ECORYS

Big data and competition

Client: Ministry of Economic Affairs

Rotterdam, 13 June 2017

Big data and competition

Client: Ministry of Economic Affairs

Authors:
Harry van Til
Nicolai van Gorp
Katelyn Price

Reviewing expert group:
Paul de Bijl (Radicand Economics)
Inge Graef (University of Tilburg)
Alexandre de Streel (University of Namur, CRIDS and CERRE)

Rotterdam, 13 June 2017

ECORYS

Table of Content

Table of contents

Nederlandstalige samenvatting (summary in Dutch)

1. Inleiding

Er komen steeds meer data beschikbaar en de mogelijkheden om die nuttig te gebruiken nemen toe

Door digitalisering is in de afgelopen decennia de hoeveelheid en verscheidenheid aan beschikbare data sterk toegenomen. Daarnaast zijn ook de mogelijkheden toegenomen om uit data bruikbare informatie te halen en door toepassing van kunstmatige intelligentie en *machine learning* op basis van data taken uit te voeren. De term "big data" wordt niet alleen gebruikt om te verwijzen naar de grote hoeveelheid en verscheidenheid aan data maar ook naar die toegenomen gebruiksmogelijkheden. Big data stelt bedrijven in staat om de kwaliteit van hun producten of diensten sterk te verbeteren, om consumenten verbeterde of nieuwe informatie aan te bieden en om producten en diensten te personaliseren. Tegelijk zijn er in het publieke debat zorgen dat partijen die toegang hebben tot veel data een machtige positie krijgen ten koste van consumenten en andere bedrijven. Tot nu toe ontbrak een systematische evaluatie van de bestaande literatuur met aandacht voor de specifieke Nederlandse beleidscontext. Het ministerie van Economische Zaken heeft Ecorys daarom gevraagd om de invloed van big data op de mededinging vanuit een economisch perspectief te onderzoeken.

2. Marktmacht

De specifieke economische kenmerken van data zijn van belang voor een mededingingsanalyse

Om de invloed van big data op mededinging te kunnen bepalen is eerst inzicht nodig in de economische kenmerken en functies van data en big data in het bijzonder. Data kunnen beschouwd worden als input in productieprocessen. Het vormt een bijzonder economisch goed omdat het non-rivaal is, het gebruik door de één sluit het gebruik door een ander niet uit. Als data ook non-exclusief zouden zijn, dat wil zeggen beschikbaar voor iedereen, zouden er geen problemen voor concurrentie kunnen ontstaan. In dat geval vormen data namelijk een zogenaamd "publiek goed" en kan geen bedrijf er op zichzelf een concurrentievoordeel uithalen. Echter, alhoewel de marginale kosten van het dupliceren van data bijna nul zijn, betekent dat niet dat iedereen toegang kan krijgen tot alle data. Bedrijven kunnen de toegang tot data namelijk beperken of er kunnen juridische beperkingen zijn die het delen van data onmogelijk maken.

Wij hebben vijf factoren geïdentificeerd waarmee risico's van datagebruik voor de mededinging kunnen worden bepaald

Data kan bijdragen aan het ontstaan van marktmacht. Marktmacht betekent dat een bedrijf zich in zekere mate onafhankelijk kan gedragen van concurrenten. Dit fenomeen komt in veel markten voor en is niet per definitie een probleem, dat is het pas als er sprake is van langdurige marktmacht waarbij er geen dreiging is van concurrentie of misbruik van marktmacht, waarover in de volgende paragraaf meer. Op basis van literatuurstudie en casestudies hebben we vijf factoren geïdentificeerd die van invloed zijn op de mate waarin (tabel 1).

Tabel 1 vijf factoren om risico's van datagebruik voor mededinging te identificeren[1]

	Factor	Toelichting
1	**In welke mate zijn de data exclusief?**	Als concurrenten ook over data kunnen beschikken dan kunnen data geen bron zijn van marktmacht. Bedrijven kunnen er zelf voor kiezen om data niet te delen of kunnen daar door regelgeving toe gedwongen zijn, privacy-wetgeving beperkt bijvoorbeeld de mogelijkheden voor handel in persoonsgegevens. Hierdoor hebben bedrijven die wel over de data beschikken een voordeel ten opzichte van concurrenten. Het kan ook zo zijn dat het voor een concurrent wel mogelijk is om de benodigde data te verzamelen maar dat daar aanzienlijke kosten mee gemoeid zijn waardoor er sprake is van een toetredingsbarrière.
2	**Levert het gebruik van een product data op die gebruikt kunnen worden om het product te verbeteren?**	Data kunnen een input zijn om de kwaliteit van een product of dienst te verbeteren. In bijna elk productieproces treedt een leereffect op waarbij meer data resulteert in betere producten of diensten. Als dit gebeurt in combinatie met een netwerk van gebruikers dan kunnen zogenoemde (indirecte) datanetwerkeffecten optreden: meer gebruikers leidt tot meer bruikbare data leidt tot betere producten leidt tot meer gebruikers, en zo voort. Denk bijvoorbeeld aan een e-commerce aanbieder die aankoopgegevens gebruikt om betere suggesties te doen. Hoe meer aankopen, hoe meer data, hoe beter de suggesties, hoe meer aankopen en zo voort. Kernvraag is hoeveel data nodig is om het product te verbeteren en in hoeverre de opbrengsten van meer data dalen bij een hoger volume. Als er sprake is van dalende meeropbrengsten (wat vaak het geval is) kunnen ook bedrijven die beschikken over een kleiner marktaandeel data gebruiken om hun producten of diensten te verbeteren.
3	**Worden data gebruikt voor het orkestreren van interacties tussen gebruikers van een netwerk?**	In sociale netwerken zoals Facebook worden data gebruikt om directe en indirecte netwerkeffecten tot stand te brengen. Een netwerk kan door concurrenten moeilijk gerepliceerd worden omdat een beginnend netwerk voor gebruikers minder aantrekkelijk is vanwege het beperkte aantal andere gebruikers. Hierdoor kunnen er sterke toetredingsdrempels zijn.
4	**Zijn er andere assets die complementair zijn aan de data? Zijn deze exclusief of zijn er substituten voor?**	Andere data, of andere kapitaalgoederen kunnen noodzakelijk zijn om data nuttig te kunnen gebruiken. Als er bijvoorbeeld bepaalde algoritmes nodig zijn om data nuttig te kunnen gebruiken dan kan een bedrijf marktmacht hebben ook al is de data ook voor anderen toegankelijk. Als een concurrent bijvoorbeeld zou beschikken over alle zoekdata van Google Search dan zou het nog zelf algoritmes moeten ontwikkelen om een vergelijkbaar product aan te bieden.
5	**Zijn er bedrijven die met een vergelijkbaar of ander business-model concurrentie vormen voor het bedrijf dat toegang heeft tot de data?**	Vergelijkbare diensten kunnen vaak op basis van verschillende databronnen worden aangeboden (denk bijvoorbeeld aan file-informatie die zowel op basis van GSM-informatie als op basis van sensordata aangeboden kan worden). Om te bepalen hoe essentieel specifieke data zijn moet geanalyseerd worden met welke andere data dezelfde producten of diensten (of andere diensten die een concurrentiedreiging vormen) kunnen worden aangeboden. Voor de concurrentiedruk is ook van belang of afnemers "multi-homen", dat wil zeggen dat zij vergelijkbare diensten naast elkaar gebruiken.

[1] We gaan hier uit van data die als productiefactor waardevol zijn voor het aanbieden van een product of dienst.

De vijf geïdentificeerde factoren zijn zowel van toepassing op analyse van marktmacht door "data" als "big data". Door big data is het belang van het combineren van diverse (complementaire) data (factor 4) echter toegenomen. Het voordeel van diverse data neemt verder toe als een groot volume data nodig is om een product of dienst te verbeteren (factor 2). Bovendien zijn er digitale platforms ontstaan, die aan data versterkte netwerkeffecten ontlenen (factor 2 en 3).

De grootste potentiele risico's van big data voor marktmacht zijn daarom te vinden bij bedrijven die netwerkeffecten tot stand brengen en toegang hebben tot veel en diverse data. Ook deze bedrijven kunnen mogelijk beconcurreerd worden door bedrijven met een ander business model (factor 5).

3. Schadelijke gevolgen voor mededinging van data-gebruik

Data kunnen in theorie schadelijke gevolgen hebben voor de mededinging als gevolg van (misbruik van) marktmacht en collusie

(Misbruik van) marktmacht

Marktmacht kan in datagedreven markten dezelfde nadelen hebben als in andere markten. Er ontstaat benadeling van consumenten en welvaartsverlies wanneer er als gevolg van marktmacht bijvoorbeeld te hoge prijzen zijn of te weinig innovatie is. Bedrijven kunnen ook misbruik maken van marktmacht door consumenten of andere bedrijven uit te buiten of door concurrenten tegen te werken. Daarnaast zijn er nadelen die buiten het domein van het mededingingsrecht vallen, zoals verminderde prikkels tot innovatie. Het is onduidelijk hoezeer deze nadelen anders van aard of ernst zijn in datagedreven markten. Niet zozeer omdat er weinig mededingingszaken zijn geweest waarin misbruik van marktmacht is gerelateerd aan data – dat zou in theorie ook kunnen betekenen dat het mededingingsrecht niet toereikend is – maar wel omdat er geen empirisch inzicht is in de mate waarin die negatieve gevolgen zich voordoen.

Wel een specifiek datagerelateerd risico betreft privacy. In markten waarin bedrijven persoonsgegevens verzamelen kan het zijn dat bedrijven de privacy van consumenten onvoldoende beschermen. Hoewel dit risico niet uniek is voor markten waarin bedrijven marktmacht hebben, kan het wel groter zijn bij marktmacht. De reden is dat prikkels tot privacybescherming lager kunnen zijn als consumenten geen of weinig alternatieven hebben als zij de privacybescherming onvoldoende vinden. Dit impliceert dat consumenten inzicht nodig hebben in de wijze waarop bedrijven gegevens gebruiken, wat vaak niet het geval is, aangezien veel consumenten de voorwaarden ongezien accepteren.

Gepersonaliseerde prijsdiscriminatie

Big data stelt bedrijven in staat om hun prijzen af te stemmen op de specifieke kenmerken van elke klant. Op zich is dergelijke prijsdiscriminatie geen nieuw fenomeen, zoals de prijshandelaren op bazaars laten zien. De schaal en precisie is in online, datagedreven markten echter vele malen groter. Dit is een gevolg van de mogelijkheid om uit verkregen gebruikersgegevens meer te leren over de betalingsbereidheid van gebruikers. Deze informatie, in combinatie met de lage kosten op internet van prijsdifferentiatie, stelt bedrijven in staat elke klant een prijs te rekenen die hij of zij er precies voor over heeft.

Gepersonaliseerde prijsdiscriminatie kan voor- en nadelen hebben. Als gebruikers geen goedkoper alternatief hebben als gevolg van marktmacht, kunnen bedrijven prijzen per klant in theorie opdrijven tot precies de prijs die de klant maximaal bereid is te betalen. Hiermee zou dan alle welvaart van consumenten (het consumentensurplus) verschuiven naar producenten (het producentensurplus) wat een vorm van misbruik van marktmacht zou kunnen zijn. Tegelijk kan gepersonaliseerde prijsdiscriminatie echter positieve gevolgen hebben voor afnemers met een relatief lage betalingsbereidheid die een lagere prijs gerekend krijgen en voor innovatie voor zover de producent zijn surplus aanwendt voor innovatiegerichte investeringen. In de praktijk lijkt pure,

gepersonaliseerde, prijsdiscriminatie waarbij voor alle consumenten de precies de prijs wordt opgedreven tot de prijs die ze maximaal bereid zijn te betalen, momenteel echter nog niet of vrijwel niet voor te komen.

Kartelvorming

Data voegen een nieuwe manier toe om kartels te vormen. In media en wetenschappelijke literatuur is gesuggereerd dat data en slimme algoritmes het makkelijker maken om een kartel te vormen. Dit kan door algoritmes expliciet zo te programmeren maar ook onbedoeld doordat een algoritme zichzelf leert dat de optimale prijs bereikt wordt door de prijs af te stemmen met concurrenten. Overtuigend bewijs van het bestaan van dergelijke kartels hebben wij niet in de literatuur aangetroffen maar ze zouden in de toekomst wel kunnen ontstaan.

4. Mededingingsbeleid en -toezicht in data-intensieve markten

Mededingingsautoriteiten kunnen mededingingsregels ook toepassen in data-gedreven markten

Onze analyse suggereert dat bestaande mededingingsregels veelal goed toepasbaar zijn in datagedreven markten. Dit betreft de wettelijke instrumenten die de Autoriteit Consument en Markt (ACM) en de Europese Commissie hebben om het ontstaan van marktmacht door fusies en misbruik van marktmacht tegen te gaan. Zij kunnen bijvoorbeeld fusies blokkeren of maatregelen opleggen aan bedrijven die marktmacht misbruiken. De mededingingsregels zijn generiek geformuleerd, mededingingsautoriteiten kunnen daardoor hun instrumentarium aanpassen voor toepassing in casussen waarin data centraal staan, daar is geen wetswijziging voor nodig. Een uitzondering vormen de omzetgrenzen die van toepassing zijn bij het melden van concentraties. Als de waarde van een overname een bepaald bedrag overstijgt als gevolg van waardevolle data, analysetechnieken of datatoepassingen, dan zouden fusies ook gemeld moeten worden.

Ex-ante maatregelen op het gebied van datadeling en dataportabiliteit kunnen marktmacht beperken maar vragen een zorgvuldige verkenning en afweging

Waar data bijdragen aan markmacht vormen datadeling en dataportabiliteit mogelijke mitigerende maatregelen.[2] Hierdoor is de data niet langer exclusief (factor 1) en kunnen, indien van toepassing, ook concurrenten de data gebruiken voor het genereren van leereffecten (factor 2) en innovatie. Dataportabiliteit kan een instrument zijn om het voordeel van netwerkeffecten (factor 3) te verminderen, dat geldt in het bijzonder als derde partijen toegang krijgen tot het netwerk. Dergelijke maatregelen behoeven echter eerst zorgvuldige verkenning, omdat deze de prikkels tot innovatie met data kunnen verminderen en hoge uitvoeringslasten mee kan brengen.

Meer in het bijzonder kan worden verkend of verruiming van de criteria voor vaststelling van een zogeheten essentiële faciliteit gewenst kan zijn. In het mededingingsrecht moeten bedrijven onder bepaalde condities toegang geven aan concurrenten tot een “essentiële faciliteit”. De door het Europese hof vastgestelde zijn hiervoor binnen het kader van het mededingingsrecht criteria om te voldoen aan een “essentiële faciliteit” zijn hoog, mede vanwege de mogelijk negatieve gevolgen voor toekomstige innovatie. Omdat de kosten van dataverzameling laag zijn (volgens sommigen zijn data zelfs een bijproduct) zijn maatregelen die gericht zijn op datadeling of dataportabiliteit mogelijk minder schadelijk voor innovatie dan soms verondersteld. Zelfs als bedrijven andere bedrijven toegang zouden moeten geven tot data zouden zij hierdoor nog een prikkel kunnen hebben om nieuwe producten te ontwikkelen. Dit kan mogelijk een argument vormen voor rechtbanken voor een verruiming van de criteria voor het vaststellen van een essentiële faciliteit.

[2] In de Algemene Verordening Gegevensbescherming die in 2018 van kracht wordt is het recht op dataportabiliteit voor persoonsgegevens opgenomen, het is nog onduidelijk in hoeverre dit recht het in de toekomst gaat vereenvoudigen om over te stappen naar een andere aanbieder.

Naast het mededingingstoezicht zijn er andere instrumenten om consumenten (en bedrijven) te beschermen

Veel van de problemen die samenhangen met de data-economie zoals onduidelijke of onevenwichtige algemene voorwaarden en verkeerd gebruik van persoonsgegevens vallen onder de reikwijdte van regels op het gebied van consumentenbescherming, privacy en databescherming. Deze regels zijn van toepassing op alle bedrijven, of zij marktmacht hebben of niet. Daardoor is toepassing en aanpassing van deze regels over het algemeen eenvoudiger te realiseren als toepassing van de mededingingsregels of de introductie van sectorspecifieke ex-ante regelgeving. Voor toezichthouders en beleidsmakers zijn er dus meerdere aangrijpingspunten. Omgekeerd zijn mededingingsregels minder geschikt om problemen op het gebied van privacy en databescherming aan te pakken.

Een specifiek aandachtspunt vormen jonge, snelgroeiende bedrijven omdat zij tegengewicht kunnen bieden aan marktmacht

Jonge bedrijven die beschikken over een netwerk of data worden soms overgenomen door gevestigde partijen. Hoewel zij vaak niet beschikken over veel omzet kunnen waarderingen hoog zijn door de waarde van toegang tot de data, het netwerk en toekomstige groei. Deze fusies kunnen economische voordelen opleveren maar ook mogelijke concurrentie uit de markt halen. Mededingingsautoriteiten zullen in de beoordeling van dergelijke fusies rekening moeten houden met de ontwikkeling van de markt, wat niet eenvoudig is omdat het moeilijk is om een goede inschatting te maken van de concurrentiedruk die uitgaat van huidige en toekomstige concurrenten. Dit vraagt om aanscherping van het instrumentarium en criteria om bijvoorbeeld de omvang van de markt en het marktaandeel vast te stellen. Die dienen namelijk niet statisch beschouwd te worden maar rekening te houden met technologische en marktontwikkelingen.

De overheid heeft ook enige invloed op factor 5 (concurrentie vanuit alternatieve businessmodellen). Ze kan bedrijven die een mogelijke concurrent vormen of kunnen gaan vormen ondersteunen. Een bedrijf met een dominante positie kan die positie verliezen aan een uitdager die met andere data of een ander bedrijfsmodel een beter product of dienst aanbiedt. De overheid kan uitdagers faciliteren door bijvoorbeeld *start-ups* te ondersteunen en door te voorkomen dat wet- en regelgeving, ook maar niet uitsluitend ten aanzien van datagebruik, onnodig belastend is voor kleinere bedrijven. Experimenteerruimte in een *regulatory sandbox* kan dergelijke bedrijven mogelijk een steun in de rug geven.

5. Conclusies en aanbevelingen

Uit ons onderzoek blijkt dat het gebruik van data gevolgen kan hebben voor de mededinging. Op basis van de zich nog ontwikkelende literatuur is het onze indruk dat het ontstaan en misbruik van marktmacht het belangrijkste aandachtspunt vormt. Het aantal mededingingszaken waarin misbruik van marktmacht is gekoppeld aan het gebruik van data is weliswaar nog beperkt maar data kunnen wel bijdragen aan het ontstaan van marktmacht. Kartelvorming door het gebruik van data en algoritmes lijkt nog niet of nauwelijks voor te komen maar dergelijke kartels zouden in de toekomst wel kunnen ontstaan en vragen dus onderzoek van toezichthouders.

De grootste potentiele risico's voor de mededinging zijn te vinden bij bedrijven die beschikken over marktmacht doordat zij netwerkeffecten tot stand brengen en toegang hebben tot veel en diverse data. Hoewel de marktmacht van deze bedrijven niet onbedreigd hoeft te zijn vereisen markten waarop deze bedrijven actief zijn wel de aandacht van toezichthouders en beleidsmakers.

ACM en de Europese Commissie beschikken over instrumenten om op te treden tegen misbruik van marktmacht en om schadelijke fusies tegen te houden, ook in data-intensieve markten. Samenwerking tussen toezichthouders kan nodig zijn om negatieve effecten van marktmacht tegen te gaan. ACM zou daar goed toe in staat moeten zijn omdat het in tegenstelling tot autoriteiten in

andere landen zowel toeziet op het mededingingsrecht als consumentenrecht en ook al samenwerkt met de Autoriteit Persoonsgegevens. Vanaf 2018 kiest een bedrijf in welk EU-land het onder toezicht staat van een privacy-toezichthouder. Hierdoor kan samenwerking tussen toezichthouders in verschillende EU-landen noodzakelijk zijn.

De ontwikkelingen op data-gebied en in het bijzonder kunstmatige intelligentie en *machine learning* staan niet stil. Die ontwikkelingen vereisen niet gelijk een aanpassing van het regelgevend mededingingskader maar het heeft wel gevolgen voor de uitvoering van het toezicht. Voor toezichthouders en beleidsmakers is het daarom van belang om kennis in huis te hebben om bedrijven die veel gebruik maken van data en algoritmes te doorgronden. Die kennis is namelijk nodig om de rol van big data in specifieke casussen te kunnen beoordelen.

Summary

1. Introduction

More and more data is becoming available, as are the applications based on data analysis

The digitization process in the last few decades has resulted in an increase in the amount of available data. Not just the volume and variety of data, but also the opportunities to analyse the data have multiplied. Moreover, technological developments in the field of artificial intelligence and machine learning have increased the possibilities to perform tasks based on data. “Big data” refers to this process. Big data provides companies with opportunities to improve the quality of their product or service, to offer new information services to consumers, and to personalise products and services. Although there are many benefits of big data, concerns have been expressed in the public debate that big data contributes to a dominant position for some market players. Such a dominant position can be detrimental to consumers and other companies. The Dutch Ministry of Economic Affairs has commissioned Ecorys to study the relationship between big data and competition from an economic perspective as this relationship has not yet been evaluated in a systematic way taking into consideration the specific policy context in the Netherlands.

2. Market power

Data has some special economic characteristics that are relevant for an analysis of market competition

In order to analyse the relationship between big data and competition, insight is needed in the economic characteristics of data and big data. Data can be considered a production factor or an input in the production of goods and services. A distinctive characteristic from an economic perspective is that it is a non-rivalrous good which means that if someone is using data, it does not prevent others from using the same data. If data would be non-exclusive it would be impossible to prevent others from using it. A good that is both non-rivalrous and non-exclusive is a public good. If data would be a public good it would be accessible to everyone. That would mean that data could never constitute a competitive advantage or barrier to enter a market. However, although data is non-rivalrous and the marginal costs of reproduction are very low, it can be exclusive, parties can be practically or legally excluded from access to data.

We have identified five factors that can influence the extent to which the use of data can result in market power

Data can contribute to the creation of market power. Based on the literature and case studies we have identified five factors that can influence the extent to which the use of data can result in market power (table 1). If a company has market power it can behave independently from competitors. Market power is a natural phenomenon in most markets, it becomes problematic if a company with market power is not challenged in the long run or if it abuses its dominant position, see the next paragraph.

Table 1 five factors to identify risks of the use of data for competition[3]

	Factor	Explanation
1	**Exclusivity** – Is the data exclusively available to one company or can other companies obtain access as well?	If competitors have access to similar data it cannot be a source of market power, but if a company has exclusive access to data it can form a competitive advantage. Companies can choose to restrict access to data or they may be legally forced not to share it, privacy rules for example pose limits on the trade of personal data. There can also be practical barriers or transaction costs involved in the acquisition of data which may result in an entry barrier to a market.
2	**Learning effects** - Does the use of data contribute to learning effects that can be used to improve the product or service?	Data can be used as an input in the production of products and services. If economies of scale or learning effects exist more data results in better products and services. When learning effects arise in a network they are sometimes referred to as indirect data network effects. These effects arise if more users in a network lead to more data which a better product, which then in turn leads to more users, etc. An example is an e-commerce supplier that uses sales data to improve product recommendations. If sales increase more data becomes available which results in better data. Two key questions are: "how much data is needed to deliver the product?" and "to what extent there are diminishing returns of additional data?". If there are no increasing returns to scale, not only the market leader but also companies with a small market share may benefit from learning effects.
3	**Orchestration of interaction on a network** - Is data used to bring together various types of users on a platform?	In social networks such as Facebook, data is used to orchestrate direct and indirect network effects between users on the platform. For competitors it is difficult to replicate a network as they do not benefit from the size of the network, which results in a barrier to entry.
4	**Complementary assets** - Are there any assets that can be considered complementary to the data? Are they exclusive or are substitutes available?	Other data or other capital goods may be needed to apply data in a business model. If for example a specific algorithm is needed for a use case, a company that has exclusive access to an algorithm may have market power, even when the data is not exclusive. Competitors that would obtain access to all of the search data of Google Search would for example still have to develop algorithms to become a viable competitor to Google Search.
5	**Competing business models** - Are there any companies that use a different business model but compete with the company considered?	Identical services can sometimes be based on different data sources (e.g. traffic information can be based on data from smartphones but also on sensor data). In order to determine how essential a particular data source is in a business model, one needs to consider what other data can be used as a substitute. The competitive pressure from alternative business models also depends on the extent to which consumers "multi-home" (practice of using multiple competing services simultaneously).

The five factors can be applied in both "data" and "big data" use cases. The big data revolution has increased the relevancy of combining various data sets (factor 4). The benefits of combining multiple data sets further increase if large amounts of data are needed to improve a product or service (factor 2). Big data has also contributed to the rise of digital platforms, these platforms facilitate interactions between users based on data (factor 3). The biggest potential risk of the use of data for competition can be found in markets in which companies have access to a large volume and variety of data and use it to orchestrate network effects. However, even the market power of

[3] We assume that the data concerned is an essential production factor to supply a product or service.

such companies can be constrained by competitive threat from existing or new alternative business models (factor 5).

3. Data and competition – theories of harm

The use of data can theoretically result in consumer harm due to (abuse of) a dominant position and collusion

(Abuse of) a dominant position

Market power results in a welfare loss if it leads to, for example, excessive prices or a lack of innovation. This is not a unique feature of data-driven markets, abuse of a dominant position can take place in any market with a dominant supplier. Companies can also abuse a dominant position by, for example, preventing rivals from accessing critical assets and achieving scale.
Not all harm that results from a dominant position falls within the scope of competition law. For example, market power theoretically can result in diminished incentives to innovate. It is uncertain if the disadvantages of market power in data driven markets are larger than or different from other markets because there is a lack of empirical insight into the extent to which market power results in harm.

A particular risk in markets in which companies collect personal data is that they do not protect the privacy of consumers sufficiently. This risk is not unique to markets with a dominant player, but at least in theory the risks may be higher as consumers have limited opportunities to use an alternative supplier if they consider privacy protection insufficient. Therefore, having a dominant player in the market may result in diminished incentives to invest in privacy protection. Note that this requires that consumers have insight into the data protection practices of the company concerned, many consumers are not aware of those practices.

Price discrimination

Big data provides companies information on the willingness to pay of their users which allows them to apply personalised prices. Price discrimination can be a particular form of abuse of a dominant position but does not necessarily harm consumers and is common in many markets, see for example merchants on a traditional bazaar. If there are no alternatives in the market, companies can theoretically charge a price equal to the maximum price a consumer is prepared to pay. As a result there would be no consumer surplus, but just producer surplus. Price discrimination can also result in an increase in consumer surplus when consumers with a low willingness to pay are charged their low reservation prices or if the surplus is used to fund investments in innovation. In practice there seem to be no (or almost no) examples of perfect price discrimination based on personalized prices.

Collusion

In the media and scientific literature it is suggested that data and smart algorithms can facilitate cartels. Cartel agreements limit competition between companies. Data-driven cartels might be the intentional result of programming efforts, but might also result from the autonomous decisions made by an artificially intelligent algorithm. We have not encountered convincing real-life examples of such cartels in the literature but they could emerge in the future.

4. Competition policy in data-driven markets

Competition authorities can apply the competition rules in data-driven markets.
Our analysis shows that existing competition rules can be applied in data-driven markets. The Dutch competition authority (Autoriteit Consument en Markt (ACM)) and the European Commission have the necessary instruments needed to take measures against companies that abuse a dominant position or to block mergers that would result in a dominant position.

Competition rules are generically formulated, this allows competition authorities to adapt their instruments to the particular circumstances in data-driven market which does not require changes to the law. An exception are the notification thresholds, these should not be based on revenues alone, but should also consider the transaction value.

Ex-ante measures such as data portability and data sharing can constrain market power but require careful consideration

In markets in which data is a source of market power, data sharing and data portability may be mandated in order to constrain the company with market power and to enable other companies to compete effectively.[4] With data sharing and data portability, data is no longer exclusive (factor 1) and competitors can use the data to generate learning effects (factor 2). Data portability can be an instrument to enable competitors to orchestrate network effects on their own platform (factor 3). Such measures require careful consideration, as regulations can distort the incentives to invest and to innovate and are difficult to implement.

