
Consumer data in online markets

Prepared for
Which?

5 June 2018

www.oxera.com

Contents

1	Individual data and consumer choice	1
2	The impact of consumer data on both existing and new services	2
2.1	The impact of consumer data on existing services through better matching	2
2.2	The impact of consumer data on new services	3
3	Economic characteristics of data	5
4	Competition and consumer impact	7
4.1	A framework for assessing the impact of consumer data on consumer outcomes	7
4.2	Messaging apps	9
4.3	Health and fitness apps	10
4.4	Retail ecommerce	11
5	The tension between competition and privacy in online advertising	13
5.1	Industry structure	13
5.2	Consumer outcomes	13
6	Conclusions	18

Figures and tables

Figure 3.1	Characteristics of types of consumer data	6
Figure 4.1	Framework for assessing the impact of consumer data on consumer outcomes	7
Figure 5.1	Online advertising value chain	13

Oxera Consulting LLP is a limited liability partnership registered in England no. OC392464, registered office: Park Central, 40/41 Park End Street, Oxford OX1 1JD, UK; in Belgium, no. 0651 990 151, registered office: Avenue Louise 81, 1050 Brussels, Belgium; and in Italy, REA no. RM - 1530473, registered office: Via delle Quattro Fontane 15, 00184 Rome, Italy. Oxera Consulting GmbH is registered in Germany, no. HRB 148781 B (Local Court of Charlottenburg), registered office: Rahel-Hirsch-Straße 10, Berlin 10557, Germany.

Although every effort has been made to ensure the accuracy of the material and the integrity of the analysis presented herein, Oxera accepts no liability for any actions taken on the basis of its contents.

No Oxera entity is either authorised or regulated by the Financial Conduct Authority or the Prudential Regulation Authority within the UK or any other financial authority applicable in other countries. Anyone considering a specific investment should consult their own broker or other investment adviser. Oxera accepts no liability for any specific investment decision, which must be at the investor's own risk.

© Oxera 2018. All rights reserved. Except for the quotation of short passages for the purposes of criticism or review, no part may be used or reproduced without permission.

1 Individual data and consumer choice

Data about us is spreading rapidly, with digital applications now covering many aspects of our everyday lives. New ways of collecting and using our data are arising continuously—one example being the location history tool available through Google Maps.¹

What is the nature of this data, how is it changing existing business models and creating new ones, and what does this mean for competition and consumers? In this paper, we delve into these questions, exploring how the use of consumer data affects consumers across a broad range of markets through competition and privacy outcomes.

As more data becomes available, business models have begun to rely on it to facilitate transactions or generate revenues through targeted advertising. This has enabled widespread innovation, which has delivered wider choice or lower prices to consumers. For example, many social media platforms are free to consumers because revenue is generated through effective (targeted) advertising.

At the same time, the use of this very same data has created competition and privacy concerns. Consumers do not always know or understand where or how their data is being used. A high concentration of data residing with a few firms could represent a barrier to entry, limiting competition. Firms might also fail to provide consumers with adequate transparency and control over how data is collected and used. The recent high-profile case of Cambridge Analytica, and how it may have accessed and used Facebook data, is an example of poor privacy outcomes for consumers.² Policymakers are acting to improve privacy outcomes—with the introduction of the EU General Data Protection Regulation, which came into force on 25 May, and changes to the Privacy and Electronic Communications Regulations on the horizon.³

This paper is a contribution to a Which? project on the collection and use of consumer data,⁴ and is structured as follows:

- section 2 discusses the impact of consumer data on both existing and new services;
- section 3 summarises the economic characteristics of data;
- section 4 develops a framework for assessing the competition and consumer impact of data use, and applies this framework to messaging apps, health and fitness apps, and retail ecommerce;
- section 5 explores the tension between competition and privacy in online advertising;
- section 6 concludes.

¹ Google Inc. (2018), '[Explore your timeline](#)'. All links referenced in this paper were accessed on 23 March 2018.

² See Information Commissioner's Office (2018), '[ICO statement: investigation into data analytics for political purposes](#)'.

³ Regulation (EEA) 2016/769 and The Privacy and Electronic Communications (EC Directive) Regulations 2003.

⁴ 'Control, Alt or Delete? Consumer Research on Attitudes to Data Collection and Use'.

2 The impact of consumer data on both existing and new services

The digitisation of our day-to-day lives has led to unprecedented quantities of data about us being generated which can be used to make existing services more efficient and create new services.

2.1 The impact of consumer data on existing services through better matching

Many economic transactions involve significant costs of searching and matching. In the past, consumers booking a holiday might have walked up and down the high street from one travel agent to the next looking for the best deal. Now they can quickly compare deals through one website, such as Expedia or Skyscanner.⁵

Access to consumer data has made it easier to match consumers with products or services—indeed, most online platforms and services rely on consumer data in their matching processes.⁶ With location data, Uber can identify taxis that are closest to consumers; data about spare rooms allows Airbnb to match hosts with guests; and credit score data enables peer-to-peer (P2P) lenders such as Zopa to match lenders with borrowers.⁷

Better matching reduces the costs for new firms to build their customer base. For example, price comparison websites help new firms in a market to acquire consumers quickly. Price comparison websites typically charge firms a fee for this service. However, this usually costs the firms less than it would do to acquire customers themselves.⁸ Some of these cost savings can then be passed on to consumers in the form of lower prices (although the exact amount of saving passed on will vary).⁹

Better matching also brings direct benefits because it reduces the time people spend searching for their ‘match’. For example, by entering their preferences into a dating app, those seeking romance can spend less time finding their ideal partner than they would through more traditional methods. However, better matching may not always result in more choice for the consumer if all customers gravitate to one supplier in a ‘winner takes all’ game, with the result that some providers do not survive and leave the market.¹⁰ Better matching, in some situations, can also lead to the ‘hollowing out’ of products or services. This occurs when providers compete on only one salient dimension, usually price, and reduce the other aspects of the product that are less salient, such as quality.¹¹

⁵ Expedia (2018), [website homepage](#); Skyscanner (2018), [website homepage](#).

