraud prevention[66], targeted advertising[67] and dynamic pricing even to automatically color black and white pictures and to create sounds to silent movies[68].

In specific sectors, algorithms and machine learning have different applications[69]. In the health sector, we can mention the use of algorithms to detect cancer[70], the use of machine learning algorithms to quantify knee osteoarthritis severity from X-ray images[71] and the use of deep learning in pre-natal care[72].

Another sector that significantly relies on algorithms and machine learning is the financial sector. Financial markets players are increasingly adopting algorithmic trading strategies[73], using machine learning technology[74]. Self-learning algorithms are better than humans to react rapidly to a market event. The so-called "algorithmic trading", according to a news report, "*is responsible for nearly 70% of the trading done in Wall Street*"[75].

One cannot deny how fast algorithms and machine learning technologies are changing the way companies do business. In various industries, algorithms and machine learning have become not only a competitive advantage, but also a requirement to compete in the market.

The rise of the competitive significance of algorithms, machine learning and its input (Big Data) has changed and will continue to change the economy. Therefore, market regulators and antitrust authorities shall recognize these changes and adapt to a data-driven economy.

---

66 PALASH, Nandy. *Online fraud prevention using genetic algorithm solution*. U.S. Patent No. 7,657,497. 2 Feb. 2010.
67 EZRACHI; STUCKE. Op. Cit, p. 17.
68 BROWNLEE, J. *8 Inspirational Applications of Deep Learning*. Machine Learning Mastery. 2016. Available at: http://machinelearningmastery.com/inspirational-applications-deep-learning/.
69 OECD, Algo. p 10.
70 HEMSOTH, N. *The Next Wave of Deep Learning Applications*. The Next Platform. 2016. Available at https://www.nextplatform.com/2016/09/14/next-wave-deep-learning-applications/.
71 SHAMIR, L.; LING, S. M.; HOCHBERG, W.; SCOTT; M; FERRUCCI, L.; GOLDBERG,I. G. Early detection of radiographic knee osteoarthritis using computer-aided analysis. Osteoarthritis and Cartilage, vol. 17, no. 10. 2009.
72 O'SHEA, K.; REID, S.;CONDOUS, G.; LU, C. *Deep neural networks for predicting pouch of Douglas obliteration based on transvaginal ultrasound sliding sign videos*. Ultrasound in Obstetrics & Gynecology, Vol. 48, No. S1. 2016. p. 381. Available at http://onlinelibrary.wiley.com/doi/10.1002/uog.17163/full.
73 CHABOUD, Alain P. et al. *Rise of the machines: Algorithmic trading in the foreign exchange market*. The Journal of Finance, v. 69, n. 5. 2014. p. 2045-2084.
74 AZOFF, E. Michael. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc. 1994.
75 BUSINESS INSIDER. *Black Box Trading: Computers Taking Over Wall Street?*. Published on 15.08.2011. Available at http://www.businessinsider.com/black-box-trading-computers-taking-over-wall-street-2011-8

In the next chapter, the impacts of this new economic reality in the antitrust analysis performed by competition authorities will be analyzed.

III – IMPACTS ON ANTITRUST ANALYSIS

Traditional antitrust tools were created in a different context. The competition policy was thought to address brick-and-mortar business practices, like old-fashioned cartels or resale price fixing.

Traditional antitrust analysis may need to adapt to the new economic reality explained in the previous chapter. The implications of Big Data and pricing algorithms for competition policy are still being debated internationally. For instance, the OECD Competition Committee held a meeting about Big Data and competition policy in 2016[76] and one about algorithm-driven collusions in 2017[77].

Yet, there is no consensus about the role competition authorities shall have in this field. Some advocate for increased antitrust intervention, arguing that privacy concerns should be considered in antitrust analysis. They affirm that the collection and use of data can foster harmful price discrimination and that privacy is an important non-price element of competition[78]. Others believe that antitrust shall not incorporate privacy concerns in its analysis. In this sense, Geoffrey A. Manne and R. Ben Sperry argue that:

> Privacy advocates have thus far failed to make their case. Even in their most plausible forms, the arguments for incorporating privacy and data concerns into antitrust analysis do not survive legal and economic scrutiny. In the absence of strong arguments suggesting likely anticompetitive effects, and in the face of enormous analytical problems (and thus a high risk of error cost), privacy should remain a matter of consumer protection, not of antitrust.[79]

[76] OECD. Big Data. Op. Cit.
[77] OECD. Algorithms. Op. Cit.
[78] For example, STUCKE; GRUNES. Op. Cit.
[79] MANNE, Geoffrey A.; SPERRY, Ben. *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*. CPI Antitrust Chronicle. May 2015. Available at SSRN: https://ssrn.com/abstract=2617685

However, it is possible to say that relevant bodies, such as the OECD[80], as well as the French and the German Competition Authorities[81], agree that some antitrust tools can be adapted to fully consider the implication of the new economic reality on competition policy.

This perspective is growing and, possibly, in the near future, competition authorities around the world will begin to deepen their knowledge and their scrutiny on data-driven markets. As stated in the joint paper made by the French and the German Competition Authorities:

> The collection, processing and commercial use of data is often seen not as a competition law issue but rather as an issue which concerns data protection enforcement. However, several recent proceedings point to the fact that competition authorities have begun to look at possible competition issues arising from the possession and use of data, even if, in the end, none were ascertained in the specific cases[82]

In this chapter, we will address possible implications of the new economic reality for antitrust analysis, dividing the chapter into two sections. In the first section (A), we will examine the impacts of the new economic reality on mergers and abuse of market power, due to the overlap in some analysis tools. In the next section (B), we will discuss the implications for collusive practices, in special regarding algorithm-led collusions.

a) MERGER REVIEW AND ABUSE OF MARKET POWER

a.1 Definition of the relevant market

According to CADE, the definition of the relevant market is one of the most important tasks in competition analysis, being the starting point for case analysis in various competition authorities' guidelines, such as Brazil, Europe and the United States.[83] In the words of CADE´s guideline to horizontal mergers:

---

80 OECD, Big Data. p 15

81 AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. *Competition Law and Data*. 2016. Available at: http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.html.

82 Idem. p. 3.

83 Documento de Trabalho No. 001/10. Delimitação de Mercado Relevante. Versão Pública. Departamento de Estudos Econômicos (DEE). Grupo de Trabalho de Métodos em Economia (GTME - GT n.º 3). Brasília,

> The relevant market is defined as the smallest group of products and the smallest geographic area necessary for a hypothetical monopolist to be in conditions to impose a small but significant non-transitory increase in prices.[84]

The definition of the relevant market is not always an easy task in data-driven markets. These markets are often multi-sided markets. As explained in the section about the Big Data ecosystem, those markets are deeply interconnected and may have complex relations between them.

The theory and the tools to define the relevant market when dealing with multi-sided markets is not new, although there are still vivid debates regarding many issues in this field.

Nevertheless, the traditional tools may need to be rethought in the case of digital platforms, in multi-sided markets, that engage in non-monetary transactions with customers, receiving compensation in form of personal data.

Traditional tools, such as the Small but Significant and Non-transitory Increase in Price (SSNIP) test, shall be adapted to markets where it is harder to identify price. For instance, a player like Google provides numerous "free" services to customers, such as search engine, cloud storage of files, e-mail accounts, translation services, GPS navigation, video uploading (YouTube), social network, etc.

The SSNIP test, also known as "hypothetical monopolist test", assesses if consumers would switch to another product due to a small but significant (5-10%) non-transitory increase in the price of the product in question.[85]

Nonetheless, in order to identify the relevant market in "free" online markets, one should not focus only on monetary transaction, but instead consider the exchange of services or products for personal data.

Therefore, the famous SSNIP test shall be adapted to this new reality. Competition authorities should not apply tests that primarily rely on prices when dealing with data-driven markets. In a roundtable on online data collection, targeting and profiling, the former European

---

November 2010. Available at: http://www.cade.gov.br/acesso-a-informacao/publicacoes-institucionais/dee-publicacoes-anexos/delimitacao_de_mercado_relevante.pdf

[84] Guia para Análise Econômica de Atos de Concentração Horizontal (Portaria Conjunta SEAE/SDE nº 50/2001). Free transation. In the original version: "Mercado relevante é definido como o menor grupo de produtos e a menor área geográfica necessários para que um suposto monopolista esteja em condições de impor um 'pequeno porém significativo e não transitório' aumento de preços". Available at: file:///C:/Users/joao.lacerda/Downloads/2001portariaConjunta50-1_guia_para_analise_economica_de_atos_de_concentracao.pdf.

[85] OECD, *Market Definition*. DAF/COMP(2012)19. 2012. p. 3. Available at http://www.oecd.org/daf/competition/Marketdefinition2012.pdf.

