DIGITAL DOMINANCE

The Power of Google, Amazon, Facebook, and Apple

Edited by Martin Moore

   and

Damian Tambini

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

The Evolution of Digital Dominance

How and Why We Got to GAFA

PATRICK BARWISE AND LEO WATKINS

Competition is for losers. If you want to create and capture lasting value, look to build a monopoly

—Peter Thiel, cofounder of PayPal and Palantir

Apple, Alphabet (Google), Microsoft, Amazon, and Facebook are now the five most valuable public companies in the world by market capitalization.[1] This is the first time ever that technology ("tech") companies have so dominated the stock market—even more than at the end of the 1990s' Internet bubble. They are a large part of everyday life in developed economies and increasingly elsewhere. They wield enormous power, raising difficult questions about their governance, regulation, and accountability. This chapter is about how and why this came about.

These tech giants vary in many ways. For instance, Apple is primarily a hardware company and Amazon has a huge physical distribution network, while Google, Microsoft, and Facebook are mainly "weightless" online businesses. Nevertheless, they share several features:

---

1. As of June 28, 2017 (see Table 1.1). A public company's market capitalization is its value to its shareholders (share price times number of shares).

- A US West Coast base;
- Dominant founders: Steve Jobs (Apple), Larry Page and Sergey Brin (Google), Bill Gates (Microsoft), Jeff Bezos (Amazon), and Mark Zuckerberg (Facebook) (Lex 2017);
- Significant control of the digital markets on which consumers and other companies depend;
- A business model to "monetize" this market power by charging users and/or others, such as advertisers, leading to sustained supernormal profits and/or growth;
- A hard-driving, innovative corporate culture epitomized by Facebook's former motto "Move fast and break things."

They have combined annual revenue of over $500bn, net income of over $90bn, and market capitalization of over $2.8 trillion (Table 1.1). Microsoft has been one of the world's most valuable companies since the 1990s, but the other four—"GAFA" (Google, Apple, Facebook, Amazon)—are relative newcomers to the list.

A 60-YEAR PATTERN: DOMINANT TECH PLAYERS CAN BE ECLIPSED, BUT NOT DISPLACED

This is the latest stage of a 60-year pattern, with the emergence of increasingly important new technology markets. These typically start as highly contested but soon become dominated by one (or two) US companies:

- 1960s mainframes (IBM)
- 1980s PCs (Microsoft and Intel)
- 1990s the World Wide Web, creating multiple new online markets including search (Google), e-commerce (Amazon), and social networking (Facebook)
- 2010s the mobile Internet (Apple and Google/Android) plus numerous mobile apps/services (GAFA and others, mostly based in the United States and China).

These companies operate in markets with important winner-take-all features such as cost and revenue economies of scale, scope, and learning, and often high switching costs, locking users in. Their strategies typically include creating proprietary standards and platforms; gathering and exploiting vast quantities of user data; product bundling; building large-scale infrastructure, some of which is then rented to other companies; strategic acquisitions; branding and intellectual property (trademark and,

Table 1.1. THE BIG FIVE US TECHNOLOGY COMPANIES

	Founded	Based	Main Product	Revenue (2016)	Market Capitalization 28.6.17	After-Tax Profit (Net Income) (2016)	P/E Ratio
Apple	1976	Cupertino, CA	Hardware	$216bn	$749bn	$45.7bn	16
Alphabet (Google)	1998	Mountain View, CA	Search	$90bn	$656bn	$19.5bn	34
Microsoft	1975	Redmond, WA	PC Software	$85bn	$534bn	$16.8bn	32
Amazon	1994	Seattle, WA	E-commerce	$136bn	$467bn	$2.4bn	195
Facebook	2004	Menlo Park, CA	Social network	$28bn	$436bn	$10.2bn	43
Total	–	–	–	**$555bn**	**$2,843bn**	**$94.6bn**	**30**

Source: Company reports and Dogs of the Dow (2017).
P/E (Price/Earnings) ratio = share price/latest earnings per share = market capitalization/net income.

especially, patent) litigation; regulatory and tax arbitrage; and political lobbying.

The result is dominance of at least one product or service category—in some cases several—leading to sustained high profits, which are then invested in (1) protecting and enhancing the core business and (2) high-potential new markets, especially where the company can use the same technology, infrastructure, or brand.

Because of these markets' winner-take-all features, it is extremely hard to *displace* a dominant, well-managed tech business from its leadership of its core product market. Instead, the greater risk is that they will be *eclipsed* by another company dominating a new, eventually bigger, adjacent market with similar winner-take-all qualities. The new market may then over-shadow the previous one, without necessarily destroying it (Thompson 2014). For instance, IBM still dominates mainframes and Microsoft still dominates PC software, but these are both mature markets that have been surpassed by newer, larger ones for online, mobile, and cloud-based services.

To head off this threat and exploit the new opportunities, dominant tech companies invest heavily in high-potential, emerging product markets and technologies, both organically and through acquisitions. Current examples include the augmented and virtual reality (AR/VR) platforms being developed by Apple, Google, and Facebook; the race between Google, Apple, Uber, Tesla, and others to develop self-driving car technology; and the creation of connected, voice-activated home hubs such as Apple's HomePod, the Amazon Echo, and Google Home.

The rest of the chapter is in three sections: the theory, the five company stories (Microsoft and GAFA), and the question: will the market end the tech giants' digital dominance?

THE THEORY: WHY TECH MARKETS ARE WINNER-TAKE-ALL

Traditional economics goes some way toward explaining these companies' market dominance. In particular, most tech markets exhibit extreme economies of scale. Software and digital content have high fixed development costs but low-to-zero marginal (copying and online distribution) costs. Unit costs are therefore almost inversely proportional to sales volume, giving a big competitive advantage to the market leader.

Digital products are also (1) "nonrivalrous"—unlike, say, pizzas, cars, or haircuts, they can be used simultaneously by a limitless number of people—and (2) "experience goods"—users need to try them and learn

about them (from personal experience, experts, and peers) to judge their quality.[2] Their nonrivalrous nature often leads to business models based on advertising (free services, maximizing reach) and/or continuing customer relationships rather than one-off sales.

The fact that these products are "experience goods" (1) increases the value of strong, trusted brands to encourage trial and (2) creates switching costs for existing users, further benefiting the market leader. The tech giants have some of the most valuable brands in the world: leading marketing company WPP now ranks Google, Apple, Microsoft, Amazon, and Facebook, in that order, as its top five global brands, with a combined value of $892bn (Kantar Millwood Brown 2017).[3] These estimates are of the shareholder value of *consumer* brand equity. These companies also have significant *employee* brand equity, helping them attract the best technical, managerial, and commercial talent—another winner-take-all factor.

Crucially, however, digital markets also have two other important characteristics that further encourage market concentration:

1. Many digital services serve communication or linking functions, generating both direct (within-market) and indirect (cross-market) network effects. These also occur in other markets but are especially prevalent and important in digital markets.
2. Digital technology enables large-scale real-time collection and automated analysis of usage data, which can be exploited both tactically and strategically, especially through continuous product improvement and personalization. The result is a recursive relationship between adoption and usage, product/service quality, and further adoption and usage, further reinforcing the winner-take-all dynamic.

Tech companies' strategies aim to exploit these winner-take-all market characteristics as well as classic sources of competitive advantage: product quality and design; marketing and branding; brand extensions and bundling; and various forms of customer lock-in. Increasingly, the companies also operate in multiple product markets, often with products and services offered free or below cost as part of a wider strategy to protect and extend

---

2. Economic analysis of these features predates the Internet: the literature on nonrivalrous (and, in the first instance, nonexcludable) "public goods" like defense and free-to-air broadcasting goes back to the 1950s (Samuelson 1954; Coase 1959) and the pioneering paper on experience goods is Nelson (1970).

3. The other two main valuation companies, Interbrand (2016) and Brand Finance (2017), also value them all in their top ten apart from Interbrand's #15 ranking for Facebook in 2016.

their core market dominance and capture more data. Examples include Amazon's Kindle and Google's Maps and Gmail.

We now discuss these distinctive winner-take-all characteristics of digital markets in more detail under four headings: direct network effects; indirect network effects ("multisided markets"); big data and machine learning; and switching costs and lock-in.

Direct Network Effects

In 1974, Jeffrey Rohlf, an economist at Bell Laboratories, published a seminal paper "A Theory of Interdependent Demand for a Telecommunications Service." Bell Labs' then-owner AT&T was contemplating the possible launch of a video telephony service, and Rohlf was researching how this should be priced if it went ahead. His mathematical model was based on the key qualitative insight (Rohlf 1974, 16) that "[t]he utility that a subscriber derives from a communications service increases as others join the system," enabling each person to communicate with more others (although some adopters are more influential than others in driving network externalities, see Tucker [2008]). Economists call this effect a direct network externality (Katz and Shapiro 1985).[4] In the context of Rohlf's paper and this chapter, the relevant network effects are positive ("revenue economies of scale"), but they can be negative, as with congestion in transport and communication networks. There can also be both positive and negative "behavioral" direct network effects if other consumers' adoption of a product makes it either more, or less, acceptable, fashionable, or attractive.

Indirect Network Effects ("Multisided Markets")

Most tech companies are, at least to a degree, "platform" businesses, creating value by *matching customers with complementary needs*, such as software developers and users (Microsoft's MS-DOS and Apple's App Store), publishers and book buyers (Amazon), drivers and potential passengers (Uber), and, in many cases including Google and Facebook, advertisers and consumers.

These network effects are called "indirect" because—unlike with the direct, single-market, externalities discussed previously—the value to

---

4. "Externality" because it involves external third parties in addition to the individual firm and customer. We interchangeably use the less technical term "network effect."

participants in each market (e.g., diners) depends on the number of participants in the *other* market (e.g., restaurants), and vice versa. Once a platform dominates the relevant markets, these network effects become self-sustaining as users on each side help generate users on the other.

Most indirect network effects are, again, positive, although they too can be negative for behavioral reasons if some participants are antisocial or untrustworthy, for example, posting malicious reviews on TripAdvisor or fake news on Facebook, or overstating the size and quality of their homes (or, conversely, throwing a noisy, late-night party as a guest) on Airbnb. Platforms often incorporate governance processes to limit these behaviors (Parker, Van Alstyne, and Choudary 2016, Chapter 8).

The need to appeal to both buyers and sellers simultaneously has been known since the first organized markets. But there was no formal modeling of two-sided markets until the late 1990s, when Rochet and Tirole (2003) noted structural similarities between the business models of payment card businesses, telecommunication networks, and computer operating systems. All exhibited network effects under which the value of the service for one group (e.g., payment card users) depended on how many members of the other group (e.g., merchants) were in the system, and vice versa.[5]

More recent work uses the term "multisided"—rather than two-sided—markets because some platforms facilitate interaction between more than two types of participant. For instance, Facebook connects six distinct groups: friends as message senders, friends as message receivers, advertisers, app developers, and businesses as both message senders and receivers (Evans and Schmalensee 2016a, 110).

Digital devices with compatible software, such as Microsoft's Xbox video games player, exhibit indirect network effects because (1) each device's installed user base constitutes an addressable market for software developers and (2) the range and quality of software available for the device are key to its user appeal (Nair, Chintagunta, and Dubé 2004; Lee 2013). Similarly, automated online marketplaces such as Amazon, Airbnb, and Uber operate in multisided markets with indirect network effects.

All businesses that depend on indirect network effects face the "chicken-and-egg" challenge of achieving critical mass in both or all the key markets simultaneously. Until the business reaches this point, it will need to convince investors that early losses will be justified by its eventual dominance of a large and profitable multisided market. Most start-up tech businesses,

---

5. These effects were also modeled independently by Parker and Van Alstyne (2005), who had noticed that most successful 1990s Internet start-ups had a two-sided market strategy.

such as Twitter, Uber, Snapchat, and Pinterest, are heavily loss-making for years and the casualty rate is high.

Achieving critical mass is easier if the product or service offers immediate benefits independent of network effects. For instance, at its 2007 launch, the iPhone already offered 2G mobile (voice, texts, e-mail, and web browsing) and music, with a market-leading touch-screen interface, driving rapid adoption. The App Store (2008) then created a virtuous circle of further adoption and app development.

Hosting a large digital platform requires massive infrastructure—servers, data storage, machine learning, payment systems, and so forth. Most of these have marked economies of scale and scope, enabling the business to take on other markets and to rent out capacity to other firms, further increasing its efficiency and profitability. The preeminent example is Amazon—both its logistics arm and its cloud computing business Amazon Web Services (AWS). Google, too, sells cloud storage, machine learning, data analytics, and other digital services that have grown out of, or complement, its core search business, while Microsoft is building its cloud services business, Azure.

Big Data and Machine Learning

The Internet enables tech companies to collect extensive, granular, real-time usage data at low cost. The resulting "big" datasets are challenging for traditional software to process because of their size, complexity, and lack of structure, but new data analytics techniques, increasingly automated ("machine learning"), can use big data to drive relentless improvement in products, services, pricing, demand forecasting, and advertising targeting. For instance, Netflix constantly analyzes viewing and preference data to inform its content purchases and commissions and to automate its personalized recommendations to users.

The more detailed the data, the wider the range of transactions, the bigger the user sample, and the greater the company's cumulative analytics experience, the better: quantity drives quality. Data and machine learning therefore offer both cost and revenue economies of scale, scope, and learning, encouraging digital businesses to offer free or subsidized additional services, at least initially, to capture more data.

The business benefits of big data are both tactical (continuous improvement) and strategic. These are interlinked: over time, continuous improvement can give the dominant provider an almost unassailable strategic advantage in service quality, customization, message targeting, and cost

reduction. Subject to privacy regulations (currently being loosened in the United States, see Waters and Bond [2017]), the data can also be sold to other, complementary companies, enabling them to obtain similar benefits. Finally, data can be analyzed at a more aggregate level to provide strategic insight into market trends. An important example is AWS's and other cloud companies' access to aggregate data on their many start-up clients, giving early intelligence on which are doing well and might be a competitive threat and/or investment opportunity.