If data would be considered as a basis for a theory of harm, the criteria for access to an "essential facility" as set by the European Court of Justice are relevant. If an asset qualifies as an essential facility, the result is that its owner is under a duty to deal with rivals and has to share access to the asset at issue. The criteria to be an essential facility are high, one of the reasons being that it can diminish incentives to invest and to innovate. However, if data is just a by-product and can be processed against low cost because it is machine generated and processed, the negative effects of data access or sharing may be lower in data-driven markets than the negative effects of accessing (non-data) assets in other markets. With data sharing there would still be sufficient incentives to invest. This may provide an argument to apply criteria in data-driven markets that are less stringent than the criteria for an essential facility as established by the European Court. This argument could also provide a basis for ex-ante legislation regarding data sharing and data portability.

Competition law is just one of the instruments that can be used to protect consumers (and companies)

Many potential problems related to the data economy such as unclear or unbalanced terms and conditions and misuse of personal data fall within the scope of consumer law and privacy and data protection law. Consumer and data protection rules are applicable to all companies, regardless of whether they are dominant or not. It is generally easier to apply or adjust those rules than to use competition rules or to introduce sector specific (ex-ante) regulation. Thus, there can be multiple ways to address competition problems. Competition rules in contrast are less suitable to deal with issues in the field of privacy and data protection.

Young, fast growing companies require attention as they can challenge incumbents with market power

Young companies that have developed a network or have access to data are sometimes acquired by an incumbent. Even when revenues are still low the valuations for such companies can be high based on the value of access to the data, the network and growth potential. Such takeovers can have positive economic effects, but they can also result in less (potential) competition in the market. Competition authorities have to consider market dynamics in their assessment of such mergers, which is not straightforward to do as it is difficult to assess the competitive threat of existing and new alternative business models. This assessment of mergers requires a constant renewal of the

[4] The General Data Protection Regulation (GDPR) that will come into force in 2018 contains provisions regarding the portability of personal data. It is uncertain to what extent data portability will enable consumers to switch to a different provider.

toolbox used by competition authorities and new criteria to define the market and market shares, considering technological and market developments.

Governments can influence factor 5 (competition from alternative business models) to some extent. A company with a dominant position can lose that position to a challenger that is able to offer a superior product based on other data or a different business model. Governments can facilitate challengers by offering support to start-ups. Compliance with data protection and other legislation can be burdensome to small companies, especially when they have an innovative business model. These companies may benefit from a "regulatory sandbox" in which they are allowed to experiment.

5. Conclusions and recommendations

Our analysis shows that the use of data can have consequences for the competitive process. Based on the developing literature, it is our impression that the main concern from a public policy perspective is dominance and abuse of a dominant position facilitated by the use of data. Although there are few examples of competition cases in which the abuse of a dominant position was related to the use of data, data can contribute to the creation of market power. We have not encountered any real-life examples of data-driven cartels but such cartels might emerge in the future, competition authorities should remain vigilant.

The main potential competition risks can be found in markets in which companies orchestrate network effects and have access to a large volume and variety of data. Although those markets may be contestable they require attention from regulators and policymakers.

The Dutch competition authority and the European commission have the instruments to fight the abuse of a dominant position and to block mergers, these instruments can also be applied in data-driven markets. Cooperation between regulators may be necessary to address competition problems. The Dutch competition authority should be adequately equipped to do so as both competition and consumer protection law are within its realm, moreover it already cooperates with the privacy and data protection regulator (Autoriteit Persoonsgegevens). From 2018 onwards, companies can select the EU Member State that is responsible for oversight, which means that cross-border cooperation between regulators may be necessary.

There are many technological developments in data science and artificial intelligence and machine learning in particular. These developments do not necessarily require changes in competition rules, but they may require changes in the application of the rules by competition authorities. The number of cases in which data plays a role will likely increase, to really understand the business model of the companies concerned it is necessary that regulators obtain expertise in data and computer science. That expertise is needed to assess the role of big data in specific cases.

1 Introduction

1.1 Big data revolution

Over the last ten to twenty years huge amounts of data from many different sources are becoming available and the technology to generate, process, store and analyse data has improved significantly. "Big data" is the buzzword that is frequently used to refer to this phenomenon.

Big data has now become so ubiquitous that it is no longer mentioned in Gartner's "Hype Cycle report", although related technologies such as machine learning, personal analytics and smart robots are.[5] Examples of big data applications are online search engines, targeted advertising on social media, medical tools that intelligently combine different data sources for diagnostic purposes and autonomous driving cars which rely on machine learning and a large volume of data. The many applications of big data bring many benefits to society. It provides companies for example with valuable information that can be used to improve products and services, while consumers have access to more and often very targeted or tailored information which can help them in making better decisions.

Although the benefits of big data are clear, there are also concerns expressed in the public debate that companies that have access to data become too powerful. In the past two years, a number of reports have been published by competition authorities and academic researchers on how big data can impact competition. Some of this literature is quite alarmistic. Authors warn for example that data and algorithms can facilitate collusion and that companies that have exclusive access will "tip" markets. Other authors argue that there is no reason to worry because data is often freely available, easy and inexpensive to collect and for these reasons cannot provide a competitive advantage.

The Dutch Ministry of Economic Affairs has engaged Ecorys to study the relationship between big data and competition. The Ministry did so for three reasons. Firstly, because the diverse economic literature on the relationship has not yet been evaluated in a systematic way. Secondly, the debate lacks much needed empirical insights. Thirdly, so far not much attention has been paid to the particular policy context of the Netherlands.

The main research question is formulated by the Ministry as: *what are the implications of big data for competition?* The ministry is specifically interested in the risks of big data for market power, consumers and competitors and a framework that can be applied to analyse those risks in specific cases. Naturally, the ministry is also interested in implications of the analysis for public policy.

1.2 Approach

In order to answer the main and underlying research questions, we use four sources of information:

1. Literature – there is a small but rapidly developing literature on big data and competition. We have reviewed this literature;[6]
2. Competition cases – in a few merger reviews the European Commission has considered the role of data. We have reviewed and highlighted these cases as they provide examples of how

[5] Forbes.com (15 March 2017, LINK).

[6] Cyril Ritter maintains a bibliography of materials relevant to the interaction of competition policy, big data and personal data that we have thankfully used in our literature review.

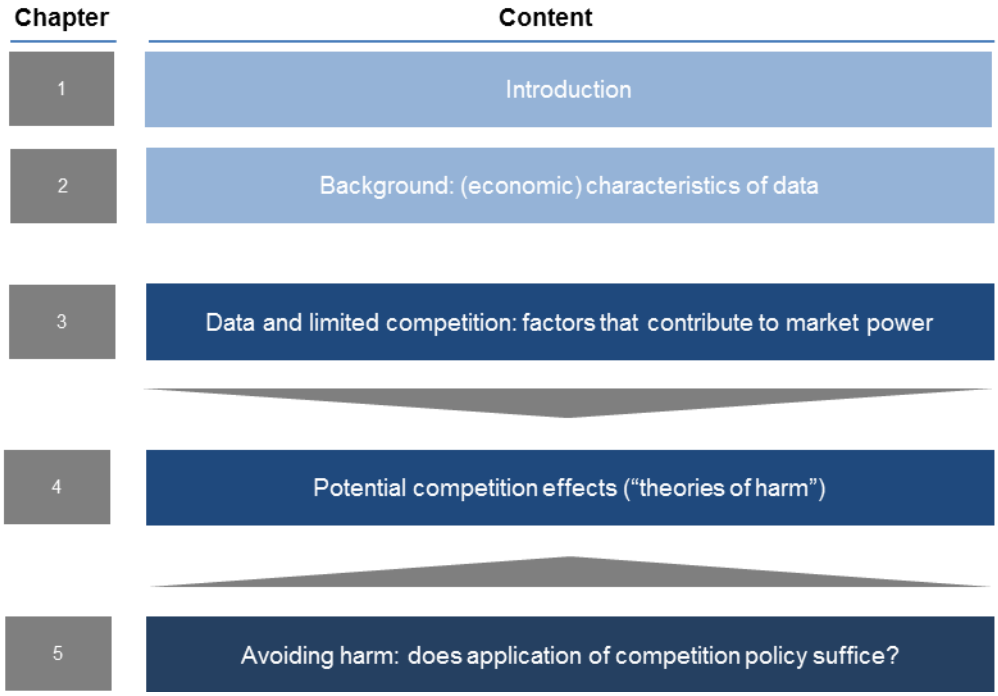
data-driven markets can be analysed. Specific competition cases are also relevant to answer the question if competition authorities are equipped to deal with competition issues in data-driven markets;

3. Case studies – the number of competition cases where big data is considered is, although growing, still limited and the literature is often of a theoretical nature. We have therefore analysed the role and usage of data through four case studies (appendix A).
4. Expert working group – on the 17th of February 2017 we hosted a panel of academic experts from various disciplines for a 'working group' on big data and competition.[7] In a full day session we discussed, in depth, three case studies (on Facebook, Google and Amazon, see Appendix A) and policy implications. We have incorporated the insights we obtained from the experts in the report without attributing them directly to the individual experts (although we do reference academic articles written by the experts). Any views expressed in this report are our own.

1.3 Structure of this report

Figure 1.1 shows the structure of this report. Each chapter finishes with a short summary, for a full summary we refer to the management summary. Appendix A incudes the four case studies.

Figure 1.1 Structure of the report



[7] The following experts participated in the working group: Olga Batura, Paul de Bijl, Nico van Eijck, Inge Graef, Jens Prüfer, Lapo Filistrucchii Alexandre de Streel and Pierre Sennelart.

2 Characteristics of data

2.1 Introduction

In this chapter we discuss some basic economic characteristics of (big) data. We do so, because these characteristics determine the ways data, and "big data" in particular, might affect competition. We start the chapter by describing the economic characteristics of data and big data in particular.

Contrary to many other goods, multiple persons can simultaneously use the same data or information. Some information is freely and abundantly available, other information is more difficult to obtain or only available for a price. The marginal costs of reproducing data are generally negligible. These are some of the characteristics of data that distinguish it from other goods.

In the previous paragraph we have loosely used the words data and information as synonyms as most dictionaries do.[8] However, there is a subtle difference between the two. Data becomes useful when it is processed, structured and interpreted. Usually, it is only when data becomes useful that one speaks of information.[9] The technological capabilities to do this with large volumes and variety of data in a short period of time have increased tremendously, programming models such as "MapReduce" allow for the quick processing of massive amounts of data. "Big data" refers to this process but is difficult to define (WRR, 2016), there is no consensus on a definition in the literature. In this report we describe some of the characteristics of big data but we do not propose a definition. In the final paragraph of this chapter we discuss if big data is conceptually different from "regular" data in terms of economic characteristics. We argue that although "big data" is a relevant development for competition policy it is not necessary to assess if the use of data in a specific case can be considered "big data".

Examples of big data applications on which there are case studies in this report (Appendix A):

- Product recommendations by e-commerce companies such as Amazon and Coolblue (online retailer based in the Netherlands) based on transaction data and feedback that is voluntarily provided by users;
- The processing of smart meter electricity data to forecast demand response to an increase in the price of electricity;
- Targeted advertising offered by Facebook to advertisers based on Facebook's user data. This data is often gathered by tracking consumers online;
- The provision of machine-generated results to a search query by a search engine such as Google.

The big data applications considered in the case studies are just some examples, there are many other applications. Other examples of big data applications are medical tools that intelligently combine different data sources for diagnostic purposes and autonomous driving cars which rely on machine learning and a large volume of data.

The examples make clear that there are many benefits of big data applications and that big data can also contribute to competition in markets. Furthermore, in most markets the benefits of big data

[8] The online version of the Merriam-Webster dictionary gives the following definition of data:
1. factual information (as measurements or statistics) used as a basis for reasoning, discussion, or calculation;
2. information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful;
3. information in numerical form that can be digitally transmitted or processed.

[9] In this report we use the singular construction when we refer to data.

will by far exceed harm due to distortions of the competitive process. Nevertheless this report focuses on harm to competition as a result of the use of data. The reason is that this harm is most interesting from the perspective of public policy, since this might provide (possible new) reasons for intervention.

2.2 The characteristics of data

2.2.1 Rivalry and exclusivity

Data is non-rivalrous but can be made exclusive

Data is non-rivalrous, which means that if someone is using data, it does not prevent others from using the same data. If data or information is known, it is difficult to exclude others from using it (which makes it non-exclusive), unless it is protected by a patent or copyright (which makes it exclusive).[10]

If data would be non-exclusive it would be impossible to prevent others from using it since the duplication of data has almost zero marginal costs. A good that is both non-rivalrous and non-exclusive is a public good. Sokol and Comerford (2016) argue that this applies to data as it is ubiquitous, inexpensive, easy to collect and non-exclusive. If data would be a public good it would be accessible to everyone. That would mean that data can never constitute a competitive advantage or barrier to enter the market. The discussion on data and competition would end here, as data cannot contribute to market power if it is a pure public good.

However, although data is non-rivalrous, it can be non-exclusive: parties can be practically or legally excluded from access to data. For example, in order to obtain access to data it may be necessary for a company to build a sufficiently large customer base. There can be economic barriers that prevent a company from obtaining users due to network and experience effects as well as scale economies (Autorité de la concurrence and Bundeskartellamt, 2016). Also, personal data are subject to special data protection rules, which limit the gathering, processing and usage of such data. Data protection rules also limit the possibilities for third parties ("data brokers") to trade data. Box 1 provides some background on the rules that apply to personal and non-personal data.

An analysis on a case-by-case basis is needed to determine if the data used by a particular company is exclusive or is accessible by others as well.

Companies can make the decision not to provide access to data or may be forced not to provide access to data

A certain degree of exclusivity is required for private actors to invest in the production of scarce goods such that supply and demand can result in a price covering the costs of production. With regards to the production of knowledge and information, exclusivity is sometimes arranged legally via patents or copyrights.[11] There may also be other practical and legal barriers to accessing data which contribute to its exclusivity. For example, Autorité de la concurrence and Bundeskartellamt (2016) mention a couple of reasons why access to personal data may be exclusive:

- In order to obtain access to data it may be necessary for a company to build a sufficiently large customer base. There can be practical (but not legal) barriers that prevent a company from obtaining users due to network and experience effects as well as scale economies;
- Personal data are subject to special data protection rules, which limit the gathering, processing and usage of such data. Data protection rules also limit the possibilities for third parties ("data

[10] Christiaans (1998) makes a distinction between excludability and exclusivity, "[when] non-excludability is no intrinsic property of the [good] but arises from institutional arrangements, it should not be called non-excludable but non-exclusive."

[11] The Database Directive provides protection for the database but it does not protect the content of the database, the scheme of the database is protected (CERRE, 2017).

brokers") to trade data. Box 1 provides some background on the rules that apply to personal and non-personal data.

Box 1 – Legal constraints on data collection and use

For an economic analysis of data, not only the technical characteristics of data are relevant but also the legal constraints on data collection and use. This is important as not all the uses of data that are technically possible are allowed. The General Data Protection Regulation (which will come into force in 2018) provides a legal typology of data. The main categories that the Regulation distinguishes are personal and non-personal data. Personal data is defined as 'any information relating to an identified or natural person (data subject)'.

CERRE (2017) provides an overview of rules that are applicable to personal and non-personal data, distinguishing:

Rules applicable to non-personal data and personal data:

- consumer protection rules;
- the protection of intellectual property and trade secrets; and
- competition rules.

Rules applicable to Personal data:

- general data protection law ;
- ePrivacy Directive (sector specific, only applicable to the providers of publicly available electronic communications services); and
- privacy protection (European Convention on Human Rights).

The rules mentioned above pose limits on the collection and use of data. This is especially the case for personal data. Personal data cannot be used for example for purposes for which the '"data subject" has not provided consent. The General Data Protection Regulation that will come into force in 2018 extends the existing rules by introducing a right to erasure and a right to data portability. This can (at least theoretically) allow data subjects to switch between service providers. Contrary to personal data, there is no legislation that mandates the portability of non-personal data.[12] However, data portability may also result from the application of competition law (CERRE, 2017). Competition rules can apply to both personal and non-personal data.

2.2.2 *Substitutability and complementarity*

Data often needs to be combined with other data, in many cases multiple datasets can serve similar purposes

A common way for economists to describe goods and services is in terms of substitutability and complementarity. Goods or services are complementary when an investment in one increases the marginal return of another. When an investment in one reduces the value of another they are substitutes. Data can be both a substitute and complement.

Complementarity relates to the "volume" and "variety" characteristics of big data (that will be discussed in the next section) because one can often create more "value" when combining different types of data. Different types of data can be substitutes and complements at the same time. Data about search behaviour may for example be substitutable for data about buying habits in

[12] The European Commission started a consultation on 'Building the European data economy' in January 2017. In this consultation it collects information on the extent that digital non-personal machine-generated data are traded and exchanged and the barriers to access such data. It also looks into ways to tackle entry barriers. This may result in rules regarding data sharing and portability in the future to some extent but it is unlikely that the difference in treatment of personal and non-personal data will disappear.

developing a marketing proposition vis-à-vis providers of goods and services; but the combination of data about search and buying behaviour can yield an even stronger proposition.

2.2.3 *Perishability (depreciation time)*

The value of data decreases in general over time

Most capital goods depreciate over time. Data can be a resource that perishes (depreciates) rather quickly. The degree of decay depends on what type of analysis one wants to do. With "*nowcasting*" for example (think of providing information on traffic jams in Google Maps), data loses most of its value almost instantly. However, to develop nowcasting algorithms one needs historical data for "diagnostic analyses" (identifying correlations by systematic analysis of the history), so the value of the data does not depreciate to zero.

2.3 The distinction between "big data" and "data"

Technological developments have decreased the costs of processing and analysing large volumes and variety of data

In the previous sub-section we have discussed the characteristics of data. Here we examine if big data is conceptually different from "regular" data in terms of economic characteristics.

In the literature, various definitions of big data have been proposed. The characteristics of big data are often described by the "three V's"[13]: *Volume* (one needs large datasets), *Velocity* (one needs speed at which data is collected, processed, and disseminated), and *Variety* (one needs multi-dimensional datasets). Some add other V's such as *Value* (the value of data depends on its use) to describe other characteristics (Ezrachi and Stucke, 2016). Rubinfeld and Gal (forthcoming) define big data by the characteristic that it cannot be analysed by traditional methods within an appropriate time frame, but rather that it requires the establishment of a unique platform that can manage the collection and analysis of such data sets in a reasonable timeframe. As such, Rubinfeld and Gal argue that besides the composition of the dataset, the ability to synthesise and analyse the data is equally important for the transformation of big data into value. Consequently, they argue that the concept of 'big data' is a moving target given that developments in collection, storage and analytical capabilities have exponentially increased the volume and variety of and the velocity at which data that can be collected and analysed. The Dutch scientific council (WRR, 2016) reaches a similar conclusion, big data is the interaction of a number of developments rather than a phenomenon that can be precisely defined.

The discussion on definitions of big data mentioned above point to some of the economic characteristics that sets "Big data" apart from "data". Thanks to the advances in computer technology and data science, more and more data can be combined with other data. Or in the terminology used in this chapter, data can be considered a complement to other data as well as to the platforms that are used to collect and analyse data sets. Variety in big datasets allows for the substitutability of big data. If a company does not have access to a particular dataset there may be other data that can be used for the same objective (e.g. travel patterns in public transport can be analysed by using smartphone data as a substitute for payment data). The definitions of big data also point to the economies of scale and scope of combining a large volume and variety of data. For the analysis of the effects of the use of data on competition an important question is if there are diminishing returns to economies of scale and scope. If there would be constant returns to scale and scope there would be a tendency to a monopoly in the market but in most use cases it is likely

[13] Gartner analyst Doug Laney introduced the 3Vs concept in a 2001 publication, "*3D data management: Controlling data volume, variety and velocity*".

that the returns of additional data diminish at some point, this will be discussed in more detail in chapter 3.

As there is no formal definition that can be used to distinguish ”data” from “‘big data” there is, in principle, no difference between the analysis of ”data” and “big data” use cases. Therefore, it is not necessary to consider in a specific case if the data concerned can be considered big data or not. That does not mean that the big data revolution has no consequences for the analysis of competition in specific markets as the ubiquity of (personal) data and technologies such as machine learning have increased the importance of data as a factor of production.

2.4 Conclusion

In this chapter, we defined big data loosely. We defined it in terms of the interaction of developments in the volume velocity and variety of data, and in terms of the new computer techniques and capabilities to effectively process and analyse data. Furthermore, we noticed that “big data” is not fundamentally different from “data”, although economies of scale and scope are more relevant. We discussed some of the characteristics of data, the most important being that data is always non-rivalrous, but can be exclusive. If data is exclusive, companies that have access to the data can use it to obtain a competitive advantage vis-à-vis companies that cannot use the data in their business model. The next chapter introduces a framework to analyse factors that contribute to a competitive advantage and market power. In this framework we use the concepts introduced in this chapter.

3 Data and market power

3.1 Introduction

In this chapter, we discuss how the use of data can contribute to market power. An understanding of the relationship between data and market power is needed before we discuss theories of how particular acts of a company negatively can affect the competitive process in the next chapter. Many "theories of harm" are based on the premise that a company has market power or dominance.

Market power is not necessarily undesirable; it can reflect economies of scale for example that are beneficial to consumers. More generally, healthy competition implies that firms strive to obtain market power, and that those firms offering better value are able to exercise market power. The theoretical notion of perfect competition views market power as a market failure, while in reality, it is a natural phenomenon. Few industries fully show the characteristics of the theoretical model of perfect competition. It is more often the case that one or more companies have a competitive advantage over others. It is *persistent* market power and the *abuse* of that power which are problematic aspects.

Access to data may be one of the reasons that a company has a competitive advantage over its rivals. This advantage can be persistent if data results in entry barriers. This is the case when new entrants are unable either to collect the data or to buy access to the same kind of data, in terms of volume and/or variety, as established companies, so that they are unable to realistically duplicate the benefits of the strategy or input (Lambrecht and Tucker, 2015). Hence, data, if it results in an entry barrier, obstructs effective competition.

Data can be an important input contributing to market power but, as follows from the defined characteristics in chapter 2, there can be other inputs that are important as well. Besides, in many cases, data can be non-rivalrous, ubiquitous, with low barriers to entry. In other cases, data it is at least to some extent exclusive and without substitutes. This makes a case by case analysis necessary whereby the characteristics of the data and the business model in which it is used are considered. This chapter provides a simple framework to analyse if the use of data can contribute to a competitive advantage and market power.

3.2 Framework for analysis of relationship between data and market power

Based on the literature we have identified five data-related factors that influence market power

We have identified five mechanisms through which data contributes to market power (see table 3.1.).

Table 3.1 Five factors to identify risks to competition of the use of data

Factor		Effect on market power
1	**Exclusivity** – Is the data exclusively available to one company or can other companies obtain access as well?	+
2	**Learning effects** - Does the use of data contribute to learning effects that can be used to improve the product or service?	+
3	**Orchestration of interaction on a network** - Is data used to bring together various types of users on a platform?	+
4	**Complementary assets** - Are there any assets that can be considered complementary to the data? Are they exclusive or are substitutes available?	+
5	**Competing business models** - Are there any companies that use a different business model but compete with the company considered?	-

It is relevant to note that these five factors may be linked or reinforce each other. For example network effects, which refer to the idea that it is generally better to be connected to a bigger network, may reinforce learning effects which benefit from economies of scale. Obviously, data has to have some added value as a factor of production. If data cannot be used in the production of goods and services it can never be a source of market power.

3.2.1 Factor 1 - Is the data exclusive?

Exclusive access to data can contribute to a competitive advantage and market power

Exclusivity of data may give rise to market power if that data is an essential input to produce a product or service. Data is an essential input and can form an entry barrier for new entrants when rivals are unable to realistically compete without access to the data.

Not all data is available to all companies in a market. Companies spend considerable money and effort to acquire data and to maintain a data-related competitive advantage, countering the argument that data is ubiquitous, low cost and widely available (Stucke and Grunes, 2016). There are a number of reasons why access to data can be restricted.

Firstly, data can be exclusive because companies are not allowed to share data. If data sharing is not allowed companies cannot acquire the data in a "data market". Secondly, companies can choose not to share the data with third parties. Thirdly, companies may also have a temporary advantage when they have instant access to data when others do not have a similar advantage.

1. Legal barriers on data sharing and "data markets"

New competitors to, for example, Facebook are endowed with a smaller number of users and these smaller companies will typically collect less data from its users (first-party data) than larger established companies. A solution to this problem might be to buy the data from a third party (a data broker). However in practice, it is often not possible to obtain the necessary data this way. For example, as we discuss below in Box 1 there are legal obstacles to data brokerage, especially in the case of personal data. In practice consumers might provide consent for data sharing without knowing it as most consumers will not read terms and conditions in detail (or at all).

Box 3.1 Data brokers

The OECD defines data brokers as "companies that gather and merge aggregated information on individuals that is then sold for various uses" (Rieke et al., 2016).

In Europe, the term data broker is less common than in the US. European commentators use a variety of different terms to refer to data brokers, including "information resellers", "data vendors,", "information brokers," "consumer data analytics", and "data warehousing".

Personal data is subject to data protection rules. These rules ensure consumer privacy but limit the gathering, processing and usage of data. Article 2 (a) of the Data protection directive (DPD) defines personal data as "any information relating to an identified or identifiable natural person". According to article 6 data can only be collected "for specific, explicit and legitimate purposes and not further processed in a way incompatible with those purposes". With the "unambiguous consent" of users data can be shared.

Currently, the DPD applies to companies, including companies headquartered outside of the EU that have an establishment or use equipment in a European country in their handling of personal data. In 2018, the General Data Protection Regulation (GDPR) will come into force, replacing the DPD. The GDPR more explicitly "applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not".

Thus, companies that collect valuable personal data have limited possibilities to share the data due to contractual and legal restrictions. Privacy rules limit the provision of data to third parties for commercial purposes but it is not impossible, users can provide their consent to share data and they may not always be aware of data sharing, even if they have provided consent. The French and German competition authorities (2016) conclude that despite the legal limitations operating in the EU brokerage is possible and that the role played by data brokers may vary depending on the country and the type of data under consideration as well as on the extent of privacy protection enjoyed by individual users pursuant to national regulations.

2. Companies can choose to restrict access to data

If a company has access to data, it can be impossible for rivals to collect similar data. A company that already has a dominant position in a market (due to a natural monopoly or other sources of market power) may for example restrict access to user data that is essential to provide services in the market concerned but also in related markets. In this way, market power in one market (that is not necessarily related to the use of data) can be used to obtain a dominant position in other markets. Manufacturers of products may for example choose not to provide access to data that is essential for other products or for repair and maintenance purposes.

3. Temporary exclusive access

When data perishes quickly, but at the same time is scarce because it is difficult to produce and has few substitutes, market power does not stem from controlling access to data as such, but from controlling *instant* access to data. An example is the PeopleBrowsr's use of the Twitter Firehose (a service providing real time access to the data stream). The value was not the tweets itself (which are publicly available on the Twitter website), but from the velocity in processing the volume and variety of tweets (see Stucke and Grunes, 2016).[14]

It is not easy to assess whether data should be considered exclusive. Especially in the case when companies choose to restrict access this is not always straightforward. Other companies may obtain access to the same (or similar) data but that may require an investment. Therefore, a case by case analysis is needed to assess to what extent exclusive access to data can be considered a barriers for competitors to enter the market. Lambrecht and Tucker (2015) argue that if a market for

[14] PeopleBrowsr paid Twitter 1 million USD for annual access to streaming data on every tweet posted on Twitter (a service called the 'Twitter Firehose'). In 2013 Twitter was sued after cutting off PeopleBrowsr's access to the Firehose (see Stucke and Grunes, 2016).

data exists, it is unlikely that big data is inimitable. However, such markets do not always exist and sometimes companies have exclusive control over data. The case studies of Amazon, Facebook and Google in appendix 1 show that they all have data that competitors do not possess. Google for example has the largest share of search queries, while its competitors are only able to collect a fraction of the data that Google gathers. Facebook has detailed data about its users. Although there may be other companies that have access to similar data or at least a part of it (such as Google), Facebook does not provide access to user data and is not allowed to do so based on privacy regulation.

Appendix A also contains a case study on smart (electricity) meter data where data can be considered non-exclusive. In the Netherlands, just as in many other countries, operators of smart meters are obliged to share data with third parties if a consumer gives consent to do so. As the data is non-exclusive there are many companies that can try to build products and services based on smart meter data.