Consumer Commissioner Meglena Kuneva stated that personal data is the "*new currency of the digital world*".[86]

A possible solution is the use of quality, instead of price, in order to define the relevant market. For instance, the OECD, in the Background note for the roundtable on the role and measurement of quality in competition analysis held by the Competition Committee in June 2013, stated that:

> A more contentious issue, however, is the application of quantitative tools for market definition that focus primarily on quality effects. The SSNDQ test is posited as one means by which a quantitative focus on quality might be realised in relation to market definition. This measures the impact of a "small but significant non-transitory decrease in quality" in a manner equivalent to the SSNIP test's assessment of price increases. The SSNDQ test faces criticism that in practice it is unworkable, however, given the inherent difficulties of measuring quality alongside the existing complications of the applying the SSNIP test itself within real market situations.[87]

The SSNDQ test (small but significant, non-transitory decrease in quality) is a proposed solution to define the relevant market when price-based tests cannot be applied. However, there are criticisms to this solution. The Canadian Competition Authority understands that it may be difficult to measure quality. Moreover, it argues that even when there is a method to quantify quality, consumers have different tastes and perceptions about it. Furthermore, many products have multiple aspects of quality, which makes it harder to access using consumers' surveys.[88]

Stucke and Grunes argue that the decrease in quality in free services can be understood as a deterioration of privacy policy.[89] In the Facebook/WhatsApp merger, the European Commission acknowledged that personal data protection is an important non-price competition factor.[90] According to Stucke and Grunes:

> For many data-driven mergers involving free products, the agency might attempt to assess what a small, but significant, non-transitory decrease in quality would entail. Privacy protection would be part of this assessment.[91]

---

[86] KUNEVA, Meglena. *Roundtable on Online Data Collection, Targeting and Profiling.* 31 March 2009. SPEECH/09/156. Available at http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm

[87] OECD. The Role and Measurement of Quality in Competition Analysis. 2013. Available at http://www.oecd.org/competition/Quality-in-competition-analysis-2013.pdf

[88] COMPETITION BUREAU. Competition Bureau submission to the OECD Competition Committee roundtable on The Role and Measurement of Quality in Competition Analysis. 2013. Available at http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03769.html

[89] STUCKE; GRUNES. Op. Cit. p.118-119.

[90] Facebook/WhatsApp (Case Comp/M.7217), Commission Decision C(2014) 7239 final, 3 October 2014, para 87.

[91] STUCKE; GRUNES. Op. Cit. p.118-119.

Therefore, since data is the new currency of online markets, the growth in the collection of personal data and the increase in its use by the collector shall be understood as a price increase, for competition assessment purposes. In addition, in markets without monetary transactions, the protection of personal data shall be viewed as an aspect of the quality of the product or service.

The European Commission officials Eleonora Ocello, Cristina Sjödin and Anatoly Suboč, in a merger brief entitled “What's Up with Merger Control in the Digital Sector? Lessons from the Facebook/WhatsApp EU merger case”, stated that:

> A second way in which data may be relevant in the competitive assessment of mergers relates to privacy. Privacy could be regarded as a non-price parameter of competition which may be degraded by the merged entity post-merger. In two-sided markets, where products are offered to users for free and monetised through targeted advertising, personal data can be viewed as the currency paid by the user in return for receiving the 'free' product, or as a dimension of product quality. Hence, a website that, post-merger, would start requiring more personal data from users or supplying such data to third parties as a condition for delivering its 'free' product could be seen as either increasing its price or as degrading the quality of its product. In certain circumstances, this behaviour could arguably amount to an infringement of competition law (irrespective of whether or not it also constitutes an infringement of data protection rules). However, while technically viable, this theory of harm could only be relevant in those cases where privacy is an important factor in the decision to purchase a product or service, i.e. a key parameter of competition.[92]

Nevertheless, some authors believe that privacy should not be a concern for competition authorities. For instance, Geoffrey A. Manne and R. Ben Sperry state that the argument for incorporating privacy concerns into competition analysis does not survive legal and economic scrutiny. They argue that privacy should be addressed by consumer protection rules and authorities, not antitrust.[93]

Another critic of the introduction of privacy concerns into antitrust analysis is James Cooper. The author mentions three major problems in the inclusion of privacy in antitrust:

> The first concern is conceptual. The analogy between privacy and quality begins to break down once we recognize that, as opposed to selecting lower quality levels to enjoy lower costs, firms invest in collecting and analyzing data to improve content and to enhance matching between sellers and consumers who have heterogeneous tastes for privacy. The second concern goes to fundamental rights to speak. An antitrust rule that limits firms’ ability to collect and analyze consumer data is likely to trigger some form of First Amendment scrutiny.

---

[92] OCELLO; SJÖDIN; SUBOČS. Op. Cit.

[93] MANNE, Geoffrey A.; SPERRY, Ben. *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*. 2015.

> Third, allowing antitrust enforcers to consider privacy would inject an undesirable level of subjectivity into antitrust enforcement decisions, which is likely to attract socially wasteful rent-seeking expenditures and to deter beneficial data collection efforts.[94]

In the Facebook/WhatsApp merger, the European Commission determined that privacy concerns do not fall within the scope of the antitrust analysis:

> For the purposes of this decision, the Commission has analysed potential data concentration only to the extent that it is likely to strengthen Facebook's position in the online advertising market or in any sub-segments thereof. Any privacy-related concerns flowing from the increased concentration of data within the control of Facebook as a result of the Transaction do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules.[95]

In any way, competition authorities shall take into consideration non-price competition factors when analyzing a merger or a conduct, especially in markets where services are not exchanged for money, but for personal data. The understanding that data collection can be equivalent to the price paid by the consumer, allows competition authorities to apply market definition techniques such as SSNDQ, having data protection as a measurable aspect of quality.

a.2 Assessment of market power

According to Richard Whish and David Bailey, market power is the ability one firm has to profitably raise prices over a period of time, or to restrict output and limit consumer choices.[96] As mentioned in the previous chapter, Big Data is increasingly relevant in today's economy. Many crucial services and products are based on the analysis of Big Data. For example, search engines are crucial tools to almost everyone. Therefore, the control of Big Data by one company can lead to significant power in the market, which shall be taken into consideration by competition authorities.

Nevertheless, some argue that usually the control over Big Data will not result in market power by the firm. They argue that data is ubiquitous, cheap, widely available and non-

---

[94] COOPER, James C. *Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity*. George Mason Law Review, Forthcoming. George Mason Law & Economics Research Paper No. 13-39. 2016. Available at SSRN: https://ssrn.com/abstract=2283390
[95] Facebook/WhatsApp (Case Comp/M.7217), Commission Decision C(2014) 7239 final, 3 October 2014, para. 164.
[96] WHISH, Richard; BAILEY, David. *Competition Law*. 7th Edition. Oxford University Press. 2012.

rivalrous. Darren S. Tucker and Hill B. Wellford[97], for instance, believe that big data is increasingly available to smaller companies, arguing that data is everywhere, since the sources of data are increasing – for example, we can mention the Internet of Things, where various objects are equipped with networked sensors that collect data. According to the FTC Chairwoman Edith Ramirez: "Big data is no longer the province of a few giant companies."[98]

The argument goes on to say that the costs of collecting and analyzing Big Data is also low. They state that most companies collect vast amounts of data as a by-product, as an accidental result of their usual activity, such as sales transactions. Moreover, the cost of software that analyze Big Data in order to extract valuable information is also decreasing. It is also possible, according to the authors, to buy data cheaply in the market.[99]

Finally, they state that data is non-rivalrous, that is to say, if one company collects a piece of data from an individual, it does not prevent another company to collect the same data, or to get the same information from the analysis of other data. It is increasingly common for people to use multiple providers of the same service (known as "user multi homing"), which allows different companies to collect the same information. In the example given by Darren S. Tucker and Hill B. Wellford:

> if one ad network determined that the user of a particular mobile device lived in Connecticut, liked to travel, and owned a dog, there is nothing to prevent another ad network from learning the same information—indeed, for a frequent Internet user, it is likely that dozens of firms will create a similar profile. Redundant data are so common as to cause problems for data brokers.[100]

Therefore, the authors argue that since Big Data is ubiquitous, cheap, widely available and non-rivalrous, the control over data by one firm will unlikely result in competition concerns. In this sense, antitrust authorities shall proceed cautiously when analyzing claims that Big Data is a source of market power and competition problems.[101]

Nonetheless, other authors refute this view. For example, Stucke and Grunes argue that if Big Data was, in fact, ubiquitous, cheap and widely available, firms would not invest large

---

[97] TUCKER, Darren S.; HILL, Wellford B. *Big Mistakes Regarding Big Data*. Antitrust Source, American Bar Association. December 2014. Available at SSRN: https://ssrn.com/abstract=2549044

[98] Edith Ramirez, Chairwoman, Federal Trade Commission, The Privacy Challenges of Big Data: A View from the Lifeguard's Chair 3 (Aug. 19, 2013). Available at https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard%E2%80%99s-chair/130819bigdataaspen.pdf

[99] TUCKER; HILL. Op. Cit.

[100] Idem.