Big data and machine learning can powerfully reinforce network effects, increasing the dominant companies' returns to scale and helping to entrench incumbents and deter market entry. However, economic theory has not yet caught up with this. For instance, Evans and Schmalensee (2016a) do not mention big data, analytics, algorithms, or machine learning. Parker, Van Alstyne, and Choudary (2016, 217–20) do list leveraging data as one of the ways in which platforms compete, but their discussion of it is barely two pages long and gives no references, reflecting the lack of relevant economic research to-date. There has been some broadly related work. Chen, Chiang, and Storey (2012) edited a special issue of *MIS Quarterly* on the use of big data analytics in business intelligence, while George, Haas, and Pentland (2014) and Einav and Levin (2014) explore its potential in management and economics research, respectively. But overall, although data and machine learning are key drivers of the tech giants' market and civic power, existing economic theory provides an insufficient framework for making this power accountable and regulating it to sustain effective competition (Feijoo, Gomez-Barroso, and Aggarwal 2016; Kahn 2017).

Switching Costs and Lock-In

Finally, all these companies use multiple ways to lock users in by increasing the cost or effort of switching to a rival product or service. As already noted, it takes time and effort to learn how to use unfamiliar systems and software. The greater the amount of such learning ("brand-specific consumer human capital"), the greater is the switching cost (Klemperer 1987; Ratchford 2001; Huang 2016). Often, there are also incompatibility issues locking users into a particular company's ecosystem (Iansiti and Levien 2004) or "walled garden": for instance, apps bought on iOS cannot be carried over to an Android device. Similarly, users' personal data archives may not be portable to another platform.

Some services' utility also increases with use by allowing for customization by the user (e.g., creating playlists on iTunes or Spotify) and/or the

company (based on the individual's usage data) or enabling the user to accrue, over time, a reputation or status (e.g., Amazon marketplace ratings) or to accumulate content they do not want to lose (e.g., Facebook message histories), all of which reinforces lock-in.

Conclusion: Digital Markets Are Winner-Take-All and the Winners Are Hard to Dislodge

In this section, we have discussed several structural reasons why digital markets tend to be winner-take-all: economies of scale; important user and employee brands; direct and indirect network effects; big data and machine learning; and other factors that enable strategies based on switching costs and lock-in.

The tech giants' market dominance is strengthened by their corporate cultures. They are all ambitious, innovative, and constantly on the lookout for emerging threats and opportunities, exemplifying Grove's (1998) view that "only the paranoid survive." This makes them tough competitors. Finally, their tax avoidance further increases their net income and competitive advantage.

Given all these factors, once a tech platform dominates its markets, it is very hard to dislodge. For a rival to do so, it would need to offer a better user experience, or better value for money, in both or all the markets connected by the platform, in a way that the incumbent could not easily copy, and over a sufficient timescale to achieve market leadership. For example, Google dominates both user search and search advertising. To dislodge it—as several have tried to do—a rival would need to offer users better searches and/or a better overall experience than Google, or some other incentive to switch to it (since Google searches are free, it cannot be undercut on price), long enough to overcome their habitual "googling" for information. Only by attracting more high-value users than Google would the challenger then be able to overtake it in search advertising revenue, although it could perhaps accelerate this (at a cost) by offering advertisers lower prices to compensate for its lower reach until it overtook Google. The overall cost would be huge—tens of billions—and with a high risk of failure, given Google's alertness and incumbency advantages: search quality, superior user interface, brand/habitual usage, dominant reach and scale in search advertising, leadership in big data and machine learning, and deep pockets.

However, competitive platforms can coexist if: (1) users can "multihome," that is, engage with more than one platform (for instance, many consumers use several complementary social networks) and/or

(2) developers can create versions of their products for several platforms at little incremental cost.

Having discussed the drivers of tech market concentration in generic and theoretical terms, we now turn to the five company stories and the extent to which some combination of these factors has, in practice, enabled each of them to achieve market dominance.

THE FIVE COMPANY STORIES

We here summarize the five companies' individual histories, strategies, business models, and current market positions and concerns. Their stories have been much more fully documented elsewhere, for example, Wallace and Erickson (1992), Isaacson (2011), Auletta (2009), Kirkpatrick (2010), and Stone (2013).

Microsoft

Microsoft was founded by Bill Gates (19) and Paul Allen (22) in 1975 as a supplier of microcomputer programming language interpreters.[6] Its big break came in 1980, when IBM gave it a contract to supply an operating system for the forthcoming IBM PC. Microsoft bought the software for $75,000 from another company, hired the programmer who wrote it, branded it MS-DOS, and licensed it to IBM and all the PC clone manufacturers, receiving a licence fee on every sale. It then acquired and developed a series of PC software products: Word (1983), Excel (1985), Windows—MS-DOS with a graphical user interface emulating that of the Apple Mac (1985), PowerPoint (1987), and Office—combining Word, Excel, PowerPoint, and other applications (1989). In 1995, Windows 95, a major upgrade using faster Intel processors, was bundled with Internet Explorer, which soon eclipsed Netscape as the dominant web browser.

Users familiar with both the Apple Mac and the Windows/Intel PC generally preferred the Mac. But the PC, widely marketed by IBM and multiple clone manufacturers, outsold the Mac and soon became the standard, first in the corporate world and then across the whole market apart from niche segments such as desktop publishing, where the Mac's superiority won out. Every PC came with MS-DOS and, later, Windows and

---

6. Allen left in 1983 after being diagnosed with Hodgkin's lymphoma.

Office, making Microsoft the dominant PC software supplier. Shapiro and Varian (1999, 10–11) described the Microsoft-Intel approach as a classic strategy based on network effects, contrasting it with Apple's strategy of controlling and integrating both the hardware and the software: "In the long run, the 'Wintel' strategy of strategic alliance was the better choice." Today, Microsoft remains the dominant PC software supplier with a global market share of 89%, versus 8% for Apple's OS X and 3% for all others (Netmarketshare 2018).

However, Microsoft has struggled to replicate this success elsewhere. Efforts under Steve Ballmer (CEO 2000–2014) to extend Windows to mobile devices repeatedly foundered, especially after the launch of Apple's iPhone and iOS (2007) and Google's Android mobile operating system (2008). Microsoft tried again to create a Windows mobile ecosystem based around Nokia's handset division, acquired for $7.9bn in 2013, but this too failed. Only 15 months later, under new CEO Satya Nadella, it took a $7.5bn impairment charge on the acquisition plus $2.5bn in restructuring costs.[7] Ballmer's resignation caused Microsoft's stock price to jump over 7% (Reisinger 2013).

Since the 2008 launch of Google Chrome, Microsoft has also lost share in the web browser market, despite bundling Internet Explorer with Windows since 1995. In search, its estimated cumulative losses were $11bn by 2013 (Reed 2013). However, its Bing search engine finally turned a profit in 2015 (Bright 2015), mainly as the default for Windows 10, iOS, Yahoo!, and AOL.

Historically, Microsoft's most successful move away from PC software was into video game consoles. This was initially a defensive move prompted by fears that Sony's PlayStation 2 would lure games players and developers away from the PC, but Microsoft's Xbox, launched in 2001, succeeded in its own right. Since 2012, Microsoft has also marketed PCs, laptops, and other devices under the Surface brand name, with some success.

Microsoft's challenge today is that the PC is no longer most users' main device—and Apple Macs and Google Chromebooks are also eating into its installed PC base. In response, it has set about transforming itself into a major player in cloud computing and office productivity services. It bought Skype in 2011 for $8.5bn, giving it a communications tool to integrate with other products like Office 365, the Lync enterprise phone platform, and real-time translation software (Bias 2015; Tun 2015). With this

---

7. Microsoft does, however, receive an estimated $2bn a year in patent royalties from Android device manufacturers (Yarow 2013), the only positive legacy of its expensive 15-year effort to build a significant mobile business.

combination (Skype for Business), it aims both to shore up its core PC software business and to create new office service opportunities, especially in the enterprise market.

Its biggest gamble to-date is the $26.2bn acquisition of the loss-making professional networking site LinkedIn in June 2016. Nadella claimed that the main aim was to exploit the data on LinkedIn's 433m users to "reinvent business processes and productivity" (Waters 2016). More prosaically, salespeople using Microsoft software could download LinkedIn data on potential leads to learn about their backgrounds, interests, and networks. Another aim may be to improve Microsoft's reputation and network in Silicon Valley (Hempel 2017).

Microsoft remains a powerful, highly profitable force and is undergoing rapid change under Satya Nadella. Nevertheless, since the millennium it has been increasingly overshadowed by the GAFA companies.

Apple

Apple began as a personal computer company, but, as discussed earlier, lost out to Microsoft and Intel in that market. Its subsequent success, making it the world's most valuable public company today, stems from its mobile devices and ecosystem, especially the iPod and iTunes (2001), iPhone and iOS (2007), App Store (2008), and iPad (2010).

The launch of the App Store created a classic two-sided market. Consumers bought iPhones because iOS had the best apps, and developers prioritized iOS because it offered the best addressable market: compared with users of other platforms, iOS users spent more on apps and the devices they owned were more uniform, reducing app development costs.[8] Underpinning all this was Apple's aesthetic and technical design edge, distinctive branding, and positioning as user-friendly rather than nerdish. The iPhone is also a personal device, not aimed at companies, as PCs were initially, increasing the scope for premium pricing.

Since 2010, Apple has sustained and extended its ecosystem by constantly adding new products (e.g., Siri and Watch) and features, driving repeated user upgrades to the latest device version. The breadth and quality of the user experience is also encouraging some PC users to switch to Macs. Finally, Apple's store network gives it a direct route to market, protects

8. Also, because iOS was based on the Mac operating system, Mac developers were able to write software for it with minimal retraining.

it from being squeezed by other retailers, boosts its brand exposure, and enables it to provide a superior, walk-in customer service.

Neither the iPod, nor the iPhone, nor the iPad was the first product in its category, but each met real consumer needs and delivered a much better user experience than the competition. Together with Apple's design edge and relentless incremental innovation (Barwise and Meehan 2011, 99–100), this has enabled the company to charge premium prices and turn its products into status symbols. Some, such as the Watch, have struggled to justify their premium prices, but the recent addition of contactless technology to the iPhone is encouraging retailers to adopt contactless payment terminals: Apple aims to use the scale of iPhone ownership to create an interactive environment for the Watch, justifying its high price, as the iPod and iTunes prepared the ground for the iPhone.

Apple is the world's most profitable public company and still dominates the premium end of the smartphone and tablet markets. However, as the rate of iPhone improvements slows and it runs out of new markets to conquer, it is increasingly turning toward its services to drive profits, including its commissions on app sales and in-app purchases in free-to-play games (Thompson 2017a). Meanwhile, it is constantly fighting the threat of hardware commoditization. The main company behind that threat is Google.

Google

Because the Internet is unimaginably vast, its value depends crucially on users' ability to find what they are looking for. In the early 1990s, the number of websites became too large for a simple index. By 1994, there were dozens of commercial search engines aiming to meet this growing need, using the relative "density" of the search terms (keywords) on different sites—a simple measure of relevance—to rank the results. They had a range of business models, all directly or indirectly based on display advertising.

Google began in 1996 as a research project by Stanford PhD students Larry Page and Sergey Brin. Page and Brin's key insight was that, from a user perspective, search results should be ranked by each site's importance as well as its relevance, reflected in the number and importance of other sites that linked to it. The resulting PageRank technology (named after Larry Page) was a big driver of their subsequent success, but far from the whole story. Page and Brin incorporated Google in 1998 with funding from angel investors including Amazon founder Jeff Bezos. In early 1999, Excite

turned down an offer to buy it for $750,000, but by June that year, it had attracted $25m in venture capital (VC) funding.

Its initial business model was based on sponsorship deals sold by sales reps on Madison Avenue. The breakthrough came in October 2000, when it started selling search advertising using its AdWords system, with advertisers bidding for keywords in real time. This auction, combined with cookie-based personalization, still determines which adverts each user sees and their ranking on the page.[9]

From the launch of AdWords in 2000, Google was a textbook success based on network externalities—literally: that same year it hired as chief economist Hal Varian, who coauthored the key book, Shapiro and Varian (1999). It succeeded by meeting the needs of both markets better than the competition. Users received the most relevant and important search results quickly and at no cost, on an attractive, uncluttered page with no distracting pop-up or banner ads. The only advertisements were short, text-based, relevant, and clearly distinguished from the natural search results. Meanwhile, advertisers received an efficient, highly targeted way of reaching potential customers actively looking for information using specific keywords. They could pay per click or even per customer acquired, increasing accountability and reducing risk. Marketing investment rapidly shifted from other media like print classifieds, leading to dramatic revenue and profit growth. Page and Brin hired Eric Schmidt as CEO in 2001. Three years later, Google's initial public offering raised $1.67bn for about 7% of the company, giving it a market capitalization of over $23bn.

Big data and machine learning lie at the heart of Google's strategy. The more data it has about each user, the better it can understand the context and intention behind every search and serve relevant results and well-targeted advertising. Thanks to its expertise in artificial intelligence (AI) and natural language processing, users can now input direct questions rather than just search terms, and receive increasingly intelligent answers.

To support its core business, Google has developed many other free services such as Chrome, Android, and Gmail, with Google Accounts unifying each user's activity. The data generated by each service is used to enhance all of them and to improve advertising targeting, while the services also direct users to each other. Google further exploits its data by buying display advertising inventory from third party sites, adding its own data on

---

9. Google did not invent this approach. Overture (originally GoTo), another start-up, had successfully launched a version of real-time bidding for keywords in 1998 (Battelle 2006, 125).

those sites' visitors, and selling the integrated data—at a premium—to advertisers looking to reach those users. Through Google Cloud Platform (GCP), it also sells infrastructure capacity to other businesses.

Google's ability to create superior, free, widely accessible services creates a high barrier to market entry, as Microsoft and others have discovered. A rival has to run large initial losses and encourage users to switch to it despite its initial inferiority. Apple Maps is one recent attempt, only possible because Apple made it the default on iOS.