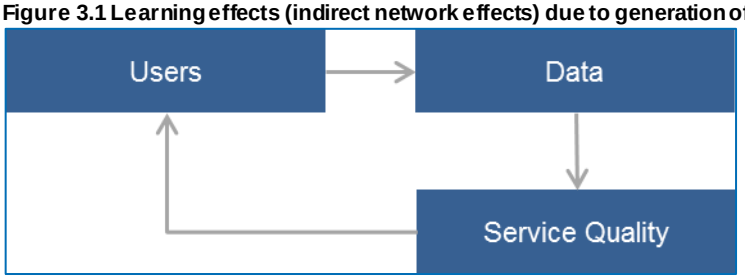
3.2.2 Factor 2 - Does the use of data contribute to learning effects that can be used to improve the product or service?

Access to data may be needed to improve products and services

Many production processes exhibit a learning effect or dynamic economies of scale. When production increases more data is generated, with the generated data the production process can be improved. In many cases there will be diminishing returns to additional data. This is the case in standard statistical analysis and also in advanced machine learning applications. It is a well-known fact for everyone who has followed an introductory statistics course that the more data available (higher volume), the more accurate are the estimated results, but in this regard there are decreasing returns to data. Learning effects can be considered a specific class of (indirect) network effects (see factor 3): individual benefits increase when others use the same service and quality increases (Stucke and Grunes, 2016).

A first mover advantage might be relevant here: a company that has a lead in the number of users will benefit from the data generated by its “launching” users. Once the dominant company is on the “plateau”, its advantage may be incontestable (OECD, 2016; Prüfer en Schottmueller, 2017). In other words, there is a reverse causality between the production and the use of data. Access to data can result in better services, which in turn can attract more users. Smaller companies have less data and less users and do not benefit from this “snowball effect” (see figure 3.1.).

Figure 3.1 Learning effects (indirect network effects) due to generation of user data



First-mover advantages which initially places a company ahead on the learning curve may decrease over time due to diminishing returns to data, allowing new entrants to catch up (Varian, 2016; Lerner, 2015). That said, the point at which the marginal benefit of data starts to decrease, as well as the intensity of such a decrease, depends on the type of data and the type of algorithmic applications. It may be the case that diminishing returns to data applies mostly to a specific kind of data that is used for a specific purpose, where "more data" means adding more of the same data (i.e. more records). However, more data can also mean that the richness of the data is increased

(i.e. new type's records), or that from combining data sets, new inferences become feasible. Overall, such additional data may lead to innovations and new business models. In such cases there may not be diminishing returns to additional data. Therefore, an analysis of learning effects requires an analysis of how much data is needed for data analytics and what the marginal contribution is of additional data. If the costs of obtaining data are small, it is easier for competitors to walk the learning curve. This means that also an analysis of the costs of data collection and processing is needed.

Learning effects and machine generated user data

Learning effects may be especially strong when the costs of obtaining additional data are low. Prüfer and Schottmüller (2017) provide a theoretical model of how "data-driven network effects" can result in the "tipping" of the market in data-driven markets.[15][16] Prüfer and Schottmüller assume that the costs of machine generated use data are zero. Their model shows a strong first-mover advantage in data-driven markets, which leads towards market tipping and monopolisation. This can cause a domino effect whereby a firm can repeatedly leverage its "data advantage" in one (data-driven[17]) market to enter and become a large (perhaps even dominant) player in other markets. Certain conditions must be satisfied for this to occur, multi-homing must be relatively costly or unattractive for customers for example and the costs of data collection have to be negligible.

3.2.3 *Factor 3 - Is data used to orchestrate interactions on a network?*

A network is difficult to replicate, if data is used to orchestrate interactions between users in a network there can be strong entry barriers

Many online markets are so-called "multi-sided markets" or "platforms", where companies serve more than one group of users/customers. A common use-case of data is to orchestrate interactions on a platform. The orchestration of interactions involves matching different users (e.g. consumers with retailers - like with Amazon; consumers with advertisers - like with Facebook; users with content providers - like with Google search or YouTube; or users with each other - like with Facebook, Linked-In, Snapchat etc.) and optimising the user-experience of the interactions following that match.

A distinguishing characteristic of platform markets is the presence of network effects. Network effects refer to the idea that it is generally better to be connected to a bigger network.[18] Network effects may be direct or indirect:

- (positive) **direct network effects (or same side effects)** occur when members of a group profit directly from more members of the same group (for example more users on a social media network). The benefit of consumers comes from the ability to communicate with other consumers via the network;
- **indirect network effects (cross-side effects)** occur when the members of group A profit indirectly from more members of their group joining the platform. Users can profit for example from many people using the same operating system, as it becomes more attractive for software developers to develop software for this system. Users in turn benefit from an increase of choice.

---

15 Prüfer and Schottmüller argue that data-driven network effects are fundamentally different from dynamic economies of scale or learning-effects. In contrast to dynamic economies of scale, data-driven indirect network effects cannot easily be copied by competitors or destroyed by the arrival of a new technology. However, traditional learning effects in for example manufacturing also benefit, at least to a certain extent from user data/feedback.

16 Prüfer and Schottmüller define a data-driven market as: "markets where the cost of quality production is decreasing in the amount of machine-generated data about user preferences or characteristics". The machine-generated data is an inseparable and costless by-product of using services in such markets which gives rise to data-driven indirect (supply side) network effects.

17 Prufer and Schottmüller (2017) divert from the definition of data-driven markets as commonly known in the literature, defining the term as; "*markets where the cost of quality production is decreasing in the amount of machine-generated data about user preferences or characteristics, which is an inseparable by-product of using services offered in such markets*".

18 Belleflamme, P., & Peitz, M. (2015). Industrial organization: markets and strategies. Cambridge University Press.

Due to network effects, access barriers to platform markets can be high. Consumers often use multiple competing services in parallel they "multi-home", e.g. they use both a Google Gmail email-account and a Microsoft Outlook account. Multi-homing may counter a potential tipping-effect and hence market power caused by network effects because users seek for interactions on multiple platforms at the same time. However, a platform with a larger market share generates more data and may be able to offer more information or better services to one or both sides of the platform than a platform with fewer users. More information and better services in turn keep attracting users to the platform and prevents them from seeking interactions elsewhere. Thus data collection and data analysis can further strengthen network effects and decrease the competitive threat from multi-homing in a sort of feed-back loop. This is a feedback loop related to feedback loops based on learning effects, where the collection and analysis of data amplifies network effects.

Appendix A contains a case study on Facebook. Facebook uses data to orchestrate interactions between users and advertisers on its social media platform. In this case study we argue that a "hypothetical data broker" that would obtain access to all Facebook's data would still not benefit from direct network effects. This example shows that network effects can also be considered a stand-alone factor that contributes to the market power of a company.

3.2.4 Factor 4 - Are there any assets that can be considered a complement to the data? Are they exclusive or are substitutes available?

The value of big data lies in the ability of companies in processing that data

To extract useful information from data it has to be processed and analysed. Certain algorithms may be necessary to extract value from data. If the data is useless without the platform or algorithms to analyse it, the company that controls the analytics platform or algorithms can have market power if no alternatives are available. It is not certain for example whether a hypothetical data broker that would obtain all of Google's search data, would be able to compete effectively with Google. Google's algorithm, or a substitute for it, is needed to provide the service. Human capital (e.g. skilled data scientists) may also be a source of a competitive advantage. However, competitors can obtain access, at least to a certain extent, to similar resources by training or recruiting staff.

Scale economies of data infrastructure become less of a competitive advantage with the transformation of fixed cost into marginal costs

Scale economies due to data infrastructure are mentioned in the literature as a possible source of market power (OECD, 2016). The capability to store and process data can in theory form a competitive advantage. Evidently, big data infrastructure can be expensive and new entrants may face barriers to finance investments in infrastructure. It is reasonable to assume that the average costs of infrastructure (needed for collection, storage or analytics) decrease if output (e.g. number of users which results in more data) increases. However, infrastructure and data analytics are increasingly offered "as a service". When infrastructure and analytics can be bought as service this transforms fixed costs into marginal costs meaning that entrants on the market do not have to make upfront investments in infrastructure, reducing scale advantages that larger companies might have.

Companies not only benefit from having access to a lot of data (economies of scale) but also from being able to combine a variety of data (economies of scope)

There is often value in combining data with other data. Facebook is for example able to track users on multiple devices (laptop, phone). This provides Facebook with a more complete user profile, which can be used by advertisers to target specific users or user groups. Similarly, Google combines geolocation data of its users with Google Maps to show users the opening hours of s hops

and restaurants. Amazon owns IMDb, a site with user reviews of movies, and uses IMDb's data to inform shoppers on the Amazon website.

3.2.5 *Factor 5 – Are there any companies that use a different business model but compete with the company considered?*

Competition can come from sometimes unexpected directions

In many digital markets consumers multi-home, that is, they use several providers to get the same kind of service (for example users may use Amazon, eBay and Bol.com to search for and purchase products or users may use Facebook, Twitter and Instagram to post and connect with their network). Autorité de la concurrence and Bundeskartellamt (2016) note that while the potential for multi-homing may decrease the market power of established undertakings (by making substitution much easier as it has been recognised by the General Court of the EU in the *Microsoft/Skype* merger case[19]), this potential multi-homing is not necessarily always sufficiently relevant. For example, multi-homing may only make a difference if end-users use rival providers sufficiently frequently. If users multi-home, but do not use rival products enough, rival companies may not gain the necessary data to offer a comparable service (think for example of real-time traffic data in navigation).

In any market, market leaders face the risk that another company finds a way to produce the good or service more efficiently or to displace the market leader by offering superior products and services. This company may already be on the market or be a new entrant. Amazon is for example not an obvious competitor to Google Search but as it has search functionalities, apps and devices, it can take traffic and users away from Google. As a result, Amazon can, at least in the US market, offer similar features to advertisers and consumers as Google does and hence be considered a competitor (see the case studies in Appendix A).

An often-mentioned argument is that established market power on digital markets can be especially vulnerable to displacement by innovative products and new business models. Evans (2015) supports this view arguing that the risk of entrenched monopolies in platform markets is very limited, evidenced by the historical disruption changes where incumbents have been ousted by new entrants (e.g. MySpace by Facebook, Nokia/Symbian by Google's Android and Apple's iOS). It is for example possible that current social media platforms may be decimated by completely new types of digital innovations such as wearable devices, virtual reality headsets/glasses or something else that we do not yet envisage. Autorité de la concurrence and Bundeskartellamt (2016) argue that while dynamic competition could be strong enough to mitigate concerns related to static market power on at least some markets, this balance should be evaluated on a case-by-case basis.

Finally, dynamic competition may be stifled if dominant companies are able to acquire new entrants before they mature and become a real competitive threat, a so called pre-emptive takeover. However, takeovers are not necessarily bad for competition, such acquisitions can also have pro-competitive effects and they may for example result in the faster adoption of new products and services. (Appendix A includes a case study of Facebook, including a discussion of the acquisition of Whatsapp in 2014. The effects of such pre-emptive takeovers were heavily debated in the context of this acquisition).

[19] Case T-79/12, *Cisco and Messagnet v. Commission*, ECLI:EU:T:2013:635, para 79 et sq.

3.3 Conclusion

This chapter presented a framework to analyse case by case how the use of big data can contribute to market power. Data-driven markets can have a tendency towards high market concentration due to a number of factors. The framework reflects that data is one input, but that other inputs may also be important sources of market power. Markets that are especially vulnerable to competition issues exhibit learning effects (or indirect network effects) and direct network effects. However, even for companies with a business model characterised by network and learning effects, there may be a competitive threat from competitors and new entrants that constrains market power.

The next chapter discusses how acts of companies, especially companies with market power, may negatively affect the competitive process and harm consumers.

4 Theories of harm

4.1 Introduction

In competition cases, it has become standard practice to describe a "theory of harm". This is a theory that sets out how particular acts of a company negatively affect (or may affect) the competitive process. Note that the potential harm is not necessarily limited to the scope of competition law. There can also be negative effects of insufficient competition, such as diminished incentives to innovate, without any infringement of competition law. Likewise, most forms of price discrimination are common phenomena in many markets and are allowed.

Based on the literature we distinguish three possible conduits through which data may affect competition and ultimately consumers. We introduce the three possible conduits below and elaborate on each of them in the remaining sections of this chapter:

1. **(Abuse of) a dominant position (Section 4.2)**
   Market power — a natural phenomenon in almost all markets — is not necessarily undesirable. Abuse of a dominant position is not allowed. There can also be negative effects of market power, such as diminished incentives to innovate, without any infringement of competition law which we also discuss in this section.

2. **Price discrimination (Section 4.3)**
   Price discrimination can be a particular form of abuse of a dominant position but as noted before price discrimination does not necessarily harm consumers and is common in many markets. Data can be a vehicle for price discrimination. By collecting data about their users, a company receives better information about their purchasing habits and is in a better position to assess their willingness to pay for a given good or service. If it has market power, the company would then be able to use that information to set different prices for the different customer groups it has identified. Theoretically, "personalised" pricing can be used to extract all consumer surplus.

3. **Collusion (Section 4.4)**
   In the literature it has been suggested that the developments in the collection and processing of big data, specifically data on competitors' pricing, may be used by companies in ways that could limit competition. Namely, pricing algorithms may be used by companies to constantly monitor competitors' prices and adjust their own prices accordingly in real time (Ezrachi and Stucke, 2016).

4.2 (Abuse of) a dominant position

Market power does not necessarily result in harm

Companies with market power can behave in ways that harm consumer welfare. However, market power can also bring benefits. Companies in industries with economies of scale for example have some degree of market power which is efficient from a welfare perspective, since costs per product of service are lowered. More generally, in many markets, companies have a certain amount of market power, and entrepreneurial activity often corresponds to developing a unique business proposition which (temporarily) results in market power. Thus, market power is a natural characteristic of the way markets work. Problems may occur however if market power becomes

substantial or is sustained for too long (due to entry barriers). In such situations, companies may become dominant, which, in turn, may give rise to abuse of dominance. Assessing harm arising from market power has however become more difficult in data-driven markets which are often multisided markets. This is because in multisided markets, prices reflect ways to generate network effects, possibly irrespective of underlying marginal cost levels. As such, it has become more difficult to assess if a price is above the "competitive price", because such a reference point no longer exists in these markets.

In the discussion on the impact of market power presented in the following sub-sections, we provide examples of how the more general theories of harm in abuse of dominance may apply to "data-driven" markets.[20]

4.2.1 Exclusionary abuse

Companies with market power can use their position to exclude competitors

Exclusionary abuse is conduct by a company with a dominant position which has the object or effect of excluding a competitor from a market. In such cases companies do not compete on the merits of the products and services they provide. This can take for example the form of refusing to supply to a competitor, a margin squeeze and tying and bundling. Other examples are conditional rebates and exclusive purchasing conditions that require customers on a market to purchase exclusively from the dominant company.

One view is that data-driven companies can engage in data-related exclusionary conduct. Stucke and Grunes (2016) list possible data-related practices that companies with market power can use to tip the market in their favour or to maintain their dominant position. Note that there these are theoretical possibilities. To date there have been few cases in which big data played a role.

1. **Exclusive dealing to prevent rivals form accessing critical data and achieving scale**
   A monopoly or company with market power can, through exclusive dealing, foreclose a rivals' access to critical data. This concern is discussed in some of the competition cases (for example the Microsoft/Linkedin merger in 2016) in chapter 5.

   By unfairly preventing smaller rivals and entrants from accessing data, a dominant company can use network effects to widen the quality gap over rivals, attracting more users and advertisers.

2. **Dominant company leverages its data-advantage in a (regulated) market to another market**
   Stucke and Grunes mention the example of GDF Suez (now known as ENGIE), a French regulated energy monopoly. ENGIE was found to have abused its dominant position by using historical data it acquired from its monopoly activities to compete in markets where it faced competition.[21]

3. **Increasing consumers' switching costs**
   To maintain its data-advantage and prevent rivals from attaining scale, a dominant company can increase the switching costs for consumers, thus making it harder for consumers to leave to an alternative service. One way to do this is by reducing the interoperability with other systems

[20] With "Data-driven markets" we mean markets where the use of data is an important element of competitive dynamics. Note that this definition is broader than the one used by Prüfer and Schottmüller (2017).

[21] Autorité de la concurrence, 2017. "Press Release 22 March 2017: Energy sector", available at http://www.autoritedelaconcurrence.fr/user/standard.php?id_rub=663&id_article=2963&lang=en.

or platforms and ensuring data portability. Means to increase switching costs can also be subtle, for example setting an app or service as the default option.

4. **Foreclose through vertical integration by a dominant platform operator**
   Dominant companies can obtain a dominant position in adjacent markets by integrating competing platforms or features into products they control. Stucke and Grunes (2016) refer for example to a statement by Facebook in its annual report of 2012. Facebook warns that *"Competitors, including Google, could use strong or dominant position in one or more markets to gain competitive advantage against us in areas where we operate…"*

   For smaller independent app developers, the risk that they are forced out of the market by a "super-platform" or "platform of platforms" is even higher. Stucke and Grunes (2016) argue that platforms such as Google rely on personal data for maintaining a competitive advantage for advertising. In order to secure sources of data it could introduce its own applications and foreclose access of rival applications. Ezrachi and Stucke (2016) refer to this as the "frenemy" relationship between platform operators and application developers.

   A related possible conduct is that companies use their dominant position (due to data) to sell the product or services in a bundle with other products (tied or bundled sales).

4.2.2 *Exploitative abuse*

Prices

In all markets, market power can result in prices above the competitive level

Market power offsets the balance of bargaining power between a company and its users, and thereby may lead to prices above the competitive market price. Based on competition law, competition authorities can address "excessive prices" of dominant firms but there are no guidelines for when prices can be considered "excessive" (there is limited experience with cases concerning excessive prices).

In markets where companies use data to orchestrate interactions between user groups, there are often prices equal to zero for one user group (at least in monetary terms). But in most of such cases, another user group is paying a monetary price. Thus for two or multi-sided markets, it is important to look at prices in both sides of the market as the price for one user group may include a high profit margin (or mark-up) in the presence of market power.

Choice, quality and innovation

Market power may reduce incentives to offer quality and to innovate

In the previous paragraph we discussed how market power (regardless if it is caused by the use of data or not) can result in prices above the competitive level. Similarly, market power may also result in suboptimal quality, both high prices and low quality can result in excessive profit margins for dominant firms. Besides high prices or low quality market power can also result in a lack of choice. If market power results in tipped markets with just one or a few suppliers, there is less choice for consumers between alternative providers. In a market with entry barriers challengers might not be able to enter market based on differentiation, responding to the heterogeneity of consumer preferences.[22]

Theoretically, market power may result in diminished incentives to innovate. There is an extensive literature on the relation between competition on the one hand and quality and innovation on the other. Some authors, notably Arrow have argued that this relation is positive because stronger

[22] See Van Gorp and Batura (2015).

competition is an incentive for suppliers to innovate and to diversify in quality dimensions as to "escape competition". Other authors, notably Schumpeter have stressed that companies need the financial means and incentives to innovate. Where monopolies exist, high profits attract challengers trying to enter the market with innovative ideas to 'steal' some of the profits of the incumbent. Aghion et al. (2002) have advanced the idea that both classical views hold some truth and that the relation between competition and quality and innovation is likely inverse U-shaped, concluding that some degree of market power may be desirable as it induces innovation. These theories are still subject to discussion among economists and its relevance differs from case to case.

Privacy protection

Privacy projection is a particular concern in data-driven markets

Where companies collect and process personal data there is a concern over the possibility that companies may not sufficiently protect the privacy of consumers (which may, for instance, give rise to unanticipated or unnoticed abuse of personal data (De Bijl, 2017; NIST, 2014[23]). In more concentrated markets, the theory of harm is that companies may have less incentive to protect the privacy of users.

There is no easy way to define privacy as there are different interpretations among cultures and individuals, the online version of the Merriam-Webster dictionary defines it as the state of being apart from company or observation and freedom from unauthorized intrusion. Magi (2011) mentions the following ways in which privacy protection prevents harm to individuals:

- Privacy prevents intrinsic loss of freedom of choice;
- Privacy helps prevent sorting of people into categories that can lead to lost opportunities and deeper inequalities;
- Privacy preserves the chance to make a fresh start; and
- Privacy protects from power imbalance between individuals and governments/organisations.

In Europe, privacy protection is considered a fundamental right. Article 8 of the EU Charter of Fundamental Rights ensures the protection of personal data while Art. 8 of the Convention of Human Right (concluded within the Council of Europe, which is broader than the EU) protects for one's "private and family life, his home and his correspondence".

The preferences regarding privacy protection differ between consumers. In that sense, privacy protection is a concern that companies may account for as a quality feature. In the discussion below we use this narrow perspective on privacy which views privacy as a "regular" consumer harm. This should be interpreted with great care as privacy protection is more than that. In particular, consumers may not be able to anticipate how important privacy of their personal data is in the future, or what the technological possibilities will be to analyse and monetise data in the future. Privacy agreements will then have characteristics of incomplete contracts, which undermines the feasibility for consumers of viewing privacy as a quality dimension that they can take into account in market interactions. This report recognises but abstracts from this aspect.

If a company considers privacy as a quality feature and not as a fundamental right, it may regard personal data (i.e. giving up privacy) as a currency to be paid by users.[24] It has been argued that in data-driven markets, a lack of competition does not necessarily result in higher consumer prices

[23] In particular, NIST (2014) points at the risk of non-contextual use of data that may expose an individual in unexpected ways.

[24] When we consider privacy as a currency, it should be noted that this is a much less transparent currency. Contrary to monetary currencies, there is no clear budget constraint: first, there is no rivalry in sharing personal data; and second, it is often not clear what kind of personal information one is sharing and with whom. In addition, since consumers may be unaware of the amount and scope of personal data that digital companies collect, they are often ignorant about the ways in which their privacy might be threatened.

(considering most of these services are provided for free), but rather in less privacy. Privacy protection can be considered a dimension of non-price competition (Ohlhausen and Okuliar, 2015).

Companies may compete by offering tighter or more transparent privacy policies (Evans, 2009; Savage and Waldman, 2015) or by introducing business models that are substantially less harmful for privacy. An example of the latter can be observed in the market for online search, where various search engines (e.g. DuckDuckGo) compete with Google Search) by offering a search service without collecting user data. In some concentrated markets, business models that offer better privacy protection may not be a feasible alternative for consumers because of other quality dimensions. For search, it may be the case that requesting more privacy decreases the quality of the services while giving up privacy to some extent may allow the search engine to be more personalised and hence be of better quality.

As with other forms of non-price competition (e.g. quality and innovation), the relation between competition and privacy protection is not necessarily linear, but may also be inverse U-shaped (this is speculative, we are not aware of any research that provides evidence). This would imply that companies in monopolistic market structures would have less incentive to protect privacy[25], but also that companies in highly competitive markets would have less incentive to protect privacy. More competition may require companies to compete more in terms of user experience. For this, companies may need more personal data. In theory, there could be a race to the bottom in terms of privacy protection in highly competitive markets.

4.2.3 *Unfair business or trading practices*

Unfair trading practices are not necessarily a result of market power but their incidence is likely to be higher in concentrated markets

In the literature on data and competition, limited attention is paid to "unfair business or trading practices" in business-to-consumer and business-to-business relationships. Examples of such practices are unjustified or disproportionate terms and conditions and a lack of redress possibilities. Another issue related to unfair business or trading practices is the risk of unanticipated and unnoticed abuse of personal data, which is not only a privacy concern, but also a consumer protection issue. These practices do not belong to the domain of competition law. The rules concerning unfair trading practices are applicable to all companies in their relationship with consumers, not just to companies with a dominant position.[26] However, these practices can be a side-effect of limited competition. In a market where companies do not face competitive pressure the financial return of engaging in unfair practices may be higher than in competitive markets as there are no alternatives available for consumers.

4.3 Price discrimination

Personalised pricing (first-degree price discrimination) is scarcely used. The welfare effects are ambiguous

Data can be a vehicle for price discrimination. By collecting data about their users, a company receives better information about their purchasing habits and is in a better position to assess their willingness to pay for a given good or service. If it has market power, the company would then be able to use that information to set different prices for the different customer groups it has identified.

[25] As this goes at the expense of profits, while users have no alternative options anyway.

[26] Note that there are Member States that have competition rules on unilateral conduct which, for example, prohibit or impose sanctions on abusive behaviour towards economically dependent undertakings and/or abuse of superior bargaining position. (see EC (2013), "On unfair trading practices in the business-to-business food and non-food supply chain in Europe").

There are at least two kinds of differential pricing. The first, "risk-based pricing," occurs when a business prices a product based on the cost of providing it to different groups of buyers. Risk-based pricing is common in the insurance and credit markets, and has been used for many years. The second, "value-based" pricing, occurs when a business prices a product based on buyers' willingness to pay. By doing so, the company can skim more of the consumer-surplus for the benefit of its own profits. (In the case off first-degree price discrimination or personalised pricing the full consumer surplus can be skimmed away, with second or third-degree price discrimination companies are only able to capture a part of consumer-surplus). We note however, that discriminatory pricing is not always negative for users. According to economic theory (first-degree) price discrimination can either increase or decrease consumer welfare, depending on demand conditions. Generally, applying discriminatory prices allows for more users to benefit from a service than applying the same price for all users (OECD, 2016b).

There are many examples where companies are able to charge different prices to different types of consumers but personalised pricing seems rare (Kerber, 2016). As indicated in CERRE (2017, p. 40), personalised prices have not been observed in practice according to the European Commission[27] as well as to the reports of OFT (2013) for the UK and to the CNIL-DGCCRF for France (2014). For example, the CNIL-DGCCRF report (2014) found no evidence of personalised prices based on IP addresses in France in e-commerce websites. On the same topic, Vissers et al. (2014) ran a three-week experiment with 66 user profiles connecting 25 airlines twice a day, and found no evidence of price targeting, though prices were observed to be very volatile.[28]

Although companies do not seem to apply pure personalised pricing often they can apply other pricing strategies to achieve similar outcomes (e.g. discounts targeted at specific groups or prices based on search queries). In sum, price discrimination is not necessarily harmful to consumers but it can be so, at least for individual consumers. Based on the available evidence, it does not seem necessary that personalized pricing should be a major concern for regulators and policymakers. However, the "big data revolution" has increased the technological possibilities to apply personalised pricing as online retailers know much more of their consumers than traditional brick and mortar shops which means that regulators should be vigilant and force companies to be transparent about their pricing policies towards consumers (CERRE, 2017).

4.4 Collusion

Data and algorithms may facilitate collusion

Some authors have expressed the concern that the use of big data and algorithms can result in the coordination of the behaviour of companies in an industry. Cartel agreements are prohibited in EU law, for example competing companies are not allowed to coordinate prices. In many cases the evidence for the existence of cartel agreements is discovered through cartel members submitting information to a cartel authority or by evidence gathered by competition authorities in "dawn raids".

Such evidence is not available if there is no involvement of humans in the coordination of behaviour between companies. When smart algorithms are the mechanism that set prices, it may be possible that the algorithm "acts" in such a way that prices collude without the knowledge of competing companies. With the application of artificial intelligence it is often even not possible to retrieve on what basis the algorithm decided to make particular decisions.

[27] Commission Guidance SWD(2016) 163, p. 147.

[28] Other types of practices may explain the high variability of online prices. In particular, it may be the case that firms use the possibility to change their prices online frequently to explore the demand curve (and estimate price elasticities).

According to the Bundeskartellamt (2016), the developments in the collection and processing of big data may be used by companies in ways that could limit competition. More market transparency can enhance the stability of collusion.[29] Stucke and Ezrachi (2015) mention four ways in which big data can be used to facilitate collusion:

1. Companies may use real-time data to monitor compliance with an explicit agreement. Note that the differences with traditional cartels are limited as monitoring of the behaviour of competitors is necessary in all cartels. As Bundeskartellamt (2016) explains, the greater information resulting from data collection may improve the stability of cartels. In a more transparent market, it is easier to detect deviations from the cartel agreement. This can limit the incentives to deviate from (implicit or explicit) collusion;
2. Companies may share identical pricing algorithms that simultaneously adjust prices. Such a cartel would be very similar to a classic hub-and-spoke cartel with a single company that acts as the leader of the cartel;
3. Companies may use big data to facilitate collusion, for example by programming immediate and automatic reactions to price changes. Again, this is not a totally new form of cartel formation. Without big data similar tactics are possible but big data may facilitate better identification of price changes and faster reactions;
4. The fourth way in which big data may facilitate collusion would form a new development compared to traditional cartels. Companies that use artificial intelligence to maximise profits develop algorithms that, through machine learning, may result in collusion without the explicit intent of the programmer to reach such an outcome. In Ezrachi and Stucke's book 'Virtual Competition' they refer to this scenario colourfully as the 'God view' scenario.