[101] Idem.

sums of money to collect or acquire data, they would simply harvest the publicly available data.[102] However, firms do incur in significant costs in order to collect useful data. They mention the case of Uber, which offered in 2015 several billion dollars to buy Nokia's mapping business "Here"[103]. In the words of one analyst commenting the offer: '*it's extraordinarily difficult to get this type of mapping data. [...] Other than Google, here is one of the few companies that can offer this data right now*".[104]

Furthermore, the value of the data collected is not absolute. It depends on the amount of the data the company already has. In other words, the value of data has increasing return to scale. According to the OECD:

> The monetary, economic and social value of personal data is likely to be governed by non-linear, increasing returns to scale. The value of an individual record, alone, may be very low but the value and usability of the record increases as the number of records to compare it with increases. These network effects have implications for policy because the value of the same record in a large database could be much more efficiently leveraged than the same record in a much smaller data set. This could have implications for competition and for other key policy items such as the portability of data.[105]

Therefore, competition authorities shall not analyze the value and competitive significance of data in absolute terms. Authorities shall assess the value, and the resulting power conferred by data in the context of the owner of the data. In the words of the OECD, "*it is crucial not to isolate personal data from the underlying context of the business model in question*".[106]

In addition, Stucke and Grunes refute the argument that data is non-rivalrous and, therefore, the fact that one company has a piece of data does not hinders another one from collecting it. The authors explain that non-rivalrous and non-excludable goods are different concepts. Non-rivalrous means that the use of the good by one person does not decreases the total stock of the good. Non-excludable means that it is not possible to stop other people consuming the good. The authors affirm that if data were non-rivalrous and non-excludable it

[102] STUCKE; GRUNES. Op. Cit. p. 42.
[103] Idem, p. 43.
[104] SCOTT, Mark; ISAAC, Mike. *Uber May Drop Google Maps, Buy Nokia's Here Mapping Business*. Financial Review. 8 May 2015. Available at http://www.afr.com/technology/apps/uber-may-dropgoogle-maps-buy-nokias-here-mapping-business-20150508-ggwv1t.
[105] OECD. *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value*. OECD Digital Economy Papers, No. 220, OECD Publishing, Paris. 2013. p. 34. Available at http://dx.doi.org/10.1787/5k486qtxldmq-en
[106] Idem. p. 19

would be, by definition, a public good. Thus, companies like Facebook would never invest resources to collect data if any competitor were able to capture it, as a free-rider.[107]

However, according to Stucke and Grunes, this is not the case. Big Data is excludable and do provide a significant competitive advantage to its owner.[108] According to the OECD, valuable personal data collected by companies are kept as trade secrets and extracting valuable information from this data requires investment. Hence, the OECD argues, the collected data will only become replicable when the technology and ability to collect and analyze such data become generic.[109]

Moreover, the authors refute the idea that the value derives not from the data, but from the algorithm that can analyze and extract information from it. They argue that if this was true, important internet players, such as IBM, Facebook and Google would not open-source any of their algorithms, which they do.

Although the algorithm is important, the data is the input for the algorithm to improve, as explained in the section regarding algorithms and machine learning. The vast amount of data is fundamental to improve continuously the algorithms and to allow better outputs.

According to Lukas Biewald, co-founder and CEO of CrowdFlower:

> First off, let's talk about training data. There's a reason that those big players I mentioned above open-sourced their algorithms without worrying too much about giving away any secrets: it's because the actual secret sauce isn't the algorithm, it's the data. Just think about Google. They can release TensorFlow without a worry that someone else will come along and create a better search engine because there are over a trillion searches on Google each year.[110]

Therefore, Big Data indeed can lead to competitive advantage that cannot be easily replicated by competitors. First, data is not as widely available as argued by some authors, such as Darren S. Tucker and Hill B. Wellford. In addition, the value of the data depends on the amount of existing data owned by the company, which creates incentives for market power concentration. Moreover, big players can indeed exclude competitors from accessing its own databases. Even if data is a non-rivalrous good, it is excludable, which makes it difficult for

---

[107] STUCKE; GRUNES. Op. Cit. p. 45.
[108] Idem, p. 45
[109] OECD. *Supporting Investment in Knowledge Capital, Growth and Innovation*. OECD Publishing, Paris. 2013. p. 232. Available at
http://dx.doi.org/10.1787/9789264193307-en
[110] GUTIERREZ, DanieL. *Human-in-the-Loop is the Future of Machine Learning.* insideBIGDATA, 11 January 2016. Available at http://insidebigdata.com/2016/01/11/human-in-the-loop-is-the-future-ofmachine-learning/.

entrants to acquire the minimum scale necessary to start operating in some data-intensive markets.

Another factor that shall be considered when analyzing market power in data-driven markets is the network effect usually present in multi-sided markets.

As shown in chapter II, the Big Data ecosystem is characterized by multi-sided markets. The French and the German competition authorities, in a joint study, when talking about data and multi-sided markets affirmed that:

> "Furthermore, so-called "network effects" are often to be found here. The term "network effects" refers to how the use of a good or service by a user impacts the value of that product to other users. Such effects may be "direct", when the benefit that users of one group get from a specific service depends on the number of other users from this group using the service. Telecommunication networks are the classic example. The more people use them and can be reached, the more useful they are. Network effects can also be "indirect", when the benefit that users of one group get from the service depends on the number of users from a different group using the service. A dating platform bringing together men and women can serve as an example here. Direct and indirect network effects may also coexist in some cases. For instance, the value of a social network for a given user is likely to increase with the total number of users of that network (direct network effects). Meanwhile, a higher number of users of a social network also increases the value for advertisers (indirect network effects)."[111]

Network effects are often understood as a barrier to entry and a cause of market power concentration. The dominant player, with a large user base, is able to collect more data from its users. With this data, its machine learning algorithms can learn better and faster and, therefore, improve the service rendered to users and advertisers, for example. With the improved service, the user base increases, reinforcing the process.

In some cases, nevertheless, network effects can foster competition. As mentioned by the joint study made by the French and the German competition authorities, network effects can provide an entrant with the possibility to rapidly grow its used base. In the words of the study:

> However, network effects may also be beneficial to new market participants if they are able to attract a high number of users for other reasons (e. g. because of an innovative feature), thereby increasing their attractiveness to future users thanks to network effects. Therefore, network effects can also stimulate competition by giving an entrant the potential for a rapid growth of its consumer base. Depending on various parameters such as the level of fixed costs or the differences in the undertakings' market shares, network effects could thus reinforce or attenuate competition.[112]

---

[111] AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. Op. Cit.

[112] Idem.

In order to measure the market power in a data-driven market, Graef proposes to focus on the ability of the firm to monetize the collected data.[113] The turnover generated by a firm who offers targeted advertisement or paid services using as an input the data collected would be a good indicator of its market power. According to the author, the market power can be examined by looking at the turnover generated by the firm who uses the data, because the value of the data is not intrinsic, but depends on the owner and on how it is used. In the words of the author:

> Since the value of a dataset depends in particular on how it is employed by its owner and not merely on its sheer volume, market shares can be calculated in a reliable way by looking at the share of the total turnover earned by undertakings active in a potential market for a specific type of data. This way the analysis of dominance does not only take into account the value of the dataset in itself but also the success of a provider in putting in place relevant resources and technologies for monetizing the data.[114]

Nevertheless, Graef acknowledges that the proposed approach is not adequate in cases where the market player does not offer paid products and services and does not monetize its dataset.[115] For instance, WhatsApp offers a "free" communication service and does not monetize its users' data with advertising. Therefore, since WhatsApp has no data-related revenue, no value could be attributed to its dataset using turnover as a criteria.

However, even in non-monetized markets, competition authorities must assess possible competition concerns involving data dominance. In this sense, Graef proposes that authorities focus on potential competition instead of actual market shares. The author states that the European Commission is increasingly considering potential competition when assessing market power in dynamic markets.[116]

In conclusion, Gref suggests that the following factors shall be considered when assessing market power in data-driven markets:

> The presence of the following non-exhaustive circumstances may in particular be considered to point towards market power in a market defined around data: (1) data is a significant input into the end products or services delivered on online platforms; (2) the incumbent relies on intellectual property law to protect its dataset as a result of which competitors cannot freely access the necessary data; (3) there are few or no actual substitutes readily available on the market for the specific information needed to compete on equal footing with an incumbent; (4) it is not viable for a potential competitor to collect data

---

[113] GRAEF, Inge. *Market Definition and Market Power in Data: The Case of Online Platforms*. World Competition 38, no. 4. Kluwer Law International BV, The Netherlands. 2015. p. 502.
[114] Idem. p. 502.
[115] Idem. p. 502.
[116] Idem. p. 503.

> itself in order to develop a new dataset with a comparable scope to that of the incumbent (for example due to network effects).[117]

a.3 Merger notification thresholds

Another impact that the data-driven economy can have on competition policy is about the turnover notification thresholds for mergers. In many jurisdictions, the criteria to determine whether a transaction is subject to mandatory notification to the antitrust authority is the turnover of the merging companies.

For example, Brazil's Competition Law (Law nº 12.529/11 and Ministries of Finance and Justice Joint Resolution No. 994/2012) establishes that concentration acts must be notified to the competition authority (CADE) if one of the economic groups involved in the merger had revenues of at least 750 million reais in the last fiscal year and if the other group had at least 75 million reais in revenues in the last fiscal year.