Google's video platform, YouTube, is a big business in its own right, with estimated annual revenue of $4bn. But it is still reckoned to be loss-making because of its high costs: uploading, indexing, and storing over 50 hours of new video every minute; supporting several billion video views each day; paying content partners; plus R&D, advertising sales, and so forth (Winkler 2015). YouTube is a long-term investment aimed at capturing viewing and revenue from both traditional broadcasters and online-only players such as Netflix. Meanwhile, it too generates valuable data.

Since 2000, Google's most important move has been the 2008 launch of Android, aimed at ensuring that neither iOS nor Windows Mobile became the dominant operating system in a world of billions of mobile devices. Google made Android open source and collaborated with technology and service companies to make it the main global standard, giving Google an even bigger lead in mobile search (a 95% share in May 2017) than in desktop search, where Microsoft (Bing), Baidu, and Yahoo each have shares of 5%–8%—still an order of magnitude less than Google's 78% (Netmarketshare 2017).

In 2015, Google reorganized as Alphabet, a holding company with the core business as its main subsidiary. Alphabet's triple-class share structure enables Page, Brin, and Schmidt to take a long-term view, ignoring investor pressure for short-term returns. Other Alphabet subsidiaries include Waymo (self-driving cars), Nest (home automation), DeepMind (AI), Verily and Calico (life sciences), Sidewalk (urban infrastructure), and two VC funds. Alphabet aims to maximize synergies between these businesses. For instance, DeepMind provides cutting-edge machine-learning capabilities across the group and is also made available to others through GCP (Google Cloud Platform) and Google Assistant. Recently, Google's core business has also sought to develop new revenue streams that reduce its dependence on search advertising, launching devices such as the Pixel smartphone and the voice-activated Google Home hub.

Overall, Google remains unassailable in search and is making big bets in a wide range of other, mostly new, product markets.

Facebook

Facebook began in 2003 as Thefacebook.com, undergraduate Mark Zuckerberg's online version of Harvard's printed "facebook" of student mugshots. It drew on ideas from other early social networking sites such as Friendster and Myspace but, unlike them, accepted only people who registered in their own names and with a Harvard.edu web address. It was soon rolled out to other US colleges, funded through online advertising and investment by Zuckerberg's friends and family.

In July 2005, NewsCorp bought Myspace, the early market leader with 21 million users, for $580m. Arguably, Myspace was already vulnerable because of its cluttered interface and other weaknesses, but NewsCorp then failed to invest in it and overloaded it with advertising, allowing Facebook to overtake it in unique global visitors in April 2008 (Albanesius 2009). Facebook kept growing, while Myspace went into decline: NewsCorp sold it for an estimated $35m in 2011.

Facebook has two key features as a social network. First, for someone to add a "friend," both sides must agree. Second, its default assumption is that content posted by users is visible to all their "friends" unless one or both parties opts out. By creating engaging content at little cost to the company, users themselves generate the audience, which Facebook then monetizes by inserting targeted advertising among the posts. This model is highly scalable because variable costs are relatively low—mainly just more data centers and servers. Users' interactions and other behavior on the platform also generate extensive data for service improvement and advertising targeting.

Facebook's success has created its own challenges, however. As users' networks expand, content from their close current friends can be swamped by posts from "friends" who mean less to them, creating a need for algorithms to match users with the content most likely to engage them and with the most relevant advertisements. Adding "friends" from different personal networks (such as school, work, and—notably—parents) can also lead to self-censorship, further reducing the consumer value. To manage this tension, Facebook now has ways for users to post to user-defined groups within their networks and is reducing its dependence on user-generated content (UGC) by increasing the flow of professionally generated content (PGC)—news articles, opinion pieces, videos. Facebook is an increasingly important channel for PGC, although many producers are in a tug-of-war with it: they want engagement on Facebook to lead users onto their sites; Facebook wants to keep them on Facebook.

Facebook's pitch to advertisers is based on its huge reach and usage, highly targeted display advertising, and measurable short-term responses. By filling out "profiles" and following things they find interesting, users generate key targeting information. Facebook also increasingly enables social and psychological targeting: identifying which users are most central and influential within their social networks and when they are most likely to be receptive to specific advertising messages. However, both Facebook and Google have been criticized by advertisers for their unreliable, unaudited audience measures and other problems (Barwise 2017).

In March 2016, 79% of online US adults were active Facebook users, well ahead of Instagram (32%), Pinterest (31%), LinkedIn (29%), and Twitter (24%) (Chaffey 2017). But Facebook's market leadership is less secure than Google's because, as already noted, users can be members of several social networks ("multihoming") and many younger users prefer newer sites such as Snapchat.

Other social media range from message platforms (e.g., Apple's iMessages, Facebook Messenger, and WhatsApp, acquired by Facebook for $19bn in 2014), to specialist professional (LinkedIn, now owned by Microsoft) and short message networks (Twitter), to social photo- and video-sharing platforms such as Flickr, Instagram (also acquired by Facebook, in 2012, for $1bn), Pinterest, and Snapchat (which Facebook also reportedly tried to buy, but was turned down). These alternatives all threaten to draw valuable users away from Facebook by offering slightly different services. For instance, Snapchat is designed for more private, intimate, and fun interactions: the audience is selected-in and the default is that messages auto-delete. Where Facebook is unable to buy out a promising rival, it usually tries to copy its features: recent examples are Instagram "Stories," Facebook "Messenger Day," and WhatsApp "status," all emulating Snapchat "Stories" with growing success.

Amazon

In 1994, Jeff Bezos quit his well-paid job as a 30-year-old high-flier at a Wall Street hedge fund to found Amazon. Bezos, who remains chairman, president, and CEO, chose the name Amazon because it sounded exotic and started with and A—an advantage if it appeared in an alphabetical list—but also because the Amazon is the world's biggest river in terms of water flow and he wanted his business to be the world's biggest online retailer, which, in revenue terms, it is.

His core strategy was—and is—to build a dominant market share and brand in the consumer markets most suited to e-commerce; squeeze suppliers' prices; and reinvest the profits in price cuts, marketing, customer retention, transaction handling, and physical and digital distribution. In line with this, Amazon has consistently prioritized long-term growth over short-term profit: the prospectus for its 1997 IPO specifically said that it would "incur substantial losses for the foreseeable future" (Seglin 1997).

Bezos started with books because they were a good fit with online retailing: a huge number of low-ticket, standardized, easy to distribute products with a preexisting inventory, enabling him to launch quickly and offer many more titles, and at much lower prices, than even the largest physical bookshop. Bookselling also generated data on affluent, educated shoppers (Packer 2014). Over time, more and more product categories have been added as Amazon has refined its seamless online shopping experience and increasingly efficient distribution system.

Amazon's customer loyalty scheme Prime, first launched in 2005 in the United States and currently reaching 64% of US households (Hyken 2017), is now central to its business model. For a fixed fee, currently $99/year or $10.99/month in the United States, it offers subscribers unlimited free one- or two-day delivery (depending on the area), Amazon Video, Prime Music, unlimited photo storage, and other services. Rapid delivery encourages users to switch purchases from other retailers. Both Prime and the digital devices it sells at or below cost (the Kindle, Kindle Fire, Fire TV, and Echo home assistant) are aimed at making Amazon consumers' default e-commerce option. Amazon also advertises on TV, Google, and Facebook, and on many smaller websites through its affiliate link program. It has also acquired consumer guide sites such as Goodreads and IMDb, in which it has embedded "buy from Amazon" links and from which it also collects user rating data.

All this reinforces its core business model: relentless retail sales growth leading to increasing economies of scale in R&D, procurement, machine learning, marketing, and logistics. It then uses its superior capabilities not only to acquire more retail business but also to rent out infrastructure to other businesses: marketplace sellers pay to use Prime to deliver their goods, and businesses of all types buy cloud-based computing from AWS. Amazon Web Services is the most profitable part of the company: in the three months to March 31, 2017, it had an operating income of $890m, 24% of its $3.66bn revenue (Amazon 2017). Amazon Web Services sells both to Amazon itself (it grew out of a 2005 restructuring of the company's backend technology) and, increasingly, to others, making it the leading

supplier in the fast-growing cloud services market, followed by Microsoft (Azure), Google, IBM, and Oracle (Columbus 2017).

Amazon has substantial and still-growing market power as both a buyer and a seller. As the range of products it sells expands, users are now going straight to it to search for them, bypassing Google and enabling it to sell search advertising. Although the volume of searches is relatively small, they have the potential to generate disproportionate advertising revenue as they increasingly replace Google's most valuable searches, where consumers are actively looking for products. Amazon has more first-party consumer purchase data than any rival, to improve targeting, and can link both search and display advertising (e.g., on Amazon Prime Video) to actual purchases. Although still a relatively small player in digital advertising, it may challenge Google and Facebook in the longer term (Hobbs 2017).

Closely linked to Amazon's strategy and business model is its ultracompetitive company culture. Bezos's annual letter to shareholders always includes a copy of his first such letter in 1997, which famously said, "This is Day One for the internet." The aim is to keep behaving as if every day were still Day One. Amazon's distribution centers are nonunionized and increasingly automated, and it is testing drones and self-driving vans to reduce delivery costs. Accusations of exploitative labor management in its warehouses find their corollary in office staff also constantly monitored and required to work under unrelenting pressure. Those who survive this "purposeful Darwinism" receive few perks but benefit from a financial package heavily weighted toward stock options (Kantor and Streitfeld 2015).

Amazon has also been accused of anticompetitive activities including price discrimination and delisting competitors' products, such as Google Chromecast and Apple TV in 2015 and Google Home in 2016. Khan (2017, this volume) gives several examples of Amazon allegedly exploiting its market power in anticompetitive ways: predatory pricing of best-selling e-books; using its buying power and Fulfillment-by-Amazon (FBA) and its extensive data to create unfair advantage over retail competitors.

Amazon's dominance of consumer e-commerce outside China looks unstoppable. Its leadership in cloud-based computing, through AWS, seems almost as secure. As already noted, AWS's inside view of its clients' businesses gives it a strategic competitive advantage, especially in deciding which tech start-ups represent significant threats or investment opportunities. With the easiest product categories already covered, core revenue growth has slowed and the remaining categories are by definition harder, but Amazon is betting on game-changing innovations like drone delivery to reduce distribution barriers for these categories.

Amazon in 2017 announced a $13.7bn takeover bid for the upmarket US grocer Whole Foods. This was its largest ever acquisition. Analysts disagree about the strategy behind this move and its chances of success, but it clearly represents a move toward integrated "omnichannel" retailing combining on- and offline channels and covering even more product and service categories including perishable groceries—an extremely challenging category. The shares of US store groups fell sharply on the announcement.

WILL THE MARKET END THE TECH GIANTS' DIGITAL DOMINANCE?

In the first section of this chapter, we discussed a range of generic factors that make the tech giants' markets winner-take-all:

- Economies of scale;
- Strong user brands and habitual usage;
- Attractiveness to talent ("employee brand equity");
- Direct (within-market) network effects;
- Indirect (cross-market) network effects;
- Big data and machine learning;
- Switching costs and lock-in;
- Corporate strategies and cultures.

In the next section, we showed how each company has indeed come to dominate its market(s) in ways that reflect these winner-take-all factors.

Evans and Schmalensee (2016b) partly dispute this view. They argue that "winner-takes-all thinking does not apply to the platform economy," at least for Google and Facebook, on the grounds that—although they dominate consumer search and social networking, respectively—in the advertising market they have to compete with each other and with other media. We disagree. Google and Facebook do, of course, have to compete for advertising. But advertising media are not homogeneous: advertisers use different channels for different purposes. Google completely dominates search advertising and Facebook has a dominant, and still growing, share of online, especially mobile, display advertising. Because marketing budgets are finite, they do compete indirectly against each other and against other advertising media—and other ways of spending marketing money (promotions, loyalty schemes, etc.)—just as all consumer products and services indirectly compete for consumers' expenditure. But advertisers have no credible substitutes of comparable scale and reach as Google in search and

Facebook in online display advertising. The fact that they continue to use them despite the numerous problems that have been highlighted (fraud, audience measurement, etc.) reflects this lack of choice. Leading marketing commentator Mark Ritson (2016) described the emergence of the "digital duopoly" as the single biggest UK marketing issue in 2016—adding that he expected it to become even worse in 2017.

It is hard to see another company any time soon overtaking Google in search, Microsoft in PC software, or Amazon in e-commerce and cloud computing. Facebook's lead in social networking looks almost as strong, despite the potential for users to "multihome" and its recent problems with audience measurement and so forth. This bullish view is reflected in these companies' high Price/Earnings (P/E) ratios in Table 1.1, showing that the financial markets expect their earnings not only to withstand competitive pressures but also to continue growing faster than the market average for the foreseeable future. Some of this expected future growth presumably relates to the perceived long-term potential of their noncore activities, perhaps especially in the case of Alphabet, but it is hard to see how P/E ratios of 30-plus could be justified if their core businesses were seen as being under significant competitive threat.[10]

Apple's lower P/E of 16 reflects its lower expected future growth rate as Samsung and other Android manufacturers gradually catch up with the quality and ease of use of its devices and ecosystem, boosted by the growing superiority of Google services such as Assistant, reflecting the high penetration of Android and Google's lead in AI (Thompson 2017a). As Apple is increasingly forced to include Google's services in its ecosystem, its price premium over Android devices—the big driver of its high margins—is likely to be eroded.

Of course, whether—and if so, how soon—this happens will depend on Apple's continuing ability to come up with new, better products, content, and services to reinforce its dominance of the market for premium-priced mobile devices. In the wider mass market for mobile devices, Android is already the global standard, accounting for 82% of new smartphones shipped in 4Q16, versus 18% for iOS (Vincent 2017). On the plus side, Apple has an outstanding track record in product quality, ease of use, design, and branding. As the number of different types of device continues to proliferate—PCs (where Apple's share is growing); mobile, wearable, and smart home devices; virtual and augmented reality (VR/AR); automotive,

10. Amazon's P/E of 195 also reflects its strategy of reinvesting most of its profit to achieve additional long-term growth. This leads to a double whammy: artificially low short-term profits and high long-term growth expectations.

and so forth—Apple may be able to keep exploiting its ability to integrate devices and services into a superior, seamless user experience at a premium price.