Actual evidence of "data-driven collusion" is scarce

OECD (2016) notes that there is little discussion in the literature about the implications of big data for the detection and investigation of cartels. Possibly, this can be explained by the very few cases that have been investigated to date. In the literature we have not encountered any real-life examples of scenario 3 (labelled "Tacit collusion on steroids: the predictable agent" by Ezrachi and Stucke (2016) and scenario 4 ("God view").

Mehra (2016) mentions an example of algorithms that result, without human intervention, in higher prices. This example can be seen as a (very) rudimentary form of the algorithm of scenario 3. The example concerns a twenty-year-old book on fruit flies which was listed in 2011 on Amazon for USD 23.7 million. This book was sold by two sellers. Both had set their prices as a function of the other stores' prices. This resulted in an upward price hike. In the box below we provide examples of scenario 1 and scenario 2.

Box 4.1 – "data-driven collusion", two real-life example

Scenario 1:

A number of papers refer to a competition case in the United States as an example of a cartel where real-time data was used to monitor compliance with a cartel agreement.

In 2015 the US Department of Justice (DoJ) started the prosecution of sellers fixing prices for posters in the Amazon marketplace.[30] An executive charged by the DoJ had developed a pricing algorithm reactive to consumer preferences. This algorithm was shared with other sellers and implemented simultaneously by them to make price coordination possible. The former executive of an e-commerce seller of posters, prints and framed art has agreed to plead guilty for conspiring to fix the prices of posters sold online.

[29] Of course increased transparency can also make the detection of collusive agreements easier. Competition authorities can use big data to monitor markets.

[30] https://www.justice.gov/atr/division-operations/division-update-2016/innovative-prosecutions-21st-century-schemes.

Importantly, coordination between natural persons was necessary for the collusion ("*During those conversations and communications, TOPKINS and his co-conspirators agreed to fix, increase, maintain and stabilize prices of agreed-upon posters*").

Scenario 2:
As with the example in scenario 1 there is no clear example of a scenario 2-cartel at the moment.

Ezrachi and Stucke (2016) mention the taxi-app Uber as an example of a potential scenario 2-cartel if its business model evolves. Uber drivers do not compete on price (but note that Uber fixes a maximum price at the moment, not a minimum price). Some drivers might be willing to offer a discount, but Uber's algorithm determines the fare and surcharges. According to Ezrachi and Stucke this by itself is legal. But as the platform's market power increases, this cluster of similar vertical agreements may in their view evolve into a classic hub-and-spoke cartel. In this case Uber as the algorithm developer would be a hub and lead a cartel of taxi drivers.

According to Graef (2016) the key issue for determining whether the Uber platform gives rise to price fixing under EU competition law would be the existence of an anticompetitive object or of anticompetitive effects. To answer this question, competition authorities may be required to analyse the working of the algorithms to see if it indeed facilitates anticompetitive collusion. In this analysis it will be also relevant that there are some (pro-competitive) efficiencies in the mechanisms used by Uber.

"Data-driven collusion" should not be the main worry of competition authorities at the moment but may become so in the future

Coordination that results from the application of artificial intelligence in particular, may pose challenges for competition policy in the future. According to OECD (2016) there is no legal basis to attribute liability to a computer engineer for having programmed a machine that eventually "self-learned" to coordinate prices with other machines. OECD (2016) makes the claim that "self-learning algorithms might be one of the biggest challenges that competition law enforcers have ever faced, and whose solution may involve artificially making market conditions more unstable and less prone to implicit collusion".

As admitted by Stucke and Ezrachi (2016), these solutions are not yet studied in detail and further research is clearly needed. In our view it is important to remember that there are no examples yet of "artificial intelligent cartels". Theoretically it is of course possible that such cartels already exist but they are not yet detected. Chen, Mislove and Wilson (2016) have empirically analysed algorithmic pricing on Amazon Marketplace. They found that algorithmic sellers can be detected using a target price time series, and they identify over 500 such sellers in their data set. They did no observe any market distortions. Based on the limited evidence so far, we are not convinced that artificial intelligence really is the biggest challenge for competition law as OECD (2016) boldly claims. Such a claim requires more research and evidence of the existence of artificial intelligence cartels in the real world.[31]

An expert that participated in the workshop that we organised as part of this study argued that the only way that the risk of intended or unintended collusion can be avoided is by mandating companies to program their algorithms in a way that makes a collusive market outcome impossible. In a speech in March 2017 Commissioner Vestager referred to this as "antitrust compliance by

[31] This research would fit into the broader discussion of the future opportunities and challenges of machine learning, for example in the context of automated driving. If machine learning results in socially suboptimal outcomes it is unlikely that this is limited to competition policy. Based on this broader debate on the role of algorithms in society, it is imaginable that in the future the current scope of liability for collusion needs to be re-evaluated.

design". In her view, this is something that businesses must do to ensure that they are in compliance with competition law.[32]

Scenario 1 and 2 are certainly feasible based on current technology. As Inge Graef (2016) notes, no cases alleging algorithmic price fixing seem to have been reported in the EU yet. Nevertheless competition authorities should be prepared to look beyond traditional forms of collusion.[33]

4.5 Conclusion

In this chapter, we discussed theories of how the use of big data can harm competition. For most of the theories, it is not clear if there is actual harm as there is a lack of empirical evidence. Market power may result for example in less incentives for innovation or excessive prices , but such a theory is difficult to test, which is not a unique feature of data-driven markets. Other harm can be observed (data-driven cartels or personalised pricing) , however, as there are not many real-life examples it seems that such harm is not widespread, that does of course not necessarily mean that it is not there.

Some (but not all) of the theories fall within the scope of competition law. To date there have been few cases in which big data played a role and no cases that have found big data to be a basis for a theory of harm on antitrust grounds for mergers or conduct cases. This may imply that big data is not a concern from a competition perspective. However, the lack of cases in which data played a role could also be a signal that competition rules or its application are inadequate. This will be discussed in the next chapter.

[32] Speech by M. Vestager at Bundeskartellamt 18th Conference on Competition, Berlin, 16 March 2017 (LINK).

[33] Graef, I. (2016), Blog, "Algorithmic price fixing under EU competition law: how to crack robot cartels?" (LINK).

5 Role and effectiveness of competition policy

5.1 Introduction

Chapter 4 discussed several potential "theories of harm" of market power in data-driven markets.

Competition law provides competition authorities instruments to address harm by punishing the abuse of a dominant position, the participation in a cartel and by forbidding mergers that are considered harmful.

In this chapter we first discuss if competition policy is adequate to handle cases in data-driven markets as this seems the most logical way to address competition concerns. We also discuss the role of other instruments, such as privacy protection and consumer protection law. We conclude that (1) competition rules can be applied in data-driven markets but also (2) that often the application of privacy/data protection and consumer protection law is more efficient and effective than interventions based on competition law. As such, competition, consumer protection and data protection law have to go hand in hand in order to create a well-functioning market.

Because of its relatively strong enforcement mechanism, data protection advocates have started to look at competition law as a way to enhance the effectiveness of the application of data protection rules.[34] In this chapter we discuss arguments pro and contra such a role for competition law and conclude that the application of privacy/data protection and consumer protection law is generally more effective.

After the discussion of the application of existing rules and instruments we discuss the implementation of new ex-ante rules. Ex-ante measures such as mandatory data access or data portability are difficult to design but can have merits in specific markets. In all markets it is vital that markets remain contestable to competitors and new entrants. This chapter finishes with suggestions for what governments can do in keeping data-driven markets contestable.

5.2 Application existing rules – competition law

Only limited number of cases linked to big data have been considered by competition authorities in recent years

Big data has caught the attention of competition authorities due to two key developments. Firstly, a string of high-profile mergers and acquisitions in digital or internet markets raised the question of a possible competition impact of bringing together and gaining control over large data sets (OECD, 2016b). Secondly, there is a growing desire to better understand the possible (welfare) implications of big data for consumers and markets. A number of cases linked to big data have been considered by competition authorities in recent years. To date, there have been no cases that have found big data to be a basis for a theory of harm on antitrust grounds for mergers or conduct cases (Sokol and Comerford, 2016). Below we provide a short summary of three cases in order to give insight into how big data has been dealt with in practice in competition law enforcement thus far.

[34] 2014 EDPS Preliminary Opinion.

Box 5.1 Big data in competition cases, three examples

Microsoft/Yahoo! Search Engine (2010)[35]

Microsoft announced in 2010 plans to acquire Yahoo! Search Engine. Microsoft was active in the design, development and supply of computer software, while the Yahoo search business encompassed internet search and online search advertising.

The concern: that increased concentration would significantly impede effective competition in the relevant markets. The Commission reported that, at the time, Google had more than 90% market share of the online advertising market and that the activities of Microsoft and Yahoo in this market amounted to less than 10% market share.

The investigation: The Commission examined the potential impact of the merger on the different market players, namely internet search users, advertisers, online publishers and distributors of search technology.

Key conclusions: The Commission approved the merger, concluding the transaction was not expected to have any negative effects on competition. In a speech in 2016, Commissioner Vestager commented on this merger case, stating that "Far from undermining competition, that merger actually had the capacity to make the market more competitive, by increasing Microsoft's scale – and the amount of data it had – and improving its chance to compete with Google in that market."

Facebook/WhatsApp Merger Case[36]

Facebook announced in 2014 a plan to acquire web-based messaging platform WhatsApp.

The concern: Facebook could use WhatsApp as a potential source of user data, potentially leading to data concentration and the hampering of competition in the online advertising market.

The investigation: Both the FTC and the European Commission examined Facebook's acquisition of WhatsApp and cleared the transaction without conditions. The Commission's investigation focused on three areas: (i) consumer communications services, (ii) social networking services, and (iii) online advertising services.

Key conclusions: The European Commission concluded that Facebook´s increased access to data through acquiring WhatsApp was not to a point that it could hamper competition. This is because after the merger, there would continue to be a sufficient number of alternative providers to Facebook for the supply of targeted advertising, and a large amount of internet user data that are valuable for advertising purposes are not within Facebook's exclusive control. Furthermore, the Commission found that Facebook Messenger and WhatsApp are not close competitors and that consumers would continue to have a wide choice of alternative consumer communications apps after the transaction The Commission noted that network effects could sometimes pose a barrier to entry in communications markets, however this particular transaction was not likely to raise barriers to entry, because "consumers can and do use multiple apps at the same time and can easily switch from one to another," and because "there are currently a significant number of market participants that collect user data alongside Facebook"".

Microsoft/LinkedIn (2016)[37]

Microsoft announced its plan to acquire LinkedIn in 2016.

The concern: While the acquisition covered seven relevant product markets, the Commission was mainly concerned with the effect on professional social network services. In particular, the Commission has

35 Source: http://europa.eu/rapid/press-release_IP-10-167_en.htm.

36 Source: https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html.

37 Source: http://europa.eu/rapid/press-release_IP-16-4284_en.htm, and http://ec.europa.eu/competition/mergers/cases/decisions/m8124_1349_5.pdf.

expressed the concern that Microsoft would pre-install LinkedIn on all Windows PCs and integrated LinkedIn into Microsoft Office and combine, to the extent allowed by contract and data protection law, LinkedIn's and Microsoft's user databases.

The investigation: After conducting a market test, the Commission decided to approve the acquisition on compliance with a series of commitments.

Key conclusions: The Commission concluded that no competition concerns arise from the concentration of the parties' user data that can be used for advertising purposes. This is because a large amount of such user data will continue to be available on the market after the transaction. In addition, the transaction would not reduce the amount of data available to third parties as neither Microsoft nor LinkedIn currently makes available its data to third parties for advertising purposes. A similar conclusion was reached with regard to the availability of data for machine learning in the context of customer relationship management as well as productivity software.

Debate if competition law is equipped to deal with big data cases

To date there have been few cases in which big data played a role and no cases that have found big data to be a basis for a theory of harm on antitrust grounds for mergers or conduct cases. This may imply that big data is not a concern from a competition perspective. However, the lack of cases in which data played a role could also be a signal that competition rules or their applications are inadequate.

Below we provide a summary of the (academic) arguments for and against a role for competition law in addressing issues arising from the use of big data, drawing largely up on the comprehensive literature review on this topic presented in the aforementioned OECD background paper (OECD, 2016). Summarized in simplified form, there are three prevailing views on the role of competition law enforcement in issues related to big data:

1. Antitrust intervention would be premature and misguided, the harms resulting from big data should not be regulated under competition law and as such, competition law is not relevant to big data;
2. Competition law enforcement to prevent anti-competitive behaviour in the context of big data is urgent and might not even suffice to address possible harmful welfare effects;
3. Big data may in certain situations give rise to competition concerns, which can then be addressed on the basis of current competition law; the available competition tools may however need to be adapted to adequately analyse competition issues.

1) There is no role for competition law enforcement in issues around big data

Sokol and Comerford (2016), on the basis of a review of the existing but limited academic literature, form the conclusion that antitrust law is inappropriate for regulating big data and that the often cited problems of competition law in addressing challenges of big data are more noise than reality because such big data "challenges" do not exist. They argue that current consumer protection legislation is sufficient to take care of privacy related concerns stemming from the use of big data, adding that antitrust intervention in big data would be "premature and misguided". The authors emphasize the pro-competitive advantages bought by big data which yield innovations that benefit consumers in terms of higher quality and lower (often free) priced products.

This view is founded on a number of theoretical lines of thought and the absence of real-life cases. Theoretically, the argument is that the characteristics of big data; non-rival, non-exclusive, of short-lived value and easy and inexpensive to collect; preclude any one company from having a large enough concentration of big data to permit anti-competitive behaviour. Adding to this, is the insight that possession of the data does not in itself provide a competitive advantage, but that the strength of the underlying product is the true source of competitive advantage. The absence of real-life

cases where big data was found to be a source of competition problems is used to further strengthen this view. An explanation given for the absence of real life evidence of competition problems linked to big data is that the feedback loops do not have as strong of an effect as what is often claimed.

2) Competition law enforcement in issues around big data is urgent and might not suffice to address consumer harm

A number of academic experts hold the view that big data gives rise to considerable consumer harm, which requires immediate intervention, possibly going beyond the application of the competition rules (Newman, 2014). This view is founded on the arguments that a number of channels related to big data can give rise to market power, namely the existence of data-driven network effects and economies of scale resulting from feedback loops, which are argued to give incumbents a sustainable competitive advantage.

3) There is a role for competition law enforcement in issues around big data if competition concerns can be identified

A third and more nuanced view is that big data may potentially give rise to competition concerns which should then be addressed on the basis of existing competition law. While this may require adjustment of the available competition tools (for instance the SSNIP-test), the substance of the competition rules still holds in a big data context (CERRE, 2017; Graef, I., 2016; Van Gorp and Batura, 2015). Both the European Commission and national competition authorities are investing considerable effort in understanding how the use of big data may have implications for competition. The UK Competition & Markets Authority published a report in mid-2015 on the commercial use of consumer data, while the German Federal Cartel Office (*Bundeskartellamt*) and the French antitrust authority (*Autorité de la Concurrence*) produced a joint report on competition law and big data, published in May 2016. Most recently in November 2016, the Catalan competition authority released a report on the competition challenges posed by the data-driven economy. The views presented in each of these reports, as well as public comments by European Commissioner for Competition Margrethe Vestager, shows a growing consensus amongst these authorities that existing antitrust laws and enforcement powers are capable of addressing any competition concerns that may be identified in relation to big data.[38] In an article on the competition concerns related to digital platforms. ACM (2016) seems to agree with this view, stating that methodological challenges such as market definition will not impede ACM in pursuing a case.

Competition rules are flexible and can be applied in data-driven markets but tools may need to change

Until competition authorities run into "unaddressable" problems that require novel approaches there seems to be no basis for changing competition rules. As there have been only a limited number of cases in which big data played a role there is insufficient evidence to justify changes in competition rules.

Although we see no basis for a change of the rules there are possibilities to change the application of the rules which competition authorities can do because competition rules are flexible. An exception to this is the use of turnover thresholds to determine whether a merger has an EU dimension and thus has to be notified to the European Commission in the framework of the EU Merger Regulation.[39] Turnover thresholds may not be appropriate anymore due to the existence of other types of business models. An acquisition of a company having only a small turnover may still have a significant competitive impact, for example because the acquisition includes the transfer of

[38] Jones Day Commentary. "European Antitrust Enforcers Move on Holders of Big Data", 26 May 2016.

[39] Article 1(2) of Council Regulation (EC) No 139/2004 of 20 January 2004 on the control of concentrations between undertakings (EU Merger Regulation) [2004] OJ L 24/1 lays down the turnover thresholds that have to be met in order for a concentration to have a Community dimension.

valuable data, and thus may deserve to be scrutinized by the Commission.[40] To address this problem, the German Bundesrat adopted an amendment of competition law which includes an additional notification requirement in its national law based on the agreed transaction volume (i.e. the purchase price) of the merger ("value of transaction test").[41] Currently, the European Commission is in the process of reviewing the thresholds contained in the EU Merger Regulation.[42]

Areas where the application of rules can be changed are the following:

- Competition in data-driven markets may require other tools, tools such as the SSNIP-test (an instrument to define markets) but also profit margins and market shares may not be adequate to describe data-driven markets, see Van Gorp and Batura, 2015. The five factors introduced in chapter 3 can be a basis for the analysis of competitive dynamics in a market. Note that in recent decisions such as the merger of Microsoft/Linkedin the EC's analysis already shows many parallels with the framework to assess market power proposed in chapter 3;
- Competition cases are often complex which results in lengthy legal proceedings. Especially in dynamic markets this can be problematic. ACM (2016) offers some suggestions how procedures can be shortened, for example by continuously market monitoring by competition authorities. This is also an argument to invest in big data expertise;
- In some instances an analysis of competition in a potential market for data used for improving the services provided may be required next to the relevant end markets or services (Graef, 2015). Such an analysis is useful if companies can trade data. A merger might not have a significant effect on the relevant end markets but may result in a dominant position in the market for data;
- If data would be considered as a basis for a theory of harm, the criteria for access to an "essential facility" as set by the European Court of Justice are relevant. If an asset qualifies as an essential facility, the result is that its owner is under a duty to deal with rivals and has to share access to the asset at issue. The criteria to be an essential facility are therefore high. The following conditions need to be fulfilled to lead to the obligation to provide access to data (CERRE, 2017): (i) the data is indispensable for the downstream product, (ii) there would not be effective competition between the upstream and downstream market (iii) refusal prevents the emergence of the secondary product, (iv) there is no objective reason for the refusal.[43]

  An important question is whether these criteria are adequate in data-driven markets. An assumption underlying the doctrine is that access to the facility harms innovation. Indeed, different interests are at stake in essential facilities cases. The imposition of a duty to deal with the objective of promoting free competition affects the interests of the dominant firm consisting of the generally recognised principles of freedom to contract, including the right to choose one's trading partners, and freedom to dispose of one's property. The decision of a competition authority or court to interfere with the interests of a dominant undertaking for the purpose of protecting effective competition therefore requires a careful balancing exercise.

---

40 For example, the *Facebook/WhatsApp* merger did not have an EU dimension because the turnover of WhatsApp did not meet the notification threshold. This while Facebook acquired WhatsApp for a purchase price of 19 billion dollar. The European Commission was only able to assess the merger under the EU Merger Regulation because the transaction was capable of being reviewed under the national competition laws of three Member States and Facebook had requested the Commission, on the basis of Article 4(5) of the EU Merger Regulation, to examine the merger (see Case No COMP/M.7217 – *Facebook/WhatsApp*, 3 October 2014, par. 4 and 9-12).

41 Source: http://klgates.com/new-merger-control-thresholds-in-germany-04-10-2017/ (28 April 2017). At the time of writing the amended law was about to be signed.

42 See the publication consultation on 'Evaluation of procedural and jurisdictional aspects of EU merger control' which ran from October 2016 until January 2017, available at http://ec.europa.eu/competition/consultations/2016_merger_control/index_en.html.

43 Relevant cases at the EU level include: Judgment in *Telefis Eireann and Independent Television Publications Ltd v. Commission of the European Communities (Magill)*, Joined cases C-241/91 and C-242/91, ECLI:EU:C:1995:98; Judgment in *Oscar Bronner GmbH & Co. KG v. Mediaprint Zeitungs*, C-7/97, ECLI:EU:C:1998:569; Judgment in *IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. KG*, C-418/01, ECLI:EU:C:2004:257; Judgment in *Microsoft*, T 167/08, ECLI:EU:T:2012:323.

As the European Court of Justice made clear in the *Bronner case*, a duty to deal will increase competition in the short term but may put incentives for competitors to develop competing facilities in the long term at risk. In addition, the incentives of dominant undertakings to invest in new facilities may be reduced if competitors are given access too easily. In the long run, it therefore seems procompetitive to allow a dominant undertaking to keep facilities developed for its own business to itself.[44] However, if data is just a by-product and can be processed against low cost because it is machine generated/processed, the negative effects of data access or sharing may be lower in data-driven markets than with regard to access to (non-data) assets in other markets. With data sharing there would still be sufficient incentives to invest. This may provide an argument to apply criteria in data-driven markets that are less stringent than the criteria for an essential facility as established by the European Court. More research is needed into the implications of measures such as data sharing and portability on the (dynamic) incentives to innovate.

In practice there may be other ways to address competition concerns than the application of competition rules, for example by applying consumer protection and privacy/data protection laws. This means that competition authorities should evaluate which instrument is best equipped to tackle the specific concern.

5.3 Application existing rules – beyond competition law

Competition law generally not the preferred option to address privacy and data protection issues

Because of its relatively strong enforcement mechanism, data protection advocates have started to look at competition law as a way to enhance the effectiveness of the application of data protection rules. The instruments (privacy and data protection rules) that are designed specifically for those issues are generally better equipped. Competition authorities should not use their competences to impose a preference for strong data protection on consumers. Instead, it is the role of competition authorities to keep markets competitive in order to ensure that consumers can choose themselves between a variety of products and services with different characteristics. A prerequisite for the existence of a well-functioning market is that consumer and data protection law are effectively applied to enable individuals to exercise a genuine and well-informed choice. While competition law thus aims to ensure the availability of choice, consumer and data protection law should empower individuals to effectively exercise such a choice. As such, competition, consumer protection and data protection law have to go hand in hand in order to create a well-functioning market.

The view expressed above is also shared by the European Commission. The debate around big data and competition has taken place amidst the development and adoption[45] of the new General Data Protection Regulation (GDPR). The GDPR sets new standards for the protection of personal data in the EU, including by enhancing individuals' control over their data.[46] In a speech in January of 2016, European Commissioner for Competition Margrethe Vestager outlined some key features of the new general data protection regulation[47] and stated that with the new GDPR in place, "I don´t

[44] Opinion of Advocate General Jacobs in *Oscar Bronner GmbH & Co. KG v. Mediaprint Zeitungs*, C-7/97, ECLI:EU:C:1998:264, par. 57.

[45] The legislative process started in January 2012 with the EU Commission's GDPR proposal.

[46] Jones Day Commentary. "European Antitrust Enforcers Move on Holders of Big Data", 26 May, 2016.

[47] Namely, the new rules; (1) require data protection to be built into products and services from the start. (2) promote ways for businesses to develop innovative services using data, without compromising privacy. (3) clarify the right to be forgotten, and allow users to transfer their data between providers.

think we need to look to competition enforcement to fix privacy problems".[48] In its merger decisions *Google/DoubleClick* and *Facebook/WhatsApp*, the European Commission has consistently held that its competition analysis is without prejudice to the obligations imposed on the parties by data protection legislation and that privacy-related concerns do not fall within the scope of EU competition law. In a similar vein, the Court of Justice already noted in the *Asnef-Equifax* judgment that "*any possible issues relating to the sensitivity of personal data are not, as such, a matter for competition law, they may be resolved on the basis of the relevant provisions governing data protection*".[49]

However, scholars have argued that the entry into force of the Lisbon Treaty in 2009 has changed the institutional setting in the European Union. In particular, the Charter of Fundamental Rights, which includes the right to data protection in Article 8, gained legally binding status as a source of primary EU law. As such, data protection has been elevated to a fundamental right in the EU legal order. Interestingly, Article 51(1) of the Charter makes clear that the EU institutions are under a duty to respect and promote the application of the rights contained therein. The EU therefore does not only have a negative duty to avoid violations (i.e. respect the fundamental rights) but also a positive obligation to take action to uphold (i.e. promote the application of) the fundamental rights. Graef (2016) argues that this may imply that the European Commission is also bound by the fundamental rights of the Charter, including the right to data protection, when acting in the field of EU competition law.

Competition law, consumer protection and data and privacy protection can be used in parallel to address competition concerns

Some of the "theories of harm" 4 (unfair trading practices and some forms of abuse of a dominant position) discussed in chapter and the concerns expressed in the literature fall (at least to some extent) under the scope of consumer protection (for example rules on regarding unfair trading practices see box 5.1) and data/privacy protection rules. Those rules can be applied in data-driven markets, just as in any other market. An advantage of applying these instruments is that it is not necessary to show that a company has a dominant position but they can help in keeping markets contestable and transparent. This means that some problems may be dealt with in more than one way. In practice this may require co-operation between consumer protection, data and privacy protection and competition policy as proposed for example in the EDPS Opinion 8/2016 on coherent enforcement of fundamental rights in the age of big data, Stucke and Grunes (2016) and CERRE (2017).

In this regard, it is vital that the information requirements in consumer protection law and the conditions for valid consent in data protection law are strictly applied and enforced. It is instructive to note that the Bundeskartellamt (the German competition authority) launched an investigation into Facebook's terms of service in March 2016 to examine whether consumers are sufficiently informed about the type and extent of personal data collected. The Bundeskartellamt suspects that Facebook's terms of service are in violation of data protection law and could thereby also constitute abuse of dominance under competition law by representing an abusive imposition of unfair conditions on users.[50] Note that the meaning of this specific case is not yet clear as there very limited information is publicly available and the case is not yet decided.

Box 5.2 Directive on Unfair trading practices ("Wet oneerlijke handelspraktijken")

[48] European Commission. (2016). "Speech by Margrethe Vestager: Competition in a Big Data World". DLD 16, Munich, 17 January 2016.

[49] Judgment in *Asnef-Equifax v. Asociación de Usuarios de Servicios Bancarios*, C-238/05, ECLI:EU:C:2006:734, par. 63.

[50] Press Release Bundeskartellamt, "Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules", 2 March 2016, available at http://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html?nn=3591568.

Unfair trading practices rules concern business-to-consumer transactions. In some Member States (but not the Netherlands) there are also specific rules concerning business-to-business relationships.

The European Commission has issued a document that provides examples for how the Directive can be applied. For all services, the main message of the guidance is the same: for a practice to be fair under EU consumer protection law , it needs to be transparent.[51]

ACM has for example fined online bookstores for unclear terms and conditions and online ticket brokers for providing misleading information regarding the availability of tickets based on the unfair trading practices rules.[52]

5.4 Other (ex-ante) interventions

Ex-ante measures require a thorough analysis and should be applied carefully
In specific markets there may be a case for ex-ante regulation to avoid exploitative behaviour, and a possible lack of innovation in the market. An advantage of ex-ante regulation is that there is no need to prove dominance and a "theory of harm" in specific cases. However, there must be sufficient grounds to suspect that data contributes to market power in the specific market concerned, as regulations can be distortionary.

Ex-ante measures concerning data can be grouped into two categories:

- Data sharing (refers to sharing between competitors)[53];
- Data portability (refers to ability of a user to transfer its data).

Both measures make data less exclusive (factor 1 of the framework introduced in chapter 3). Data portability can also reduce the switching costs of using a different platform. In this way data portability diminishes the impact of network effects on market power (factor 3) and it may allow competitors to compete based on a different business model (factor 5). Data sharing (between competitors) is a more effective measure if learning effects are of particular importance (factor 1).

Ex-ante regulation in the form of data sharing and portability[54] faces substantial design challenges. It is not clear for example how a company with a search engine product could share data in a way that is useful to competitors and also complies with all the rules regarding privacy and data protection. Moreover there is a considerable risk that it results in diminished incentives to innovate for. If ex-ante regulation is used to counter market failures there is a risk that due to "government failure" consumers are worse off.

---

51 EC Staff document (2016). Guidance on the implementation/application of directive 2005/29/ec on unfair commercial practices, SWD(2016) 163.

52 See for example the decision concerning Otravo (formerly WTC) of December 2014 (ACM case number 14.0949.32).

53 To be compatible with competition law, the overall effect on consumers in the relevant markets must be favourable. It is not necessary, in principle, for each consumer individually to derive a benefit from an agreement, a decision or a concerted practice. Judgement in *Asnef-Equifax v. Asociación de Usuarios de Servicios Bancarios*, C-238/05, ECLI:EU:C:2006:734, par. 63.