⁶ Oxera (2015), [‘A fair share? The economics of the sharing economy’](#), *Agenda*, December.

⁷ Airbnb (2018), ‘What factors determine how my listing appears in search results?’; Uber (2018), [‘How Uber uses location information \(iOS\)’](#); Zopa (2016), [‘Zopa and credit scores’](#), 11 July.

⁸ Competition and Markets Authority (2017), ‘Digital comparison tools market study’, Final report, [‘Paper E: Competitive landscape and effectiveness of competition’](#), Appendix 3, 26 September.

⁹ Oxera (2014), [‘Passing game: the ongoing debate about pass-on in damages actions’](#), *Agenda*, January.

¹⁰ Note, however, that the ‘winners’ may not stay winners for long: ‘Examples of firms that saw their dominance eroded abound. MySpace and Friendster lost to Facebook; AltaVista and Lycos lost to Google; Blackberry and Nokia were displaced by the iPhone, which now competes against Android phones’, in Kennedy, J. (2017), [‘The Myth of Data Monopoly: Why Antitrust Concerns About Data Are Overblown’](#), Information Technology and Innovation Foundation, 6 March, p. 12.

¹¹ The CMA investigated hollowing out in the 2017 DCT (digital comparison tools) market study, and recommended that ‘sector regulators look to work with DCTs and suppliers to improve the effectiveness of quality metrics in order to mitigate against the risk of hollowing-out’. Competition and Markets Authority (2017), [‘Digital comparison tools: summary of final report’](#), 26 September.

The rise of ad-funded business models that provide better matching of adverts to consumers, such as Facebook, is also partly the result of greater access to consumer data.¹² These platforms provide their services free of charge to consumers, but generate their revenues from advertisers (on the other side of the market). We return to targeted advertising in section 5.

Two examples illustrate the scale of the impact of better matching on existing services.

First, the market for **online dating services** relies on consumer data such as personal details, location and preferences to match people. According to an online survey in 2016, 29% of men and 22% of women in the UK aged 18–64 use online dating sites or apps.¹³ In 2017, there were 85.5m active paying online dating accounts across Europe.¹⁴

Second, visitors are increasingly choosing to use P2P¹⁵ **accommodation platforms** such as Airbnb to find places to stay when on holiday or other trips away from home. Such platforms rely on consumer data such as personal details, location and preferences to match visitors with hosts. Airbnb's market share of the short-stay accommodation market in London (by number of overnight stays) was estimated to have more than doubled from 2015 to 2016 (from 4% to 9%).¹⁶

2.2 The impact of consumer data on new services

Many types of services rely on service providers interacting with consumers on a one-to-one basis, in order to find out more about the consumer (i.e. to access data about the consumer). For example, personal trainers meet with consumers in order to assess current levels of fitness and design suitable fitness programmes.

Technological advances mean that service providers can operate remotely (such as via a video call), but wider access to consumer data also means that the (relatively costly) one-to-one contact is sometimes no longer required at all. With access to consumer data, new service providers can provide services remotely to many consumers simultaneously (and at lower cost).

For example, fitness apps and wearables mean that consumers can track their fitness and set suitable goals without necessarily requiring a personal trainer, all because the app/wearable-provider has access to their personal data (including personal and payment details, and health and fitness details).¹⁷

These online-based services broaden consumer access (through lower prices). However, these new services may not replace the original one-to-one service provided by a professional, as there continues to be a demand for one-to-one interaction (for example, there is still demand for personal trainers to provide in-person assistance).

The growth of **health and fitness apps and wearables** illustrates the scale of the impact that access to consumer data is having. According to a 2016 survey, 21% of men and 18% of women in the UK aged 16+ monitor their

¹² Facebook (2018), '[Choose your audience](#)'.

¹³ Statista (2018), 'Share of users of dating sites or applications in Europe, by sex', original source: TNS Sofres (2016), 'Rapport d'étude Dating et convivialité', February, p. 8.

¹⁴ Statista (2017), 'eServices Report 2017', Statista Digital Market Outlook – Market report, December.

¹⁵ P2P models involve two individuals selling and buying goods or services directly with each other.

¹⁶ Colliers (2017), '[Airbnb In London](#)'.

¹⁷ Wired (2018), '[How to Manage your Privacy on Fitness Apps](#)', 30 January.

health or fitness via apps or wearables.¹⁸ In 2017, the total UK revenues of fitness apps and wearables were £250m.¹⁹ For context, in 2015 UK consumers spent £4bn on gym memberships.²⁰

In the light of the large-scale impact of consumer data on both existing and new services, we next explore the economic characteristics of different types of data. We will see that not all data has the same implications for competition and consumer outcomes.

¹⁸ Statista (2018), 'Share of respondents monitoring their health or fitness via applications, fitness band, clip or smartwatch in the United Kingdom (UK) in 2016, by gender', original source: GfK (2016), 'Health and fitness tracking', September, p. 35. Wearables were defined in the survey as 'fitness band, clip or smartwatch'.

¹⁹ Statista (2017), 'eServices Report 2017', Statista Digital Market Outlook – Market report, December. USD converted to GBP using average 2017 exchange rate.

²⁰ Statista (2018), 'Fitness industry in the United Kingdom (UK)', Dossier: 'UK: estimated growth of the wellness and fitness market 2015-2020, by category', p. 11, original source: PwC (2016), 'Capture the growth Report 2016', p. 9.

3 Economic characteristics of data

Use of Consumer data can affect consumer outcomes in terms of competition and privacy, depending on the economic characteristics of that data.