Nevertheless, in some cases, notification thresholds based solely on turnover can leave out mergers with significant competition consequences. The OECD gives the example of a merger in which an established dominant player acquires an entrant with valuable data or who is able to innovate in order to gain access to a variety of additional data sources.[118]

In the Facebook/WhatsApp merger, WhatsApp had a small turnover, which resulted in lack of notification in many countries. For instance, the transaction was not notified in Brazil, where the company has more than 120 million users.[119]

Despite the small turnover of WhatsApp, Facebook paid 19 billion US dollars for the company. The value comes from the number of users WhatsApp has worldwide and their personal data. Although the transaction did not trigger the EU notification threshold, it was ultimately analyzed by the European Commission as a "one stop shop" review, in order to avoid multiple reviews from different jurisdictions within the European Union.[120]

---

[117] Idem. p. 504.
[118] OECD. Big Data. Op. Cit. p. 20.
[119] http://link.estadao.com.br/noticias/empresas,whatsapp-chega-a-120-milhoes-de-usuarios-no-brasil,70001817647
[120] OECD. Big data. Op.Cit. p. 20.

The European Competition Commissioner Margrethe Vestager affirmed that looking only to turnover might lead to ineffective merger review policy, since important data-driven mergers would not be notified:

> The issue seems to be that it's not always turnover that makes a company an attractive merger partner. Sometimes, what matters are its assets. That could be a customer base or even a set of data. (…) Or a company might be valuable simply because of its ability to innovate. A merger that involves this sort of company could clearly affect competition, even though the company's turnover might not be high enough to meet our thresholds. So by looking only at turnover, we might be missing some important deals that we ought to review.[121]

The OECD suggests a possible solution: including an additional threshold based not on the turnover of the parties but on the total value of the transaction. The price paid for a data-intense company with little turnover indicates the business significance of the data acquired. Moreover, according to the OECD:

> Moreover, such transaction thresholds could help enable competition authorities to identify pre-emptive acquisitions intended to displace potential disruptive innovators (some of which may be data-driven innovators), as already discussed in OECD (2015b).[122]

Thresholds based on the total value of the transaction are already in place in some jurisdictions such as the US and Mexico. The use of such thresholds is also being considered in Germany, where the Federal Ministry of Economic Affairs and Energy issued a draft amendment to the Act against Restraints of Competition, in order to introduce an additional threshold based on the value of the transaction – 350 million euros – in addition to the existing turnover thresholds.[123]

In Brazil, the Law allows the Brazilian competition authority to request notification within one year after the closing date of the merger that does not meet the turnover thresholds.

In conclusion, it is important that competition authorities be watchful about data-driven mergers and acquisitions that can have harmful effects on competition and innovation.

---

[121] VESTAGER, M. *Refining the EU Merger Control System*, Speech at the Studienvereinigung Kartellrecht, Brussels. 10 March 2016. Available at: https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system_en.

[122] OECD. Big Data. Op. Cit. p. 20.

[123] Idem. p. 20.

a.4 Exclusionary practices

(i) Data as an "essential facility"

In many markets, data and the information extracted from it can have enormous competitive significance. As already explained in this dissertation, Big Data can provide a firm with considerable competitive advantage.

However, the controversial question is whether in some markets Big Data can be considered an essential input, which is necessary for a firm to compete. In other words, it shall be determined if the "essential facilities" doctrine should be applied to Big Data.

Defining the essential facility doctrine, Mats Bergman states that it is applied to compel a dominant firm to supply in non-discriminatory terms a critical or essential intermediate good or input to its downstream and, sometimes, upstream competitor.[124] The economic effect of the application of the doctrine would be similar to a price regulation of the input in order to enable downstream competition.

If data is regarded as an essential facility and the doctrine is applied, then the owner of the data will be obliged to provide access to its data to competitors in a non-discriminatory manner.

In Brazil, CADE applied the essential facility doctrine more than once. Recently, CADE addressed this issue when analyzing the Bovespa/Cetip merger.[125] CADE applied the doctrine to the central security depository service, which has the characteristics of a natural monopoly and is crucial to enable other players to operate in the stock exchange market.

Under United States Antitrust Law, Eleanor Fox describes in detail the essential facility doctrine:

> Two principles, however, qualify the basic principle that an individual firm, acting alone, has the right to choose its customers. First, there is a narrow "essential facilities" or "bottleneck monopoly" doctrine which holds that where a firm controls a facility that cannot feasibly be duplicated, where access to the facility by competitors is necessary for effective competition in the market, and where the controlling firm can give access without degrading

[124] BERGMAN, Mats A. *The role of the essential facilities doctrine*. The Antitrust Bulletin, v. 46, n. 2, p. 403-434, 2001.
[125] Concentration Act nº 08700.004860/2016-11.

> its own performance, the controlling firm must give access. Second, a firm in a monopoly position may not engage in predatory strategies including refusals to deal when, by the refusal, the dominant firm foregoes profit opportunities and imposes costs on itself in order to impose greater costs on its competitor.[126]

Albeit the essential facility doctrine was already applied in important jurisdictions, the OECD noted that the is not universally accepted and has received considerable criticism:

> Recognising that the essential facility doctrine is not universally accepted by courts or competition law practitioners, the addition of fast moving and speculative claims to application of the doctrine is particularly challenging and has received strong opposition, not only from studies sponsored by current incumbents (Lerner, 2014), but also from some antitrust practitioners (Balto and Lane, 2016) and academics (Sokol and Comerford, 2016).[127]

The critics normally argue that in innovation-driven markets, entrants who are capable of having a simple insight into customers' needs can establish themselves in the markets in spite of initial small amount of user data. Examples are Slack, Facebook, Snapchat, etc.[128]

According to Daniel Sokol and Roisin Comerford:

> Some have suggested that antitrust remedies may be appropriate where a dominant firm has misused Big Data to gain or sustain an improper competitive advantage. The imposition of such remedies presents obvious problems. From an antitrust perspective, forced sharing of information with rivals infers the essential facilities doctrine, and such forced dealing with competitors in the Big Data environment is far beyond the limits of what a duty to deal would require. If Big Data were deemed an essential facility and a duty to deal imposed, the competitive dynamics of the market would be dramatically altered. Such an extreme and far-reaching remedy is out of line with current antitrust policy (Orbach and Avraham 2014).[129]

Nevertheless, as noted by the OECD, the examples of successful entrants with small dataset occurred years ago, and the increased importance of Big Data is an extremely recent phenomenon. Therefore, the significance of Big Data when firms like Facebook entered the market is different from the importance it has nowadays.

Furthermore, besides showing that data is an essential facility, there is another requirement in order to apply the doctrine, which is showing that the indispensable facility cannot be reasonably duplicated. As already mentioned in the section regarding the assessment

---

126 FOX, Eleanor M. *Monopolization and Dominance in the United States and the European Community: Efficiency Opportunity and Fairness*. Notre Dame L. Rev., v. 61. 1986. p. 981.
127 OECD. p. 21-22.
128 SOKOL, D. Daniel; COMERFORD, Roisin E. *Does Antitrust Have a Role to Play in Regulating Big Data?* Cambridge Handbook of Antitrust, Intellectual Property and High Tech, Roger D. Blair & D. Daniel Sokol editors, Cambridge University Press, Forthcoming. 2016.Available at SSRN: https://ssrn.com/abstract=2723693
129 Idem. p. 17-18.

of market power, some argue that data cannot be an essential facility because it can be easily duplicated, since it is ubiquitous, cheap, widely available and non-rivalrous.[130]

However, the OECD refutes this argument by stating that although there are some types of data that are easily available in the market, other types of data, especially user personal data, which is normally an important input of valuable market information, is not generally available:

> Computerized information (software, databases) that is readily available in markets is easily replicable. However, the data on customers and product sales that firms gather for marketing and new product development are protected as a valuable corporate secret. Exploiting these data also requires investments in new capabilities and organizational change and therefore takes time. The data will therefore not be replicable until the technology and skill needed to capture and analyze such data become generic.[131]

In conclusion, the application of the essential facility doctrine in relation to Big Data is still controversial. In any way, we can describe certain types of exclusionary conducts that may raise competition concerns and may require antitrust authorities to be watchful. The joint report made by the French and German competition authorities list four types of exclusionary practices that can occur in a data-driven environment: (i) refusal to access, (ii) discriminatory access, (iii) exclusive contracts, and (iv) tied sales and cross-usage[132].

(ii) Refusal to access

Regarding the conduct of refusal to access, the joint report mentions that ECJ precedents have limited compulsory access to essential facilities to exceptional circumstances. Therefore, the application of the essential facilities doctrine to oblige the holder of a dataset to provide access to competitor would only occur in exceptional circumstances, where the data owned by the company is truly unique and essential to the activity and the competitor has no possibility to obtain the data by other means.[133]

---

130 TUCKER; HILL. Op. Cit.
131 OECD. Supporting. Op. Cit.
132 AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. Op. Cit. p. 17-20.
133 Idem.