In contrast, Google, Microsoft, and Amazon, like IBM before them, all fit the long-term pattern that dominant tech players are rarely displaced as market leaders in their core markets, because the winner-take-all dynamics are so powerful. Facebook's position is almost as secure. Only Apple is in significant danger of seeing its margins squeezed by a gradual process of commoditization.

Competition beyond the Tech Giants' Core Markets

For all five companies, the question remains whether, in line with the pattern discussed in the introduction, they will be *eclipsed* (as opposed to *displaced*) by a rival—either another large established player or a start-up—becoming the dominant provider of a new, important product or service that overshadows them. Microsoft has already been surpassed by Apple and Google in terms of profit and market capitalization (Table 1.1), and all five companies are acutely aware of the potential threats—and opportunities—presented by new product markets and technologies.

Major product markets currently of interest—in addition to Amazon's recent move to transform grocery retailing through its Whole Foods acquisition—are transport, home automation, entertainment, healthcare, business, and professional processes, and a wide range of applications under the broad heading the "Internet of things" (IoT) that will generate even more data—and further increase society's vulnerability to cyber attack. Key technologies include AI, voice and visual image recognition, VR/AR, cloud-based services, payment systems, and cyber security. All the tech giants are investing in several of these, both organically and through acquisition. Their access to vast amounts of user data makes them well placed to spot trends early, and their scale and profitability give them plenty of capacity to invest in and acquire new businesses and technologies.

The only national market of comparable scale to the United States is China. Chinese retail e-commerce is booming, with an estimated value already more than double that in North America: $899bn versus $423bn in 2016 (eMarketer 2016). Chinese tech companies operate under tight government controls and a constant threat of having their activities curbed, but benefit from protection from foreign competition and a somewhat cavalier view of privacy, data security, corporate governance, and intellectual property (not unlike the United States in the 19th century), although

Intellectual Property protection may improve as they build up their own patent portfolios and brands. China's "big four" tech companies are Tencent (mobile messaging and other content and services), Alibaba (e-commerce, digital entertainment, and cloud), Baidu (search and AI), and Huawei (mobile devices). Reflecting broader differences in business culture, Chinese tech companies tend to be less focused than those based in the United States, but the two are starting to converge as the top US tech groups diversify beyond their core businesses (Waters 2017).

We can expect to see more Chinese tech successes over the next 10 years, increasingly based on innovation as well as imitation and with growing international sales, in competition with the US players. However, their current activities are still largely focused on Greater China and there is no realistic prospect of their offering a serious challenge to the United States elsewhere in the next few years.

If anyone does overtake one of these companies in the next few years, it is more likely to be also based in Silicon Valley or Seattle. In *The Death of Distance* (1997), *The Economist*'s Frances Cairncross predicted a sharp reduction in the economic importance of geography. This has not happened. In addition to the top five companies by market capitalization discussed here, three of the other nine tech firms in the global 100 most valuable public companies—Oracle, Intel, and Cisco—are also based in Silicon Valley.[11] Beyond the United States, there are just four Asian companies and one European one on the list.[12] So, including the top five, eight of the world's top 14 public tech companies are based in or near Silicon Valley. No other *country* has more than one (although other Chinese tech giants will doubtless soon join the list).

Silicon Valley is also the leading cluster for tech start-ups. Of the top 50 global tech "unicorns"—companies founded after 2000 with a valuation over $1bn—at the time of writing, 21 are US-based. Sixteen of these are in Silicon Valley, including Uber, Airbnb, and Palantir (big data analytics) ranked 1, 4, and 5, respectively (CB Insights 2017). The other five are scattered around the United States: even America has only one Silicon Valley.[13]

In conclusion, with the partial exception of Apple, the tech giants seem unlikely to lose their dominance of their core market(s) any time soon,

11. The only other US company on the list is New York–based IBM.
12. Tencent (China), Samsung (Korea), Taiwan Semiconductor, Broadcom (Singapore), and SAP (Germany).
13. For the various reasons for Silicon Valley's dominance, see Hafner and Lyon (1998), Mazzucato (2015), Porter (1998), Bell (2005), Garmaise (2011), and Ben-Hahar (2016).

although they all, to varying degrees, face competitive threats at the margin. They are at greater risk of being overtaken by another company building a dominant share of a new, bigger, market. If and when that happens, the successful rival—either another tech giant or a start-up—is also likely to be based in Silicon Valley.

Do We Have a Problem?

How concerned should we be that market competition is unlikely to end Google, Microsoft, Facebook, and Amazon's dominance of their core markets in the foreseeable future? That market dominance brings many benefits to consumers and other businesses. Current competition regulation is designed to prevent firms from using their market power to charge higher prices, or offer lower quality, than would prevail in a competitive market. It is unsuited to a platform context where, in Google's case, consumers pay nothing and advertisers have a highly effective tool that did not exist 20 years ago and for which they pay a competitive, auction-based market price. Of course, incumbent industries disrupted by tech-based platforms (hotels by Airbnb, taxis by Uber, etc.) complain and highlight their real and imagined negative impacts. But much of this is just a normal part of disruptive innovation: the victims of creative destruction don't like it.

On this basis, there are good arguments for light-touch, perhaps technology-specific, regulation of platform businesses (Laffont and Tirole 2000) but not, in our view, for no regulation at all. Parker, Van Alstyne, and Choudary (2016, 239–53) list a wide range of reasons why we need "Regulation 2.0" for these markets: concerns about platform access, fair pricing, data privacy and security, national control of information assets, tax, labor regulation, and potential manipulation of consumers and markets. Similarly, Khan (2017, this volume) argues for more sophisticated regulation to address a range of anticompetitive behaviors. To this list we might add concerns about cyber security, digital advertising (fraud, mismeasurement, etc.), the impact of fake news, the decline in professional journalism, and the contribution of social media to political polarization (Barwise 2017). Finally, recent research suggests that the inequality between firms in winner-take-all markets, including tech, is one of three big drivers of growing income inequality (the other two being outsourcing and IT/automation: Bloom 2017).

The responses to-date differ between Europe and the United States. European antitrust legislation focuses on ensuring fair competition (reflected in the Commission's recent €2.4bn fine on Google for

"systematically" giving prominent placement in searches to its own shopping service and demoting rival services), whereas US legislation focuses more narrowly on whether market dominance leads to demonstrable consumer harm (Khan 2017, this volume; Thompson 2017b). Because the dominant tech platforms are all US-based, this is likely to be an area of growing transatlantic conflict.

REFERENCES

Albanesius, Chloe. 2009. "More Americans Go to Facebook Than MySpace." *PCMag.com*, June 16.

Amazon. 2017. "Amazon.com Announces First Quarter Sales up 23% to $35.7 Billion." April 27.

Auletta, Ken. 2009. *Googled: The End of the World as We Know It*. London: Virgin Books.

Barwise, Patrick. 2017. "Disrupting the Digital Giants: Advertisers and Traditional Media Push Back." *LSE Media Policy Project* (blog). April 5.

Barwise, Patrick, and Seán Meehan. 2011. *Beyond the Familiar: Long-Term Growth through Customer Focus and Innovation*. San Francisco: Jossey-Bass.

Battelle, John. 2006. *The Search: How Google and Its Rivals Rewrote the Rules and Transformed Our Culture*. London: Nicholas Brealey.

Bell, Geoffrey G. 2005. "Clusters, Networks and Firm Innovativeness." *Strategic Management Journal* 26, no. 3 (March): 287–95.

Ben-Shar, Omri. 2016. "California Got It Right: Ban the Non-Compete Agreements." www.forbes.com, October 27.

Bias, William. 2015. "This Is Why Microsoft Bought Skype." *247wallst.com*, June 3.

Bloom, Nicholas. 2017. "Corporations in the Age of Inequality." *Harvard Business Review*, digital article, April 14. https://hbr.org/cover-story/2017/03/corporations-in-the-age-of-inequality.

Bright, Peter. 2015. "Bing Profitable, but Microsoft Revenue down 12 Percent as Shift to Cloud Continues." *Ars Technica*, October 23.

Brand Finance. 2017. *Brand Finance Global 500 2017*. London: Brand Finance: 18–27.

Cairncross, Frances. 1997. *The Death of Distance: How the Communications Revolution Will Change Our Lives and Our Work*. Boston MA: Harvard Business School Press.

CB Insights. 2017. "The Global Unicorn Club," Accessed May 12, 2017. https://www.cbinsights.com/research-unicorn-companies.

Chaffey, Dave. 2017. "Global Social Media Research Summary 2017." *Smartinsights.com*, April 27.

Chen, Hsinchun, Roger H. L. Chiang, and Veda C. Storey. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact." *MIS Quarterly* 36, no. 4 (December): 1165–88.

Coase, Ronald H. "The Federal Communications Commission." *Journal of Law and Economics* 2 (October): 1–40.

Columbus, Louis. 2017. "Roundup of Cloud Computing Forecasts, 2017." *Forbes*, April 29.

Dogs of the Dow. 2017. "Largest Companies by Market Cap Today." Accessed June 28, 2017. http://dogsofthedow.com/largest-companies-by-market-cap.htm,

Einav, Liran, and Jonathan Levin. 2014. "The Data Revolution and Economic Analysis." *Innovation Policy and the Economy* 14, no. 1: 1–24.

eMarketer. 2016. "Worldwide Retail Ecommerce Sales Will Reach $1.915 Trillion This Year." April 22.

Evans, David S., and Richard Schmalensee. 2016a. *Matchmakers: The New Economics of Multisided Markets*. Boston, MA: Harvard Business Review Press.

Evans, David S., and Richard Schmalensee. 2016b. "Why Winner-Take-All Thinking Doesn't Apply to the Platform Economy." *Harvard Business Review*, digital article, May 4. https://hbr.org/2016/05/why-winner-takes-all-thinking-doesnt-apply-to-silicon-valley.

Feijoo, Claudio, Jose-Luis Gomez-Barroso, and Shivom Aggarwal. 2016. "Economics of Big Data." In *Handbook on the Economics of the Internet*, edited by Johannes M. Bauer and Michael Latzer, chapter 25. Cheltenham, UK: Edward Elgar. http://www.e-elgar.com/shop/handbook-on-the-economics-of-the-internet.

Garmaise, Mark. J. 2011. "Ties That Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment." *Journal of Law, Economics and Organization* 27, no. 2: 376–425.

George, Gerard, Martine R. Haas, and Alex Pentland. 2014. "Big Data and Management." *Academy of Management Journal* 57, no. 2: 321–26.

Grove, Andy. 1998. *Only the Paranoid Survive*. London: Profile Books.

Hafner, Katie, and Matthew Lyon. 1998. *Where Wizards Stay Up Late: The Origins of the Internet*. New York: Simon and Schuster.

Hempel, Jessi. 2017. "Now We Know Why Microsoft Bought LinkedIn." Wired.com, March 14. https://www.wired.com/2017/03/now-we-know-why-microsoft-bought-linkedin/.

Hobbs, Thomas. 2017. "Can Amazon break the 'digital duopoly'?" www.marketingweek.com, June.

Huang, Yufeng. 2016. "Learning by Doing and Consumer Switching Costs." Rochester, NY: Simon Business School working paper.

Hyken, Shep. 2017. "Sixty-four Percent of US Households Have Amazon Prime." www.forbescom, June 17.

Iansiti, Marco, and Roy Levien. 2004. "Strategy as Ecology." *Harvard Business Review* 82, no. 3 (March): 68–78.

Interbrand. 2016. *Best Global Brands 2016 Rankings*. London: Interbrand.

Isaacson, Walter. 2011. *Steve Jobs*. New York: Simon and Schuster.

Kantar Millwood Brown. 2017. *Brandz Top 100 Most Valuable Global Brands 2017*. London: WPP, 30–33.

Kantor, Jodi, and David Streitfeld. 2015. "Inside Amazon: Wrestling Big Ideas in a Bruising Workplace." *New York Times,* August 15.

Katz, Michael L., and Carl Shapiro. 1985. "Network Externalities, Competition and Compatibility." *American Economic Review* 75, no. 3 (June): 424–40.

Khan, Lina M. 2017. "Amazon's Antitrust Paradox." *Yale Law Journal* 126, no. 3 (January): 710–805.

Kirkpatrick, David. 2010. *The Facebook Effect: The Real Inside Story of Mark Zuckerberg and the World's Fastest-Growing Company*. London: Virgin Books.

Klemperer, Paul. 1987. "Markets with Consumer Switching Costs." *Quarterly Journal of Economics* 102, no. 2 (May): 375–94.

Laffont, Jean-Jacques, and Jean Tirole. 2000. *Competition in Telecommunications*. Cambridge MA: MIT Press.

Lee, Robin S. 2013. "Vertical Integration and Exclusivity in Platform and Two-Sided Markets." *American Economic Review* 103, no. 7: 2960–3000.

Lex. 2017. "Tech Entrepreneurs: Great Man Theory." www.ft.com, June 15.

Mazzucato, Mariana. 2015. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*. Rev. ed. London: Anthem Press.

Nair, Harikesh, Pradeep Chintagunta, and Jean-Pierre Dubé. 2004. "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants." *Quantitative Marketing and Economics* 2, no. 1: 25–58.

Nelson, Philip. 1970. "Information and Consumer Behavior." *Journal of Political Economy* 78, no. 2 (March–April): 311–29.

Netmarketshare. 2017. "Desktop and Mobile/Tablet Search Engine Market Share, May 2017." *Netmarketshare.com*.

Netmarketshare. 2018. "Operating System Market Share." www.netmarketshare.com/operating-system-market-share, January 29, 2018.

Packer, George. 2014. "Cheap Words." *New York Times Magazine*, February 17 and 24. https://www.newyorker.com/magazine/2014/02/17/cheap-words

Parker, Geoffrey G., and Marshall W. Van Alstyne. 2005. "Two-Sided Network Effects: A Theory of Information Product Design." *Management Science* 51, no. 10: 1494–504.

Parker, Geoffrey G., Marshall W. Van Alstyne, and Sangeet P. Choudary. 2016. *Platform Revolution: How Networked Markets Are Transforming the Economy—And How to Make Them Work for You*. New York: W. W. Norton.