The extent to which different firms are able to access similar data is important for competition, and depends on two key factors.²¹

- **The cost of acquiring data**, which in turn depends on how the data is collected (from lowest to highest cost): people actively choose to provide their data (e.g. payment details); the data is observed from their behaviour (e.g. browsing history); or it is inferred through analysis of previously acquired data (e.g. personality traits). However, people may consider some of their data to be more sensitive than others (such as payment details), and be less willing to give it away.
- **The rate at which a piece of data depreciates**. Depreciation reflects the time over which data remains relevant or may need to be ‘refreshed’, driven by the frequency with which data points may change. For example, someone’s browsing history may have an economically useful life of only a few minutes to several days, whereas their date of birth is relevant for their entire life.²²

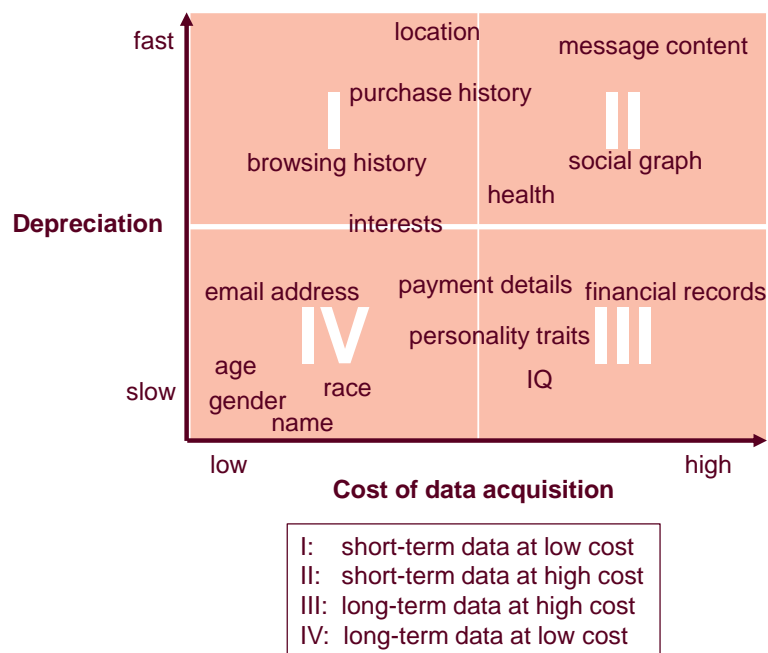
Figure 3.1 shows where different types of data lie along these two dimensions. Starting in the bottom-left corner, data on demographics such as age tends to be widely available as consumers can provide it multiple times and tend do so without much hesitation. Age also evolves in a fully predictable way, so knowing a person’s age once is sufficient for future reference. In contrast, browsing history is also being tracked by multiple firms at the same time, but it changes constantly and needs to be frequently updated to have any value.²³

²¹ Many categorisations of data exist, but these two dimensions capture many of the aspects discussed elsewhere. For example, the discussion of whether data has properties of a public good (by being non-excludable and non-rivalrous) revolves around the question of whether datasets can be replicated, but this categorisation is harder to apply in practice as it neglects the practical drivers of data collection. See, for example, European Commission (2017), ‘[The economics of ownership, access and trade in digital data](#)’, JRC Digital Economy Working Paper 2017-01.

²² See, for example, Kennedy, J. (2017), ‘[The Myth of Data Monopoly: Why Antitrust Concerns About Data Are Overblown](#)’, Information Technology and Innovation Foundation, 6 March, p. 7.

²³ Bujlow, T., Carela-Español, V., Sole-Pareta, J. and Barlet-Ros, P. (2017), ‘[A survey on web tracking: Mechanisms, implications, and defenses](#)’, *Proceedings of the IEEE*, **105**:8, 28 July, pp. 1476–1510. Englehardt, S. and Narayanan, A. (2016), ‘Online tracking: A 1-million-site measurement and analysis’, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, October, pp. 1388–1401.

Figure 3.1 Characteristics of types of consumer data



Source: Oxera.

Someone's social network and interactions (their 'social graph') is more likely to be accessible to only a few firms as it is relatively costly to collect and requires regular updating. Complex inferred data, such as personality traits, may be available (in different forms) to various firms at different costs, as this information can be inferred from a range of factors. For example, online browsing behaviour and even bank transaction data can reveal certain personality traits such as conscientiousness or extroversion.²⁴

It is likely that firms that have access to more comprehensive datasets will have more accurate data—through 'likes' on Facebook, you can get to know someone as well as their loved ones do.²⁵ The importance of the marginal impact of enhanced accuracy is likely to depend on the specific application, as we explore in section 4.

Next we develop a framework for assessing the competition and consumer impact of data. As we will see, the cost of data acquisition and the rate of data depreciation are key factors influencing the consumer impact.

²⁴ Adeyemi, I.R., Abd Razak, S. and Salleh, M. (2016), 'Understanding Online Behavior: Exploring the Probability of Online Personality Trait Using Supervised Machine-Learning Approach', *Frontiers in ICT*, 3:8, 31 May. See DataSine (2017), 'Commercial Psychographic Personalisation', blog post, 15 November.