The joint report also raises concerns regarding privacy issues that can result from the forced sharing of personal data without the consent of the consumers who provided the data with companies that have no previous relationship with the consumer.[134]

(iii) Discriminatory access

Regarding discriminatory access, the report mentions the case of Cegedim, decided by the French Competition Authority. In this case, Cegedim was a dominant player in the market of medical information data in France. The company refused to sell its main database (OneKey) to customers who used the software of one of its competitor in a different market – customer relationship management in the health sector. The French Competition Authority decided that the discrimination was unlawful.

Other types of discriminatory practices are described by the joint report. For instance, it mentions the case of online sales platforms that are vertically integrated and, therefore, also operates in the online retail market. The firm may have access to crucial information about the competitors in the retail market who use the platform. Such integrated firm would be able to gain competitive advantage from this information and restrict the data that retailers can collect from its platform about the transactions they are involved in.[135]

(iv) Exclusive contracts

The report also refers to exclusive contracts as another anticompetitive conduct that may occur in data-driven markets. A dominant company can conclude exclusivity agreement with data providers in order to prevent its competitors from having access to crucial information about the market. It may also adopt strategies to induce exclusivity on consumers who provide essential data.[136]

[134] Idem.
[135] Idem.
[136] Idem.

According to Richard Whish, competition concerns may raise from exclusive supply obligation, that is to say, a contract in which the supplier can only supply to one customer, and from exclusive purchasing obligation, where a customer is obliged to acquire products from a specific supplier.[137]

(v) Tied sales and cross-usage of data

Finally, the report addresses tied sales and cross-usage of data. It affirms that competition authorities shall be watchful when a dominant company in one market that collects strategic data tries to leverage its market power to another market. A report made by the UK Competition and Markets Authority addresses the case where a company who holds a valuable dataset ties the access to the data to its data analytics service.[138]

Moreover, the use of the data collected in one market in another one (cross-usage) can, in some cases, lead to anticompetitive results. The report mentions the French case of the oil company GDF-Suez, who had access to valuable consumption data in the context of a regulated offers providers. The French Competition Authority imposed interim-measures to the company in order to allow access to this data by the competitors.[139] Therefore, in cases where one player provides a public service activity, sometimes as a regulated monopolist, it may have access to data that represents a significant competitive advantage in another market, which cannot be obtained by its competitors. In such cases, antitrust intervention to order the sharing of the data may be a necessary measure.

In conclusion, the new economic reality, characterized by data-driven services and the high competitive significance of Big Data has important impacts on competition policy regarding merger review and analysis of abuse of dominant position.

First, the definition of market power shall take into consideration non-price variables, such as quality and privacy protection, especially in services paid with user data. In addition, competition authorities shall recognize the importance of data when measuring market power

[137] WHISH;BAILEY. Op. Cit. p. 682-683.

[138] COMPETITION AND MARKETS AUTHORITY. *The Commercial Use of Consumer data*. 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435817/The_commercial_use_of_consumer_data.pdf , p.90.

[139] AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. Op. Cit. p. 20.

of an undertaking. Moreover, competition authorities and legislatures should review merger notification thresholds to include transaction value threshold in addition to revenue-based ones, considering that some important data-related mergers would fall out the current notification criteria. Finally, authorities shall be watchful to possible exclusionary conducts related to data, such as refusal or discriminatory access to essential data, exclusive contracts and tied sales and cross-usage of datasets.

In the next section, the dissertation will address potential impacts of the new economic reality on collusive practices. The section will focus on algorithm-driven collusion.

b) COLLUSIVE PRACTICES

As explained in chapter II, algorithms and machine learning have important and beneficial applications in today's economy. In various sector, the use of self-learning algorithms is the common practice. For instance, it is hard to imagine Wall Street nowadays without high-frequency algorithmic trading.

Nevertheless, the spread of algorithms in the business practice may raise competition concerns regarding the likelihood of collusion. Some authors suggest that there is a risk that algorithms can change market structures that would traditionally be conductive to a highly competitive environment and facilitate collusion.

Therefore, these authors suggest that algorithms might facilitate tacit collusion by providing an efficient instrument to signal and implement the collusion and to detect and punish deviations.

Furthermore, this chapter will discuss the challenge posed by algorithm-driven tacit collusion to competition authorities, since tacit collusion is not normally an antitrust infringement, albeit it can have increased harmful effects due to algorithms.

b.1 Collusion

According to the OECD, the term "collusion" is normally defined in the literature as any form of agreement or coordination between competitors aiming at increasing profits to a supra-competitive level, which results in a deadweight loss.[140] It is a business strategy undertaken by firms to jointly maximize profits. In order to facilitate the collusion, there must be a structure that allows the competitors to agree on, monitor and enforce the common policy.

Collusion can be explicit or tacit. The first one refers to explicit agreements between competitors in order to fix price or production levels. Tacit collusion is a coordination between competitors that happens without any explicit agreement. In this case, the anticompetitive outcome results from every competitor acting having in mind its own profit maximization strategy.

Some market factors can make competitors realize that their pricing strategies are interdependent. Therefore, the coordination occurs without any agreement. About this topic, Marc Ivaldi, Bruno Jullien, Patrick Rey, Paul Seabright, Jean Tirole stated that:

> "Tacit collusion" need not involve any "collusion" in the legal sense, and in particular need involve no communication between the parties. It is referred to as tacit collusion only because the outcome (in terms of prices set or quantities produced, for example) may well resemble that of explicit collusion or even of an official cartel. A better term from a legal perspective might be "tacit coordination".[141]

Although both types of collusion have harmful effects to competition and, therefore, to consumers, normally, antitrust laws do not consider tacit collusion an infringement. The law normally focuses on the means used by the competing firms in order to reach the coordinated outcome. In this sense, most competition laws only punish the agreement between competitors, deeming as lawful other forms of collusion caused by independent strategies (tacit collusion). In short, competition authorities require some sort of direct or indirect contact between competitors to punish a collusion.

However, some jurisdictions[142] accept that the competition authority can infer the existence of an agreement when the parallel conduct is inconsistent with unilateral behavior due to the presence of additional factors (plus factor), such as probable communications between the parties.

---

[140] OECD. Algorithms. Op. Cit. p. 17.

[141] IVALDI, Marc; JULLIEN, Bruno; REY, Patrick; SEABRIGHT, Paul; TIROLE, Jean. *The Economics of Tacit Collusion*. Final Report for DG Competition, European Commission. IDEI, Toulouse. March 2003. p. 4. Available at http://ec.europa.eu/competition/mergers/studies_reports/the_economics_of_tacit_collusion_en.pdf

[142] For example, Brazil (Administrative Proceeding nº 08012.000677/1999-70) and the USA (ABA Section of Antitrust Law, Antitrust Law Developments. 6th ed. 2007. p. 11–16).

In the next sections, we will analyze the impacts of algorithms on the likelihood of collusion in data-intense markets.

b.2 Algorithms changing structural factors

In this section, we will address the impacts of algorithms on some structural factors that are normally regarded as facilitators of collusion.[143]

We can mention four relevant structural factors for collusion: (i) the number of competitors, (ii) the existence of barriers to entry, (iii) market transparency, and (iv) frequency of interaction.

(i) Number of competitors

First, the number of competitors is important because in a much dispersed market – with a high number of competitors – it is more difficult to coordinate everyone's behavior. In addition, it is harder to identify a parameter, or a "focal point" for coordination, as mentioned by the OECD.[144] Finally, the individual incentive to collude is also diminished when there are numerous players in the market.

The reasons why it is easier to collude when there are fewer players in the market are related to practical issues, such as making it easier to coordinate, monitor and enforce the common policy. However, algorithms can make those factors less relevant, since they can enable effective coordination to take place in less concentrated markets. Algorithms can monitor more companies and react in a faster way to deviations. Therefore, as explained by the OECD, "*the small number of firms is an important but not a necessary condition for algorithmic collusion to take place*".[145]

---

[143] OECD. Algorithms. Op. Cit. p. 18.
[144] Idem. p. 18.
[145] Idem. p. 18.

(ii) Entry barriers

In relation to entry barriers, the absence of barriers will significantly decrease the incentives to collude, since the rise in prices would attract new entrants and supra-competitive profits would not last long. Potential competition from new entrants have important power to hinder collusion. According to the OECD, it is not clear if algorithms decrease or increase entry barriers.[146]

We can empirically observe that markets in which algorithms are intensively used are normally very concentrated, such as search engines, online platforms, social networks, airlines, etc. Nevertheless, we cannot conclude that this is caused by the use of algorithms, since those markets are characterized by natural barriers to entry – network effects, economies of scale and scope – due to the already explained reinforcing relationship between the collection of data, the improvement of the algorithms and of the service rendered to customers.