54 It should be kept in mind that restrictions on the portability of data as imposed by a dominant undertaking may also qualify as abusive under competition law. As such, competition enforcement can also be used to enforce data portability. In this regard, the former Competition Commissioner argued in a 2012 speech that data portability '*goes to the heart of competition policy*' and that '*portability of data is important for those markets where effective competition requires that customers can switch by taking their own data with them*' (J. Almunia, Speech: Competition and personal data protection, Privacy Platform event: Competition and Privacy in Markets of Data Brussels, 26 November 2012, available at http://europa.eu/rapid/press-release_SPEECH-12-860_en.htm). It is instructive to note that one of the concerns previously expressed by the Commission in the *Google* investigation relates to restrictions that Google allegedly imposes on the portability of advertising campaigns in AdWords (see J. Almunia, Speech: The Google antitrust case: what is at stake?, 1 October 2013, available at http://europa.eu/rapid/press-release_SPEECH-13-768_en.htm). In addition, in the *Facebook/WhatsApp* merger decision the Commission already assessed whether data portability issues constituted a barrier to consumers' switching in the context of consumer communications apps (Case No COMP/M.7217 – *Facebook/WhatsApp*, 3 October 2014, para. 113-115 and 134).

Nevertheless, several policy initiatives have started to consider the inclusion of portability duties in legal fields other than data protection. Rules regarding data access and data portability are already in place in sectors where companies have a dominant position as for example in the energy sector. In this study we have included a case study of smart meter electricity data in the Netherlands (see Appendix A). Smart meters are owned by network operators, energy suppliers and other market participants can obtain access to meter data with the consent of consumers. Likewise the Payment Services Directive will allow third parties access to bank account data. In the car market manufacturers must ensure that independent operators have easy, restriction-free, and standardised access to information on the repair and maintenance of vehicles.[55]

In the field of (personal) data protection there are already rules concerning data portability. The new General Data Protection Regulation allows users to transfer their own personal data among providers. It will be interesting to see how data portability based on the Regulation develops and what form it will take. The recently published draft Guidelines on the right to data portability of the Article 29 Working Party[56] indicate that the exact scope and interpretation is still work in progress.

The Commission Communication on "Building a European Data Economy" published in January 2017 refers to the possibility of taking measures to ensure portability of non-personal data.[57] In addition, the draft Directive on the Supply of Digital Content speaks of data retrieval obligations in the context of consumer law. Article 13(2)(c) of the latter proposal requires a supplier to provide a consumer who terminates a contract for the supply of digital content '*with technical means to retrieve all content provided by the consumer and any other data produced or generated through the consumer's use of the digital content to the extent that data has been retained by the supplier*'. The provision goes on to state that the consumer is '*entitled to retrieve the content free of charge, without significant inconvenience, in reasonable time and in a commonly used data format*'.[58] It is important to note that, unlike the right to data portability of the General Data Protection Regulation, the proposal for a Digital Content Directive would not entitle consumers to have their digital content directly transmitted to a new provider. It only entitles consumers to retrieve data in a commonly used format. On the other hand, the proposal for a Digital Content Directive has a broad scope as it enables a consumer to retrieve any other data, to the extent that it has been retained by the supplier, generated through the use of the digital content which is not as such provided by the consumer. The right to data portability of the General Data Protection Regulation only covers personal data provided by the data subject.

Aim of public policy should be to keep markets contestable

In the economic literature, there is a long debate on the dynamics between competition and innovation. On the one hand market power can provide an incentive to innovate as if offers the perspective for monopoly rents (Schumpeter). On the other hand the need to outperform competitors in a competitive market can provide an incentive for innovation (Arrow). The more recent literature takes a nuanced view and poses that both too little and too much competition can be negative for innovation.

---

55 Commission Regulation (EU) No 566/2011 on access to vehicle repair and maintenance information.

56 Article 29 Working Party, draft Guidelines on the right to data portability, 13 December 2016, available at http://ec.europa.eu/information_society/newsroom/image/document/2016-51/wp242_en_40852.pdf.

57 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'Building a European Data Economy', 10 January 2017, COM(2017) 9 final, p. 15-17.

58 Article 16(4)(b) of the proposal for a Directive of the European Parliament and of the Council on certain aspects concerning contracts for the supply of digital content (proposal for a Digital Content Directive), 9 December 2015, COM(2015) 634 final, provides for a similar obligation for suppliers with regard to long term contracts for the supply of digital content.

Shapiro (2012) has pointed out that both the Schumpeter and Arrow view result in a similar policy recommendation in that competition policy should protect the process of innovation by keeping the market contestable (multiple firms are vying to win profitable future sales) (see also De Steel, A. and Larouche, P., 2015).

Policies that are intended to protect consumers and privacy may have the side-effect that they make entry to a market more difficult. Such entry barriers can have a negative effect on dynamic competition, incumbents are not challenged as much by new disruptive business models as they would with be without those barriers. A new entrant can for example not buy personal data on a data-brokering market as this is not allowed without user consent. Compliance with consumer and privacy protection legislation requires investments and can make it difficult for small companies to operate on a market and to compete effectively with companies with a larger scale. It is desirable that in the design of consumer and privacy protection rules the effects on competition and specifically entry barriers are considered to maximize dynamic efficiency.

Policies to assist start-ups can create the challengers of tomorrow and are in that sense also an instrument to keep markets competitive. "Expertgroep big data en privacy" (2016) discusses the idea of a *regulatory sandbox* where companies are allowed to experiment. In the sandbox companies interact with regulators about the application of rules regarding for example privacy protection for a specific product or services and formalize this in an agreement. Companies obtain (ex-ante) assurance from the regulator that it will not start an infringement procedure if it complies with the agreement. Such a sandbox might be especially beneficial for companies that try to compete with dominant companies with innovative business models.

5.5 Conclusion

Until competition authorities run into "unaddressable" problems that require novel approaches there seems to be no basis for changing competition rules. As there have been only a limited number of cases in which big data played a role, there is insufficient evidence to justify changes in competition rules.

Note that not all potential harm (e.g. lack of innovation) can be addressed by competition authorities. In specific markets there may be a case for ex-ante regulation. An additional advantage of ex-ante regulation is that there is no need to prove dominance and a "theory of harm" in specific competition cases. However, ex-ante regulation in the form of data sharing and portability faces substantial design challenges. If ex-ante regulation is used to counter market failures there is a risk that due to "government failure" consumers are worse off. But if data is just a by-product and can be processed against low cost because it is machine generated/processed, the negative effects of data access or sharing may be lower in data-driven markets than the negative effects of accessing (non-data) assets in other markets. This can provide an argument for mandatory data sharing or portability in specific markets.

Some, but not all, of the harm that results from a dominant position can be addressed by other rules outside of competition law; notably consumer, privacy, and data protection rules. This means that some problems may be dealt with in more than one way. However, competition rules are not designed for privacy protection and in general, privacy protection rules are better equipped to protect consumers than competition rules. The coordination of consumer protection, data protection and competition policy requires coordination between authorities.[59]

[59] In the Netherlands, consumer protection and competition regulation are already the responsibility of a single agency (Autoriteit Consument en Markt). This should facilitate cooperation. There is also an agreement on cooperation between

A *Regulatory sandbox* might be beneficial for companies that try to compete with dominant companies with innovative business models. Such a sandbox in which small companies can experiment could be an instrument to support the challengers of tomorrow and to keep markets contestable.

ACM and the Dutch DPA (Autoriteit Persoonsgegevens) (Samenwerkingsprotocol tussen Autoriteit Consument en Markt en Autoriteit Persoonsgegevens (Staatscourant Nr. 58078).

References

- ACM (2016), 'Grote platforms, grote problemen?
- Autoritat Catalana de la Competència (2016), "The Data-Driven Economy. Challenges for Competition";
- Autorité de la concurrence and Bundeskartellamt, "Competition law and data", 2016;
- Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. "Competition And Innovation: An Inverted-U Relationship," Quarterly Journal of Economics, 2005, v120(2,May), 701-728;
- Belleflamme, P. and Peitz, M. (2015), *Industrial Organization: Markets and Strategies*, 2nd ed, Cambridge University Press;
- Bundesskartellamt (2016), "Market power of platforms and networks";
- Christiaans, T. (1998). "A Note on Public Goods: Non-Excludability Implies Joint Consumability". No. 68-98. *Volkswirtschaftliche Diskussionsbeiträge*;
- De Bijl, P.W.J. (2017), "Digital technology and public policy", speech at ACM Academy Lecture, February 21, ACM, The Hague;
- Evans, D., 2009, "The Online Advertising Industry: Economics, Evolution, and Privacy." *Journal of Economic Perspectives*, 23(3): 37-60;
- Ezrachi, A. and Stucke, M. (2016), *Virtual competition: the promise and perils of the algorithm - driven economy*, Harvard University Press;
- Expertgroep Big data en privacy (2016), 'Licht op de schaduw, verantwoord innoveren met big data' (rapport van de expertgroep big data en privacy aan de minister van Economische Zaken);
- Federal Trade Commission (2016), "Big data: A Tool for Inclusion or Exclusion? Understanding the issues";
- Gal, M. and Rubinfeld, D. (2016), "The Hidden Costs of Free Goods: Implications for Antitrust Enforcement", Antitrust Law Journal 521;
- Gawer, A. (2016) "Competition Policy and Regulatory Reforms for Big Data: Propositions to Harness the Power of Big Data while Curbing Platforms' Abuse of Dominance - Note by Annabelle Gawer" DAF/COMP/WD (2016) 74;
- Graef, I. (2015), "Market Definition and Market Power in Data: The Case of Online Platforms", World Competition 473;
- Graef, I. (2016), *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility*, Kluwer Law International;
- Kerber, W. (2016), 'Digital Markets, Data, and Privacy: Competition Law, Consumer Law, and Data Protection', MAGKS, Joint Discussion Paper Series in Economics, No. 14-2016. Available at SSRN: https://ssrn.com/abstract=2770479 or http://dx.doi.org/10.2139/ssrn.2770479;
- Lambrecht, A. and Tucker, C. (2015), "Can Big Data Protect a Company from Competition?", available on SSRN;
- Lerner, A. (2014), "The Role of 'Big Data' in Online Platform Competition", Available at SSRN: https://ssrn.com/abstract=2482780 or http://dx.doi.org/10.2139/ssrn.2482780;
- NIST (2014), "Privacy Engineering Objectives and Risk Model - Discussion Deck", CSRC, National Institute of Standards and Technology, Gaithersburg, MD, retrieved at http://csrc.nist.gov/projects/privacy_engineering/documents.html;
- Bourreau, M., De Streel, A. and Graef, I. (2017), 'Competition Policy and Big Data: Market Power, personalised pricing and advertising', CERRE Project Report, available at http://cerre.eu/publications/big-data-and-competition-policy;
- Mehra, K. S. (2016), 'Antitrust and the robo-seller: competition in the time of algorithms', Minnesota law review;
- Newman, "Search, Antitrust, and the Economics of the Control of User Data", *Yale Journal on Regulation* 2014, vol. 31, no. 2, (401);

- OECD (2016a), "Price Discrimination", 2016;
- OECD (2016b), "Big Data: Bringing Competition Policy to the Digital Era. Background note by the Secretariat", Directorate for Financial and enterprise affairs competition committee, DAF/COMP(2016)14, 29-30 November 2016;
- Ohlhausen, M. and Okuliar, A. (2015), "Competition, Consumer Protection, and the Right (Approach) to Privacy", Antitrust Law Journal, Forthcoming, Available at SSRN: https://ssrn.com/abstract=2561563;
- Prüfer, J. and Schottmüller (2017), "Competing with Big Data", TILEC Discussion Paper No. 2017-006, available at SSRN: https://ssrn.com/abstract=2918726;
- Rieke, A., Yu, H., Robinson, D., Von Hoboken, J. (2016) 'Data brokers in an open society', Upturn report prepared for the Open Society Foundations;
- Rubinfeld, D. and Gal, M. (2017) "Access barriers to big data" to be published in the *Arizona Law Review* forthcoming;
- Savage, Scott J. and Waldman, Donald, (2015), Privacy tradeoffs in smartphone applications, Economics Letters, 137, issue C, p. 171-175;
- Sokol, S. and Comerford, R. (2016), "Does Antitrust Have a Role to Play in Regulating Big Data?", available on SSRN in 2016, also in *Cambridge handbook of antitrust, intellectual property and high tech*, Roger Blair and Daniel Sokol, editors, Cambridge University Press, forthcoming, 2017;
- Stucke, M. and Grunes, A. (2016), *Big Data and Competition Policy*, Oxford University Press;
- Varian, H. (2016) "Big Data, Personalization, and Competition", presentation at the OECD competition policy roundtable on Big Data, November 2016;
- Varian, H. (2016) "Use and Abuse of Network Effects", presentation;
- Vezzoso, S. (2016), "Competition Policy in a World of Big Data", in *Research handbook on digital transformations*, F. Xavier Olleros and Majlinda Zhegu, editors, Edward Elgar;
- Wetenschappelijk raad voor het regeringsbeleid (WRR, Dutch scientific council for government policy) (2016), "Big Data in een vrije en veilige samenleving".

Appendix A: Case studies

Introduction

We have developed the framework presented in chapter 0 partly by doing a number of case studies:

1. E-commerce – Amazon;
2. Social network – Facebook;
3. Super platform or platform of platforms – Google.

In these case studies we analyse the role of data in the business model of the respective companies. We also analyse if and how data contributes to a competitive advantage and market power. In each case we discuss the potential impact of market power. Our goal in these case studies is not to assess if a particular company has a dominant position or to establish if a theory harm is applicable. Rather we try to distil the characteristics of the company and the use of data that are relevant in the analysis of the relationship between big data and competition.

The three cases follow a similar structure. We have used the platform characteristics mentioned in TNO (2015) to describe the relevant aspects of the business model. After the analysis of the business model we discuss the role of data and how data can result in a competitive advantage and market power. For each case we go deeper into some theories of harm that may be especially relevant for that particular business activity.

Besides the three cases that concern a particular business model we have studied a fourth market that sheds light on the relationship between data and competition from a different perspective:

4. Smart electricity meters.

In contrast to the data used in the business models in the first three case studies data on electricity consumption is non-exclusive, with the consent of users any (certified) company can obtain access to the data. Our rationale for including this case is that it may offer lessons for industries where data is exclusive and data sharing and portability non-existent or still in its infancy.

Case study 1: E-commerce - Amazon

Introduction

This case study analyses one of the largest e-commerce platforms in the world: Amazon. Amazon is involved in a wide variety of activities (cloud services, home automation, handsets, etc.)[60]. This case study focusses on Amazon's core business (operating an e-commerce platforms as well as retailing of consumer goods[61] [62]) and analyses the role of data as a contributor to Amazon's market power vis-à-vis other retailers and vis-à-vis other e-commerce platforms.

This case study explores the causalities between data and market power; i.e. we answer the question (how) does a data advantage give rise to market power? In the case of Amazon, there may be data-driven network effects which contribute to Amazon's competitive advantage over others, but also Amazon's (data-driven) operational efficiency throughout the distributional value chain contributes to its competitive position.

Furthermore, we analyse certain specific concerns of anti-competitive (or unfair) behaviour within the context of e-commerce services. More specifically, we address the concern that operators of an online market place who act as a retailer on their own platform (like Amazon) may engage in exclusionary conduct. Exclusionary conduct may be in the form of excluding other retailers from the online market place, or in the form of excluding other retailers from accessing certain data with the intent of raising rivals' costs / lowering rivals' quality. Asymmetric access to data may also give an advantage in pricing strategies; e.g. asymmetric information about prices would allow Amazon to respond to price changes more swiftly than others, notably when dynamic pricing algorithms are used[63].

Company Description[64]

General description of activities

When Amazon started in the 1990s, its ambition was to become the largest global e-commerce website, earth's most customer centric company, and cost-leader among retailers. Today, Amazon operates a website that functions like a global online market place, allowing retailers from all over the world to set up shop. The company is also the largest retailer on its own website. Amazon also engages in retailing and distribution of digital content via its own e-reader (Kindle) and its online video services *Prime Video*.

Moreover, Amazon is integrating more and more of the logistical activities of physical products in the value chain. Within the near future, the company aims to control the entire end-to-end distribution network, including warehousing (storage and order picking) and distribution, both long-

[60] In the Dutch context, it would make more sense to take Bol.com as a topic of the case study. However, we choose Amazon for practical reasons because this company is much better documented. Bol.com is a company that resembles Amazon in many ways, and many of the conclusions drawn for Amazon may also apply to Bol.com; noting that Amazon's market share is considerably higher in some other countries (such as the US) than Bol.com's market share in the Netherlands.

[61] There is a distinction between retail of physical goods and retail of digital content because of the differences in the underlying distribution network used, however, market boundaries are blurring (as we will see in the case). The retail of physical goods is (for the time being) considered the core business of Amazon, judged by the fact that the majority of Amazon's assets are related to its logistical system.

[62] Despite the focus on retail of consumer goods, this case study will discuss, where necessary, the interrelatedness between Amazon's core business and its other activities (Including the provision of cloud services, home automation, handsets, etc.) for the assessment of market power and the role of data; notably when discussing dynamic competition.

[63] We briefly touch upon whether dynamic pricing algorithms may facilitate collusion. However, this topic is more elaborately discussed in a separate case study.

[64] This section draws strongly on the following sources:
- CBS: *60 minutes: Amazon's Jeff Bezos looks to the future* http://www.cbs.com/shows/60_minutes/video/lAZRiF8BTzazhjKy9ZIV5camUOze_71_/amazon-s-jeff-bezos-looks-to-the-future/;
- ColdFusion: *How BIG is Amazon?* https://www.youtube.com/watch?v=tCUuvyVwbJs;
- ZDF: *ZDFzoom – Die Macht von Amazon* - https://www.youtube.com/watch?v=BpKQSSiEoRk.

haul using its own airplanes and last-mile distribution using drones and an Uber-like package delivery platform[65]. The backbone for the distribution of digital content is part of a wider B2B service *Amazon Web Services (AWS)* which includes cloud computing services, as well as a Content Delivery Network[66]. Amazon has incorporated its logistical activities as an integral part of its platform under the brand names *Prime* and *Fulfilment By Amazon (FBA)*. Prime is a subscription for consumers which includes free delivery of Amazon products (non-Prime users only get free delivery if the total purchase is over 25 euros) and free access to digital content on *Prime Video*. FBA is a is service aimed at retailer allowing them to fully outsource logistical activities (storage, picking, shipping and payment) to Amazon which makes their products eligible for *Prime free shipping*.

Value creation, network effects and revenues

Amazon has two types of users: consumers and retailers. Each user group has its own needs which determine the user experience and value proposition offered by Amazon. The users - experiences of these two user groups are interrelated, creating demand externalities between different user groups, causing so-called indirect network effects. This means that the user experience of one user group increases as the number and variety of users in the other user-group increase. Moreover, there are also some direct network effects within a user group which means that the user experience of one user group increases as the number and variety of users in that same user-group increase. With Amazon, direct network effects result from reviews (written by consumers) and suggestions based on if-then-algorithms (based on what other consumers have bought in the past).

Amazon generates revenues (profits) through retail margins made on own sales and margins made on sales by other retailers[67]. Amazon's revenues are thus determined by the number of transactions taking place on its platform, which is a function of the number of users. Amazon realises additional revenues with the *Prime* and *FBA* brands. Consumers can get a subscription to *Prime* and retailers pay a user fee when outsourcing logistical activities to *FBA*.

Assets

Amazon's key assets are related to: the *channels* through which they reach users (Website, App, and Kindle); the *data* on search and buying behaviour on its website and data on product offerings and prices on its website; the *algorithms* used for analysing this data[68]; and its *investments in the logistical system* used for storage and delivery of physical and digital content (Prime, FBA, and AWS).

Costs: economies of scale and scope, and operational efficiency

The Amazon business case is characterised by two sources of scale economies: 1) high upfront technology investments in the channels to reach users and in the logistical system, and 2) bargaining power vis-à-vis suppliers. Economies of scope means that the average costs are decreasing if the variety of services/goods goes up. Amazon realises economies of scope by diversifying product offerings and opening its website and logistical system for other retailers. Operational efficiency is constantly improved by automation of processes in warehousing and distribution activities, driving Amazon's tendency to vertically integrate throughout the distribution value chain.

---

65 The last mile distribution is still in development. Amazon is experimenting with setting up its own delivery network using self-employed licensed drivers (http://www.bizjournals.com/seattle/blog/techflash/2016/02/amazon-begins-to-hand-off-last-mile-delivery.html) as well as drones and air based distribution centres (http://fortune.com/2016/12/29/amazon-floating-warehouses/).

66 AWS is rented out to parties like Netflix and Dropbox, but also forms the backbone for *Amazon Prime Video*.

67 Amazon also offers retailers more prominent product placement across its websites, services and devices (including Kindle, Prime Video, etc.) in return for an advertisement fee, but we exclude this activity from this case for reasons of scope.

68 Amazon's algorithms are developed in Amazon's Software Development Centres located on various continents.

Competitive environment

User power

Consumers can choose among a wide variety of e-commerce platforms ranging from pure market places (where the platform operator is not active as a retailer: like eBay), to pure online shops (where the platform operator is the only retailers and does not allow other retailers to sell via its web shop: like CoolBlue in the Netherlands), to hybrid forms (acting as a retailer on their own market place: like Amazon and Bol.com). Retailers can also multi-home. In fact, 77%[69] of the retailers using Amazon make use of the other market places, or set up their own web shop and use other channels to reach end-users like search engines and comparison websites.

In this context of multi-homing, Amazon's ability to fulfil users' needs is essential for persuading them to complete transactions on Amazon and not on other e-commerce platforms. Because of indirect network effects, it may suffice to focus on maximising the user experience of one group to prevent the other group from using other platforms. In other words, Amazon can increase its bargaining power vis-à-vis retailers by increasing the 'stickiness of consumers' (i.e. lowering their incentives to multi-home) through optimal fulfilment of consumers' needs for instant gratification and value for money[70]. Moreover, by getting consumers subscribed to free home-delivery via *Prime*, they are less likely to shop at other e-commerce platforms where they still face delivery charges. By including free video in the *Prime* subscription, more consumers are persuaded to subscribe to *Prime* (and less willing to multi-home).

Next to consumer stickiness, retailers' abilities and incentives for multi-homing are reduced further by persuading retailers to make use of *FBA*. While FBA greatly contributes to retailers' experience (as it takes care of shipping, storing, picking, and payment processes) it also prevents the retailer from multi-homing because once the product is stored with Amazon, it can no longer be offered on eBay. Today, 79% of the retailers on Amazon use FBA[71] [72].

Bargaining position vis-à-vis suppliers

Companies like Amazon and Bol.com can have strong bargaining positions vis-à-vis some **producers** of content, products or services (CPS), enabling them to negotiate large rebates [73]. However, other producers can counter this power with their own strong brand name and by choosing other channels to reach end-users[74].

All e-commerce platforms experience some degree of dependency on other platforms. Notably search engines, comparison websites and app stores are important channels to reach end-users. For a strong brand name like Amazon, this dependency is often less strong [75] or even mutual[76]. Moreover, there is a dependency of e-commerce platforms on providers of payment systems (e.g. credit card platforms).

Amazon owns and controls most of the activities in its distribution chain, except for last-mile delivery. For this end-node in the distribution network, Amazon is largely dependent on large

---

69 http://www.webretailer.com/lean-commerce/amazon-sellers-survey-2016/#/.

70 By making use of ranking algorithms, by orchestrating price and quality competition among retailers, by letting consumers use all the major payment systems, and by striving for cost-leadership in end-to-end distribution.

71 http://www.webretailer.com/lean-commerce/amazon-sellers-survey-2016/#/.

72 A retailer making full use of the entire FBA service package also outsources its stock keeping to Amazon. By placing its stock in the Amazon warehouses, the retailer can no longer multi-home with other platforms for that particular product. Some retailers use different platforms for different types of products. See https://www.youtube.com/watch?v=seTBVbbNTX4.

73 https://www.nrc.nl/nieuws/2016/04/26/wordt-bolcomzo-dominant-als-amazon-1612998-a1113983.

74 By using other e-commerce platforms and/or by opening their own webshop.

75 People would find their way to Amazon without the use of Google.

76 Amazon is such a popular website that blocking the Amazon App from the App store would harm experience of the users of the App store.

carriers, like UPS, DHL, DPD and national postal services. Because of scale economies, these carriers have certain degrees of bargaining power which motivates Amazon to vertically integrate using innovative solutions, such as using self-employed licensed drivers and drones[77].

Competitive pressures from direct competitors

From the description of users' ability to multi-home it shows that Amazon competes with a diverse set of e-commerce platforms.[78] Some of these platforms are like market places (like eBay and Marktplaats); others are webshops of a single retailer (like Mediamarkt and Coolblue); some are both at the same time (like Amazon and Bol.com). Some of these platforms are like department stores offering a wide range of product categories (like Wehkamp), others only sell one type of product (like Zalando and H&M).

A key part of Amazon's competitive position is related to cost leadership in distribution for which Amazon has invested quite a lot in physical assets. Amazon vertically integrates more and more nodes in the physical distribution chain with the aim of further optimising the operational efficiency of its end-to-end logistical system through automation and robotisation[79]. Other e-commerce platforms such as eBay attain more value to scalability and agility of their business models by outsourcing warehousing and distribution functions to their network of retailers (see Rogers, 2016).

Dynamics (competitive pressures from outsiders)

Next to other e-commerce platforms, Amazon experiences increased competitive pressures from many other platforms acting as a channel to reach end-users (such as search engines, social networks, and app stores) are trying to attract transactions to their platforms. Notably, social media websites and search engines are evolving their revenue models (which was typically based on advertisement) from pay-per-click/show to pay-per-buy, which makes them a direct competitor of e-commerce platforms. Compared to Amazon, they are often more dependent on the network of retailers for taking care of warehousing and distribution, but this is not necessarily a competitive disadvantage (see Rogers, 2016).

Entrants may typically try to enter the market by offering digital content. The channels through which end-users are reached are very different: App stores (Apple and Google), handsets (Kindle and Kobo), or dedicated Apps (Netflix, Videoland, and Prime Video). The retail of digital content could be regarded a market separate from the retail of physical products because of the different nature of the underlying distribution networks[80]. However, retailing digital content is a powerful first step towards becoming an all-round e-commerce platform[81]. Similarly, Amazon is using retailing of digital content (Kindle and Prime Video) as to strengthen its current position.

---

77 If we allow for some speculation, it may very well be that Amazon will operate its last-mile network of licensed drivers as a two-sided platform for consumers and retailers to use. In the Netherlands, we already such a business model by another party: keendelivery.com (formerly known as JetVerzendt).

78 The number and size of Amazon's direct competitors differs between countries. This may be explained by differences between countries in the degree to which Amazon enjoys first-mover advantages. A key part of Amazon's competitive position is related to cost leadership in distribution for which Amazon has invested quite a lot in physical assets.

79 A notable acquisition by Amazon was Kiva, a company delivering unique highly advanced warehouse robots to many retailers and warehouses. After the purchase of Kiva, Amazon kept the technology for itself, putting competitors at a (temporal) disadvantage. See https://www.bloomberg.com/news/articles/2016-06-29/how-amazon-triggered-a-robot-arms-race.

80 For the distribution of digital content, one needs a so-called Content Delivery Network; i.e. store the content on a network of servers which are located near the end-users, such that the end-users experience as little as possible interruptions caused by congestion.

81 In fact, Amazon itself started as a distributer of content (books), which was a deliberate choice in the strategy to become "the world's largest retailer". Books were a rather simple product to catalogue, search, suggest and to send via mail order, yet highly heterogeneous from the perspective of content. As such, books were a relatively easy product from the logistical perspective and at the same time attracted a very heterogeneous audience. Moreover, books had proven to be a beloved topic for discussion among users, as was clear from many on-line book fora where people published and discussed book reviews.

Looking forward, competition is shifting toward the Internet of Things (IoT), in which home automation platforms (like Amazon's *Echo* and *Alexa*, and similar innovations by Facebook, Google, Apple, and Microsoft) are aiming to become the dominant gateway for consumers towards instant gratification. Amazon's webshop and distribution systems may be a competitive advantage in this competition for tomorrow's gateway.