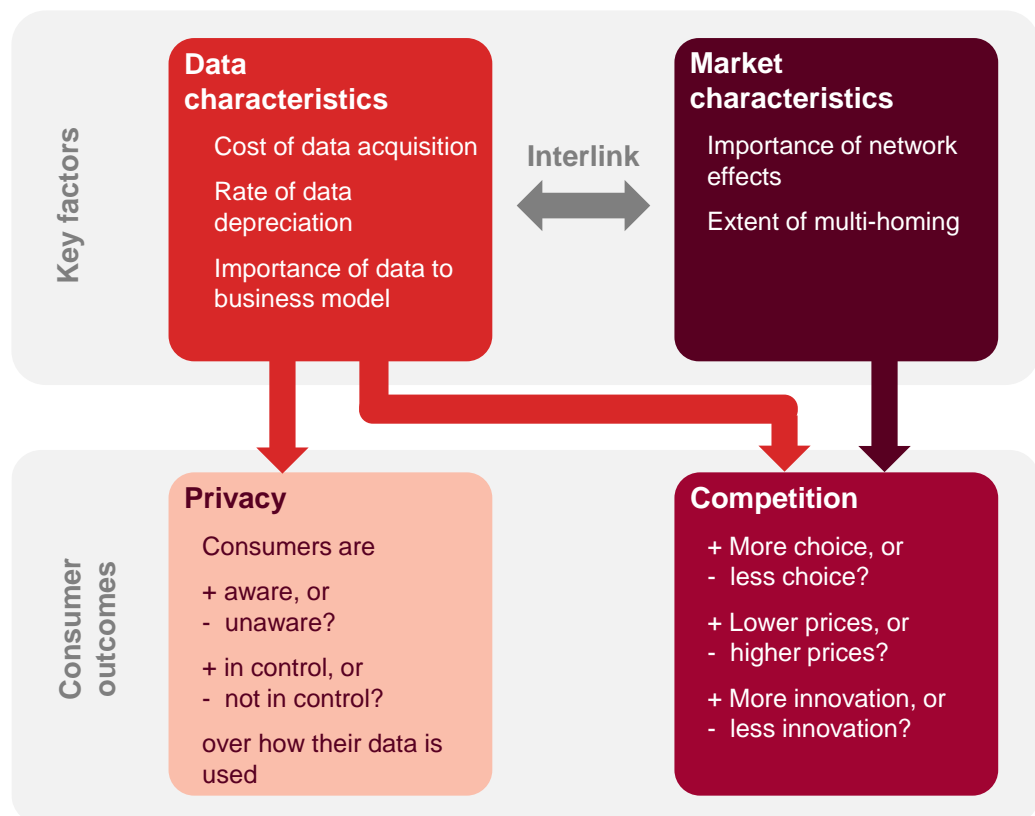
²⁵ Youyou, W., Kosinski, M. and Stillwell, D. (2015), 'Computer-based personality judgments are more accurate than those made by humans', *PNAS*, 112:4, 12 January, pp. 1036–40. See also the seminal paper by Kosinski et al.: Kosinski, M., Stillwell, D. and Graepel, T. (2013), 'Private traits and attributes are predictable from digital records of human behavior', *PNAS*, 110:15, 12 February, pp. 5802–05.

4 Competition and consumer impact

4.1 A framework for assessing the impact of consumer data on consumer outcomes

Figure 4.1 describes the effect of the use of data and market characteristics on outcomes for consumers, in terms of competition and privacy.

Figure 4.1 Framework for assessing the impact of consumer data on consumer outcomes



Source: Oxera.

Consumers enjoy greater choice through firms competing for customers by innovating and lowering prices. However, choice can also come from competition on non-price factors, such as privacy.

The impact of consumer data on competition is driven by the characteristics of the data itself and the importance of the data for the business model in question. Low-cost slow-depreciation data is likely to be easily acquired by many firms, suggesting that they can effectively compete, while high-cost slow-depreciation data may enable a longer-term advantage in a specific data segment. These characteristics of specific pieces of data interact with the characteristics of the market such as network effects and multi-homing.²⁶ For example, a lower cost of data acquisition is likely to be associated with more multi-homing.

²⁶ A network effect is where the benefit that one consumer receives from a network product is affected by how many other consumers also use it. Network effects are a form of economies of scale driven by the demand characteristics of a product rather than the supply side. See Oxera (2013), '[Snowball effects: competition in markets with network externalities](#)', *Agenda*, December. Multi-homing is where consumers use multiple platforms/websites/apps/providers for the same purpose. Using multiple messaging apps is an example of multi-homing.

If a dominant firm has access to more consumer data than other firms, it may well see opportunities sooner than rivals do. However, whether it decides to invest in that innovation, and how quickly, will depend on its potential to grow its market through that innovation, or if the innovation threatens its existing business models.²⁷

The impact of the use of consumer data on consumer outcomes is broader than competition concerns, as competition alone may not deliver good privacy outcomes. Privacy concerns typically centre on two market failures:

- consumers may not know that their data is being collected or how it is being used. This failure may be addressed through greater transparency by the party collecting the data;
- consumers may be unable to prevent their data being used or shared in ways they dislike. But giving consumers 'control' over how their data is used may undermine existing business models, so any remedy would need to be carefully considered. For example, if consumers did not allow social media platforms to use their data for advertising, the platforms might have to charge consumers a fee for their service or limit their services (as they would raise less revenue from the other side of the market).

Consumers are concerned about privacy. However, privacy preferences and consumers' definition of privacy vary greatly across individuals and contexts—so pinpointing consumer valuations of privacy is notoriously difficult.²⁸ For example, the research undertaken by Which?/BritainThinks (2018) shows that people are often pragmatic about data collection and use if they see that the data is relevant and they understand the benefit that they (or society) get from sharing their data. The research suggests that it is the *use* of data and the impact on their lives, not the *collection* itself that informs whether people think it is acceptable.

In addition, people do not always act upon their privacy preferences in a consistent way, because of behavioural biases.²⁹ For example, they value their privacy more when they already have it than when they don't (this is due to loss aversion).³⁰ Many people also think that it may be too late to regain control of their data because organisations already have it; see Which?/BritainThinks (2018).

The variation in preferences might suggest that any policy or regulatory interventions need to be aimed at helping consumers select the right services and settings for their preferences (despite their biases). Food labelling regulation may provide a useful template: the aim is to make it clear and easy for consumers to understand the implications of different choices, while not reducing the amount of choice available.³¹

Below we apply this framework to three markets. For each one, we analyse the data characteristics, the market characteristics, the implications for competition, and finally the implications for privacy.

²⁷ Stucke, M.E. and Grunes, A.P. (2015), 'Debunking the Myths Over Big Data and Antitrust', *CPI Antitrust Chronicle*, May, University of Tennessee Legal Studies Research Paper 276, p. 8.

²⁸ For an overview, see section 3.8 in Acquisti, A., Taylor, C. and Wagman, L. (2016), 'The Economics of Privacy', *Journal of Economic Literature*, **54**:2, pp. 442–492.

²⁹ Oxera (2014), '[Too much information? The economics of privacy](#)', *Agenda*, October.

³⁰ Acquisti, A., John, L.K. and Loewenstein, G. (2013), '[What is privacy worth?](#)', *Journal of Legal Studies*, **42**:2, pp. 249–74.

³¹ See Department of Health, Food Standards Agency, Welsh Government and Food Standards Scotland (2016), '[Guide to creating a front of pack \(FoP\) nutrition label for pre-packed products sold through retail outlets](#)', November.