Regarding the effect of the use of algorithms on the likelihood of entry the OECD believes that:

> Also the impact of algorithms on the likelihood of entry is not univocal. On the one hand, as discussed in OECD (2016a), algorithms can be used to identify any market threats very fast, for instance through a phenomenon known as nowcasting, allowing incumbents to pre-emptively acquire any potential competitors or to react aggressively to market entry. On the other hand, the increasing availability of online data resulting from the use of algorithms may provide useful market information to potential entrants and improve certainty, which could reduce entry costs.[147]

(iii) Market transparency

Transparency in the market is a factor that facilitate collusion, since it makes it easier for firms to monitor the behavior of others and makes it harder for a competitor to deviate from the collusion. Algorithms normally enhance market transparency, which can make the market more prone to collusion.

[146] Idem. p. 19.
[147] Idem. p. 19.

The business model of algorithm-driven markets is based on the collection and processing of large amount of data about the market. For instance, now-casting algorithms, explained in the previous chapter, are based on the rapid collection and processing of market information to predict future outcomes and allow better decision-making by the company. This can also be applied to pricing strategies. Many companies have algorithms that monitors competitors' prices and use this information to set its own.

Therefore, when dominant players start using this sort of strategy to set prices, smaller firms have the incentive to follow. This fact, according to the OECD, may facilitate collusion:

> The result is an industry where all market participants constantly collect and observe in real-time rivals' actions, consumers' choices and changes in the market environment, creating thus a transparent environment that is prone to collusion.[148]

(iv) Frequency of interaction

As explained in chapter II, the new economic reality is characterized by the increased speed at which information is collected, processed and in which business decisions are made.

Algorithms allowed instant pricing. In data-driven markets, the cost of adjusting prices in response to market events are substantially lower than in traditional brick-and-mortar business.

This increased frequency of market interaction makes it easier to retaliate deviation from the collusion, which makes the collusion much more stable.

Collusions are inherently unstable because the dominant strategy of a colluding competitor is to deviate from the collusion. However, the instant retaliation makes the individual incentive to deviate disappear. The reaction to the deviation would be so fast that the deviating firm would make no profit. According to the OECD: "*in fact, the combination of machine learning with market data may allow algorithms to accurately predict rivals' actions and to anticipate any deviations before they actually take place*".

In the words of the joint report made by the French and German competition authorities:

> Even though market transparency as a facilitating factor for collusion has been debated for several decades now, it gains new relevance due to technical

[148] Idem. p. 20.

> developments such as sophisticated computer algorithms. For example, by processing all available information and thus monitoring and analyzing or anticipating their competitors' responses to current and future prices, competitors may easier be able to find a sustainable supracompetitive price equilibrium which they can agree on.[149]

b.3 Algorithms that facilitate or lead to collusion

In the previous section, we addressed some ways in which the use of algorithms can affect market structural factors that can facilitate collusion. In this section, we will analyze the direct use of algorithms to facilitate or lead to collusion, even without any human collusive intent.

We will divide the section in three parts. First, we will address the use of algorithms to implement, monitor and enforce cartels. In this first case, there is human agreement, which makes it less challenging for competition authorities.

Secondly, we will analyze the use of a single algorithm to set prices for multiple competitors, even when they have legitimate reasons to do that. This scenario is called the "hub-and-spoke scenario".

Finally, we will address the most challenging scenario for competition authorities: self-learning algorithms that collude without human intervention.

(i) Algorithms that help humans collude

In this case, humans collude and use algorithms to execute their agreement. Ezrachi and Stucke call this case the "*messenger scenario*", since the algorithm is merely the messenger that help the cartelists to implement, monitor and enforce the unlawful agreement.[150]

There are some examples of cartels that used algorithms as intermediaries. A famous case is the Poster Cartel case, which was prosecuted by the US Department of Justice.[151] In this

---

[149] AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. Op. Cit. p. 14-15.
[150] EZRACHI; STUCKE. Op. Cit. p. 39.
[151] Department of Justice press release 15-421. Former E-Commerce Executive Charged with Price Fixing in the Antitrust Division's First Online Marketplace Prosecution. 06/04/2015. Available at

case, David Topkins, the founder of Poster Revolution, and other conspirators jointly adopted a pricing algorithm that collected price information from competitors with the intent to coordinate the prices of their poster sold on the Amazon platform.

From an enforcement and legal perspective, the usage of algorithms to facilitate cartels does not pose many challenges. There is human agreement to fix prices. In the words of Assistant Attorney General Bill Baer of the US Department of Justice's Antitrust Division:

> We will not tolerate anticompetitive conduct, whether it occurs in a smoke-filled room or over the Internet using complex pricing algorithms. American consumers have the right to a free and fair marketplace online, as well as in brick and mortar businesses.[152]

Although from a legal perspective the result is the same, Ezrachi and Stucke believe that the use of algorithms as a messenger can have psychological impacts on colluders. The authors believe that by increasing the distance between the cartelists, the algorithm may mitigate the guilt of wrongdoing, in comparison with secret meetings in "smoke-filled rooms".

(ii) Hub-and-Spoke

In the hub-and-spoke scenario, algorithms are used as a central hub to coordinate pricing strategies of competitors. According to Barack Orbach:

> In antitrust law, a hub-and-spoke conspiracy is a cartel in which a firm (the hub) organizes collusion (the rim of the wheel or the rim) among upstream or downstream firms (the spokes) through vertical restraints. Such a conspiracy may be illegal per se under antitrust law where the horizontal agreement among the spokes (the rim) is per se unlawful, such as fixing prices or allocating territories or customers among competing spokes.[153]

The existence of hub-and spoke cartels precedes the creation of computer algorithms. This type of cartel can occur in different markets and does not need an algorithm.

However, in some cases an algorithm can function as the hub of a hub-and-spoke cartel to facilitate collusion between competitors.

---

https://www.justice.gov/opa/pr/former-e-commerce-executive-charged-price-fixing-antitrust-divisions-first-online-marketplace

[152] Idem.
[153] ORBACH, Barack. Hub-and-Spoke Conspiracies. The antitrust source. April 2016. Available at https://www.americanbar.org/content/dam/aba/publishing/antitrust_source/apr16_orbach_4_11f.authcheckdam.pdf

For example, online retailers may outsource its pricing decision to a specialized company who may have the required data and analytics to price in the optimum level to maximize profits. With the rise of Big Data and Big Analytics, this practice might become more common, especially for non-dominant firms.

It may happen that many players of a market outsource its pricing decision to the same company, who would function as a hub. As a result, prices would be aligned because all the companies use the same pricing algorithm, even though there was no agreement in this sense. In the words of Ezrachi and Stucke: "*collusion may be the consequence, but not necessarily the original aim, when each competitor opts for the same third-party pricing algorithm*".[154]

This may also be the case of platforms like Uber, who defines the prices of the services offered. Uber provides a platform between passengers and drivers using a single dynamic pricing algorithm. There is no price competition between the suppliers of the service: the drivers. Therefore, drivers are part of a tacit collusion without any agreement or intent to collude. They only have a similar relationship with a hub.[155]

Another example is the Eturas case, decided by the European Court of Justice. The Court found an online hub-and-spoke agreement. In this case, the administrator of a Lithuanian online travel booking system (the hub) sent a message to the agents informing about a new technical restriction that limited discount rates.[156] The Court decided that the agents who had knowledge of the message could be presumed to be participants of the conduct, unless they publicly distanced themselves from the message. The decision indicates that the knowledge requirement is relevant in this sort of conduct.

Nevertheless, it is still a matter of debates in the literature what would be the criteria to identify if an algorithm-driven hub-and-spoke cartel would be anticompetitive.

(iii) Self-learning algorithms

[154] EZRACHI; STUCKE. Op. Cit. p. 49.
[155] Idem. p. 50-51
[156] ECJ Case C-74/14. Available at http://curia.europa.eu/juris/document/document.jsf?text=&docid=173680&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=523288

This scenario is the most challenging way algorithms can lead to collusive outcomes, since there is no human intent and programmers do not need to program the algorithm to collude. This is the situation in which pricing algorithms with powerful predictive and self-learning capacity collude by unilaterally setting prices on the optimal level, without any human intervention.

According to the OECD, it is still unclear if self-learning algorithms can, indeed, reach collusion.[157] However, in collusive-friendly market conditions, algorithms that learn faster than humans through trial and error could eventually realize that the optimal strategy is to collude, reaching a cooperative equilibrium.

With the rise of Big Data and Big Analytics, more and more firms in different industries will adopt pricing algorithms. With more players using pricing algorithms, more market data will be digitized and accessible to others, which would enhance market transparency. The self-learning pricing algorithm would perform predictive analysis to find the optimal price, using real-time, historical and third party data to forecast the market reaction to a change in the price.[158]

It may be the case that all algorithms will predict that collusion will lead to an optimal price equilibrium. The reaction to a price decrease would be so fast that deviating from the cooperative equilibrium would not be the individual dominant strategy for any firm.

From the enforcement perspective, this is the most challenging scenario. Both in the first scenario, where humans used algorithms to execute unlawful agreements, and in the hub-and-spoke case, there was some sort of human action.