Overall conclusion

Although consumers and retailers can multi-home, Amazon's first-mover advantage and costs leadership result in a competitive advantage over many of its competitors. These fist mover advantages differ per country. Furthermore, the global scale of Amazon allows the company to invest in high tech solutions in distribution as well as in new channels to reach end-users. Direct competitors who lack the global scale, often do not have these abilities, but dynamic competitors providing other services at global scale do (e.g. Google, Apple, Facebook) and form a threat for Amazon as their revenue models evolve towards pay-per-buy models. Amazon is responding to this threat by investing in new channels to reach end-users (i.e. devices and technologies).

Role of data (use cases)

Data and algorithms play an important role in Amazon's business case for optimising user experience, where consumer experience is measured in terms of instant gratification and value for money and retailer's experience in terms of optimising sales. Moreover, data and algorithms play an important role in realising operational efficiency in the distribution chain.

Orchestration of transactions

Data that Amazon uses for the orchestration of transactions is personal data (on search and buying behaviour, user reviews) in combination with data on prices and inventory of retailers. The data on search and buying behaviour is used for developing user profiles containing information about individual consumer needs at a specific point in time. Subsequently, the data on inventory of retailers is used for finding a match between products offered and consumer needs at that point in time. The match between consumers' needs and products is produced by an automated filter based on some form of if-then algorithm. Subsequently, for each product, an algorithm suggests a default seller whose prices are lowest and ratings are best. Consumers can select an alternative seller, but only a few do[82]. Data on user reviews of products and of retailers can be used to add sentiments to this equation[83], but user reviews can also directly be used by users themselves as a human controlled filter for identifying the most appropriate product and a trustworthy retailer[84].

Being a retailer on its own platform, Amazon uses data for strategic choices about new product lines to be included in its own assortment. For this purpose, Amazon mainly uses data on transactions (i.e. data on sales and prices), possibly in combination with user generated data (notably user reviews). Retailers on Amazon's platform also have access to transaction data through APIs via the Amazon Marketplace Web Service (*MWS*) [85]. Next to assortment decisions, data can also be used for design decisions. A related example is how movie-review data (such as published on Amazon's IMDB website) can be analysed for making new blockbuster movies or

---

82 82% of the shoppers buys from the default seller. See Chen at al (2016) referring to Taft, D. K. Amazon buy box: The internet's $80 billion sales button. eWeek, October 2014. http://www.eweek.com/enterprise-apps/slideshows/amazon-buy-box-the-internets-80-billion-sales-button.html.

83 Bauman, Konstantin, Bing Liu, and Alexander Tuzhilin. "Recommending Items with Conditions Enhancing User Experiences Based on Sentiment Analysis of Reviews." *CBRecSys 2016* (2016): 19.

84 Similarly, user reviews provide Amazon with sentimental information about the added value to its market place of specific retailers, which can be used to identify retailing partners with whom to cooperate more closely or to ban from the store (a retailer who gets too many bad reviews is banned from the store).

85 See Chen, Le, Alan Mislove, and Christo Wilson. "An empirical analysis of algorithmic pricing on Amazon marketplace." *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016.

series[86]. Moreover, pricing data is used for orchestrating algorithm based dynamic price competition among retailers, ensuring that the Amazon platform is competitively priced vis-à-vis other e-commerce platforms and webshops.

Economic characteristics of the data

Data on buying behaviour is treated as *exclusive* by Amazon (it does not actively share this information). However, because users multi-home, other e-commerce platforms avail of similar (but not identical) data about buying behaviour. Search behaviour is *less exclusive* because, in the presence of multi-homing, people are likely to search for products via multiple e-commerce channels and share the same information with other webshops, market place, review websites, or search engines. Moreover, for the purpose of developing user profiles, dynamic competitors from adjacent markets have many other types of data that can *substitute* for search data[87]. These different forms of user sourced information are *substitutes* and *complements* at the same time. Data about search behaviour is partly substitutable for data about buying habits when analysing consumer needs, but the combination of data about search and buying behaviour also allows for identifying which needs have already been fulfilled, resulting in more relevant rankings and suggestions. Data on ratings and user reviews is *complementary* for this purpose as it allows for the inclusion of sentiments in the equation, and for users to manually refine the selection of relevant products. Data on user reviews is to mostly *exclusive*. Indeed, it is on public display on Amazon's website and can be analysed by third parties[88], but it cannot simply be copied due to intellectual property rights[89] and it will be practically challenging to re-use Amazon data for ranking results on eBay.

For the purpose of orchestrating transactions, data on retailer's inventory is a *perfect complement* to the user profiles and needs and *non-substitutable*. The data is *non-exclusive* as long as retailers are multi-homing. Today most retailers share information about their inventory with multiple e-commerce channels[90] but not with dynamic competitors from adjacent markets. Moreover, Amazon's strategy behind Prime and FBA aims to reduce multi-homing by retailers which would make data on inventory *more exclusive*.

Transaction data is the main input for assortment or product line decisions, possibly *complemented* with user generated data. Transaction data is in principle exclusive, but Amazon discloses the data through APIs, making it *non-exclusive*, and there are many third parties offering data on and analysis of sales volumes and prices on Amazon[91]. There are also many third-parties providing dynamic pricing algorithms to retailers based on having direct access to the Amazon data[92]. Furthermore, for the purpose of assortment decisions, competitors can also purchase *substitutable data* from several third-party data providers like Gfk, and other e-commerce platforms (like eBay) also provide market data through APIs[93].

---

86 http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html.

87 E.g. social graphs, communication data, viewing behaviour, location data, 'likes and dislikes', etc. These data are gathered through services other than e-commerce, but can also be used for orchestrating transactions if retailers actively share information about their inventory or assortment. The latter is not always the case.

88 It can be gathered by third parties using web scraping technologies.

89 Intellectual copy rights rest with the author of the review who consents (via terms of use of the Amazon website) to Amazon using the review. As such, third parties may not copy reviews for publishing on their own website, without the consent of the original author.

90 At least 77% of the retailers on Amazon http://www.webretailer.com/lean-commerce/amazon-sellers-survey-2016/#/.

91 See https://www.google.nl/search?q=amazon+sales+data+analysis.

92 See e.g. https://www.pricespectre.com/faq.shtml.

93 http://ebaydeveloper.typepad.com/dev/2006/12/tapping_into_eb.html and http://cgi6.ebay.com/ws/eBayISAPI.dll?SolutionsDirectory&page=results&subgroup=15.

Operational efficiency

The objective of operational efficiency is about containing costs of inventory and distribution and minimising the time between order placement and order delivery[94].

Types and characteristics of data

Forecasting short-term demand is based on user sourced data (searches, purchases, and reviews) and transaction data (sale and prices). Amazon uses data on size and shape of products in the automation of storage processes as to optimise storage capacity of warehouses. This data is most likely retrieved from electronic files produced by the manufacturers of the products. The data is *non-exclusive*. In the automation of order picking processes Amazon uses transaction data (on items sold and delivery addresses) and sensor data throughout the logistical system. The sensor data is retrieved through sensors of the growing number of robots, that again use the data as an input for managing Amazon's warehousing and logistics. So far, artificial intelligence has not played a significant role in warehousing and distribution, but this is likely to change in the (near) future[95].

Market power

The role of data provides Amazon with a competitive advantage by lowering the mobility of users across platforms, as it:

- enhances the network effects on the platform;
- supports Amazon's pricing position vis-à-vis other web shops and e-commerce platforms; and
- (indirectly) supports Amazon's cost leadership in distribution by realising the scale necessary for generating the necessary ROI to investments made in automation.

(Data-driven) network effects

The orchestration of transaction is characterised by so-called data-driven network effect. It follows that more users, leads to better data, leads to a better functioning algorithm in orchestrating transactions, leads to more instant gratification and sales, leads to more users; and the circle is full. Data involved in this circular process are data about search and buying behaviour, ratings and reviews, and data about retailers' inventory. Data about search and inventory are largely non-exclusive and/or substitutable; as such, there is only a partial data-advantage. Furthermore, Amazon may also be superior in attracting transactions simply because its algorithms are better[96]. Assuming that other e-commerce platforms have managed to develop algorithms of equal quality, Amazon may have some data related first-mover advantage over direct competitors which is related to exclusive user data on purchasing behaviour and reviews on the Amazon platform.

Amazon's position is contested by dynamic competitors – as evidenced by the strategic responsiveness of Amazon to the threat of dynamic competitors (with Kindle and Alexa). Like direct competitors, dynamic competitors lack the same volume and variety of purchasing data, but they make up for this with many other types of data which Amazon does not have and with more effective (or a wider set of) channels to reach end-users. For example, while Amazon realises highly relevant suggestions by combining search data with purchasing date, Facebook has valuable information about social graphs and your communication with peers, which Amazon does not have. Moreover, using different types of data, dynamic competitors challenge Amazon also with different channels to reach consumers which may be more successful in continuously drawing people's attention[97].

---

94 Which simultaneously serves 'instant gratification of consumer needs'.

95 https://www.bloomberg.com/news/articles/2016-06-29/how-amazon-triggered-a-robot-arms-race.

96 Considering that Amazon was among the first to develop if-then suggesting algorithms in the 1990's it is possible that Amazon is ahead of the learning curve.

97 The Facebook App constantly pushes for your attention which makes that the potential buyers are more often exposed to suggestions made by Facebook. Similarly, Google's multi-service business model makes that, whatever people are doing online, they are likely to use a service which is powered by Google.

Dynamic pricing

The facilitation of algorithmic or dynamic pricing on Amazon's market place ensures that offers on the Amazon website are competitively priced vis-à-vis other webshops and e-commerce market places. Other e-commerce market places facilitate dynamic pricing in a same way and, while Amazon's data used for algorithmic pricing is also available to competing e-commerce platforms, dynamic could in theory also occur across platforms. As such, Amazon does not gain a competitive advantage due to a data-advantage.

Operational efficiency

Data seems not to play a direct decisive role in realising operational efficiency. Operational efficiency is realised through automation, which enlarges the importance of scale economies. Amazon leverages the scale realised through (data-driven) network effects on the market place into its logistical activities. The question is therefore not if competing retailers have enough volume, variety and velocity of data, but whether competing distribution systems can also realise the minimum efficient scale in order to make profitable investments into automation of logistics. Scale in warehousing and distribution can also be realised by third party providers of warehousing services pooling transactions of multiple retailers[98]. Moreover, in last-mile delivery, there are other non-integrated courier platforms offering their services to competing online retailers and market places[99].

Operational efficiency and cost leadership contribute to delivering 'value for money' which makes consumers (and indirectly retailers) more sticky. But Amazon's cost leadership also allows the company to offer low-cost warehousing and distribution services (FBA) to retailers who subsequently cannot offer the same products elsewhere. At the same time, cost leadership allows Amazon to offer home delivery subscriptions (Prime) to consumers who subsequently will not shop at other e-commerce platforms where they face additional delivery charges.

Theories of Harm

Data is one, but not the only factor determining market power

As we have seen above Amazon's business case is characterised by some data-driven learning and network effects, but the associated data is only partially exclusive and only partially non-substitutable. Notably, multi-homing makes that essential data (e.g. on retailer's inventory) is non-exclusive. Multi-homing is in general an important driver of countervailing buying power. As such, cost-leadership seems equally (or perhaps even more) important for the competitive advantage of Amazon. In a wider and more long-term context, dynamic competitors from outside the e-commerce sector contest Amazon's position with new channels to reach end-users and/or by integrating e-commerce in existing channels[100]. These dynamic competitors can often leverage a differentiated dataset.

Exclusive or discriminatory behaviour

Being a retailer on its own platform, Amazon may have incentives to exclude other retailers from the platform. For example, Amazon may choose to close the platform for (a selection of) other retailers or maintaining an information asymmetry, the effect of which may be equivalent to raising rivals' costs / lowering its quality. Alternatively, Amazon could structurally display its own offers more prominently in the Buy Box.

When analysing this potential behaviour, one should keep in mind that the incentives of Amazon as a retailer may conflict with the interests of the company as the operator of an e-commerce platform. The platform (rather than the retail service) is characterised by network effects which may give rise

98 An example is Quiet Logistics, a company that provides 'fulfilment services' in the US.

99 Such as Keendelivery.

100 Like social media websites, search engines, navigation services, app stores, and online broadcasting/streaming services.

to exponential growth paths (but which may also result in exponential paths of decline [101]). It seems therefore logical to conclude that, as long as there are competing platforms and people are multi-homing, the interests of the platform come before the interests of the retail service. In a competitive setting, there may thus be good arguments for Amazon as a platform not to engage in exclusive behaviour because this would go at the expense of diversity in product offerings and network effects [102]. This also explains why Amazon shares data on sales and prices through APIs with other retailers. Similarly, Amazon would run the risk of losing diversity in product offerings if it would structurally discriminate against other retailers in selecting the default seller in the Buy Box [103].

Pricing strategies

Amazon can potentially undercut most of the prices charged by other retailers because of scale and operational efficiency. An aggressive pricing algorithm implementing this strategy could either scare other retailers away or have a same impact as 'lowest price guarantees', reducing rivals' incentives to cut prices and thereby leading to supra-competitive prices [104]. Either way, this aggressive pricing strategies (possibly supported by having a data advantage) may seem beneficial from the perspective of Amazon as a retailer, but it may not be in line with interests of Amazon as the operator of a platform as it would not contribute to delivering 'value for money', nor to delivering a wide variety of choice. In the context of multi-homing, people may start looking for products elsewhere. The need to balance Amazon's dual objectives (as a retailer and as a platform) makes Amazon's pricing algorithms more complex than pricing algorithms of other retailers on Amazon's platform who do not need to balance these objectives.

Personal profiling makes price discrimination easier. Because of personal profiling it is easier to identify differences in the willingness to pay of people and to adjust prices accordingly. There have been early examples in 2000 that Amazon was applying differentiated prices for different customers [105] [106]. While price discrimination is made easier because of data and personal profiling, it is also made more difficult because of the ease of multi-homing [107]. Moreover, when detected, it may set up people and result in negative reviews which may do more harm than good in the context of multi-homing and network effects [108]. Again, we notice that there may be different interests for Amazon as a retailer and Amazon as a platform.

---

101 For example, according to Van Gorp and Batura (2015), My Space had more US visitors than Google in 2006. However, Facebook caught up to My Space in 2008 and the number of active users on My space has declined ever since.

102 When a product competes with the Amazon platform or the wider Amazon ecosystem (e.g. like Google's Chromecast, AppleTV, or Google Home), it may also be in the interest of Amazon as a platform operator to ban these products from the store (On October 1, 2015, Amazon announced that Google Chromecast and Apple TV products were banned from sale on Amazon.com by all merchants; In December 2016, Google Home smart speakers, a direct competitor with Amazon Echo, were blocked from sale on Amazon.com.).

103 We note that, even if Amazon is selecting the default seller in a non-discriminatory way on the basis of objective criteria, this may eventually result in a bias to a particular retailer because of data-driven circular causal relations: retailer A appears in the Buy Box → more people buy from retailer A → retailer A gets more reviews → retailer A appears more likely in the Buy Box.

104 See Hviid and Shaffer (2010) "Matching own prices, rivals' prices or both?", *journal of industrial economics*, Volume LVIII September 2010 No. 3, pp. 479-505.

105 http://edition.cnn.com/2005/LAW/06/24/ramasastry.website.prices/.

106 After some negative public feedback the company said it would stop applying such pricing strategies; see: http://www.bizjournals.com/seattle/stories/2000/09/25/daily21.html.

107 Interestingly in this context, between 2008 and 2010 reporting about price discrimination by Amazon kept popping up over time, particularly about customers being charged higher prices based on their *Prime* membership (i.e. those customers who are less likely to multi-home). see http://crookedtimber.org/2008/12/22/amazons-price-discrimination/ and http://federalism.typepad.com/crime_federalism/2009/02/amazoncom-pricing-scam-price-goes-up-after-book-put-into-wish-list.html and https://brightviolet.wordpress.com/2010/05/18/amazon-coms-price-discrimination-how-your-amazon-prime-membership-isnt-such-a-good-deal-after-all/.

108 For this reason Netflix said "We like our simple pricing" in response to a study indicating that Netflix could increase its profits by tens of millions per year if it applied differentiated prices. http://www.forbes.com/sites/adamtanner/2014/03/26/different-customers-different-prices-thanks-to-big-data/#2609dec7f31c.

There are examples of collusive pricing on Amazon[109]. One related insight from the analysis above is that it would be in the interest of Amazon to safeguard consumers from such practices.

Conclusion

- In some regions, Amazon clearly has a first mover advantage which has resulted in a certain degree of market power. The first mover advantage results from data-driven network effects, where data plays an important role in increasing user stickiness in the context of multi-homing. However, multi-homing is not prevented and there is constant competition over 'channels to reach end-users' which makes that Amazon's market position is constantly challenged. Scale from the website is leveraged to realise operational efficiency in the vertically integrated distribution activities;
- An important lesson is that exclusivity of data which cannot easily be duplicated and which is characterised by reverse causalities with mediation functionalities of the platform (transactions in the case of Amazon), contributes to what is known as data-driven network effects (and hence market power). A company may use additional data to support mediation functionalities. But such data is not exclusive or can easily be duplicated, and there are no reverse causalities between the production of data and the mediation functionalities, the data is likely not a source of market power;
- While (data-driven) network effects lead to scale, a platform with many users can more easily adopt high-tech solutions that are characterised by scale economies. As such, data may indirectly contribute to market power;
- Another insight is that with platform based business models information advantages are not (always) associated with competitive advantage. Amazon has invested in sharing data on volumes and prices with platforms users (other retailers) such that the platform can perform optimally in its intermediation functionalities, and compete with other platforms.

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

Case study 2: Social Media - Facebook

Introduction

This case study analyses the largest[110] social media platform in the world, Facebook, in order to explore the causalities between big data and market power in the context of social media. Only the Facebook social media platform, and no other Facebook-owned products and services,[111] is examined. The Facebook-owned social media application Instagram, and Facebook-owned communications service Whatsapp, are discussed in the context of the social media competitive environment.

This case study aims to provide insights and learnings on the links between big data and market power that can be applied to social media in general. Specifically, this case study explores two [112] potential consequences of data-driven market power in the context of social media: lowered incentives to protect the privacy of users, resulting in suboptimal privacy protection and, data-driven efficiencies allowing improvement of services. Findings and learnings of this case study may also be relevant in contexts outside of social media where the type of big data used is personal data.

Social Media

Social media is a "*group of Internet-based applications that [...] employs mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content*".[113] More generally, social media falls under the category of online platforms which, other than social media, encompass a wide range of activities including online advertising platforms, marketplaces, search engines, communications services, payment systems, and platforms for the collaborative economy.[114] The key characteristics of online platforms, and thereby of social media, are the use of information and communication technologies to facilitate interactions between users, collection and use of data about these interactions and network effects (platform value increases with more and more users).

Company Description

General description of activities

Facebook *"is a mobile application and website that enables people to connect, share, discover, and communicate with each other on mobile devices and personal computers"*. [115] The most important social media feature[116] of Facebook is News Feed "*which displays an algorithmically-ranked series of stories and advertisements individualized for each person".*[117] In understanding the importance of News Feed, it is essential to recognise that Facebook is a two-sided platform. On one side of the platform are consumers who use Facebook for its social media functionality, posting and seeing posts within News Feed. On the other side are advertisers who use the Facebook News Feed (among other features) to publish advertisements (ads). For consumers, there is no monetary cost for using Facebook. Advertisers pay Facebook for publishing their ads based on the number of

---

110 Based on number of active users and geographical spread of users. Some regional social media platforms in Russia, China or Japan may be larger based purely on numbers of users.

111 Such as Facebook M (virtual assistant) or Oculus (virtual reality glasses).

112 One competition concern relevant to social media that will not be discussed in this case study due to unavailability of data relates to market power leading to suboptimal innovation.

113 European Commission, Digital Single Market, Glossary of Terms. Accessed 7 March 2017. https://ec.europa.eu/digital-single-market/en/glossary#letter_s.

114 European Commission, Digital Single Market, Online Platforms. Accessed 7 March, 2017. https://ec.europa.eu/digital-single-market/en/online-platforms-digital-single-market.

115 Facebook, (2015), Facebook 2015 Annual Report.

116 Other social networking features in Facebook include: friend requests, "like" of posts/pages, "pokes" to friends, creating and rsvp´ing to events, and messaging friends via Facebook messenger.

117 Facebook, (2015), Facebook 2015 Annual Report.

clicks made by people, the number of actions taken by people, or the number of impressions delivered.[118]

Revenues, value creation, and network effects

In 2016, 98% of Facebook´s revenue was generated from selling ad placements.[119] Facebook generates advertising revenue once a user takes an action in response to a particular ad. The ads that are published on the Facebook platform and affiliated websites (collectively referred to as "Facebook Audience") are tailored to the user, based on an algorithm which matches the personal data Facebook has on a user to the advertiser´s defined audience profile. On average, Facebook made $4.73 of revenue per user from advertising in Q4 2016.[120] Users located in US and Canada generated the most revenue, both on a per-user ($19.28 revenue per user) and total basis (52% of total revenue in Q4 2016).

For consumers, the value of the News Feed is in the ability to share (via "posts"), and see shared content ("posts") from their connections (Facebook friends); engaging, communicating and staying up to date with the "news" in their network. Given the social element of Facebook (sharing, connecting, communicating), Facebook becomes more valuable to an individual user as more and more people use Facebook. This is known as direct network effects. As at the end of 2016, Facebook reported having 1.23 billion daily active users[121] (i.e. via website and mobile) and 1.15 billion mobile daily active users[122] (i.e. via mobile only).[123] For advertisers, the value of the Facebook platform and Facebook Audience is in having a digital space to post their ads and have these ads seen by their desired audience. Hence, the size of Facebook´s user base and the level of engagement of its users on the one side of the platform, directly relates to the attractiveness of Facebook to advertisers on the other side of the platform. This is known as indirect network effects.

Assets and Partners

Facebook's key assets are related to: 1) the technological channels through which they reach users (Website and Mobile App); 2) the personal data of their users; 3) the algorithms used for analysing this data; and 4) the data centres in over 30 countries which support the geographically spread Facebook user base. Facebook has three categories of key partners: 1) System partners, including operating systems such as Android, iOS, and Blackberry OS, as well as payment systems such as credit cards and bank payment systems allowing consumers to purchase virtual and digital goods; 2) Product partners with whom Facebook works together to offer complementary or integrated services, for example working with Skype for video calling; and 3) Third party websites which add the Facebook like button, allow log in to their services via a user´s Facebook account, or which are a member of Facebook Audience.

---

118 Facebook makes revenue from the delivery of click-based ads in the period in which a person clicks on the content, from action-based ads in the period in which a person takes the action the advertiser contracted for, and from the display of impression-based ads in the contracted period in which the impressions are delivered. Impressions are considered delivered when an ad is displayed to people (Facebook 2015 Annual Report).

119 The remaining 2% was generated from what Facebook terms "Payments", a fee-based service Facebook provides to developers which allows people to buy virtual and digital goods from developer´s applications, primarily social games. Facebook receives a fee from developers when a person engages in a payment transaction.

120 Average revenue per user (ARPU) is total revenue in a given geography during a given quarter, divided by the average of the number of monthly average users (MAUs) in the geography at the beginning and end of the quarter. ARPU includes all sources of revenue. The number of MAUs used in this calculation only includes users of Facebook and Messenger.

121 Facebook define a daily active user as a registered Facebook user who logged in and visited Facebook through the website or a mobile device, or used the Messenger application (and is also a registered Facebook user), on a given day.

122 Facebook define a mobile DAU as a user who accessed Facebook via a mobile application or via mobile versions of the website such as m.facebook.com, whether on a mobile phone or tablet, or used the Messenger application (and is also a registered Facebook user) on a given day.

123 Facebook acknowledge difficulties in accurately measuring the number of Facebook users, estimating that around 7% of user accounts may be duplicate or false (e.g. accounts of businesses/organisations or accounts used for spamming).

Costs: economies of scale and scope, and operational efficiency

Facebook´s scale economies relate to the fact that Facebook has expanded its "production" of its social media platform from wholly within the US to worldwide, with 85.2% of Facebook users located outside of the US and Canada.[124] Facebook´s scope economies relate to Facebook offering more and more products both within the realm of social media (Instagram) and broader online platform products including Whatsapp (communications services), Oculus (virtual reality), Virtual Assistant M (artificial intelligent), MarketPlace (ecommerce/market places), Workplace (business/professional collaboration and communication services) and Atlas (services for managing and measuring ad performance on and off the Facebook platform).

Facebook realises operational efficiencies through horizontal integration with non-Facebook-owned websites and apps both in terms of increasing data collection and in terms of increasing revenue from publishing ads. With respect to data collection, Facebook utilises the Facebook "like" button as well as cookies installed on third party websites to collect the browsing data of their users. With respect to ad publishing, Facebook offers the ability to publish ads not only on the Facebook platform, but also on third party affiliated websites and mobile apps (Facebook Audience).

The competitive environment

Given the two-sided nature of the Facebook platform, Facebook competes both as a social media platform, and as an online advertising platform. These two competitive environments have different dynamics, yet overlap in terms of the players involved and influence each other's success. The success of Facebook as a social media platform has a direct influence on the success of Facebook as an online advertising platform; the more active users, the more attractive for advertisers. Equally, decisions on the online advertising side have a direct influence on the social media side; too many, or irrelevant, annoying or offensive ads makes Facebook less attractive for users.

Competition - social media platforms

Facebook competes with other social media platforms to attract, engage and retain users. There are currently no products that replicate the full range of the Facebook platform capabilities on a similar geographical scale. Prior to Facebook, MySpace (launched in 2003) was the most popular social media platform internationally, while Hyves (launched 2004) was popular in the Netherlands. Google has integrated social media functionality into a number of its products since early 2004 in an attempt to enter the social media platform market, with the latest attempt being Google+ (also referred to as Google Plus).[125] However, Facebook overtook MySpace in 2008 and Hyves in 2010, while Google never gained a substantial active user base on any of its attempted social media platforms. In more recent years, there are several start-ups trying to enter the social media market with a similar product to Facebook[126], but do not seem to be able to capture market share. There are a number of regional social media platforms that have strong positions in particular countries, for example China's WeChat and Russia's VK, as well as more niche social media platforms targeting differentiated segments or services such as LinkedIn (professional networking and employment) and Twitter (news and social networking).

Products in the communications services market, such as WhatsApp, Telegram Messenger, Viber, Facebook Messenger, Wire and LINE are considered dynamic competitors to the Facebook social media platform. In their investigation into the Facebook acquisition of WhatsApp in 2014, the European Commission disagreed with arguments that WhatsApp was already a social network and therefore a competitor to Facebook, finding the two to be distant competitors in particular due to a

[124] Facebook, (2016), "Facebook Q4 2016 Results", slide 3.

[125] Google+ is the company's fourth attempt at a social networking product, following Google Buzz (launched 2010, retired in 2011), Google Friend Connect (launched 2008, retired by March 1, 2012), and Orkut (launched in 2004, as of 2013 operated entirely by subsidiary Google Brazil – retired in September 2014).

[126] Notably Diaspora and Ello.

substantially richer experience offered by Facebook. The lines between social media and communications services within the world of online platforms are however blurring and evolving quickly. While (mobile based) applications in the communications services market offer functionality limited to one-to-one text, call, photo or video communication and not the full capabilities of social media, there are examples of communication apps which are increasingly building in social media type functionality. With Snapchat for example, it is now possible to share a (time-limited) photo or video to a user's entire Snapchat network – much like a "post" on Facebook. WhatsApp launched in February 2017 "status", the ability to post (time-limited) photos and videos for contacts to see, exactly the same functionality as Snapchat stories, which Instagram copied earlier.

Similarly, products which originally began as web-based platforms within a niche segment are increasingly adding functionality where a user can create a profile, share content, connect and follow friends, and like or comment on content shared by others, for example Facebook-owned Instagram (photo sharing platform), Tumblr (blogging platform) and Pinterest (idea sharing platform). Web- and mobile-based information and entertainment products including online games such as PokemonGo, are also increasingly incorporating social media functionality. YouTube (video sharing and vlogging) already allows users to "subscribe" to video channels much like the "following" functionality of Twitter and Instagram is exploring community features that allow content creators to share text and posts.