4.2 Messaging apps

Messaging apps (such as WhatsApp and Snapchat) are typically provided free of charge, as the providers get access to consumer data, which can be sold to third parties and used for targeted advertising. There are also messaging apps that compete by providing greater privacy (such as Telegram and Signal).

Messaging apps typically obtain the following consumer data as a minimum:³²

- **personal details** such as name and phone number—this data is likely to be cheap to acquire and depreciates slowly;
- **social graph** (i.e. the consumer's contacts)—this data is likely to be costly to acquire and depreciates slowly. However, the use of existing contact lists stored on a device may assist in the acquisition of the social graph data required by messaging apps;
- **message contents**—this data is likely to be expensive to acquire and depreciates quickly.

While network effects are strong in this market, and there are therefore some apps with large market shares, these network effects are limited by the high degree of multi-homing in messaging apps, as users face little constraint in using multiple competitors that all provide their services free of charge. For example, 83% of 18–24-year-olds in the UK use Facebook (closely associated with Facebook Messenger), 62% use Snapchat, and 60% use WhatsApp. These apps alone account for an average of two messaging apps per person.³³

While mobile devices enable the convenient aggregation of notifications from multiple messaging apps (for example, on the home screen), there appear to be psychological reasons for using multiple messaging apps. For example, one survey found that consumers use multiple apps to segregate and separate different groups of connections.³⁴ Furthermore, participants would have 'felt deprived of their privacy if everyone would be able to contact them on any communication channel'.³⁵

Multi-homing means that consumers are willing to try out competing apps, reducing the barriers for new providers to enter this market and the cost of acquiring a customer base.

Innovation has recently come from apps providers, such as Signal and Telegram, that offer privacy as their primary selling point.³⁶ Over 2014–17, Telegram doubled its monthly active users to 80m worldwide, although this is still small compared with Facebook Messenger and WhatsApp, which each have 1bn+ monthly active users.³⁷

³² For example, see Snap Inc. (2018), '[Privacy Policy](#)', 5 March.

³³ YouGov (2017), 'YouGov Reports: The Power of the Image Online 2017', cited in Ofcom (2017), '[Communications Market Report](#)', 3 August, Fig. 1.18: 'Which if any, of the following online social media/messaging apps or sites do you personally belong to/use? Please choose all that apply'.

³⁴ Arnold, R. and Schneider, A. (2017), '[An App for Every Step – A psychological perspective on interoperability of Mobile Messenger Apps](#)', 28th ITS European Conference of the International Telecommunications Society, Passau, 30th July – 2nd August 2017.

³⁵ *Ibid.*, p. 12.

³⁶ Signal (2018), [website homepage](#); Telegram (2018), [website homepage](#).

³⁷ Statista (2018), 'Number of monthly active Telegram users worldwide from March 2014 to December 2017 (in millions)', original source: Telegram Messenger, Bloomberg. Statista (2018), 'Number of monthly active Facebook Messenger users from April 2014 to April 2017 (in millions)', original source: VentureBeat. Statista (2018), 'Number of monthly active WhatsApp users worldwide from April 2013 to December 2017 (in millions)', original source: Facebook (2017), 'Q4 2017 Earnings Call', p. 4.

New market entry and innovation at this time are primarily due to costless multi-homing, which is a result of some consumer data being cheap and long-lasting. However, there may be issues around privacy in messaging apps, as highlighted by the intervention of the UK Information Commissioner regarding the possibility of WhatsApp sharing data with Facebook.³⁸

4.3 Health and fitness apps

Health and fitness wearables and apps are increasingly popular with consumers (see section 2), with many firms, such as Garmin and Fitbit, providing apps and devices.

Health and fitness apps collect data and allow consumers to track their progress over time, as well as provide coaching. Low (or zero) consumer prices are also a function of consumer data, as it enables (targeted) advertising.

The consumer data includes:

- **personal and payment details**—this data is likely to be cheap to acquire and depreciates slowly (although the degree to which the consumer is willing to hand over their payment details may vary widely depending on the brand's reputation and consumer trust);
- **health and fitness details** (such as number of steps, heart rate, sleep patterns)—this data is likely to be expensive to acquire as multi-homing is not common in this market. The data may depreciate slowly;
- **location**—this data is likely to be cheap to acquire, and depreciates quickly.

In relation to the characteristics of the market, multi-homing is not common, with 96% of consumers using only one health or fitness app, and there are limited network effects, although comparisons with other people are encouraged by providers.³⁹

The market does not currently appear to be heavily concentrated, and at this point innovation is rapid, with new apps and wearables regularly coming onto the market (e.g. smart socks).⁴⁰ However, there are likely to be considerable risks for privacy around health and fitness details, which are considered very personal. Indeed, the sharing of this data has raised concerns in the media (around both consumer awareness and control).⁴¹

In particular, health and fitness details have considerable value to insurers. For example, health insurer, Vitality, offers customers subsidised wearables and lower insurance prices as a reward for taking exercise, which reduces the cross-subsidy in health insurance between those who engage in healthy behaviour and those who do not.⁴²

In principle, this programme does not raise obvious privacy concerns, because it is transparent and opt-in—users have control (as they choose to opt in) and

³⁸ Information Commissioner's Office (2018), '[A win for the data protection of UK consumers – WhatsApp signs public commitment not to share personal data with Facebook until data protection concerns are addressed](#)', 14 March.

³⁹ Flurry Analytics (2017), '[Health & Fitness App Users Are Going the Distance with Record-High Engagement](#)', blog post on tumblr.com. Strava (2018), '[Filtered Leaderboards](#)'.

⁴⁰ Statista (2018), '[Market share of wearables unit shipments worldwide by vendor from 1Q'14 to 4Q'17](#)'. Sensoria Fitness (2018), [website homepage](#).

⁴¹ *Financial Times* (2013), '[Health apps run into privacy snags](#)', 1 September. Shape.com (2018), '[This Is What Your Phone Does with Your Personal Health Data](#)', 23 August.

⁴² Vitality (2018), '[Activity tracking](#)'.

know that Vitality has access to some of their data, even though they may not know exactly what data is being accessed (assuming that consumers engage with the potential long-term impact of their decisions). The sharing of health and fitness details with insurers would raise privacy concerns if there were no transparency or *meaningful* control.