Here, each firm is acting independently to maximize their own profits. In most jurisdictions, antitrust authorities cannot punish mere parallelism without an agreement. As mentioned, the law normally focus on the unlawful means – agreement between competitors – and not on the supra-competitive outcome. According to the OECD:

> Indeed, by relying on machine learning to move business decisions from humans to computers, managers do not only avoid any explicit communication during the initiation and implementation stages of collusion, but are also released from the burden of creating any structures, such as signalling mechanisms, that could be seen by authorities as facilitating practices of collusion.[159]

[157] OECD. Algorithms. Op. Cit. p. 30.
[158] EZRACHI; STUCKE. Op. Cit. p. 61.
[159] OECD. Algorithms. Op. Cit. p. 31

The following section will address the challenges regarding algorithm-led collusion without human agreement and mention possible solutions given by the literature.

b.4 Collusions without agreements

The most challenging scenario from a competition enforcement perspective is the collusion that take place without any agreement or communications between competitors. As mentioned, pricing algorithms might enable supra-competitive output without the need of human intervention.

This situation is complex because it escapes the current legal framework to fight collusion: punishing anticompetitive agreements between competitors.

As explained, algorithms can change structural market factor in a way that facilitates collusive outcome. In addition, pricing algorithms can also directly lead to collusion.

It is possible to identify similarities between the competition concerns raised by self-learning price algorithms and the oligopoly problem.[160] According to the OECD:

> The “oligopoly problem”, an expression sometimes attributed to Posner (1969), refers to the concern that high interdependence and mutual self-awareness in oligopolistic markets might result in tacit collusion, an outcome which is socially undesirable but that falls out of the reach of competition law.[161]

As already explained, algorithms can influence structural factors that amplifies the oligopoly problem, such as the number of competitors, the existence of barriers to entry, market transparency and frequency of interaction.

The competition literature has not provided many effective answers to the oligopoly problem, in part due to the difficulties to find a collusion-prone oligopoly in real markets.[162]

In addition, pricing algorithms can directly reach a collusive outcome without any agreement, which could harm consumers and escape antitrust enforcement.

[160] Idem. p. 34
[161] Idem. p. 34.
[162] Idem. p. 34-35.

Therefore, one can question whether the requirement of agreement to allow antitrust intervention or the concept of agreement shall be revisited in light of the potential competition concerns amplified by algorithms.

In most jurisdictions, the law requires the existence of an agreement to punish collusion. Nevertheless, the notion of agreement can be modified in order to adapt antitrust enforcement to reality. For instance, we can mention the "plus factor" doctrine, in which the existence of an agreement can be inferred even without direct evidence.

In Europe, the law requires the presence of an agreement or a concerted practice, but it does not precisely define those concepts. The European Court of Justice defined as:

> a concurrence of wills between economic operators on the implementation of a policy, the pursuit of an objective, or the adoption of a given line of conduct on the market, irrespective of the manner in which the parties' intention to behave on the market in accordance with the terms of that agreement is expressed[163]

In the United States, in its turn, the Supreme Court established that it is necessary to prove "*a unity of purpose or a common design and understanding, or a meeting of minds*".[164] According to the OECD:

> This definition is, in principle, very broad and could potentially cover parallel conduct. In practice, courts have required evidence that observed parallel conduct is indeed the result of co-ordination among the parties and not mere oligopolistic interdependence (so-called "plus factors"). An example of a "plus factor" required by courts is that the parties have communicated their intentions to act in a certain way.[165]

Referring to the requirement of an agreement instead of focusing on the economic outcome, Kaplow presents critics. The author believes that the current doctrine is too formalistic and treats similar situations in different ways. For him, tacit and explicit collusion can both harm consumers and competition. Therefore, he defends a broad concept of "agreement".[166]

There is no consensual solution to this problem. It would be hard to promote legislative change in order to punish tacit collusion. The key point is that it is difficult to sustain punishment for a conduct performed by algorithms when if it was performed by humans it would be considered legal.

---

[163] Case T-41/96, Bayer AG v Commission, [2000] ECR II-3383
[164] Interstate Circuit Inc v United States, 306 US 208, 810 (1939)
[165] OECD. Algorithms. Op. Cit. p. 36.
[166] KAPLOW, L. On the Meaning of Horizontal Agreements in Competition Law. California Law Review, Vol. 99, No. 3. 2011. p. 683-818. Available at: https://www.jstor.org/stable/23014697?seq=1 - page_scan_tab_contents.

In other words, how can mere parallelism without agreement be illegal when conducted by algorithms and legal when conducted by humans? The effect-based approach might sound as too much intervention in the free trade for many jurists. A firm has the right to independently price its goods in the market seeking profit maximization. The fact that algorithms and not humans perform this task does not change that conclusion.

A possible solution suggested by the OECD is to apply something like the "unfair competition" prohibition, present on Section 5 of the Federal Trade Commission Act, instead of relying on the prohibition of anticompetitive agreements (cartels). Under the current US law, to deem a conduct an unfair competition method, the authority must prove that the conduct is unfair because it causes or is likely to cause substantial injury to consumers; it cannot be reasonably avoided by consumers; and is not outweighed by countervailing benefits to consumers or to competition.[167]

This rule could potentially be applied to the pricing strategies performed by an algorithm, without the need to revisit the concept of agreement or to introduce legislative changes.

In any way, the challenges posed by tacit collusion led by pricing algorithms do not have definite answers yet. This is a rich field of study for the antitrust community and for policy-makers.

IV - CONCLUSION

This dissertation defined the concept of Big Data, adopting the famous definition that focus on the volume and the velocity in which data is collected and processed, on the variety of information aggregated and on the value extracted from it – the “4Vs definition”.

In addition, it showed why Big Data is so important in today's economy, mentioning various application in different industries. It also briefly explained the Big Data ecosystem, showing that Big Data is often collected and traded in multi-sided markets.

[167] OECD. Algorithms. Op. Cit. p. 37.

Then, it defined algorithms and machine learning, showing their economic significance in different sectors of the economy, such as finance and health.

After that, in chapter III, the dissertation analyzed the impacts of the new economic reality explained in chapter II on the traditional antitrust analysis. It first addressed the impacts on merger review and on abuse of market power.

Regarding definition of relevant market, it showed that traditional tools, such as the SSNIP test, should be adapted when analyzing markets where services are not exchanged for money, but for personal data. It explained that privacy protection may be an important factor of non-price competition and presented critics made by some authors concerning the introduction of privacy concerns into antitrust analysis.

In relation to the assessment of market power, it showed that the market power generated by the ownership of Big Data is a controversial discussion in the literature. Some respected authors believe that data is ubiquitous, cheap, widely available and non-rivalrous and, therefore, does not confer market power. Others, on the other hand, question these assumptions and argue that Big Data can be a relevant source of market power and a barrier to entry.

Then, it showed that relevant data-driven mergers might fall out current notification thresholds. It mentioned the billionaire Facebook/WhatsApp acquisition as an example of relevant merger that was not notified in some jurisdictions, such as Brazil, due to the absence of a transaction threshold in addition to revenue-based thresholds.

After that, we addressed the issue of data-related exclusionary practices. We first examined whether data could be viewed as an essential facility. The dissertation showed that there is an open debate in the literature about that. Then, we analyzed four types of exclusionary conducts that can be related to data: refusal to access, discriminatory access, exclusive contracts and tied sales and cross-usage.

Next, the dissertation analyzed the impacts of the new economic reality – in particular of the use of algorithms – on the fight against collusion. First, it describe basic concepts about collusion. Secondly, the dissertation addressed the issue of algorithms that change relevant market factors that influence the likelihood of tacit collusion. Then, it addressed the use of algorithms that facilitate or lead to collusion. Finally, it talked about the challenges posed by the possibility of algorithm-driven collusion without any human agreement from the enforcement perspective.

In conclusion, we can affirm that we are witnesses of a relevant change in the economy. Data and algorithms are increasingly important and present in our daily lives. In various sector we can see an increasing importance of the collection of data and of the extraction of useful information from it.

This new economic reality has important impacts on competition policy. Traditional antitrust tools were developed in a completely different context in order to deal with brick-and-mortar business practices. Therefore, some concepts and tools traditionally used by antitrust authorities should be revisited in order to address the new reality. In short, competition policy shall adapt to a data-driven economy.

BIBLIOGRAPHY

AUTORITÉ DE LA CONCURRENCE; BUNDESKARTELLAMT. *Competition Law and Data*. 2016. Available at: http://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Berichte/Big%20Data%20Papier.html.

AZOFF, E. Michael. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc. 1994.

BERGMAN, Mats A. *The role of the essential facilities doctrine*. The Antitrust Bulletin, v. 46, n. 2, p. 403-434, 2001.

BROWNLEE, J. *8 Inspirational Applications of Deep Learning*. Machine Learning Mastery. 2016. Available at: http://machinelearningmastery.com/inspirational-applications-deep-learning/.

BUSINESS INSIDER. *Black Box Trading: Computers Taking Over Wall Street?*. Published on 15.08.2011. Available at http://www.businessinsider.com/black-box-trading-computers-taking-over-wall-street-2011-8

CARNEGIE MELLON UNIVERSITY. *Carnegie Mellon Artificial Intelligence Beats Top Poker Pros*. Published on January 31, 2017. Available at https://www.cmu.edu/news/stories/archives/2017/january/AI-beats-poker-pros.html

CHABOUD, Alain P. et al. *Rise of the machines: Algorithmic trading in the foreign exchange market*. The Journal of Finance, v. 69, n. 5. 2014. p. 2045-2084.