In understanding the competitive environment of social media it is important to recognise the importance of network effects and multi-homing. Users multi-home, meaning that they use multiple social media platforms. However companies competing head-on to Facebook by providing a very similar product do not seem to be able to capture adequate market share. The problem may be that for social media platforms very similar to Facebook, multi-homing is not an option for consumers. Heterogeneous user preferences have allowed a differentiation strategy to bring success for competitors in the social media market whereby platforms positioned themselves to address specific groups of users or targeted a unique perspective (niche markets). There is a trend towards the existing differentiated platforms blending into one another. That is, once niche platforms have obtained a critical mass, they extend their functionality to include those offered by other social media platforms. In addition, products or services which may seem to be distant competitors may become direct competitors in the future by adding social media functionality onto an existing product or service and taking advantage of an existing user base. With existing platforms becoming increasingly similar and communications platforms expanding into social media, the extent of user multi-homing may decrease as social media platforms converge and offer a less unique and differentiated product. Thus it appears that social networks that which come from different directions (e.g. Instagram, Snapchat) are able to become serious challengers to Facebook (multi-homing is common practice for such social media platforms).

It may be the case therefore that social media platforms for which single-homing is common practice, there is little scope for competition. Relevant rivals (in this case more regional rivals) have similar business models and in this sense, competition is more static. For other platforms, where multi-homing is common practice, there is room for competition. Relevant rivals have different business models and in this sense, competition is more dynamic (by developing new business models, innovation, etc.). Head-on competitors have little chance to become successful due to network effects and customer lock-in, but alternatives based on different concepts might at some point make FB superfluous.

Competition - online advertising platforms

The online advertising market is broadly split into two main categories [127]; search advertising and display advertising.[128] In 2015, search advertising represented slightly more of the US online advertising market than display advertising in terms of revenue (representing 49% of revenue compared to 42% for display advertising and 9% for other forms of advertising [129]). Google, Microsoft (Bing), Yahoo, and AOL are the dominant players in search advertising, while Facebook leads the display ads business (32.5% market share), followed by Google (13.8%), Twitter and Yahoo.[130] Overall, Google is the largest player in the US online advertising market, with the company's revenue accounting for a little over 50% of total market revenue in 2015. Facebook, with close to 14% of total US online advertising market revenue, is the next largest player. The companies ranked 3 to 10 account collectively for 10% of total market revenue in 2015, while the companies ranked 11 to 25 account for a further 10%, demonstrating that Google and Facebook are the two heavy weights in the online advertising market.

Important trends in the online advertising market include a growing shift towards mobile[131] as opposed to desktop advertising, and sustained and significant growth in advertising delivered on social media platforms. Specifically with respect to social media advertising, social media ad spending is likely to exceed $35 billion in 2017, representing 16 percent of all digital ad spending globally.[132] Facebook accounts for two-thirds of this total ad spending. Facebook's closest competitors are Twitter, Instagram (accounts for 5%), LinkedIn, YouTube, Pinterest and Snapchat. Social media platforms generally add advertising once they have matured, with Facebook, Twitter, LinkedIn and YouTube having long been active in social media advertising and Instagram, Snapchat and Pinterest being newer players.

Role of data (use cases)

Facebook uses collected user data as an input of production to improve their services, on the one side by increasing the relevance and quality of the functionalities provided to users, on the other side by offering better targeted advertising services. These two user cases are discussed in turn, explaining the role of data in each user case.

Optimising the Facebook experience for users

On the user side, the quality of the functionalities offered to users can be enhanced by using the collected data to increase the relevance of suggested social network stories, suggested contacts and social networking interactions.[133]

Types of data and how it is used

The type of data used to optimise the user experience of the Facebook platform for the user includes "volunteered data" and "observed data". Volunteered data includes information the user provides themselves, for example demographic information in their profile (age, gender, location and more), photos, lists of friends in their contacts, likes of pages or organisations, and search queries entered into the search box within the Facebook platform. Observed data is data collected or created by Facebook through analysing the (browsing) behaviours and habits of its users. The means by which Facebook obtains this behavioural data is through the use of "cookies". Information

127 With a minor "other" category including lead generation and classified online advertising.

128 Display advertising, as defined by the IAB, encompasses the advertising categories: banner ads, digital video, digital audio, sponsorships, and rich media advertising served to mobile devices.

129 PWC, 2016. "IAB internet advertising revenue report 2015 full year results - An industry survey conducted by PwC and sponsored by the Interactive Advertising Bureau (IAB)".

130 Figures are based on total ad spending in the search and display advertising markets respectively. Source: eMarketer.

131 Google is the leader in mobile advertising, capturing 32% of the mobile ad market, with Facebook the closest competitor with 22% of the mobile market. Source: eMarketer.

132 Source: eMarketer.

133 Graef, Inge. 'Market Definition and Market Power in Data: The Case of Online Platforms'. World Competition 38, no. 4 (2015): 473–506.

about the user's interests and preferences is stored by the web browser in a text file that is sent back to the server every time the user accesses a server's page using the same web browser.[134]

How this data is used involves an algorithm selecting which pieces of information are most relevant for a specific user.[135] More specifically, this is a self-learning algorithm which ranks stories in order of importance, selecting only the most relevant and engaging stories to display in a user's Facebook News Feed. Facebook gives highest priority to those stories a user is most likely to engage with (i.e. Like, comment on, share, click and read), because the more a user engages with the content, the higher the value Facebook is delivering to them, and thus the more they will use Facebook in the future. In order to choose the stories most likely to illicit engagement from a particular user, the algorithm assigns each story a personalised relevancy score that differs for each user, putting the most relevant first. The relevancy score assigned to any individual story is calculated based on thousands of different factors, with the most important being [136]; who posted it (the higher a user's past engagement with that poster, the higher the score), when it was posted (more recent the higher the score), interactions with the post (the more others have engaged with the post, the higher the score) and the type of content (the more that user engages with a particular type of content – status update, link, photo, video, event, job change – the higher the score). Importantly, the more a user engages, the more Facebook's algorithm learns about what that user cares about, evolving that understanding if a user's behaviour changes.

In addition to the implicit ways in which a user can influence the stories that appear in their News Feed as just explained (i.e. by engaging with the content), Facebook offers explicit ways for a user to teach the algorithm which stories to choose; via drop-down options on each and every story, and via News Feed settings. The drop-down allows a user to hide a post, unfollow the author, save the story for later, report the post or turn off notifications for that post. The settings include a "See First" option whereby a user can choose people or pages whose posts will always be shown at the top of their News Feed. On a more global scale, Facebook's algorithm learns to show stories less if many users hide or report them, or show them more if many users share the same story. Facebook also constantly works to improve News Feed, for example by showing fewer posts that ask users to like or share or that include spam links or language to trick people into clicking, as well as showing more posts from friends and from high quality articles.

Optimising the effectiveness of online advertising

On the online advertising side, the collected data can be used to increase the relevance of ads to an individual user.

The type of data and how it is used

The type of data used to achieve the tailoring of ads includes "volunteered data" and "observed data", the same data as in the user case just discussed. There are four ways in which Facebook collects information about its users in order to match users and advertisements; Activity on Facebook apps and services, other online activity, information shared with a business, and location. Each is further explained as follows[137]:

- Activity on Facebook apps and services: activity on the Facebook platform including likes of pages by the user or the user's friends, Facebook and Instagram profile information (current and past places of residence, date of birth, gender, email address, schools attended, employment history, job title, relationship status, language spoken, family members and life events) and places a user has checked into. The Facebook Like button is a means by which

[134] Graef, Inge. 'Market Definition and Market Power in Data: The Case of Online Platforms'. World Competition 38, no. 4 (2015): 473–506.

[135] Facebook, 'How News Feed Works', available at https://en-gb.facebook.com/help/327131014036297/ (accessed 8 Feb. 2017).

[136] https://techcrunch.com/2016/09/06/ultimate-guide-to-the-news-feed/.

[137] https://www.facebook.com/ads/about/?entry_product=ad_preferences.

Facebook can collect data about which products or articles a user has liked, even when a user is not on the Facebook platform but clicks on the Facebook Like button elsewhere on the web;

- Other online activity: this includes data on website pages viewed where that website uses a "Facebook pixel"[138] (an image that triggers a cookie), data on if the user downloaded the advertiser's app, and data on if the user made a purchase or added an item to a shopping cart of the advertiser's webshop;
- Information shared with a business: includes information on whether a user belongs to a customer list. A user may be added to a customer list whenever they share their phone number or email address with a business, for example for loyalty programs and retail purchases. One way users can share their mobile number or email address is by logging in/signing up to a website using their Facebook login;
- Location: includes sources such as where the user connects to the Internet (based on IP address), where the user uses their phone (based on GPS and location services) and the location the user has entered into their Facebook or Instagram profile.

How this data is used involves an algorithm selecting which ads are most relevant to show in a specific user's News Feed. This algorithm is a similar, but separate ranking algorithm to the one which determine the stories to appear in a user's News Feed. The algorithm chooses the ads a Facebook user is most likely to be interested in based on matching the information Facebook knows about the user with the advertiser's audience definition. It is important to understand that advertisers do not gain access to Facebook's user data itself. Instead, advertisers define who they want to reach based on factors such as interests, age, location, and more. The more Facebook knows about a user, the more relevant the ads will be to that user. Facebook limits the number of ads a user sees in order to maximise the likelihood that the ad it shows to a user will result in the user engaging with that ad.

The algorithm matching the ads to the users is self-learning, continuously improving based on implicit and explicit signals from users. Implicit signals are if a user engages with the ad itself, while explicit signals can include the user giving direct feedback on an ad, managing their ad preferences or managing their ad settings. Giving direct feedback on an ad is where the user clicks on a drop-down menu in the ad whereby they can tell Facebook that the ad is useful, they can hide the ad, or they can ask why they are seeing this ad.

A user can influence the "advertising profile" Facebook has built on them via Facebook's "Advert Preferences" by removing any interests, "categories" (e.g. late technology adaptor, housemate-based household, living away from family, frequent traveller), past advertisement interactions or certain profile information (specifically: relationship status, employer, job title and education) from their advertising profile. A user can also turn three advert settings on or off; i) whether Facebook is allowed to use a Like of a user as a form of advert to their friends, essentially a kind of recommendation advert, ii) whether Facebook can use a user´s ad preferences to tailor ads to that user on other services and apps outside the Facebook platform and apps (the Facebook Audience Network), and iii) whether Facebook can use a user´s web browsing activity outside the Facebook platform to tailor ads inside the Facebook platform.

In addition to collecting data on users, Facebook offers services for advertisers which collects data on the performance of their ads.

[138] The Facebook pixel is a piece of JavaScript code which an advertiser can install on their website in order to obtain data about the people who visit their website and see how their customers are moving between devices before they convert. Advertisers use the collected data to create ads for publishing on Facebook based on the products people have visited on their website, find more Facebook users who are similar to their best customers and create custom audiences of Facebook users who took specific actions on their website (such as visited a product page, added to cart or purchased a product).

Market power

Facebook competes in two markets that are inextricably linked, whereby competition in the social media market for "eyeballs" (aka users), drives the competition for advertising spend in online advertising. In terms of eyeballs, the Facebook social media platform has by far the largest market share of any social media or communication service (whether calculated based on numbers of users or on number of site visit). Data (US only) on number of site visits shows Facebook with a 42% market share, followed by YouTube (25%), Twitter (5%) and then Pinterest, Instagram, Tumblr and LinkedIn (all between 1% and 2%).[139]

Facebook, as do other social media platforms with a large user base, enjoy direct network effects. In addition, Facebook enjoys a (big) dataset of significantly more *volume* and *variety* than any other competitor given the scale, scope and operational efficiencies Facebook realises in collecting user data (as previously discussed). The data generated by users drives a learning effect whereby Facebook can improve and develop its social media platform to offer a superior service for users, retaining users and keeping direct network effects in place. This in turn enhances the indirect network effects for advertisers on the platform.

However, there are still many competitors that limit Facebook's market power in both markets. A further constraint to market power is multi-homing, users can switch easily to alternative platforms. The fact that Facebook implements functionality that is the same as Snapchat, shows that they consider them to be a competitor.

Thus, although data clearly contributes to Facebook's strong position in the market data is clearly not the only contributing factor. Facebook's market power is based on an incumbent position plus Facebook's ability to attract new and retain existing users based on the strength of their product offering. Big data helps them keep improving their product offering. But if a "hypothetical data broker" would be able to offer all of Facebook's data to a competitor it is unlikely that Facebook's market power would disappear as it would still benefit from its direct network and indirect network effects.

Theories of harm

The theory of harm specifically relevant for social media services is that market power may lessen non-price competition in terms of the array of privacy protections offered to consumers. Privacy competition, like other facets of non-price competition, already exists in other industries, but some dominant companies do not face the competitive pressure to improve quality along this dimension. Facebook, which is free to users in exchange for personal data of the user, may offer a suboptimal level of privacy protection in the absence of non-price competition (users may not even be aware of the suboptimal level of privacy protection).

This report (see chapter 6) suggests that there are in general two ways to mitigate "harm" due to a dominant position. One is competition law (The German Bundeskartellamt (BKartA) is attempting to go down this path, see box 1) and the other is consumer, data and privacy protection law.

Box A.1 - Facebook Privacy Breach Investigation (2016)

In March 2016, Bundeskartellamt (the German Competition Authority) initiated proceedings against Facebook on suspicion that the social network provider had abused its market power by infringing data protection rules.

[139] Data is from the month November 2016, http://www.dreamgrow.com/top-10-social-networking-sites-by-market-share-of-visits-august-2016/.

The concern: that Facebook is violating data protection law by not properly informing individuals of its data collection and use practices.

The investigation: Bundeskartellamt will examine, among other issues, to what extent a connection exists between the possibly dominant position of Facebook and the conditions of use for users (terms and conditions relating to data protection provision).

Key conclusions: The proceedings remain ongoing however, the case represents the first attempt by a European competition authority to integrate data protection interests into competition analysis, and raises interesting questions about the interface between these two areas of law.

Competition law requires that a company has a dominant position. As we have seen in the discussion on market power this is not obvious in the case of Facebook (it would require a detailed analysis of opportunities for multi-homing and dynamics in the market). If an inquiry would find that Facebook has a dominant position a second step would be to prove that Facebook abused that dominant position and that it resulted in harm for consumers. These two steps are incredibly difficult to come to a conclusion and find a company in abuse of market power, in practice it is very rare that such a case is brought forward and analysed.

For privacy protection authorities, the only relevant question would be if the behaviour is in breach of privacy regulation and if consumers are sufficiently informed about the type and extent of data collected. This would require an investigation into the way Facebook uses data, its terms and conditions and its compliance with privacy protection rules. It is interesting to observe that in its beginnings, Facebook appeared to view privacy as something that no longer existed. In more recent years Facebook has increasingly added privacy options for users, possibly due to regulatory scrutiny or competitive pressure.

Findings relevant for big data and competition

Facebook has a large market share and likely some degree of market power in the market for social media services and (targeted) online advertising. Data generated by users is the core of the network. Facebook earns its revenues by allowing advertisers to target users. Facebook is one of the few firms that has access to personal data (geolocation, age group, profession etc. etc.) of millions of internet users. One of Facebook's selling points to advertisers is that it has multiple data points about a user. Facebook does not sell data to advertisers. Rather it provides advertisers access to its network.

The Facebook case shows that it is difficult if not impossible to disentangle the "network effects" from the "data effects". It is unlikely that Facebook's market power would disappear if "hypothetical data broker" that would be able to offer all of Facebook's data to a competitor as Facebook would still benefit from its network and network effects. Although Facebook has a strong position it is not necessarily "dominant". That conclusion would require an in-depth analysis of the opportunities to "multi-home" and the competitive threat of other players in the market.

Case study 3: Leveraging market power to other markets - Google

Introduction

This case study analyses the largest online search company in the world: Google. Google is involved in a wide variety of internet-related services such as, search, shopping, mapping, communications, social media, browser software, mobile operating systems and home automation. This case study focusses on how Google has grown to become a provider of a wide variety of services, and analyses the role of (big) data as a contributor to Google's success in these services. More specifically, we analyse to what extent Google has been able to leverage market power in the search market to other markets[140] by using personal data across different services.

In the literature 'leveraging market power to other markets' is mentioned as one of the risks of big data for competition. According to the theory of Prüfer and Schottmüller (2017)[141], a big-data advantage may translate into a cost or quality advantage when entering other markets. Network effects[142] and machine learning drive this advantage and may (according to Prüfer and Schottmüller) cause a domino effect whereby a firm can repeatedly leverage its 'data advantage' in one (data-driven[143]) market to enter and become a large (perhaps even dominant) player in other markets. This case study analyses how and to what extent Google is able to leverage its big-data advantage in such a way.

Platform business models

Like Google, companies such as Apple, Amazon, Microsoft and Facebook employ a platform business model, mostly digital platforms with the exception of Amazon which combines a digital platform with the physical element of their logistical infrastructure. Digital platforms capture, transmit and monetise data, including personal data, over the Internet.[144] As such, data is a resource of major importance in platform business models (to a lesser extent for Amazon who is not purely digital but also has key physical elements – logistics – included in their product offering). Companies employing these data-driven platform business models often offer a very dynamic scope of services and products.[145] The relevant product market(s) is thus not so pertinent in this case study. Instead the markets are viewed together as one whole and the question is whether data is the crucial element providing Google with opportunities to compete in markets outside its core search business. Thus, we are not concluding if Google has market power in the search market, but whether Google can use its data gained in its search business to compete in other markets.

Company Description

Google is a company specialising in internet-related services and products. In 2015, Google reorganised its various interests into a conglomerate (multi-industry company) called Alphabet Inc. Google remains the core business of Alphabet and is the umbrella company for Alphabet's internet services.[146] Alphabet has businesses outside of, and less related to, Google which it terms "Other

---

140 Google operates in many markets and there are countless interesting perspectives available in order to analyse the role of data in the success of Google. We will not study all markets in which Google operates in detail.

141 Prufer, J., & Schottmüller, C. (2017). Competing with Big Data.

142 Prufer and Schottmüller (2017) expand on the common definition of "indirect network effects" as commonly known in the literature, introducing the term "data-driven indirect network effects". They propose that such effects arise on the supply side of a market but are driven by user demand, explaining that demand for the services of one provider generates user information as a costless by-product which the provider can use to better adapt the product to users' preferences, thereby increasing quality in the future.

143 Prufer and Schottmüller (2017) divert from the definition of data-driven markets as commonly known in the literature, defining the term as; "*markets where the cost of quality production is decreasing in the amount of machine-generated data about user preferences or characteristics, which is an inseparable by-product of using services offered in such markets*".

144 Evans, P. C., & Gawer, A. (2016). The rise of the platform enterprise: a global survey.

145 Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). Big data for big business? A taxonomy of data-driven business models used by start-up firms. A Taxonomy of Data-Driven Business Models Used by Start-Up Firms (March 27, 2014).

146 The Economist, Feb 2016. http://www.economist.com/news/business-and-finance/21689995-worlds-largest-listed-company-has-earned-patience-investors-googles-parent-company.

Bets". These include Access (high-speed internet service named Google Fibre), Calico (research and development biotech company focused on aging and age-related diseases), CapitalG (growth equity investment fund), GV (venture capital investment formerly called Google Venture), Nest (household devices), Verily (life sciences and health care), Waymo (self-driving car company), and X (a research lab undertaking ambitious breakthrough technology projects).[147]

General description of activities

Google, founded in 1998, began as an online search engine company. Today, Google has seven core products, each with over one billion users; Google Search, Google Maps [148], YouTube, Android operating system [149], the Google Play store, Gmail and the Chrome browser.[150] This suite of products and associated user base is greater than that of any other internet company. In total, Google lists 78 consumer products on its website including Google Scholar, Google Translate, Google+, Google Drive, Google Hangouts, Google Play Music, Google Photos and Google Home, along with a further 26 business products such as AdSense, AdWords, DoubleClick by Google, Google Cloud Platform and Blogger. Many of Google's products, as well as Alphabet´s "Other Bets", rely on data as a resource of major importance.

Revenues, value creation, and network effects

Google generates the vast majority of its revenue by delivering advertising (88% of total revenues in 2016), via desktop and other devices, within its Search, Gmail, Maps, YouTube and Google Play services.[151] Advertisers pay Google per click, or in the case of YouTube, per ad when the user chooses not to skip an ad. Google also charges advertisers by impression, however this represents a small part of Google's advertising revenue base. Google generates a small portion of its revenue from its non-advertising products and services, primarily through sales of digital content, apps and cloud offerings, as well as sales of hardware products.

In Google's online products, consumers are matched with sellers or advertisers. Google offers both search (google search results page, maps and shopping) and display (YouTube, Blogger, Gmail and partnering websites across the internet) forms of online advertising. Google's multi-service business model means that, whatever people are doing online, they are likely to use a product which is powered by Google. The more products Google offers, the more it knows about individual users and the more channels it has for reaching users. Through the former, Google aims to offer online advertising that consumers find relevant, and through the latter, that advertisers find cost-effective. The more relevant (i.e. better matches with users) and cost-effective Google's online advertising is, the more value created for advertisers.

Google offers its (online) products for free[152] to online users across a range of services including search (including Maps, YouTube and Shopping), communications (including Gmail and Hangouts), Cloud (including Drive) and transaction and operating platforms (including Google Play store, Android and Chrome). The value created for users is specific to each and every product, however in general, the value Google creates for users is summarised in the company's mission statement; "*Google's mission to organize the world's information and make it universally accessible and useful*". For Google's core products, the value created for users, and the value Google aims to create in its product innovations, is for users to more quickly, easily and naturally find, access and organise information.

147 Alphabet Inc, 2016. "Annual Report for the fiscal year ending 31 December 2016".

148 Google Maps for desktop was launched in 2005 and Google Maps for mobile in 2008.

149 Launched in 2007, is the dominant mobile device operating system worldwide with over 84.82% market share https://www.statista.com/statistics/263453/global-market-share-held-by-smartphone-operating-systems/.

150 http://fortune.com/2016/08/25/facebook-google-tech-companies-billion-users/.

151 as well as advertisements served on Google Network Members' properties participating in Google's AdSense for Search, AdSense for Content and AdMob businesses.

152 Google offers an enterprise version of some products where companies pay a fee.

Google products are two-sided platforms in the sense that users on one side enjoy the product for free given that advertisers on the other side pay Google for the ability to advertise to these users. Hence, the size of Google's user base and the level of engagement of its users on the one side of the platform, directly relates to the attractiveness of Google to advertisers on the other side of the platform. This is known as indirect network effects. Direct network effects are also relevant for some of Google's products, most notably Google Maps and Google+ (its social media product). Direct network effects describe the case when the value of the product increases with the number of users. The traffic information in Google Maps become better if more users use Google Maps, and the attraction of Google+ increases for users if more people join the network.

Assets and Partners

Google's main assets relate to 1) infrastructure, 2) data management, and 3) analytics (including advanced analytics technology based on artificial intelligence). Infrastructure includes the software and hardware supporting Google's digital services. Software includes the Chrome browser, Android mobile operating system, Chrome operating system, and Daydream virtual reality platform. Hardware includes devices like the Pixel phone, Chromebook laptop and Google Home. Data management involves Google owning and operating data centres in the US, Europe, South America, and Asia.

Google's main partners include Google Network Members and distribution partners, to whom Google pays traffic acquisition costs. Google Network Members are third parties that use Google advertising programs to deliver relevant ads on their sites. Distribution partners include browser providers, mobile carriers, original equipment manufacturers and software developers who make Google's search access points and services available to users.

Costs: economies of scale and scope

Google's scale economies relate mostly to network and learning effects which allow that the quality of Google's services to increase with the number of users. To illustrate Google's scale, we use Google search engine as an example. Google´s production of search results increased from one billion per year in 1999 to over two trillion in 2016.[153] It has been estimated that Google has 80.52% worldwide market share and 64% US market share in the search engine market, far greater than any other competitor with the second largest competitor Microsoft´s Bing having an estimated 6.92% and 21.4% respectively.[154] Important in relation to economies of scale is the learning effect, whereby more users make the use of a product or service better. For example, more users of Google Search allows Google's search algorithm to gain insight into what users want based on user clicks, learning by trial and error, and therefore improving the quality of search results.[155]

Google realises scope economies in its revenue model through its Google Network, whereby non-Google-owned websites and apps can publish Google ads, increasing data collection and revenue. Additionally, Google realises scope economies through its Google Play Store in which they sell apps that can then be used as a further channel to bring advertisements to consumers. Google's scope economies also arise from the fact that big datasets can be used in a wide variety of online services. The more products where Google is able to collect user data, the further this contributes to the Volume, Variety, Velocity and Veracity of their datasets; which not only allows Google to improve existing products, but to explore new opportunities in other markets both within the online

[153] The exact number if unknown, however Google last announced in 2016 that it produces trillions of search results per year. See http://searchengineland.com/google-now-handles-2-999-trillion-searches-per-year-250247.

[154] Based on desktop searches only. Worldwide market share statistics come from NetMarketshare. US market share statistics come from comscore.

[155] Sokol, D. and R. Comerford (2016), "Does Antitrust Have a Role to Play in Regulating Big Data?", in Cambridge Handbook of Antitrust, Intellectual Property and High Tech, Cambridge University Press.

realm and beyond (i.e. hardware such as laptops and phones, as well as home automation, and self-driving cars).

Competitive environment

Each of Google's products operates in a competitive environment specific to that product market. On the whole, however, Google (in its core business)[156] competes particularly with companies that seek to connect people with online information and provide them with relevant advertising. Like Google, companies such as Apple, Microsoft, Facebook and Amazon operate an online ecosystem on which many users and advertisers depend. These companies represent a limited number of firms that have leveraged the power of the platform business model to achieve dramatic growth in size and scale and become household names (sometimes being referred to as super platforms).[157] A striking feature of the business model of these companies is that the scope of their services and products is very dynamic. Outside of their core operations, they are active in markets that are closely related, as well as (at least from an outside perspective) those hardly related to the core activities. Rather than addressing the competitive environment of the individual products offered by Google and its competitors, we examine the competitive environment of these super platforms in the context of entering new markets.

Competitive pressures from direct competitors

The world's biggest tech firms[158] and among the largest companies in the world are Apple, Alphabet, Microsoft, Facebook and Amazon.[159,160] Google leads search (80% market share[161]), Facebook on social media (42% market share[162]), Amazon on e-commerce (38% market share)[163], Apple on hardware/devices - particularly its Apple iPhone (18.3% market share[164]) and Microsoft on desktop operating systems (84.1% market share[165]). While each firm holds a strong position in its own traditional or "home" product market, these firms compete with one another in each other's traditional product markets. For example, Microsoft (Bing) competes with Google in online search, Apple and Google compete with Amazon in retailing of digital content, Google competes with Microsoft in desktop operating systems, Microsoft (LinkedIn) and Google compete with Facebook in social media, Apple competes with Google in mobile device operating systems and Google and Microsoft compete with Amazon in cloud computing.

The strong overlap in the products offered by these super platforms reflects an evolution in the business model (and developments in technology) from individual and isolated desktop-, internet- and mobile-related products towards an online, integrated ecosystem. Each of the aforementioned super platforms is competing intensely to become the dominant super platform offering a one-stop-shop online ecosystem across desktop, mobile and other web-enabled devices. Each company has

---

156 This is not the case with, for example, Cloud, Google+, or Google Home (Nest).

157 Evans, P. C., & Gawer, A. (2016). The rise of the platform enterprise: a global survey.

158 On one day in August 2016, the five biggest companies in the world by market value (as measured by market capitalisation on the US stock exchange) were all tech firms: Apple ($571B), Alphabet ($540B), Microsoft ($441B), Amazon ($364B) and Facebook ($357B). Source: Bloomberg available at https://www.bloomberg.com/gadfly/articles/2016-08-02/tech-giants-form-fab-five-to-dominate-stock-valuation-chart.

159 Within PWC's Global Top 100 Companies by market capitalisation, these five firms came within the top ten across all sectors for 2016; Apple Inc (1st), Alphabet (2nd), Microsoft (3rd), Facebook (6th), and Amazon (9th).

160 According to Forbe's Global 2000 which ranks public companies based on four equally-weighted metrics of revenue, profit, assets and market value, the top tech firms in 2016 within the entire list were: Apple (#8), Microsoft (#23), Alphabet (#27), Facebook (#188) and Amazon (#237).

161 Available at https://www.netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0.

162 Data is from the month November 2016, available at http://www.dreamgrow.com/top-10-social-networking-sites-by-market-share-of-visits-august-2016/.

163 For the US market only, based on percentage of total market revenue in 2016 holiday season. https://intelligence.slice.com/two-extra-shopping-days-make-2016-biggest-holiday-yet/.

164 Data is for Q4 2016 only. Samsung has higher market share than Apple based on the full year of 2016 (21.2% versus 14.6%) and has been the lead smartphone vendor since 2012 with a market share of 20 to 30 percent. http://www.idc.com/getdoc.jsp?containerId=prUS42268917.