Taken to the extreme, access to more consumer data could eventually undermine the functioning of health insurance markets, with more accurate health predictions reducing cross-subsidies between individuals. This may have important distributional impacts between healthy and less healthy people.⁴³

4.4 Retail ecommerce

Retail ecommerce is growing in the UK, and accounted for 19% of all retail sales in 2017 (expected to be 25% in 2021).⁴⁴ In general, it is procompetitive, but there may be scope for competition concerns due to concentration in the market, which is exacerbated by the use of consumer data as it allows greater 'personalisation' (tailored products or services).

The consumer data includes:

- **payment and delivery details** input by the consumer themselves—this data is likely to be cheap to acquire, but depreciates slowly. Inputting this data requires effort and trust in the security of the retailer's website;
- **purchase history**—this data is likely to be expensive to acquire, and depreciates quickly. Purchase history enables firms to provide consumers with better recommendations: a 2016 survey found that 58% of consumers were more likely to make a purchase when a retailer recommended options based on their past purchases or preferences.⁴⁵ A 2015 study found that 11.5% of ecommerce revenue could be attributed to personalised product recommendations.⁴⁶ The high cost of acquiring purchase history data risks greater market concentration, because better recommendations and wider user experience are provided by firms with access to more data. However, the value of purchase history data may depreciate quickly, weakening the incumbent advantage.

Despite the effort and trust required to provide payment and personal details, multi-homing is relatively common in ecommerce. According to research by Oxera, less than one-third of consumers across Germany, France, Spain and Poland used only one online marketplace in the previous month (although there was less multi-homing when 'buying products or services from "general" platforms' that offered non-specialist products).⁴⁷ Multi-homing is also encouraged by payment systems such as PayPal, which increase trust and reduce the effort of using multiple sites.⁴⁸

Concentration appears to be increasing, and there are implications for competition in this market: Amazon's UK revenues grew by 49% from 2015 to

⁴³ European Insurance and Occupational Pensions Authority (2017), '[EIOPA InsurTech Roundtable How technology and data are reshaping the insurance landscape](#)', Summary from the roundtable organised by EIOPA on 28 April 2017, 5 July.

⁴⁴ eMarketer (2017), '[UK Retail and Ecommerce: Economic, Sales and Buyer Trends for 2016–2021](#)', 7 August.

⁴⁵ Accenture (2016), '[Personalization Pulse Check 2016](#)'.

⁴⁶ Marketing Sherpa (2015), '[Ecommerce Chart: The most effective types of personalized product recommendations](#)', 11 August.

⁴⁷ Oxera (2015), '[Benefits of online platforms](#)', October, pp. 28.

⁴⁸ See PayPal (2018), '[Safety and security](#)', and <https://www.paypal.com/uk/home>.

2017, and, in 2017, 37% of all UK online spending was via Amazon.⁴⁹ At the same time, there is evidence of rapid innovation—for example, Amazon now offers two-hour delivery with Prime Now.⁵⁰ Smaller ecommerce players also appear to be innovating, with better offerings and more sophisticated pricing, such as the greater use of high-impact ‘flash sales’.⁵¹

These markets are examples of how consumer data is being used to deliver a product or service directly. One of the most contentious uses of consumer data has been its use in targeted advertising, which is discussed below.

⁴⁹ Amazon (2018), [‘Annual Report Pursuant To Section 13 or 15\(D\) of The Securities Exchange Act Of 1934 Form 10-K for the fiscal year ended December 31, 2017’](#), Commission file no. 000-22513. USD converted to GBP using annual average exchange rates. Salmon (2017), [‘King of the jungle’](#), 10 July.

⁵⁰ Amazon (2018), [‘How Prime Now Works’](#).

⁵¹ Big Commerce (2017), [‘The 19 Most Innovative Ecommerce Brands of 2018’](#).

5 The tension between competition and privacy in online advertising

One of the main uses of consumer data is in producing targeted advertising (as we saw in section 2), where it is an important element in some of the new (data-enabled) business models.

Online advertising is often designed to sell a product or service, but it can also be designed to influence opinion or behaviour, such as voting or promoting public safety. It relies on consumer data to target relevant ads to specific groups. Its value has been growing, from £2.8bn in the UK in 2007 to £9.5bn in 2016, accounting for 47% of total marketing spending in the UK.⁵²

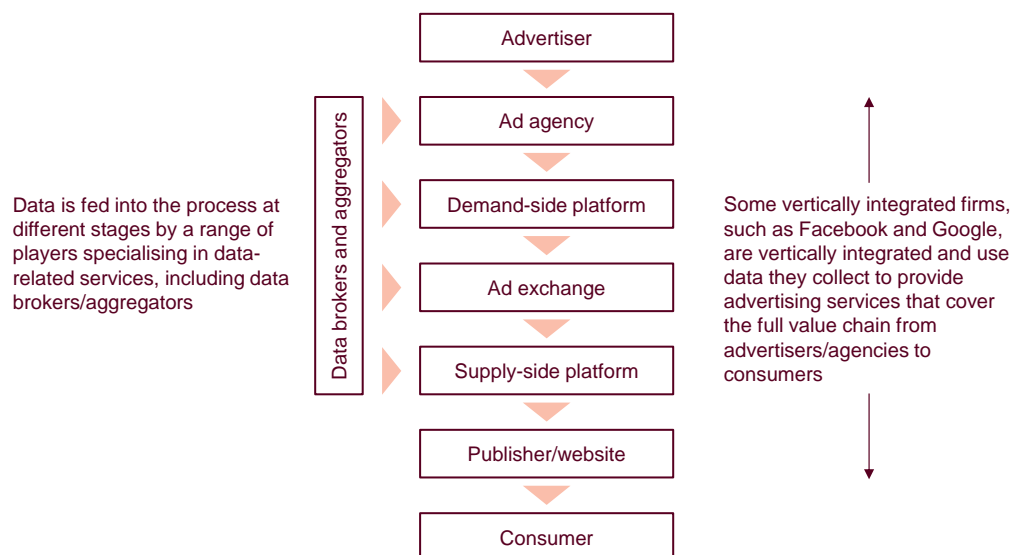
Below we describe the structure of online advertising markets and look at how they affect consumers from a competition and a privacy perspective.