Porter, Michael E. 1998. "Clusters and the New Economics of Competition." *Harvard Business Review* 76, no. 6 (November–December): 77–90.

Ratchford, Brian T. 2001. "The Economics of Consumer Knowledge." *Journal of Consumer Research* 27, no. 4: 397–411.

Reed, Brad. 2013. "Microsoft Has Lost $11 Billion Trying to Compete with Google." *BGR.com*, July 9.

Reisinger, Don. 2013. "Microsoft Shares Surge on Ballmer Retirement News." www.cnet.com, August 23.

Ritson, Mark. 2016. "The Marketing Stories That Mattered This Year." *Marketing Week*, December 15.

Rohlf, Jeffrey. 1974. "A Theory of Interdependent Demand for a Communications Service." *Bell Journal of Economics and Management Science* 5, no. 1 (Spring): 16–37.

Rochet, Jean-Charles, and Jean Tirole. 2003. "Platform Competition in Two-Sided Markets." *Journal of the European Economic Association* 1, no. 4: 990–1029.

Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics* 36, no. 4 (November): 387–89.

Seglin, Jeffrey L. 1997. "Hot Strategy: 'Be Unprofitable for a Long Time.'" *Inc.com*, September 1.

Shapiro, Carl, and Hal R. Varian. 1999. *Information Rules: A Strategic Guide to the Network Economy*. Boston, MA: Harvard Business School Press.

Statcounter. 2017. "Desktop Operating System Market Share Worldwide, May 2017." Accessed June 7, 2017.

Stone, Brad. 2013. *The Everything Store: Jeff Bezos and the Age of Amazon*. New York: Back Bay Books/Little, Brown.

Thompson, Ben. 2014. "Peak Google." www.stratechery.com, October 22.

Thompson, Ben. 2017a. "Apple's Strengths and Weaknesses." www.stratechery.com, June 5.

Thompson, Ben. 2017b. "Ends, Means, and Antitrust." www.stratechery.com, June 28.

Tucker, Catherine. 2008. "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities." *Management Science* 54, no. 12: 2024–38.

Tun, Zaw Thiha. 2015. "How Skype Makes Money." www.investopedia.com, July 9.

Vincent, James. 2017. "99.6 Percent of New Smartphones Run Android or iOS." *The Verge*, February 16.

Wallace, James, and Jim Erickson. 1992. *Hard Drive: Bill Gates and the Making of the Microsoft Empire*. New York: Wiley.

Waters, Richard. 2016. "Microsoft Recruits Help in Strategy Shift." www.ft.com, June 13.

Waters, Richard. 2017. "Chinese and US Tech Models Are Starting to Converge." www.ft.com, June 29.

Waters, Richard, and Shannon Bond. 2017. "US Moves Step Closer to Overturning Broadband Privacy Regulations." www.ft.com, March 29.

Winkler, Rolfe. 2015. "YouTube: 1 Billion Viewers, No Profit." *Wall Street Journal*, February 25.

Yarow, Jay. 2013. "Microsoft Is Making an Astonishing $2 Billion per Year from Android Patent Royalties." *Business Insider,* November 6.

Amazon—An Infrastructure Service and Its Challenge to Current Antitrust Law

LINA M. KHAN

INTRODUCTION

In Amazon's early years, a running joke among Wall Street analysts was that CEO Jeff Bezos was building a house of cards. Entering its sixth year in 2000, the company had yet to crack a profit and was mounting millions of dollars in continuous losses, each quarter's larger than the last. Nevertheless, a segment of shareholders believed that by dumping money into advertising and steep discounts, Amazon was making a sound investment that would yield returns once e-commerce took off. Each quarter the company would report losses, and its stock price would rise. One news site captured the split sentiment by asking, "Amazon: Ponzi Scheme or Wal-Mart of the Web?" (Slate 2000).

Sixteen years on, nobody seriously doubts that Amazon is anything but the titan of 21st-century commerce. In 2015, it earned $107 billion in revenue (Enright 2016), and, as of 2013, it sold more than its next 12 online competitors combined (Banjo and Ziobro 2013). By some estimates, Amazon now captures 46% of online shopping, with its share growing faster than the sector as a whole (LaVecchia and Mitchell 2016). In addition to being a retailer, it is a marketing platform, a delivery and logistics network, a payment service, a credit lender, an auction house, a major book publisher, a producer of television and films, a fashion designer, a hardware manufacturer, and a leading provider of cloud server space and computing

power. Although Amazon has clocked staggering growth—reporting double-digit increases in net sales yearly—it reports meager profits, choosing to invest aggressively instead. The company listed consistent losses for the first seven years it was in business, with debts of $2 billion (CNN Money 2002). While it exits the red more regularly now,[1] negative returns are still common.

Despite the company's history of thin returns, investors have zealously backed it: Amazon's shares trade at over 900 times diluted earnings, making it the most expensive stock in the Standard & Poor's 500 (Krantz 2015). As two reporters marveled, "The company barely ekes out a profit, spends a fortune on expansion and free shipping and is famously opaque about its business operations. Yet investors . . . pour into the stock" (Clark and Young 2013). Another commented that Amazon is in "a class of its own when it comes to valuation" (Krantz 2015).

Reporters and financial analysts continue to speculate about when and how Amazon's deep investments and steep losses will pay off (see, e.g., Manjoo 2015). Customers, meanwhile, universally seem to love the company. Close to half of all online buyers go directly to Amazon first to search for products (Moore 2015), and in 2016, the Reputation Institute named the firm the "most reputable company in America" for the third year running (Strauss 2016; see also Hoffmann 2014). In recent years, journalists have exposed the aggressive business tactics Amazon employs. For instance Amazon named one campaign "The Gazelle Project," a strategy whereby Amazon would approach small publishers "the way a cheetah would a sickly gazelle" (Streitfeld 2013). This, as well as other reporting (Soper 2015; Wingfield and Somaiya 2015), drew widespread attention (Streitfeld 2013), perhaps because it offered a glimpse at the potential social costs of Amazon's dominance. The firm's highly public dispute with Hachette in 2014—in which Amazon delisted the publisher's books from its website during business negotiations—similarly generated extensive press scrutiny and dialogue (see Krugman 2014). More generally, there is growing public awareness that Amazon has established itself as an essential part of the Internet economy (see Manjoo 2016), and a gnawing sense that its dominance—its sheer scale and breadth—may pose hazards. But when pressed on why, critics often fumble to explain how a company that has so clearly delivered enormous benefits to consumers—not to mention revolutionized e-commerce in general—could, at the end of the day, threaten our

---

1. Partly due to the success of Amazon Web Services, Amazon has recently begun reporting consistent profits. See Wingfield (2016).

markets. Trying to make sense of the contradiction, one journalist noted that the critics' argument seems to be that "even though Amazon's activities tend to reduce book prices, which is considered good for consumers, they ultimately hurt consumers" (Vara 2015).

In some ways, the story of Amazon's sustained and growing dominance is also the story of changes in our antitrust laws. Due to the Chicago School's influence on legal thinking and practice in the 1970s and 1980s, antitrust law now assesses competition largely with an eye to the short-term interests of consumers, not producers or the health of the market as a whole; antitrust doctrine views low consumer prices, alone, to be evidence of sound competition. By this measure, Amazon has excelled; it has evaded government scrutiny in part through fervently devoting its business strategy and rhetoric to reducing prices for consumers. Amazon's closest encounter with antitrust authorities was when the Justice Department sued other companies for teaming up against Amazon.[2] It is as if Bezos charted the company's growth by first drawing a map of antitrust laws, and then devising routes to smoothly bypass them. With its missionary zeal for consumers, Amazon has marched toward monopoly by singing the tune of contemporary antitrust.

This chapter maps out facets of Amazon's power. In particular, it traces the sources of Amazon's growth and analyzes the potential effects of its dominance. Doing so enables us to make sense of the company's business strategy and illuminates anticompetitive aspects of its structure and conduct. This analysis reveals that the current framework in antitrust—specifically its equating competition with "consumer welfare," typically measured through short-term effects on price and output—fails to capture the architecture of market power in the 21st-century marketplace. In other words, the potential harms to competition posed by Amazon's dominance are not cognizable if we assess competition primarily through price and output. Focusing on these metrics instead blinds us to the potential hazards.

The chapter argues that gauging real competition in the 21st-century marketplace—especially in the case of online platforms—requires analyzing the underlying structure and dynamics of markets. Rather than pegging competition to a narrow set of outcomes, this approach would examine the competitive process itself. Animating this framework is the idea that a company's power and the potential anticompetitive nature of that power

2. See United States v. Apple Inc., 952 F. Supp. 2d 638, 650 (S.D.N.Y. 2013).

cannot be fully understood without looking to the structure of a business and the structural role it plays in markets. Applying this idea involves, for example, assessing whether a company's structure creates certain anticompetitive conflicts of interest; whether it can cross-leverage market advantages across distinct lines of business; and whether the structure of the market incentivizes and permits predatory conduct.

THE CHICAGO SCHOOL REVOLUTION

One of the most significant changes in antitrust law and interpretation over the last century has been the move away from economic structuralism. Broadly, economic structuralism rests on the idea that concentrated market structures promote anticompetitive forms of conduct (see, e.g., Bain 1950; 1968; Turner and Kaysen 1959). This market structure–based understanding of competition was a foundation of antitrust thought and policy through the 1960s. Subscribing to this view, courts blocked mergers that they determined would lead to anticompetitive market structures. In some instances, this meant halting horizontal deals—mergers combining two direct competitors operating in the same market or product line—that would have handed the new entity a large share of the market. In others, it involved rejecting vertical mergers—deals joining companies that operated in different tiers of the same supply or production chain—that would "foreclose competition." Centrally, this approach involved policing not just for size but also for conflicts of interest—like whether allowing a dominant shoe manufacturer to extend into shoe retailing would create an incentive for the manufacturer to disadvantage or discriminate against competing retailers.

The Chicago School approach to antitrust, which gained mainstream prominence and credibility in the 1970s and 1980s, rejected this structuralist view. In the words of Richard Posner, the essence of the Chicago School position is that "the proper lens for viewing antitrust problems is price theory" (Posner 1979, 932). Foundational to this view is a faith in the efficiency of markets, propelled by profit-maximizing actors. The Chicago School approach bases its vision of industrial organization on a simple theoretical premise: "[R]ational economic actors working within the confines of the market seek to maximize profits by combining inputs in the most efficient manner. A failure to act in this fashion will be punished by the competitive forces of the market."

Practically, the shift from structuralism to price theory had two major ramifications for antitrust analysis. First, it led to a significant narrowing of

the concept of entry barriers. According to the Chicago School, advantages that incumbents enjoy from economies of scale, capital requirements, and product differentiation do not constitute entry barriers, as these factors are considered to reflect no more than the "objective technical demands of production and distribution" (Eisner 1991, 105).

The second consequence of the shift away from structuralism was that consumer prices became the dominant metric for assessing competition. In his highly influential work, *The Antitrust Paradox*, Robert Bork asserted that the sole normative objective of antitrust should be to maximize consumer welfare, best pursued through promoting economic efficiency (Bork 1978; see also Crane 2014). In 1979, in *Reiter v. Sonotone Corp.*, the Supreme Court followed Bork's work and declared that "Congress designed the Sherman Act as a 'consumer welfare prescription'"—a statement that is widely viewed as erroneous (see Orbach 2013, 2152). Still, this philosophy wound its way into policy and doctrine (DOJ 1982). Today, showing antitrust injury requires showing harm to consumer welfare, generally in the form of price increases and output restrictions.[3]

Two areas of enforcement that this reorientation has affected dramatically are predatory pricing and vertical integration. The Chicago School claims, "predatory pricing, vertical integration, and tying arrangements never or almost never reduce consumer welfare" (Crane 2014). Both predatory pricing and vertical integration are highly relevant to analyzing Amazon's path to dominance and the source of its power.

Through the mid-20th century, Congress repeatedly enacted legislation targeting predatory pricing. Congress, as well as state legislatures, viewed predatory pricing as a tactic used by highly capitalized firms to bankrupt rivals and destroy competition—in other words, as a tool to concentrate control. Laws prohibiting predatory pricing were part of a larger arrangement of pricing laws that sought to distribute power and opportunity.

Starting in the 1960s, Chicago School scholars criticized predatory pricing law and enforcement as misguided. They argued that predatory pricing rarely occurred in practice and that, by targeting conduct that resulted in lower prices, government was undermining competition. As the influence of these scholars grew, their thinking shaped both government enforcement and Supreme Court doctrine. In a series of cases in the 1980s and 1990s, the Supreme Court declared, "predatory pricing schemes are rarely tried, and even more rarely successful."[4] Furthermore the Court introduced

3. See, e.g., Ginzburg v. Mem'l Healthcare Sys., Inc., 993 F. Supp. 998, 1015 (S.D. Tex. 1997); Angelico v. Lehigh Valley Hosp., Inc., 984 F. Supp. 308, 312 (E.D. Pa. 1997).
4. Matsushita Electric Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 589 (1986).

a legal test requiring that plaintiffs bringing these claims demonstrate that the alleged predatory conduct would lead to higher prices, sufficient to compensate predators for the amounts expended on the predation—a requirement now known as the "recoupment test."

In placing recoupment at the center of predatory pricing analysis, the Court presumed that direct profit maximization is the singular goal of predatory pricing.[5] Furthermore, by establishing that harm occurs *only* when predatory pricing results in higher prices, the Court collapsed the rich set of concerns that had animated earlier critics of predation, including an aversion to large firms that exploit their size and a desire to preserve local control. Instead, the Court adopted the Chicago School's narrow conception of what constitutes this harm (higher prices) and how this harm comes about—namely, through the alleged predator raising prices on the previously discounted good.

Today, succeeding on a predatory pricing claim requires a plaintiff to meet the recoupment test by showing that the defendant would be able to recoup its losses through sustaining supracompetitive prices. Since the Court introduced this recoupment requirement, the number of cases brought and won by plaintiffs has dropped dramatically (Sokol 2014, 1013).