165 Global market share held by operating systems for desktop PCs, February 2017, available at Statistica https://www.statista.com/statistics/218089/global-market-share-of-windows-7/.

continuously been entering new markets; drawing upon a combination of many factors in attempting to gain a competitive advantage, including the unique value proposition of their product, the "stickiness" of their ecosystem[166], the reputation of their brand, the size of their user base and the big data they collect and analyse.[167]

Competitive pressures from outsiders (competitive dynamics)

The emergence of the Internet of Things is a critical development providing opportunities for existing and new firms to compete with Google, by developing new products and sources of revenue, as well as new ways to achieve cost efficiencies that can drive sustainable competitive advantages.[168] Competitive pressure may come from any new technology that draws consumers onto platforms outside of the Google ecosystem. Google has so far been successful in responding to the shifting of search-related online advertising revenues away from desktop and towards mobile devices[169] due to their entering and gaining of strong position in mobile and video channels.[170] However, in the emerging world of the Internet of Things, competition comes from any company that can web-enable devices, provide common platforms on which web-enabled devices can communicate, and/or develop new applications to capture new users.[171]

It looks as if the same companies that came to dominate the world of desktop and mobile internet-related products and services are also those best positioned to provide the platform-centric worlds that the Internet of Things will likely need. For example, within the emerging market for assistants (artificial intelligence), there are largely the same group of super platforms active; Microsoft's Cortona, Apple's Siri, Amazon's Alexa, Facebook's M and Google's Google Assistant. The case is similar in the emerging market for virtual reality products with Google active in developing an augmented reality head-mounted display, Microsoft a mixed reality HoloLens headset, Facebook the virtual reality Oculus glasses and Apple a mixed reality iPhone. Artificial intelligence and virtual reality products, as well as wearable devices such as smart watches, are steps towards a post-mobile world and reflect a growing focus on using sensor-based data and creating analytically rich data sets.[172] As the Internet of Things will by definition generate voluminous amounts of unstructured data, the availability of big data analytics is a key enabler giving the super platforms leverage in entering new markets associated with the Internet of Things.

With so many potential markets - connected wearable devices (i.e. smart watches), connected cars, connected homes (i.e. home energy efficiency, home comfort and security), connected cities (i.e. smart electricity grids and electric vehicle infrastructure), and the industrial internet (i.e. condition monitoring on the factory floor) - there is potential for disruptive change from new competitors. The Internet of Things is predicted to deliver the most value in solving complex logistics, manufacturing, services and supply chain problems.[173] Google is already exploring development of flying vehicles

---

166 Customers who, having fully invested in the app ecosystem and are used to the key strokes and functionality and where everything is, have costs both in switching between like-for-like products and in adopting additional products. Apple customers for example who have bought an iPhone tend to buy iPads, Apple smart watches, or MacBooks and stay in the ecosystem because of apps they have downloaded or paid for and the contacts, photos, calendar, etc. stored and interacting together in that ecosystem. Similarly, Microsoft´s enterprise software customers with Windows Office tend to purchase other Microsoft enterprise software such as cloud services.

167 Ted Wechsler, Berkshire Hathaway investment manager. Berkshire Portfolio Manager Explains Apple Investment.

168 Goldman Sachs, 2014. "The Internet of Things: Making sense of the next mega-trend" available at http://www.goldmansachs.com/our-thinking/outlook/internet-of-things/iot-report.pdf.

169 PWC, 2016. "IAB Internet Advertising Report 2015 full year results - An industry survey conducted by PwC and sponsored by the Interactive Advertising Bureau (IAB)".

170 Mobile: Google pays Apple to be the default search engine on iOS devices, plus succeeds in having Google Apps installed by default in the majority of Android mobile devices. Video: Google's YouTube product is the dominant online video site.

171 Goldman Sachs, 2014. "The Internet of Things: Making sense of the next mega-trend" available at http://www.goldmansachs.com/our-thinking/outlook/internet-of-things/iot-report.pdf.

172 https://www.forbes.com/sites/louiscolumbus/2016/11/27/roundup-of-internet-of-things-forecasts-and-market-estimates-2016/#2cc9f63b292d.

173 Goldman Sachs, 2014. "The Internet of Things: Making sense of the next mega-trend" available at http://www.goldmansachs.com/our-thinking/outlook/internet-of-things/iot-report.pdf.

(drones), similar to the Amazon's Prime Air concept, as well as investing in connected and driverless car technology. Thus Google appears to be proactive in positioning itself to remain a super platform into the future.

Role of data (use cases)

Google and Google's users generate a lot of personal and non-personal data. This data is used to 1) optimise Google's services, and 2) optimise online advertising. We consider the role of big data in entering other markets within user case 1) optimising Google services.

Optimising Google services

Google uses data as an input of production to improve their services for users (and advertisers). The following quote from Google's 2016 Annual Report captures the essence of how Google uses data to optimise its services:

> "Within Google, our investments in machine learning over a decade are what have enabled us to build Google products that get better over time, making them smarter and more useful -- it's what allows you to use your voice to search for information, to translate the web from one language to another, to see better YouTube recommendations, and to search for people and events that are important to you in Google Photos."

Types of data and how it is used

Google collects and uses user data in the delivery of its services to users. There are three main types of data that Google collects; 1) observed data 2) volunteered data and 3) user created data. Observed data is collected when users use a Google product, including search queries made, websites visited, videos watched, ads clicked on, date and time of the activity, the location of the user (based on IP address), device information and cookie data.[174] Volunteered data is information a user provides to Google when signing up for a Google account, including name, email address and password, birthday, gender, phone number and country. User created data relates to anything that a user creates when signed in to Google and using Google services, including emails sent and received on Gmail, contacts added, calendar events made, photos and videos uploaded and docs, sheets and slides created on Drive.

Google uses this data to optimise their services, making products faster, smarter and more useful. For example, query or search data is used to improve the relevance of Google Search results in the future by looking at, for example, in which language, from which geographical location, and at what time of the day a user enters a particular search query.[175] The Google Maps app retrieves location data from a user's phone to know the location of that user, combined with data from people nearby (to for example detect when a lot of vehicles are moving slowly) to recognise traffic patterns and suggest the best navigation route. Google also uses a user's search history to autocomplete searches before a user finishes typing them, based on the same or similar searches that user has made in the past. On a global level, Google employs a spelling correction model which uses data from users making spelling mistakes in the past in order to know what users are searching for even when making spelling errors in the future. Similarly, Google's YouTube uses both user-specific data on videos watched as well as global user data on which videos are popular and trending in order to recommend videos for users to watch in the future. More generally, data Google collects via cookies helps a website remember information about a user's visit and which can be used to, for example, remember search preferences, make ads relevant, count the number of visitors to a page and help sign up for services.

---

[174] Google Privacy, Your Data: We want you to understand what data we collect and use, available at https://privacy.google.com/your-data.html.

[175] Graef, Inge. 'Market Definition and Market Power in Data: The Case of Online Platforms'. World Competition 38, no. 4 (2015): 473–506.

As well as using personal data collected within a product to optimise the services of that same product, personal data collected within one of Google's product is also used in delivering services (and relevant advertising) within other Google products. If a user is logged in to their Google account, Google can combine personal data across its different services, as well as websites and apps, to improve the quality of the user experience and to increase relevance of ads shown to the user.[176] My Activity is a Google product where users can see what information Google has collected and stored about them across all Google products, showing everything a user has searched, viewed, and watched using Google services. Users can permanently delete specific activities or entire topics they don't want associated with their account. Google also offers users data portability, allowing users to take search or email data with them for example in switching to a competing product.

Optimising online advertising

Google collects and uses data (and provides products to allow advertisers to collect and use data) in the delivery of its online advertising services. Google offers many online advertising products including AdSense, AdWords, Google Analytics and DoubleClick-branded services:

- Google AdWords: an auction-based advertising program that delivers ads based on user search queries. Advertisers bid on the keywords that will trigger display of their ads. Advertisers can display their ads over Google sites or Google Network websites;
- Google AdSense: helps Google Network websites to deliver AdWords ads that are relevant to the search results or content on their pages;
- Google Analytics: offers a range of marketing analytics products for advertisers to better understand their website and app users and evaluate the performance of content and products on those websites and apps;
- DoubleClick Ad Exchange: is a real-time auction marketplace for the trading of display ad space.

Types of data and how it is used

Google states in its privacy policies that it does not sell data to third parties. Thus, advertisers do not gain access to Google's user data. Instead, Google uses the collected and stored user data to match users with relevant ads. Google uses a wide range of data to determine the ads a user sees. For example, based on a user's current or past location, a user might see ads of nearby businesses or events.[177] Sometimes the ad a user sees is based on the context of a page, on a user's app activity or activity on Google services.

Google uses cookies to collect user data about behaviour on its own products and across many websites which partner with Google to show ads. Cookies allow Google to show ads that are likely to be more relevant (such as ads based on websites a user has have visited) and to prevent a user seeing the same ad over and over again. By using cookies, Google offers services (i.e. AdSense) that let website operators target their ads to people who visited their pages.

Do others have access to the same or comparable data (for similar use cases)?

The super platforms all have to some extent overlapping or comparable data to Google. Through the use of cookies, Google and Facebook collect observed data on online browsing behaviour to determine a user's interests. In addition, Facebook gains information on user's interests through volunteered user data by means of its like button and through what a user "follows". Google, Microsoft (Bing), Facebook and Amazon all collect volunteered data on user interests based on what a user searches for, albeit that Amazon's data is specific to purchase interests and Facebook

[176] Google Official Blog, 2012. "Google's New Privacy Policy", available at https://googleblog.blogspot.nl/2012/02/googles-new-privacy-policy.html.

[177] Google Privacy & Terms, Advertising, available at https://www.google.com/policies/technologies/ads/.

data more specific to social interest in comparison to a much broader interest data that Google and Microsoft are able to collect. Apple also gains information on interests, specific to devices, music, films and apps downloaded, searched or purchased in its apps store (as does Google through its Google Play store). Due to sign ups and accounts, all super platforms have basic demographic data on users such as email address and password, age, gender and location. Facebook, Microsoft and Google likely have more specific demographic user data including employer, birthday, school and key contacts. Google, through its Google Maps product, as well as Apple and Microsoft (Bing), have geospatial data including, traffic data, points of interest, landmarks and navigation. Given Google's size, it is likely that Google has the most extensive data in terms of variety, velocity, volume and veracity of any platform. However, it is unclear whether this gives Google an advantage over competitors who have more focussed data sets. It is also unclear how much of Google's data is obsolete when entering new markets and at which point returns to additional user data begin to diminish.[178]

Market power

In this section, we start with the assumption that Google has a strong position in one market (search) and analyse the effect of that position on Google's market power in other markets. Thus, we are not concluding if Google has market power in search market, but whether Google can use its data gained in its search business to compete in other markets. Important to note for our analysis of market power in this case study is that the connection between services (and the role data plays therein) is key, more so than defining relevant markets. Market definition and structural presumptions are less relevant in fast changing internet-related markets where competition is innovation-based.[179]

Prüfer and Schottmüller study under which conditions and how a dominant company in one data-driven market can leverage its position to another market, including traditional markets that were not data-driven before its entry. They find that a dominant company in one data-driven market who therefore has a (big) dataset superior to its competitors, can use this data advantage to offer consumers a significantly higher quality level (due to learning effects) and to enjoy significantly lower marginal costs of innovation (due to scale economies fuelled by network effects) when entering other markets. Due to these data-driven advantages, companies with a dominant position in one market can thus gain a dominant position in other markets. Stucke and Grunes (2016) list data-related practices that companies with market power can use to tip the market in their favour or to maintain their dominant position including, among others, leveraging a data-advantage in one market to another market.

There is however opposition to the view that (big) data, specifically the associated learning effects and network effects, can provide a competitive advantage in gaining a dominant position. Varian (2015) argues that data alone is useless and that learning effects are subject to diminishing returns to scale. Furthermore, Varian argues that the competitive advantages brought by indirect network effects (which he argues are simply supply side economies of scale) are not unique to data-driven business models, but are the same in any business model. Lerner (2014) also holds the view that because of rapidly diminishing returns to user data any advantages of scale weaken or even disappear, adding that competitive success of online platforms is driven by much more than the amount of user data collected (such as engineering ability/talent).

In analysing the case of Google specifically, it is apparent that Google has had considerable success with its traditional Search product and subsequently, with a number of other products

[178] Shelanski, H. A. (2013). Information, innovation, and competition policy for the Internet. University of Pennsylvania Law Review, 1663-1705.

[179] Shelanski, H. A. (2013). Information, innovation, and competition policy for the Internet. University of Pennsylvania Law Review, 1663-1705.

which Google now considers as its core products. Google has also experienced less success in various markets for example in social media. In the case of Google's successes, each has specific factors which enabled success. For Google Maps [180], Google built from scratch (after first using another providers data) geospatial data and combined that with user created-content and user location data to provide real time and accurate maps to deliver a superior (and free) online mapping and navigation service. For Google's Android product, offered at zero cost to mobile device manufacturers, price and application availability were two key reasons for Android's popularity in the mobile device operating system market.[181] In the case of YouTube, which was acquired by Google, the success post acquisition was due in large part to Google's existing infrastructure and data centres which made it possible for the huge size of the video content to be uploaded. In the case of Google's less successful market entry attempt with social media product Google+ (and its earlier attempts), the product was simply not as good as that of Facebook.

User data have been a key strategic asset in Google's interactions with its competitors[182] however, it is also clear that Google's level of success is different across different services. It seems that the closer a service is to 'search', the better Google performs in optimising user experience as evidenced by having a large user base (e.g. Maps and YouTube). In services that are more distant from search (e.g. communication and enterprise software), Google has enjoyed less success. Google's lower success in markets more distant from their core business may be due to the fact they face much stronger competition from incumbents like Facebook and Microsoft, incumbents who enjoy a competitive first-mover advantage over Google (driven by network and learning effects). It may also be that Google is less successful in markets where direct network effects are present (e.g. social media). It may also be relevant if the new service lends itself for adopting the revenue model currently applied by the company of interest in adjacent markets. Whatever the reason, or combination of reasons, we see in practice that competitive forces cannot be compensated for by a data-advantage (if Google has such an advantage compared to the incumbents).

A preliminary conclusion we draw is that for companies with data-driven business models, the degree to which a big-data advantage contributes to strengthening competitive positions depends on what defines user experience in a case, what specific data is required to produce such experience (and whether that data is exclusive and non-substitutable), and what other assets are required to produce such experience (and whether those assets are exclusive and non-substitutable).

Theories of Harm

Fast changing markets, such as markets for digital goods and services, are characterised by innovation-based competition whereby competition occurs through dynamic cycles of technological change and innovation rather than through static price competition.[183] Intensive and continuous investment in research and development is a strong force in digital platform markets for companies to improve existing products and develop new platforms and applications.[184] The theory of harm analysed in this case study argues that in a market that has tipped, meaning that one company has

[180] Google offered its free Google Maps mobile navigation as a mobile app in 2009, competing with user-paid mobile navigation apps leaders Navigon, TomTom and CoPilot (ALK Technologies). Nokia and Apple offered free mobile map apps with limited success.

[181] Edelman, B., & Geradin, D. (2016). Android and competition law: exploring and assessing Google's practices in mobile. European Competition Journal, 1-36.

[182] Shelanski, H. A. (2013). Information, innovation, and competition policy for the Internet. University of Pennsylvania Law Review, 1663-1705.

[183] This is a view held by one noted school of thought, often called the "Schumpeterian School" after the Economist Josef Schumpeter.

[184] Shelanski, H. A. (2013). Information, innovation, and competition policy for the Internet. University of Pennsylvania Law Review, 1663-1705.

gained a dominant (monopoly) position, there are very few incentives for both the dominant firm and the ousted rival firms to further invest in innovation.

This theory of harm that, driven by network and learning effects offered by big data, a market will tip and result in a lack of innovation is based on certain assumptions. Prüfer and Schottmüller explain that the tendency to tip is stronger in data-driven market because of the personalisation of services that data allows, a possibility that does exist to the same extent in non-data-driven markets. In any market, network effects can lead to tipping (irrespective of data), but multi-homing and demand heterogeneity counter act this. Thanks to big data however, services can be personalised so as to address the heterogeneity of user preferences, strengthening network effects and weakening the counter-force of multi-homing, therefore leading to tipping in data-driven markets.

In practice however we are yet to see a market that has tipped. Even the online search market, where Google enjoys considerable market share, has not tipped to the point that there is no longer innovation. Microsoft's Bing remains as a competitor and both Google and Bing continue to innovate, whether through constantly revising and refining the algorithms that match user's queries to search results or by changing features and functionality. Prüfer and Schottmüller say that before the market has tipped, there is significant competition and as a result, high levels of innovation. As we do not see any market that has tipped in practice, we are in the situation prior to market tipping and therefore in a situation with innovation. Evans (2015) holds the view that the risk of entrenched monopolies in platform markets is very limited, evidence by the historical disruption changes where incumbents have been ousted by new entrants (e.g. MySpace by Facebook, Nokia/Symbian by Google's Android and Apple's iOS). Shelanski argues that it is rare to find a digital product or service that stays the same from day to day, adding that apart from company-led innovation, innovation also comes from the user side, with users creating new ways to use platforms and applications.[185]

While cases of no innovation seem yet to materialise, there may be cases of lower innovation. Lower innovation than would otherwise be the case in the presence of a more competitive market, is very difficult to prove. If we imagine that there is a case of lower innovation in a market, there may be solutions such as data sharing but these are not easy to design or implement in practice (see main report for further discussion).

Conclusion

Google has access to the most detailed and extensive database of online behaviour. Google can be considered a super-platform or platform of platforms. Its search engine, mobile operating system, browser, video platform etc. have evolved into the online ecosystem with the highest market share. This case study shows that sometimes market power on one market is an asset in other markets. Google has shown that it can use its access to data and users to enter new markets. It entered for example the market for online price comparison, maps, online data storage etc. This suggests that its access to machine generated data in one market provides a competitive advantage in other markets. However, this is not a formula for success in all markets, Google's struggles in the social media and car markets show that success is not guaranteed and there are other factors that determine market outcomes. Furthermore, such a scope advantage does not remove the need for a company like Google to make additional investment in new essential data (e.g. mapping data) or new data collectors (e.g. interactive thermostats) when entering a new market.

[185] Shelanski, H. A. (2013). Information, innovation, and competition policy for the Internet. University of Pennsylvania Law Review, 1663-1705.

Case study 4: Smart meters electricity market

Introduction

Rationale

- In the electricity market data is increasingly important. The 'smart grid' where energy consumers manage both their electricity production and consumption in real-time is more and more becoming a reality. Smart meters are a crucial component of a smart grid. Before the introduction of smart meters there was limited information available on electricity consumption. With smart meters, consumers can monitor their own consumption and they can share this information with their energy supplier and third parties;
- The electricity and gas markets are regulated markets. This is especially the case for network operators. Network operators are natural monopolies. In the Netherlands, they are all fully owned by local governments, including the biggest three (Alliander, Stedin, Enexis). When smart meters were introduced, data sharing and the avoidance of market power due to access to the data were explicitly considered.

Focus

- Contrary to the other case studies, we do not analyse in this case how data can contribute to market power but we study what the effects are of policy measures to avoid that access to data or data-exclusivity results in market power. This case study also does not concern a particular company;
- We will analyse why network operators are forced to share smart meter data and what would happen in the absence of data sharing regulations.

Scope

- In order to make the discussion easier we limit ourselves to the electricity market, although the results are also applicable to gas and heat markets.

Relation with other cases

- There are other industries with data sharing obligations. In banking for example a directive will come into force that mandates banks ("Payment Service Provider") to enable third parties to access bank account data if a consumer gives consent to do so (Second Payments Services Directive (PSD2) to be implemented by Member States in national law by 13 January 2018).

Role of smart meters in the electricity sector

Network operators

In the Netherlands, network operators are not allowed to sell energy. The sole task of network operators is the operation of networks and a number of other 'regulated tasks'. Installing and operating smart meters is one of these tasks. All Dutch households will be offered a smart meter before the end of 2020. The Dutch government made the decision to install smart meters in 2014 after a cost-benefit analysis showed that the benefits would outweigh costs. This mandatory cost-benefit analysis was part of the Third Energy Package of the European Commission.

Consumers are free to choose their own energy supplier (such as Nuon, Eneco, Essent, Oxxio, Van de Bron). Products and services based on the smart meter data are also liberalised, whereby certified companies ('onafhankelijke dienstenaanbieders') obtain access to the data if they have the consent of a consumer.

Smart meter data is not processed by the individual network operators but by a central hub (EDSN).

Revenues, value creation

Network operators are regulated monopolies, whereby maximum network tariffs are set by the regulator (Autoriteit Consument en Markt). Tariffs are not related to electricity consumption but the capacity of the connection.

Companies that use smart meter data sell services and products to consumers. Energy suppliers are an active player in the market for data products and services but there are also independent market players that offer in-home displays, web access to consumption data or other products and services. The market for energy supply and energy services is a free market. Network operators are not allowed to be active in this market.

Role of data (use cases)

The main use cases of data on electricity consumption and production are (Van Gerwen et al, 2010)[186]:

- *Billing* – energy suppliers use consumption data for billing purposes. Energy suppliers get access to bi-monthly data. Without explicit consent of consumers energy suppliers do not obtain more granular data;
- *Insight into electricity consumption and demand response* - Smart meters make it easier for consumers to monitor their energy consumption. There are two ways to do this. First, they can provide consent to third parties (energy suppliers or suppliers of energy services) to extract smart meter data ('P4 port'). Second, they can use the 'P1 port' to connect devices, such as an in-home display. The P1 port can also be used to intelligently switch appliances on and off;
- *Managing of the grid* – Detailed data on energy consumption can be used by energy suppliers (and/or network operators?) to monitor the supply/production balance. Excess supply or demand can be identified faster.

Smart meter data are considered personal data and personal data protection legislation is applicable. As users have to provide explicit consent to use data, smart meter data is to a certain extent exclusive.[187] However, it is non-rival, access of one party to the data does not preclude access by another party.

Competitive environment with regulation

In the present market set-up in the Netherlands, there is no party that has exclusive access to smart meter data. If consumers provide their consent, companies can obtain access to the data. There is no competition in the smart meter market itself as network operators are monopolists and they alone are allowed to install smart meters. Network operators are not allowed to offer products and services that use the smart meter data in the Netherlands. This limits the risk that network operators use their market power as the operator of the network and smart meter to compete in the market for energy services and products.

Competitive environment without regulation

In the hypothetical situation without any regulation it is likely that the owner of a smart meter (be it the network operator, a supplier or another party) would be hesitant to share the data with third parties. This is because it could capture the value of the data by supplying energy or energy services or selling the data. Without legislation that forces smart meter operators to share data with third parties, it is uncertain whether companies that deliver services and products that depend on smart meter data would have access to the market.

[186] Gerwen, R. van, F. et al. (2010), 'Intelligente meters in Nederland. Herziene financiële analyse en adviezen voor beleid'.

[187] In the policy debate privacy concerns of smart meter data sharing have been discussed extensively. This has resulted in an opt-out possibility for consumers.

Even without regulation on data sharing there would be factors that would limit the market power of the company that controls smart meter data. Consumers could decide to switch to other suppliers. However, without regulation on the transfer of smart meters this could involve substantial switching costs which would limit the competitive threat to suppliers.

There are alternative ways in which consumers can obtain insight into their electricity consumption. The electricity consumption of specific household appliances can be monitored for example. However, these alternatives are an imperfect substitute for smart meter data. It is likely that the threat that consumers would use other devices would be insufficient to constrain the market power of the smart meter operator.

Due to regulation, the non-exclusivity of data is ensured. At the moment, most EU Member States have in place some kind of regulation regarding smart meters. The UK is the only Member State that does not regulate the network operator or smart meter operator but the energy suppliers.[188] However, as in other Member States, there is regulation regarding data sharing whereby the UK also requires the transfer of the ownership smart meters when consumers switch to a different supplier.

With regulation, all smart meters have to comply with technical requirements. In a free market there would likely be more choice for consumers. The smart meter can for example only send data at specific time intervals (quarter of an hour for electricity), and it is not possible to measure electricity production with the smart meter. Regulation requires standardisation which is not always in line with what consumes and other market players desire.

To summarise, without regulation it is likely that access to smart meter data would be limited. This could result in harm to consumers as the incentives to offer innovative products and services based on smart meter data would be lower. However, in a free market consumers would possibly have more choice regarding the technical characteristics of the meter. Finally, the decision to offer all households a smart meter is based on the assumption that market forces alone do not suffice to bring smart meter penetration to a socially desirable level. It is likely that without regulation the number of installed meters would be lower.

Conclusion

- In the electricity market, the non-exclusivity of smart meter data is ensured by regulation. The mandatory access to data allows third parties to deliver services to consumers. The market for those services is competitive;
- If meters would not be regulated it is likely that the owner of the meter would have some degree of market power. This market power would be constrained by the ability of consumers to switch to a different supplier or smart meter operator. However, this would likely involve switching costs. There are also substitutes to smart meter data such as data on electricity consumption by household appliances that could constrain the market power of the smart meter owner with exclusive access to the data;
- The regulations regarding data access also have some clear disadvantages. Network operators are not allowed to offer services to consumers. Consumers have to decide for themselves if they use products or services to monitor their electricity consumption. This may contribute to lower energy savings than foreseen at the time of the cost-benefit analysis that formed the basis for the decision to install smart meters in all households. According to CE Delft (2015), the initial estimate of energy savings in the cost-benefit analysis by KEMA in 2010 was 3.2%.[189] The actual reduction in electricity consumption (2015) is 0.6%. If network operators would have been

[188] Source: website JRC (LINK), February 2017.

[189] CE Delft (2015), 'Slim gebruik van slimme meters, energiebesparing door grotere beleidsmatige reikwijdte.

allowed to offer in-home displays when they install smart meters, energy savings might have been higher;

- A comparison of the introduction of smart meters with the other case studies investigated in this study reveal the differences in regulatory treatment. From an early stage, policy makers have considered the competitive effects of exclusive access to smart meter data. In the policy debate, privacy concerns of smart meter data sharing have been discussed extensively. This has resulted in an opt-out possibility for consumers;
- In the online environment (for example social media), regulation on data sharing and data access has followed a more gradual approach. It seems that the main reason for this difference is that it is generally accepted that electricity distribution networks are monopolies, there is a long history of regulation in the sector. Moreover, compared to some other markets third party access to data is relatively easy to design as electricity is a commodity;
- Electricity network operators are natural monopolies and thus a textbook example of firms with market power. In this case study we showed that this market power is constrained by regulation and that 'big data and competition' is already considered in the shaping of public policy. In markets without a monopoly but with market players that have some degree of market power, lessons can be drawn from the smart meter case:
  - Big data and competition requires coordination between economic policy (competition in the market), privacy/consumer policy and other policy fields (such as energy policy);
  - The benefits of non-exclusive access to data have to be weighed against the benefits (economics of scale, scope) of having a single party offer services or products to consumers.

About Ecorys

Ecorys is a leading international research and consultancy company, addressing society's key challenges. With world-class research-based consultancy, we help public and private clients make and implement informed decisions leading to positive impact on society. We support our clients with sound analysis and inspiring ideas, practical solutions and delivery of projects for complex market, policy and management issues.

In 1929, businessmen from what is now Erasmus University Rotterdam founded the Netherlands Economic Institute (NEI). Its goal was to bridge the opposing worlds of economic research and business – in 2000, this much respected Institute became Ecorys.

Throughout the years, Ecorys expanded across the globe, with offices in Europe, Africa, the Middle East and Asia. Our staff originates from many different cultural backgrounds and areas of expertise because we believe in the power that different perspectives bring to our organisation and our clients.

Ecorys excels in six areas of expertise:
- transport and mobility;
- economy and innovation;
- energy, environment and water;
- regional development;
- public finance;
- health and education.

Ecorys offers a clear set of products and services:
- preparation and formulation of policies;
- programme management;
- communications;
- capacity building;
- monitoring and evaluation.

We value our independence, our integrity and our partners. We care about the environment in which we work and live. We have an active Corporate Social Responsibility policy, which aims to create shared value that benefits society and business. We are ISO 14001 certified, supported by all our staff.

We look forward to work together and contribute to our collective future.

Manon Janssen,
Chief Executive Officer & Chair of the Board of Management

P.O. Box 4175
3006 AD Rotterdam
The Netherlands

Watermanweg 44
3067 GG Rotterdam
The Netherlands
T +31 (0)10 453 88 00

F +31 (0)10 453 07 68
E netherlands@ecorys.com
Registration no. 24316726

W www.ecorys.nl

Sound analysis, inspiring ideas