5.1 Industry structure

The digital advertising industry has multiple layers with a wide range of players at each level, as shown in Figure 5.1. The aim of using data for targeting is to make adverts more relevant to individuals, thereby increasing the probability of triggering a consumer action (such as a purchase) in response.

In general, advertisers—often through ad agencies and demand-side platforms—bid for a suitable placement and target audience for their ad. Supply-side platforms aggregate the advertising opportunities and place them for auction on the ad exchange. The matching of adverts and audience and the price-setting ultimately happen on the ad exchange.

Figure 5.1 Online advertising value chain



Source: Oxera.

5.2 Consumer outcomes

Consumer outcomes are affected by targeted advertising in two ways:

- in terms of the **level of privacy** that consumers experience, in relation to the data collected for targeting;

⁵² Statista (2017), 'Digital advertising in the United Kingdom (UK)', Dossier, pp. 7 and 10.

- the **degree of competition** in targeted advertising affects the price and quality that consumers receive from the end product or service being advertised and the digital services that collect the data.

These are discussed in turn below, before we explore the tensions between them.

5.2.1 Privacy

Consumers value their privacy—i.e. the ability to control what information about them is shared with whom and for what purposes (Which?/BritainThinks (2018)). Certain advertising practices may come at a ‘privacy cost’ to consumers, especially where their data may be shared beyond agreed or understood limits (as seen in the Cambridge Analytica case), or if they perceive advertising as targeted and intrusive.⁵³

On the flipside, consumers may derive benefits from revealing data about themselves—directly, by finding targeted ads more useful than generic ones, or indirectly, by using services that are cheaper because of the value of the data that they are sharing (see section 2).⁵⁴ Anecdotal evidence would also suggest that poorly targeted ads can irritate consumers.

Advertisers and advertising platforms may have an incentive to take into consideration some of the privacy cost to consumers. Giving consumers the option to consent to providing their data, or reminding them of their options to control what data they share, and the uses to which it is put, may increase consumers’ responsiveness and thereby the return on the investment in that ad.⁵⁵

However, there are concerns that the markets and processes that rely on individual data may not be sufficiently transparent or easy to understand.⁵⁶ A lack of transparency or high complexity can prevent consumers from engaging with privacy and so can allow advertisers and platforms to offer less privacy than they otherwise would. Theoretical models show that monopolistic platforms can indeed ‘price out’ privacy-sensitive consumers.⁵⁷

In practice, privacy preferences vary greatly across individuals and contexts (as we saw in section 4.1). Accounting for this variation presents a challenge for both firms and policymakers, as simple rules or policies may be overly restrictive in some cases and not restrictive enough in others. However, greater transparency would appear to be beneficial in all cases.

5.2.2 Competition

Competition in online advertising has recently come under scrutiny in various jurisdictions, including in the European Commission 2017 Decision on Google Shopping, but also in sector-wide investigations in Australia, France and

⁵³ See Information Commissioner’s Office (2018), [‘ICO statement: investigation into data analytics for political purposes’](#). See, for example, Goldfarb, A. and Tucker, C. (2011), ‘Online display advertising: Targeting and obtrusiveness’, *Marketing Science*, **30**:3, pp. 389–404.

⁵⁴ For example, in April 2017, 90% of consumers in the USA stated that they found marketing content personalisation ‘very/somewhat appealing’. Statista (2017), ‘Marketing personalization dossier’, p. 23.

⁵⁵ Tucker, C.E. (2014), ‘Social Networks, Personalized Advertising, and Privacy Controls’, *Journal of Marketing Research*, **51**:5, pp. 546–62; and *Harvard Business Review* (2018), [‘Ads That Don’t Overstep’](#), January–February.

⁵⁶ For an overview see, for example, Acquisti, A., Taylor, C. and Wagman, L. (2016), ‘The Economics of Privacy’, *Journal of Economic Literature*, **54**:2, pp. 442–92, section 3.1.

⁵⁷ Bloch, F. and Demange, G. (2017), ‘Taxation and privacy protection on Internet platforms’, *Journal of Public Economic Theory*, **20**:1, pp. 52–66.

Germany.⁵⁸ Competition authorities have developed different theories on how some players could hinder competition, most of which relate to the strong and growing market positions of Facebook and Google. The two firms' combined digital advertising revenues accounted for 58% of the total digital ad spend in the UK in 2017 (Google 41%, Facebook 17%).⁵⁹

The strong position and vertically integrated nature of these platforms has raised two types of concern.

1. A dominant position may allow a large vertically integrated platform to set higher prices or offer lower-quality services to advertisers, for example by artificially creating scarcity of possible ad placements.⁶⁰ Consumers would be affected if higher advertising costs were passed on to them. However, there may be limits to any potential anticompetitive behaviour. For example, large advertisers can insist on certain quality standards being met,⁶¹ and other advertising channels (such as TV) may be converging with online advertising in the medium term, creating more competition.⁶²
2. A dominant position may allow a vertically integrated platform to foreclose competitors by refusing competitors access to its systems—for example, by limiting interoperability with upstream or downstream 'partial' competitors, thereby hindering competition.⁶³ This could allow the platform to set higher prices, again potentially increasing final prices for consumers.

Platforms would have the ability to foreclose competitors if:

- a) competitors have to incur (prohibitively) high costs to obtain the same consumer data, or cannot update it with the same frequency. However, efforts by publishers to create similar datasets suggest that such concerns may be limited,⁶⁴ or
- b) there are 'captive' consumers who can be identified or reached only through certain ad platforms.

Both concerns could cause harm for consumers through higher prices for advertised goods or lower-quality advertising. Current empirical evidence on the validity of these concerns is limited, but these theories are likely to be tested in the context of the ongoing competition authority investigations.