Analysis of vertical integration has similarly moved away from structural concerns (Cole 1952). For most of the last century, enforcers reviewed vertical integration under the same standards as horizontal mergers. Critics of vertical integration primarily focused on two theories of potential harm: leverage and foreclosure. Leverage reflects the idea that a firm can use its dominance in one line of business to establish dominance in another. Foreclosure, meanwhile, occurs when a firm uses one line of business to disadvantage rivals in another line.

Chicago School theory holds that concerns about both leverage and foreclosure are misguided. Vertical mergers antitrust law's aversion to vertical arrangements was, Bork argued, irrational. "The law against vertical mergers is merely a law against the creation of efficiency" (Bork 1978, 234). With the election of President Reagan, this view of vertical integration became national policy. In 1982 and 1984, the Department of Justice (DOJ) and the FTC issued new merger guidelines outlining the framework that officials would use when reviewing horizontal deals. The new guidelines narrowed the circumstances in which the agencies would challenge vertical mergers (DOJ 1982; 1984). Although subsequent administrations

5. See ibid., 224.

have continued reviewing vertical mergers, the Chicago School's view that these deals generally do not pose threats to competition has remained dominant.

AMAZON'S BUSINESS STRATEGY

Amazon has established dominance as an online platform thanks to two elements of its business strategy: a willingness to sustain losses and invest aggressively at the expense of profits, and integration across multiple business lines.[6] These facets of its strategy are independently significant and closely interlinked—indeed, one way it has been able to expand into so many areas is through foregoing returns. This strategy—pursuing market share at the expense of short-term returns—defies the Chicago School's assumption of rational, profit-seeking market actors. More significantly, Amazon's choice to pursue heavy losses while also integrating across sectors suggests that in order to fully understand the company and the structural power it is amassing, we must view it as an integrated entity. Seeking to gauge the firm's market role by isolating a particular line of business and assessing prices in that segment fails to capture both (1) the true shape of the company's dominance and (2) the ways in which it is able to leverage advantages gained in one sector to boost its business in another.

Foregoing Short-Term Returns for Long-Term Dominance

Recently, Amazon has started reporting consistent profits, largely due to the success of Amazon Web Services, its cloud computing business (see Bensinger 2016). Its North America retail business runs on much thinner margins, and its international retail business still runs at a loss (Bensinger 2016). But for the vast majority of its 20 years in business, losses—not profits—were the norm. Through 2013, Amazon had generated a positive net income in just over half of its financial reporting quarters. Even in quarters in which it did enter the black, its margins were razor-thin, despite astounding growth. The graph below captures the general trend (Figure 4.1).

6. I am using "dominance" to connote that the company controls a significant share of market activity in a sector. I do not mean to attach the legal significance that sometimes attends "dominance."
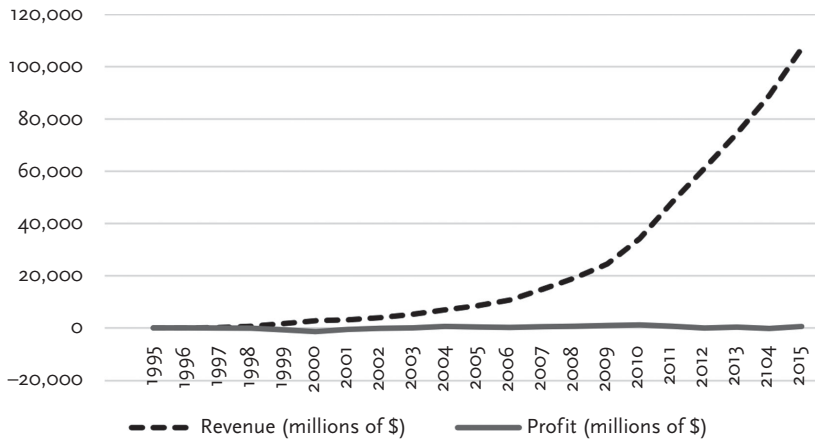
Figure 4.1: Amazon's Profits
Source: Based on Evans (2013) and Market Watch (date).

Just as striking as Amazon's lack of interest in generating profit has been investors' willingness to back the company (see Streitfeld 2013). With the exception of a few quarters in 2014, Amazon's shareholders have poured money in despite the company's penchant for losses. On a regular basis, Amazon would report losses, and its share price would soar (see, e.g., Dini 2000; Quick Pen 2015). As one analyst told the *New York Times*, "Amazon's stock price doesn't seem to be correlated to its actual experience in any way" (Streitfeld 2015; see also Elmer-DeWitt 2015).

Analysts and reporters have spilled substantial ink seeking to understand the phenomenon. As one commentator joked in a widely circulated post, "Amazon, as best I can tell, is a charitable organization being run by elements of the investment community for the benefit of consumers" (Yglesias 2013).

In some ways, the puzzlement is for naught: Amazon's trajectory reflects the business philosophy that Bezos outlined from the start. In his first letter to shareholders, Bezos wrote:

> We believe that a fundamental measure of our success will be the share-holder value we create over the *long term*. This value will be a direct result of our ability to extend and solidify our current market leadership position. . . . We first measure ourselves in terms of the metrics most indicative of our market leadership: customer and revenue growth, the degree to which our customers continue

to purchase from us on a repeat basis, and the strength of our brand. We have invested and will continue to invest aggressively to expand and leverage our customer base, brand, and infrastructure as we move to establish an enduring franchise. (Bezos 1998)

In other words, the premise of Amazon's business model was to establish scale. To achieve scale, the company prioritized growth. Under this approach, aggressive investing would be key, even if that involved slashing prices or spending billions on expanding capacity, in order to become consumers' one-stop-shop. This approach meant that Amazon "may make decisions and weigh trade-offs differently than some companies," Bezos warned (Bezos 1998). "At this stage, we choose to prioritize growth because we believe that scale is central to achieving the potential of our business model" (Bezos 1998).

The insistent emphasis on "market leadership" (Bezos relies on the term six times in the short letter) signaled that Amazon intended to dominate. And, by many measures, Amazon has succeeded. Its year-on-year revenue growth far outpaces that of other online retailers (Garcia 2016; see also BI Intelligence 2016). Despite efforts by big-box competitors like Walmart, Sears, and Macy's to boost their online operations, no rival has succeeded in winning back market share (see Wahba 2015).

One of the primary ways Amazon has built a huge edge is through Amazon Prime, the company's loyalty program, in which Amazon has invested aggressively. Initiated in 2005, Amazon Prime began by offering US consumers unlimited two-day shipping for $79 (Kawamoto 2005). In the years since, Amazon has bundled in other deals and perks, like renting e-books and streaming music and video, as well as one-hour or same-day delivery. The program has arguably been the retailer's single biggest driver of growth (DiChristopher 2015).[7] Amazon does not disclose the exact number of Prime subscribers in the United States, but analysts believe the number of users has surpassed 63 million—19 million more than in 2015 (Leonard 2016). Membership doubled between 2011 and 2013. By 2020, it is estimated that half of US households may be enrolled (Frommer 2015).

As with its other ventures, Amazon lost money on Prime to gain buy-in. In 2011 it was estimated that each Prime subscriber cost Amazon at least $90 a year—$55 in shipping, $35 in digital video—and that the company therefore took an $11 loss annually for each customer (Woo 2011). One Amazon expert tallies that Amazon has been losing $1 billion to $2 billion

---

7. It has also been a key force driving up Amazon's stock price (see Stone 2010).

a year on Prime memberships (Seetharaman and Layne 2015). The full cost of Amazon Prime is steeper yet, given that the company has been investing heavily in warehouses, delivery facilities, and trucks, as part of its plan to speed up delivery for Prime customers—expenditures that regularly push it into the red (see Weise 2015).

Despite these losses—or perhaps because of them—Prime is considered crucial to Amazon's growth as an online retailer. According to analysts, customers increase their purchases from Amazon by about 150% after they become Prime members (Stone 2010). Prime members constitute 47% of Amazon's US shoppers (Tuttle 2016). Amazon Prime members also spend more on the company's website—an average of $1,500 annually, compared to $625 spent annually by non-Prime members (Rubin 2016). Business experts note that by making shipping free, Prime "successfully strips out paying for . . . the leading consumer burden of online shopping" (Fox Rubin 2015). Moreover, the annual fee drives customers to increase their Amazon purchases in order to maximize the return on their investment (see Tuttle 2010).

As a result, Amazon Prime users are both more likely to buy on its platform and less likely to shop elsewhere. "[Sixty-three percent] of Amazon Prime members carry out a paid transaction on the site in the same visit," compared to 13% of non-Prime members (O'Connor 2015). For Walmart and Target, those figures are 5% and 2%, respectively (O'Connor 2015). One study found that less than 1% of Amazon Prime members are likely to consider competitor retail sites in the same shopping session. Non-Prime members, meanwhile, are eight times more likely than Prime members to shop between both Amazon and Target in the same session (O'Connor 2015). In the words of one former Amazon employee who worked on the Prime team, "It was never about the $79. It was really about changing people's mentality so they wouldn't shop anywhere else" (Stone 2010). In that regard, Amazon Prime seems to have proven successful (Tuttle 2010).

In 2014, Amazon hiked its Prime membership fee to $99 (Bensinger 2014). The move prompted some consumer ire, but 95% of Prime members surveyed said they would either definitely or probably renew their membership regardless (Whitney 2014), suggesting that Amazon has created significant buy-in and that no competitor is currently offering a comparably valuable service at a lower price. It may, however, also reveal the general stickiness of online shopping patterns. Although competition for online services may seem to be "just one click away," research drawing on behavioral tendencies shows that the "switching cost" of changing web services can, in fact, be quite high (see Candeub 2014).

No doubt, Amazon's dominance stems in part from its first-mover advantage as a pioneer of large-scale online commerce. But in several key ways, Amazon has achieved its position through deeply cutting prices and investing heavily in growing its operations—both at the expense of profits. The fact that Amazon has been willing to forego profits for growth undercuts a central premise of contemporary predatory pricing doctrine, which assumes that predation is irrational precisely because firms prioritize profits over growth. In this way, Amazon's strategy has enabled it to use predatory pricing tactics without triggering the scrutiny of predatory pricing laws.

Another key element of Amazon's strategy—and one partly enabled by its capacity to thrive despite posting losses—has been to expand aggressively into multiple business lines. In addition to being a retailer, Amazon is a marketing platform, a delivery and logistics network, a payment service, a credit lender, an auction house, a major book publisher, a producer of television and films, a fashion designer, a hardware manufacturer, and a leading provider of cloud server space and computing power (Amazon. com, Inc. 2016). For the most part, Amazon has expanded into these areas by acquiring existing firms.

Involvement in multiple, related business lines means that, in many instances, Amazon's rivals are also its customers. The retailers that compete with it to sell goods may also use its delivery services, for example, and the media companies that compete with it to produce or market content may also use its platform or cloud infrastructure. At a basic level this arrangement creates conflicts of interest, given that Amazon is positioned to favor its own products over those of its competitors.

Critically, not only has Amazon integrated across select lines of business but also it has emerged as central infrastructure for the Internet economy. Reports suggest this was part of Bezos's vision from the start. According to early Amazon employees, when the CEO founded the business, "his underlying goals were not to build an online bookstore or an online retailer, but rather a 'utility' that would become essential to commerce" (Mulpuru and Walker 2012). In other words, Bezos's target customer was not only end-consumers but also other businesses.

Amazon controls key critical infrastructure for the Internet economy—in ways that are difficult for new entrants to replicate or compete against. This gives the company a key advantage over its rivals: Amazon's competitors have come to depend on it. Like its willingness to sustain losses, this feature of Amazon's power largely confounds contemporary antitrust analysis, which assumes that rational firms seek to drive their rivals out of business. Amazon's game is more sophisticated. By making itself indispensable

to e-commerce, Amazon enjoys receiving business from its rivals, even as it competes with them. Moreover, Amazon gleans information from these competitors as a service provider that it may use to gain a further advantage over them as rivals—enabling it to further entrench its dominant position.

ESTABLISHING STRUCTURAL DOMINANCE

Amazon now controls 46% of all e-commerce in the United States (see LaVecchia and Mitchell 2016). Not only is it the fastest-growing major retailer, but it is also growing faster than e-commerce as a whole (Ray 2016; see generally Leonard 2016). In 2010, it employed 33,700 workers; by June 2016, it had 268,900 (Leonard 2016). It is enjoying rapid success even in sectors that it only recently entered. For example, the company "is expected to triple its share of the U.S. apparel market over the next five years" (Banjo 2016). Its clothing sales recently rose by $1.1 billion—even as online sales at the six largest US department stores fell by over $500 million.[8]

These figures alone are daunting, but they do not capture the full extent of Amazon's role and power. Amazon's willingness to sustain losses and invest aggressively at the expense of profits, coupled with its integration across sectors, has enabled it to establish a dominant structural role in the market.

In what follows, several examples of Amazon's conduct illustrate how the firm has established structural dominance. The first example focuses on predatory pricing. The other examples, Fulfillment-by-Amazon and Amazon Marketplace, demonstrate how Amazon has become an infrastructure company, both for physical delivery and e-commerce, and how this vertical integration implicates market competition. These cases highlight how Amazon can use its role as an infrastructure provider to benefit its other lines of business. These examples also demonstrate how high barriers to entry may make it difficult for potential competitors to enter these spheres, locking in Amazon's dominance for the foreseeable future. All four of these accounts raise concerns about contemporary antitrust's ability to register and address the anticompetitive threat posed by Amazon and other dominant online platforms.

8. Its clothing sales are greater than the combined online sales of its five largest online apparel competitors: Macy's, Nordstrom, Kohl's, Gap, and Victoria's Secret's parent (Banjo 2016).

Discriminatory Pricing and Fees

Under current doctrine, whether below-cost pricing is predatory or not turns on whether a firm recoups its losses. Thereby, the following paragraphs examine how Amazon could use its dominance to recoup its losses in ways that are more sophisticated than what current doctrine recognizes.