CISCO. Cisco Global Cloud Index: Forecast and Methodology, 2013–2018 (2014). Available at http://www. cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_ White_Paper.pdf.

COMPETITION AND MARKETS AUTHORITY. *The Commercial Use of Consumer data*. 2015. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/435817/The_commercial_use_of_consumer_data.pdf

COMPETITION BUREAU. Competition Bureau submission to the OECD Competition Committee roundtable on The Role and Measurement of Quality in Competition Analysis. 2013. Available at http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03769.html

COOPER, James C. *Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity*. George Mason Law Review, Forthcoming. George Mason Law & Economics Research Paper No. 13-39. 2016. Available at SSRN: https://ssrn.com/abstract=2283390

DEZYRE. *How Big Data Analysis Helped Increase Walmart's Sales Turnover*. Available at https://www.capgemini-consulting.com/resource-fileaccess/resource/pdf/walmart_pov_15_7_2015.pdf.

EZRACHI, Ariel; STUCKE, Maurice. Virtual Competition. *The promise and perils of the* algorithm-driven economy. Cambridge, Massachusetts. Harvard University Press. 2016

FOX, Eleanor M. *Monopolization and Dominance in the United States and the European Community: Efficiency Opportunity and Fairness*. Notre Dame L. Rev., v. 61. 1986.

GITTELSOHN, John. *Google Data to Help Auction.com Predict Homebuying Trends*. Bloomberg Business, 30/10/2014. Available at http://www.bloomberg.com/news/2014-10-30/google-data-to-helpauction-com-predict-homebuying-trends.html

GRAEF, Inge. *Market Definition and Market Power in Data: The Case of Online Platforms*. World Competition 38, no. 4. Kluwer Law International BV, The Netherlands. 2015.

GUTIERREZ, DanieL. *Human-in-the-Loop is the Future of Machine Learning.* insideBIGDATA, 11 January 2016. Available at http://insidebigdata.com/2016/01/11/human-in-the-loop-is-the-future-ofmachine-learning/.

HEMSOTH, N. *The Next Wave of Deep Learning Applications*. The Next Platform. 2016. Available at https://www.nextplatform.com/2016/09/14/next-wave-deep-learning-applications/.

IVALDI, Marc; JULLIEN, Bruno; REY, Patrick; SEABRIGHT, Paul; TIROLE, Jean. *The Economics of Tacit Collusion*. Final Report for DG Competition, European Commission. IDEI, Toulouse. March 2003. p. 4. Available at http://ec.europa.eu/competition/mergers/studies_reports/the_economics_of_tacit_collusion_en.pdf

KAPLOW, L. On the Meaning of Horizontal Agreements in Competition Law. California Law Review, Vol. 99, No. 3. 2011. p. 683-818. Available at: https://www.jstor.org/stable/23014697?seq=1 - page_scan_tab_contents.

KUNEVA, Meglena. *Roundtable on Online Data Collection, Targeting and Profiling.* 31 March 2009. SPEECH/09/156. Available at http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm

MANNE, Geoffrey A.; SPERRY, Ben. *The Problems and Perils of Bootstrapping Privacy and Data into an Antitrust Framework*. CPI Antitrust Chronicle. May 2015. Available at SSRN: https://ssrn.com/abstract=2617685

MCKINSEY GLOBAL INSTITUTE. Big Data: The Next Frontier for Innovation, Competition, and Productivity, June 2011, available at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

MEHRA, Salil K. *De-Humanizing Antitrust: The Rise of the Machines and the Regulation of Competition*. Legal Studies Research Paper Series. Research Paper Nº 2014-43. 11/06.2014.

NORTON, Steven. *Big Companies Rein in Data Sprawl: Whole Foods, GE and Others Merge Disparate Data Sets to Cut Costs, Learn About Customers*. Wall Street Journal, 21 October

2015. Available at: http://www.wsj.com/articles/big-companies-rein-in-data-sprawl-1445482009

OECD. *Algorithms and Collusion - Background Note by the Secretariat*. DAF/COMP(2017)4. 2017. Available at file:///Users/joao.alacerda/Downloads/Algoritms%20and%20collusion%20OECD.pdf.

______. *Big Data: Bringing Competition Policy to the Digital Era*. Background Note by the Secretariat. DAF/COMP(2016)14. Available at https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf

______. *Data-driven innovation. Big Data for growth and well-being*. Interim Synthesis Report. 2014.

______. *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value*. OECD Digital Economy Papers, No. 220, OECD Publishing, Paris. 2013. p. 34. Available at http://dx.doi.org/10.1787/5k486qtxldmq-en

______. *Market Definition*. DAF/COMP(2012)19. 2012. p. 3. Available at http://www.oecd.org/daf/competition/Marketdefinition2012.pdf.

______. *Supporting Investment in Knowledge Capital, Growth and Innovation*. OECD Publishing, Paris. 2013. p. 232. Available at
http://dx.doi.org/10.1787/9789264193307-en

______. *The Role and Measurement of Quality in Competition Analysis*. 2013. Available at http://www.oecd.org/competition/Quality-in-competition-analysis-2013.pdf

OCELLO, Eleonora; SJÖDIN, Cristina; SUBOČS ,Anatoly. *What's Up with Merger Control in the Digital Sector? Lessons from the Facebook/WhatsApp EU Merger Case*. 1 Competition Merger Brief. February 2015.

ORBACH, Barack. *Hub-and-Spoke Conspiracies*. The antitrust source. April 2016. Available at
https://www.americanbar.org/content/dam/aba/publishing/antitrust_source/apr16_orbach_4_11f.authcheckdam.pdf

O'SHEA, K.; REID, S.;CONDOUS, G.; LU, C. *Deep neural networks for predicting pouch of Douglas obliteration based on transvaginal ultrasound sliding sign videos*. Ultrasound in Obstetrics & Gynecology, Vol. 48, No. S1. 2016. p. 381. Available at http://onlinelibrary.wiley.com/doi/10.1002/uog.17163/full.

PACKER, George. *Cheap Words: Amazon is good for customers. But is it good for books?*. The New Yorker. February 2014. Available at http://www.newyorker.com/magazine/2014/02/17/ cheap-words

PALASH, Nandy. *Online fraud prevention using genetic algorithm solution*. U.S. Patent No. 7,657,497. 2 Feb. 2010.

PRAKHAR, Swarup. *Artificial Intelligence*. International Journal of Computing and Corporate Research, Vol. 2, No. 4. 2012. Available at http://www.ijccr.com/july2012/4.pdf.

SCOTT, Mark; ISAAC, Mike. *Uber May Drop Google Maps, Buy Nokia's Here Mapping Business*. Financial Review. 8 May 2015. Available at http://www.afr.com/technology/apps/uber-may-dropgoogle-maps-buy-nokias-here-mapping-business-20150508-ggwv1t.

SHAMIR, L.; LING, S. M.; HOCHBERG, W.; SCOTT; M; FERRUCCI, L.; GOLDBERG,I. G. *Early detection of radiographic knee osteoarthritis using computer-aided analysis*. Osteoarthritis and Cartilage, vol. 17, no. 10. 2009.

SOKOL, D. Daniel; COMERFORD, Roisin E. *Does Antitrust Have a Role to Play in Regulating Big Data?* Cambridge Handbook of Antitrust, Intellectual Property and High Tech, Roger D. Blair & D. Daniel Sokol editors, Cambridge University Press, Forthcoming. 2016. Available at SSRN: https://ssrn.com/abstract=2723693

STEWART, Jim. *Walmart plays catch-up with Amazon*. The New York Times. 22.10.2015. Available at https://www.nytimes.com/2015/10/23/business/walmart-plays-catch-up-with-amazon.html?mcubz=0

STUCKE, Maurice; GRUNES, Allen. *Big Data and Competition Policy*. Oxford University Press. 2016.

THE EUROPEAN DATA PROTECTION SUPERVISOR. *Towards a new digital ethics Data, dignity and technology*. Opinion 4/2015. Available at https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf

TUCKER, Darren S.; HILL, Wellford B. *Big Mistakes Regarding Big Data*. Antitrust Source, American Bar Association. December 2014. Available at SSRN: https://ssrn.com/abstract=2549044

TURING, Alan. *Computing Machinery and Intelligence*. Mind 49. 1950.

VESTAGER, M. *Refining the EU Merger Control System*, Speech at the Studienvereinigung Kartellrecht, Brussels. 10 March 2016. Available at: https://ec.europa.eu/commission/2014-2019/vestager/announcements/refining-eu-merger-control-system_en.

WHISH, Richard; BAILEY, David. *Competition Law*. 7th Edition. Oxford University Press. 2012.

XIA, F.; YANG, L. T.; WANG, L.; VINEL, A. *Internet of things*. International Journal of Communication Systems, 25(9), 1101. 2012.