5.2.3 The tension between competition and privacy in targeted advertising

Addressing privacy and competition in targeted advertising is likely to create tensions: competition leads to good consumer outcomes, but the act of

⁵⁸ Oxera (2018), '[The Google case: shop till you drop \(off the screen\)](#)', *Agenda*, March. European Commission (2016) '[Google Search \(Shopping\): Antitrust Procedure Council Regulation \(EC\) 1/2003, Case AT.3970](#)', 27 June. See Australian Competition & Consumer Commission (2018), '[Digital Platforms Inquiry](#)', 26 February; Autorité de la Concurrence (2018), '[Avis n° 18-A-03 du 6 mars 2018 portant sur l'exploitation des données dans le secteur de la publicité sur internet](#)', 6 March; Bundeskartellamt (2018), 'Bundeskartellamt launches sector inquiry into market conditions in online advertising sector', press release, 1 February.

⁵⁹ Statista (2017), 'Digital Advertising in the United Kingdom (UK)', Dossier, pp. 16 and 18.

⁶⁰ Monopolies Commission (2015), '[Competition policy: The challenge of digital markets](#)', Special Report No 68, pp. 46–47.

⁶¹ For example, see *Wall Street Journal* (2018), '[P&G Contends Too Much Digital Ad Spending Is a Waste](#)', 1 March. *The Drum* (2017), '[P&G to review all agency contracts in 2017 in four-step plan to bring transparency to media supply chain](#)', 30 January.

⁶² Monopolies Commission (2015), op. cit., pp. 42–44.

⁶³ Ibid., pp. 45–46.

⁶⁴ *Martech Today* (2017), '[Major publishers band together against the "walled gardens" in Sonobi's cookie-less advertising marketplace](#)', 21 June.

increasing competition may reduce privacy. There are two dynamics in online markets where this tension is displayed.

Dynamic 1: *greater competition between ad platforms can lead to greater privacy (and other positive consumer outcomes)*

As we saw in section 4, competition in the messaging app market has led to some apps competing on the basis of greater privacy. We see similar dynamics in device markets, where Apple markets itself on providing greater privacy than its competitors.⁶⁵

Advertising platforms of all sizes may offer low levels of privacy in terms of transparency and control to consumers.⁶⁶ However, a dominant position may allow an ad platform to impose privacy terms on consumers that would not be acceptable if there were greater competition.⁶⁷ In such a case, taking measures to increase competition between platforms could improve the privacy offering available to consumers.

However, a second dynamic must also be considered before measures are taken to increase competition in ad markets.

Dynamic 2: *some mechanisms to encourage more competition between ad platforms are not conducive to greater privacy*

Regulators have a variety of tools available to them in encouraging greater competition in online advertising markets. However, some of these tools may be counterproductive if the objective is greater privacy.

For example, regulators could reduce the cost of data acquisition by encouraging (or mandating) greater data sharing between advertisers. However, greater data sharing arguably reduces the level of privacy.

Data sharing also has an ambiguous effect on market dynamics more broadly. Data sharing between advertisers is likely to make advertisers better off, and ad platforms often benefit from giving advertisers more information on individuals when targeting them in ad auctions.⁶⁸ However, more extensive data sharing between ad platforms and advertisers might also raise prices under specific circumstances.⁶⁹

Therefore policymakers should be careful that any intervention in ad markets does not produce unintended consequences that could harm privacy.

5.2.4 Implications

Advertisers and ad platforms have an incentive to use consumer data in order to match their campaigns closely to individual consumer interests, thereby driving competition on ad technology. They also have some incentive to ensure that consumers do not perceive these ads as too intrusive, but low levels of transparency and control can lead to less privacy than would be optimal for consumers.

⁶⁵ Apple (2018), '[Privacy](#)'.

⁶⁶ For example, see Kennedy, J. (2017), '[The Myth of Data Monopoly: Why Antitrust Concerns About Data Are Overblown](#)', Information Technology and Innovation Foundation, 6 March, pp. 15–16.

⁶⁷ Bundeskartellamt (2016), 'Bundeskartellamt initiates proceeding against Facebook on suspicion of having abused its market power by infringing data protection rules', press release, 2 March.

⁶⁸ Google Inc (2018), '[Value of Sharing Data](#)'. Hummel, P. and McAfee, R.P. (2016), 'When does improved targeting increase revenue?', *ACM Transactions on Economics and Computation (TEAC)*, 5:1, p. 4.

⁶⁹ de Cornière, A. and de Nijs, R. (2016), 'Online advertising and privacy', *The RAND Journal of Economics*, 47, pp. 48–72.

Consumers have some limited tools for making the trade-off between maintaining a high level of privacy and encouraging firms to compete by sharing their data widely. For example, people can opt out of being tracked by data aggregators, or by using privacy-enhancing tools such as specific web browsers. It is unclear whether these tools can help consumers to influence data use in advertising markets more widely—and there is still a role for policymakers in striking a balance between competition and privacy.

A clear understanding of consumer preferences is important to ensure good outcomes from the use of data in advertising. A challenge is the variety in preferences, not only across consumers but also across contexts (see *Which?/BritainThinks 2018*). One way of achieving this understanding would be by making it easier for consumers to choose their preferred privacy settings. Such choices should be presented in easily interpretable ways, as consumers may find it difficult to engage with complex settings about multiple platforms on multiple devices (as we discussed in section 3).

6 Conclusions

Firms have access to much more data about us than they ever have had. Such access to consumer data has raised concerns, including about privacy. But it has also led to positive changes in many markets and sectors across the economy. It has provided consumers with new products and services, and made existing products and services better and cheaper.

Many of these business models rely on data to facilitate transactions and to generate revenues through targeted advertising, in a way that was not previously possible. This has enabled innovation and delivered benefits to consumers in the form of greater choice or lower prices.

These innovations have, however, also led to risks to privacy. In certain circumstances, competition in the market can mitigate concerns about privacy, crucially depending on whether consumers are able to understand the privacy implications of using a particular service and can exercise choice.

Messaging apps are an example where competition can bring about good outcomes. The nature of the data has led to consumers exercising choice between messaging apps, and service providers competing on privacy-related features. By contrast, with health and fitness apps, behavioural biases and a lack of engagement with terms and conditions can lead to privacy concerns, even where there appears to be competition in the market.

www.oxera.com