Most obviously, Amazon could earn back the losses it generated on bestseller e-books by raising prices of either particular lines of e-books or e-books as a whole. However, conducting recoupment analysis with Amazon is particularly challenging because it may not be apparent when and by how much Amazon raises prices.

Online commerce enables Amazon to obscure price hikes in at least two ways: rapid, constant price fluctuations and personalized pricing.[9] Constant price fluctuations diminish our ability to discern pricing trends. By one account, Amazon changes prices more than 2.5 million times each day (Ferdman 2013). Amazon is also able to tailor prices to individual consumers, known as first-degree price discrimination. There is no public evidence that Amazon is currently engaging in personalized pricing,[10] but online retailers generally are devoting significant resources to analyzing how to implement it (Khan 2014). A major topic of discussion at the 2014 National Retail Federation annual convention, for example, was how to introduce discriminatory pricing without triggering consumer backlash (Khan 2014). One mechanism discussed was highly personalized coupons sent at the point of sale, which would avoid the need to show consumers different prices but would still achieve discriminatory pricing (Khan 2014).

If retailers—including Amazon—implement discriminatory pricing on a wide scale, each individual would be subject to his or her own personal price trajectory, eliminating the notion of a single pricing trend. It is not clear how we would measure price hikes for the purpose of recoupment analysis in that scenario. There would be no obvious conclusions if some consumers faced higher prices while others enjoyed lower ones. But given the magnitude and accuracy of data that Amazon has collected on millions of users, tailored pricing is not simply a hypothetical power.

It is true that brick-and-mortar stores also collect data on customer purchasing habits and send personalized coupons. But the types of consumer behavior that Internet firms can access—how long you hover your mouse

---

9. Several journalists have tracked instances of price discrimination in e-commerce. See, e.g., Angwin and Larson (2015); Valentino-Devries, Singer-Vine, and Soltani (2012).
10. But recent reporting does suggest that Amazon manipulates how it presents pricing in order to favor its own products. See Angwin and Mattu (2016).

on a particular item, how many days an item sits in your shopping basket before you purchase it, or the fashion blogs you visit before looking for those same items through a search engine—is uncharted ground. The degree to which a firm can tailor and personalize an online shopping experience is different in kind from the methods available to a brick-and-mortar store—precisely because the type of behavior that online firms can track is far more detailed and nuanced. And unlike brick-and-mortar stores—where everyone at least *sees* a common price (even if they go on to receive discounts)—Internet retail enables firms to entirely personalize consumer experiences, which eliminates any collective baseline from which to gauge price increases or decreases.

In which product market would Amazon choose to raise prices? This is also an open question—and one that current predatory pricing doctrine ignores. Courts generally assume that a firm will recoup by increasing prices on the same goods on which it previously lost money. But recoupment across markets is also available as a strategy, especially for firms as diversified across products and services as Amazon. Reporting suggests the company did just this in 2013, by hiking prices on scholarly and small-press books and creating the risk of a "two-tier system where some books are priced beyond an audience's reach" (Streitfeld 2013). Although Amazon may be recouping its initial losses in e-books through markups on physical books, this cross-market recoupment is not a scenario that enforcers or judges generally consider (see Areeda and Hovenkamp 2010; Leslie 2013, 1720; Trujillo 1994). One possible reason for this neglect is that Chicago School scholarship, which assumes recoupment in single-product markets is unlikely, also holds recoupment in multiproduct scenarios to be implausible (Leslie 2013, 1720–21).

Although current predatory pricing doctrine focuses only on recoupment through raising prices for consumers, Amazon could also recoup its losses by imposing higher fees on publishers. Large book retailer chains like Barnes & Noble have long used their market dominance to charge publishers for favorable product placement, such as displays in a storefront window or on a prominent table (see Kennedy 2005). Amazon's dominance in the e-book market has enabled it to demand similar fees for even the most basic of services. For example, when renewing its contract with Hachette in 2014, Amazon demanded payments for services including the preorder button, personalized recommendations, and an Amazon employee assigned to the publisher (Stewart 2014). In the words of one person close to the negotiations, Amazon "is very inventive about what we'd call standard service. . . . They're teasing out all these layers and saying, 'If you want that service, you'll have to pay for it'" (Stewart 2014). By introducing fees on services

that it previously offered for free, Amazon has created another source of revenue. Amazon's power to demand these fees—and recoup some of the losses it sustained in below-cost pricing—stems from dominance partly built through that same below-cost pricing. The fact that Amazon has itself vertically integrated into book publishing—and hence can promote its own content—may give it additional leverage to hike fees. Any publisher that refuses could see Amazon favor its own books over the publisher's. It is not uncommon for half of the titles on Amazon's Kindle bestseller list to be its own (see LaVecchia and Mitchell 2016, 2).

While not captured by current antitrust doctrine, the pressure Amazon puts on publishers merits concern. For one, consolidation among book sellers—partly spurred by Amazon's pricing tactics and demands for better terms from publishers—has also spurred consolidation among publishers. Consolidation among publishers last reached its heyday in the 1990s—as publishing houses sought to bulk up in response to the growing clout of Borders and Barnes & Noble—and by the early 2000s, the industry had settled into the "Big Six" in the United States (Kachka 2013). This trend has cost authors and readers alike, leaving writers with fewer paths to market and readers with a less diverse marketplace. Since Amazon's rise, the major publishers have merged further—thinning down to five, with rumors of more consolidation to come (Kachka 2013).

Second, the increasing cost of doing business with Amazon is upending the publishers' business model in ways that further risk sapping diversity. Traditionally, publishing houses used a cross-subsidization model whereby they would use their bestsellers to subsidize weightier and riskier books requiring greater upfront investment. In the face of higher fees imposed by Amazon, publishers say they are less able to invest in a range of books. In a recent letter to the Department of Justice, a group of authors wrote that Amazon's actions have "extract[ed] vital resources from the [book] industry in ways that lessen the diversity and quality of books" (Authors United 2015). The authors noted that publishers have responded to Amazon's fees by both publishing fewer titles and focusing largely on books by celebrities and bestselling authors (Authors United 2015). The authors also noted, "Readers are presented with fewer books that espouse unusual, quirky, offbeat, or politically risky ideas, as well as books from new and unproven authors. This impoverishes America's marketplace of ideas" (Authors United 2015)

Amazon's conduct would be readily cognizable as a threat under the pre–Chicago School view that predatory pricing laws specifically and antitrust generally promoted a broad set of values. Under the predatory pricing jurisprudence of the early and mid-20th century, harm to the diversity

and vibrancy of ideas in the book market may have been a primary basis for government intervention. The political risks associated with Amazon's market dominance also implicate some of the major concerns that animate antitrust laws. For instance, the risk that Amazon may retaliate against books that it disfavors—either to impose greater pressure on publishers or for other political reasons—raises concerns about media freedom. Given that antitrust authorities previously considered diversity of speech and ideas a factor in their analysis, Amazon's degree of control, too, should warrant concern.

Even within the narrower "consumer welfare" framework, Amazon's attempts to recoup losses through fees on publishers should be understood as harmful. A market with less choice and diversity for readers amounts to a form of consumer injury. That the DOJ ignored this concern in its suit against Apple and the publishers suggests that its conception of predatory pricing overlooks the full suite of harms that Amazon's actions may cause.

Amazon's below-cost pricing in the e-book market—which enabled it to capture 65% of that market,[11] a sizable share by any measure—strains predatory pricing doctrine in several ways. First, Amazon is positioned to recoup its losses by raising prices on less popular or obscure e-books, or by raising prices on print books. In either case, Amazon would be recouping outside the original market where it sustained losses (bestseller e-books), so courts are unlikely to look for or consider these scenarios. Additionally, constant fluctuations in prices and the ability to price discriminate enable Amazon to raise prices with little chance of detection. Lastly, Amazon could recoup its losses by extracting more from publishers, who are dependent on its platform to market both e-books and print books. This may diminish the quality and breadth of the works that are published, but since this is most directly a supplier-side rather than buyer-side harm, it is less likely that a modern court would consider it closely. The current predatory pricing framework fails to capture the harm posed to the book market by Amazon's tactics.

Amazon Delivery and Leveraging Dominance across Sectors

Amazon's willingness to sustain losses has allowed it to engage in below-cost pricing in order to establish dominance as an online retailer. Amazon

---

11. At the height of its market share, this figure was closer to 90%. After Apple entered the market, Amazon's share fell slightly and then stabilized around 65% (Packer 2014).

has translated its dominance as an online retailer into significant bargaining power in the delivery sector, using it to secure favorable conditions from third-party delivery companies. This in turn has enabled Amazon to extend its dominance over other retailers by creating the Fulfillment-by-Amazon service and establishing its own physical delivery capacity. This illustrates how a company can leverage its dominant platform to successfully integrate into other sectors, creating anticompetitive dynamics. Retail competitors are left with two undesirable choices: either try to compete with Amazon at a disadvantage or become reliant on a competitor to handle delivery and logistics.

As Amazon expanded its share of e-commerce—and enlarged the e-commerce sector as a whole—it started constituting a greater share of delivery companies' business. For example, in 2015, UPS derived $1 billion worth of business from Amazon alone (Stevens and Bensinger 2015). The fact that it accounted for a growing share of these firms' businesses gave Amazon bargaining power to negotiate for lower rates. By some estimates, Amazon enjoyed a 70% discount over regular delivery prices (Clifford and Cain Miller 2010). Delivery companies sought to make up for the discounts they gave to Amazon by raising the prices they charged to independent sellers (see Stevens 2016), a phenomenon recently termed the "waterbed effect" (see Dobson and Inderst 2008, 336–37; Kirkwood 2012, 1544). As scholars have described,

> [T]he presence of a waterbed effect can further distort competition by giving a powerful buyer now a two-fold advantage, namely, through more advantageous terms for itself and through higher purchasing costs for its rivals. What then becomes a virtuous circle for the strong buyer ends up as a vicious circle for its weaker competitors (Dobson and Inderst 2008, 337).

To this twofold advantage Amazon added a third perk: harnessing the weakness of its rivals into a business opportunity. In 2006, Amazon introduced Fulfillment-by-Amazon (FBA), a logistics and delivery service for independent sellers (Phx.corporate-ir.net 2006). Merchants who sign up for FBA store their products in Amazon's warehouses, and Amazon packs, ships, and provides customer service on any orders. Products sold through FBA are eligible for service through Amazon Prime—namely, free two-day shipping and/or free regular shipping, depending on the order (Phx.corporate-ir.net 2006). Since many merchants selling on Amazon are competing with Amazon's own retail operation and its Amazon Prime service, using FBA offers sellers the opportunity to compete at less of a disadvantage.

Notably, it is partly because independent sellers faced higher rates from UPS and FedEx—a result of Amazon's dominance—that Amazon succeeded in directing sellers to its new business venture (see Cole 2012). In many instances, orders routed through FBA were still being shipped and delivered by UPS and FedEx, since Amazon relied on these firms (Wohlsen 2014). But because Amazon had secured discounts unavailable to other sellers, it was cheaper for those sellers to go through Amazon than to use UPS and FedEx directly. Amazon had used its dominance in the retail sector to create and boost a new venture in the delivery sector, inserting itself into the business of its competitors.

Amazon has followed up on this initial foray into fulfillment services by creating a logistics empire. Building out physical capacity lets Amazon further reduce its delivery times, raising the bar for entry yet higher. Moreover, it is the firm's capacity for aggressive investing that has enabled it to rapidly establish an extensive network of physical infrastructure. Since 2010, Amazon has spent $13.9 billion building warehouses (Kucera 2013), and it spent $11.5 billion on shipping in 2015 alone (Leonard 2016). Amazon has opened more than 180 warehouses (Bensinger and Stevens 2016), 28 sorting centers, 59 delivery stations that feed packages to local couriers, and more than 65 Prime Now hubs (Bensinger and Stevens 2016). Analysts estimate that the locations of Amazon's fulfillment centers bring it within 20 miles of 31% of the US population and within 20 miles of 60% of its core same-day base (D'Onfro 2015). This sprawling network of fulfillment centers—each placed in or near a major metropolitan area—equips Amazon to offer one-hour delivery in some locations and same-day in others (a service it offers free to members of Amazon Prime) (Bensinger and Stevens 2016). While several rivals initially entered the delivery market to compete with Prime shipping, some are now retreating (Soper 2015). As one analyst noted, "Prime has proven exceedingly difficult for rivals to copy" (Stone 2010; see also Mangalindan 2012).

Most recently, Amazon has also expanded into trucking. In December 2015, it announced it plans to roll out thousands of branded semi-trucks, a move that will give it yet more control over delivery, as it seeks to speed up how quickly it can transport goods to customers (Del Ray 2015; Leonard 2016). Amazon now owns four thousand truck trailers and has also signed contracts for container ships, planes (Lewis 2016), and drones (Manjoo 2016). As of October 2016, Amazon had leased at least 40 jets (Manjoo 2016). Former employees say Amazon's long-term goal is to circumvent UPS and FedEx altogether, though the company itself has said it is looking only to supplement its reliance on these firms, not supplant them (see Del Ray 2015; also Leonard 2016).

The way that Amazon has leveraged its dominance as an online retailer to vertically integrate into delivery is instructive on several fronts. First, it is a textbook example of how a company can use its dominance in one sphere to advantage a separate line of business. To be sure, this dynamic is not intrinsically anticompetitive. What should prompt concern in Amazon's case, however, is that Amazon achieved these cross-sector advantages in part due to its bargaining power. Because Amazon was able to demand heavy discounts from FedEx and UPS, other sellers faced price hikes from these companies—which positioned Amazon to capture them as clients for its new business. By overlooking structural factors like bargaining power, modern antitrust doctrine fails to address this type of threat to competitive markets.

Second, Amazon is positioned to use its dominance across online retail and delivery in ways that involve tying, are exclusionary, and create entry barriers (Elhauge 2009). That is, Amazon's distortion of the delivery sector in turn creates anticompetitive challenges in the retail sector. For example, sellers who use FBA have a better chance of being listed higher in Amazon search results than those who do not, which means Amazon is tying the outcomes it generates for sellers using its retail platform to whether they also use its delivery business (Mitchell n.d.). Amazon is also positioned to use its logistics infrastructure to deliver its own retail goods faster than those of independent sellers that use its platform and fulfillment service—a form of discrimination that exemplifies traditional concerns about vertical integration. And Amazon's capacity for losses and expansive logistics capacities mean that it could privilege its own goods while still offering independent sellers the ability to ship goods more cheaply and quickly than they could by using UPS and FedEx directly.

Relatedly, Amazon's expansion into the delivery sector also raises questions about the Chicago School's limited conception of entry barriers. The company's capacity for losses—the permission it has won from investors to show negative profits—has been key in enabling Amazon to achieve outsized growth in delivery and logistics. Matching Amazon's network would require a rival to invest heavily and—in order to viably compete—offer free or otherwise below-cost shipping. In interviews with reporters, venture capitalists say there is no appetite to fund firms looking to compete with Amazon on physical delivery. In this way, Amazon's ability to sustain losses creates an entry barrier for any firm that does not enjoy the same privilege.

Third, Amazon's use of Prime and FBA exemplifies how the company has structurally placed itself at the center of e-commerce. Already 55% of American online shoppers begin their online shopping on Amazon's platform (Del Ray 2016). Given the traffic, it is becoming increasingly clear that

in order to succeed in e-commerce, an independent merchant will need to use Amazon's infrastructure. The fact that Amazon competes with many of the businesses that are coming to depend on it creates a host of conflicts of interest that the company can exploit to privilege its own products.

The framework in antitrust today fails to recognize the risk that Amazon's dominance poses to open and competitive markets. In part, this is because—as with the framework's view of predatory pricing—the primary harm that registers within the "consumer welfare" frame is higher consumer prices. On the Chicago School's account, Amazon's vertical integration would only be harmful if and when it chooses to use its dominance in delivery and retail to hike fees to consumers. Amazon has already raised Prime prices (Bensinger 2014). But antitrust enforcers should be equally concerned about the fact that Amazon increasingly controls the infrastructure of online commerce—and the ways in which it is harnessing this dominance to expand and advantage its new business ventures. The conflicts of interest that arise from Amazon both competing with merchants and delivering their wares pose a hazard to competition, particularly in light of Amazon's entrenched position as an online platform. Amazon's conflicts of interest tarnish the neutrality of the competitive process. The thousands of retailers and independent businesses that must ride Amazon's rails to reach market are increasingly dependent on their biggest competitor.

Amazon Marketplace and Exploiting Data

As described above, vertical integration in retail and physical delivery may enable Amazon to leverage cross-sector advantages in ways that are potentially anticompetitive but not understood as such under current antitrust doctrine. Analogous dynamics are at play with Amazon's dominance in the provision of *online* infrastructure, in particular its Marketplace for third-party sellers. Because information about Amazon's practices in this area is limited, this section is necessarily brief. But to capture fully the anticompetitive features of Amazon's business strategy, it is vital to analyze how vertical integration across Internet businesses introduces more sophisticated—and potentially more troubling—opportunities to abuse cross-market advantages and foreclose rivals.

The clearest example of how the company leverages its power across online businesses is Amazon Marketplace, where third-party retailers sell their wares. Since Amazon commands a large share of e-commerce traffic, many smaller merchants find it necessary to use its site to draw buyers (Loten and Janofsky 2015). These sellers list their goods on Amazon's

platform and the company collects fees ranging from 6% to 50% of their sales from them (Loten and Janofsky 2015). More than 2 million third-party sellers used Amazon's platform as of 2015, an increase from the roughly 1 million that used the platform in 2006 (Loten and Janofsky 2015). The revenue that Amazon generates through Marketplace has been a major source of its growth: third-party sellers' share of total items sold on Amazon rose from 36% in 2011 (Bensinger 2012) to over 50% in 2015 (Halpin 2015).

Third-party sellers using Marketplace recognize that using the platform puts them in a bind. As one merchant observed, "You can't really be a high-volume seller online without being on Amazon, but sellers are very aware of the fact that Amazon is also their primary competitor" (Loten and Janofsky 2015). Evidence suggests that their unease is well founded. Amazon seems to use its Marketplace "as a vast laboratory to spot new products to sell, test sales of potential new goods, and exert more control over pricing" (Bensinger 2012). Specifically, reporting suggests that "Amazon uses sales data from outside merchants to make purchasing decisions in order to undercut them on price" and give its own items "featured placement under a given search" (Bensinger 2012). Take the example of Pillow Pets, "stuffed-animal pillows modelled after NFL mascots" that a third-party merchant sold through Amazon's site (Bensinger 2012). For several months, the merchant sold up to 100 pillows per day (Bensinger 2012). According to one account, "just ahead of the holiday season, [the merchant] noticed Amazon had itself beg[u]n offering the same Pillow Pets for the same price while giving [its own] products featured placement on the site" (Bensinger 2012). The merchant's own sales dropped to 20 per day (Bensinger 2012). Amazon has gone head-to-head with independent merchants on price, vigorously matching and even undercutting them on products that they had originally introduced. By going directly to the manufacturer, Amazon seeks to cut out the independent sellers.

In other instances, Amazon has responded to popular third-party products by producing them itself. Last year, a manufacturer that had been selling an aluminum laptop stand on Marketplace for more than a decade saw a similar stand appear at half the price. The manufacturer learned that the brand was AmazonBasics, the private line that Amazon has been developing since 2009 (Soper 2016). As one news site describes it, initially, AmazonBasics focused on generic goods like batteries and blank DVDs. "Then, for several years, the house brand 'slept quietly as it retained data about other sellers' successes'" (Soper 2016). As it now rolls out more AmazonBasics products, it is clear that the company has used "insights gleaned from its vast Web store to build a private-label juggernaut that now

includes more than 3,000 products" (Soper 2016). One study found that in the case of women's clothing, Amazon "began selling 25% of the top items first sold through marketplace vendors" (Anderson 2014).

In using its Marketplace this way, Amazon increases sales while shedding risk. It is third-party sellers who bear the initial costs and uncertainties when introducing new products; by merely spotting them, Amazon gets to sell products only once their success has been tested. The anticompetitive implications here seem clear: Amazon is exploiting the fact that some of its customers are also its rivals. The source of this power is: (1) its dominance as a platform, which effectively necessitates that independent merchants use its site; (2) its vertical integration—namely, the fact that it both sells goods as a retailer and hosts sales by others as a marketplace; and (3) its ability to amass swaths of data, by virtue of being an Internet company. Notably, it is this last factor—its control over data—that heightens the anticompetitive potential of the first two.

Evidence suggests that Amazon is keenly aware of and interested in exploiting these opportunities. For example, the company has reportedly used insights gleaned from its cloud computing service to inform its investment decisions (see Barr 2011). By observing which start-ups are expanding their usage of Amazon Web Services, Amazon can make early assessments of the potential success of upcoming firms. Amazon has used this "unique window into the technology startup world" to invest in several start-ups that were also customers of its cloud business (Barr 2011).

How Amazon has cross-leveraged its advantages across distinct lines of business suggests that the law fails to appreciate when vertical integration may prove anticompetitive. This shortcoming is underscored with online platforms, which both serve as infrastructure for other companies and collect swaths of data that they can then use to build up other lines of business. In this way, the current antitrust regime has yet to reckon with the fact that firms with concentrated control over data can systematically tilt a market in their favor, dramatically reshaping the sector.

TWO MODELS FOR ADDRESSING PLATFORM POWER

If it is true that the economics of platform markets may encourage anticompetitive market structures, there are at least two approaches we can take. Key is deciding whether we want to govern online platform markets through competition, or want to accept that they are inherently monopolistic or oligopolistic and regulate them instead. If we take the former approach, we should reform antitrust law to prevent this dominance from

emerging or to limit its scope. If we take the latter approach, we should adopt regulations to take advantage of these economies of scale while neutering the firm's ability to exploit its dominance.

Governing Online Platform Markets through Competition

Reforming antitrust to address the anticompetitive nature of platform markets could involve making the law against predatory pricing more robust and strictly policing forms of vertical integration that firms can use for anticompetitive ends. Importantly, each of these doctrinal areas should be reformulated so that it is sensitive to preserving the competitive process and limiting conflicts of interest that may incentivize anticompetitive conduct.

Revising predatory pricing doctrine to reflect the economics of platform markets, where firms can sink money for years given unlimited investor backing, would require abandoning the recoupment requirement in cases of below-cost pricing by dominant platforms. And given that platforms are uniquely positioned to fund predation, a competition-based approach might also consider introducing a presumption of predation for dominant platforms found to be pricing products below cost.

Similarly, antitrust law should be reformed to address how vertical integration may give rise to anticompetitive conflicts of interest and the fact that a dominant firm may use its dominance in one sector to advance another line of business. One way to address the concern about a firm's capacity to cross-leverage data is to expressly include it into merger review. It could make sense for the agencies to automatically review any deal that involves exchange of certain forms (or a certain quantity) of data. A stricter approach to vertical integration would place prophylactic limits on vertical mergers by platforms that have reached a certain level of dominance. Adopting this prophylactic approach would mean banning a dominant firm from entering any market that it already serves as a platform—in other words, from competing directly with the businesses that depend on it.

Governing Dominant Platforms as Monopolies through Regulation

As described above, one option is to govern dominant platforms through promoting competition, thereby limiting the power that any one actor accrues. The other is to accept dominant online platforms as natural monopolies

or oligopolies, seeking to regulate their power instead. Traditionally the United States has regulated natural monopolies through public utility regulations and common carrier duties. Industries that historically have been regulated as utilities include commodities (water, electric power, gas), transportation (railroads, ferries), and communications (telegraphy, telephones) (Wu 2010, 1616). Critically, a public utility regime aims at eliminating competition: it accepts the benefits of monopoly and chooses to instead limit how a monopoly may use its power (Wu 2010, 1643).

Given that Amazon increasingly serves as essential infrastructure across the Internet economy, applying elements of public utility regulations to its business is worth considering (see Rahman 2015; 2016). The most common public utility policies are (1) requiring nondiscrimination in price and service, (2) setting limits on rate-setting, and (3) imposing capitalization and investment requirements. Of these three traditional policies, nondiscrimination would make the most sense, while rate-setting and investment requirements would be trickier to implement and, perhaps, would less obviously address an outstanding deficiency.

A nondiscrimination policy that prohibited Amazon from privileging its own goods and from discriminating among producers and consumers would be significant. This approach would permit the company to maintain its involvement across multiple lines of business and permit it to enjoy the benefits of scale while mitigating the concern that Amazon could unfairly advantage its own business or unfairly discriminate among platform users to gain leverage or market power.[12] Coupling nondiscrimination with common carrier obligations—requiring platforms to ensure open and fair access to other businesses—would further limit Amazon's power to use its dominance in anticompetitive ways.

CONCLUSION

Internet platforms mediate a large and growing share of our commerce and communications. Yet evidence shows that competition in platform markets is flagging, with sectors coalescing around one or two giants (FCC 2017). The titan in e-commerce is Amazon—a company that has built its dominance through aggressively pursuing growth at the expense of profits and that has integrated across many related lines of business. As a result,

12. Net neutrality is a form of common carrier regime. For an exposition of why net neutrality and search neutrality should apply to major platforms, see Pasquale (2008, 263).

the company has positioned itself at the center of Internet commerce and serves as essential infrastructure for a host of other businesses that now depend on it. This chapter argues that Amazon's business strategies and current market dominance pose anticompetitive concerns that the consumer welfare framework in antitrust fails to recognize.

In particular, current law underappreciates the risk of predatory pricing and how integration across distinct business lines may prove anticompetitive. These concerns are heightened in the context of online platforms for two reasons. First, the economics of platform markets incentivize the pursuit of growth over profits, a strategy that investors have rewarded. Under these conditions predatory pricing becomes highly rational—even as existing doctrine treats it as irrational. Second, because online platforms serve as critical intermediaries, integrating across business lines positions these platforms to control the essential infrastructure on which their rivals depend. This dual role also enables a platform to exploit information collected on companies using its services to undermine them as competitors.

In order to capture these anticompetitive concerns, we should replace the consumer welfare framework with an approach oriented around preserving a competitive process and market structure. Applying this idea involves, for example, assessing whether a company's structure creates anticompetitive conflicts of interest; whether it can cross-leverage market advantages across distinct lines of business; and whether the economics of online platform markets incentivizes predatory conduct and capital markets permit it. More specifically, restoring traditional antitrust principles to create a presumption of predation and to ban vertical integration by dominant platforms could help maintain competition in these markets. If, instead, we accept dominant online platforms as natural monopolies or oligopolies, then applying elements of a public utility regime or essential facilities obligations would maintain the benefits of scale while limiting the ability of dominant platforms to abuse the power that comes with it.

As Amazon continues both to deepen its existing control over key infrastructure and to reach into new lines of business, its dominance demands scrutiny. To revise antitrust law and competition policy for platform markets, we should be guided by two questions. First, does the legal framework capture the realities of how dominant firms acquire and exercise power in the Internet economy? And second, what forms and degrees of power should the law identify as a threat to competition? Without considering these questions, we risk permitting the growth of powers that we oppose but fail to recognize.

REFERENCES

Amazon.com, Inc. 2010. "Annual Report (Form 10-K) 27." January 29, 2010. http://www.sec.gov/Archives/edgar/data/1018724/000119312510016098/d10k.htm; http://perma.cc/L27R-CHUY.

Amazon.com, Inc. 2016. "Annual Report (Form 10-K) 4." April 6, 2016. http://phx.corporate -ir.net/External.File?item=UGFyZW50SUQ9NjI4NTg0fENoaWxkSU Q9MzI5NTMwfFR 5cGU9MQ==&t=1.

Anderson, George. 2014. "Is Amazon Undercutting Third-Party Sellers Using Their Own Data?" *Forbes*, October 30, 2014. http://www.forbes.com/sites/retailwire/2014/10/30/is-amazon-undercutting-third-party-sellers-using-their-own-data.

Angwin, Julia, and Jeff Larson. 2015. "The Tiger Mom Tax: Asians Are Nearly Twice as Likely to Get a Higher Price from Princeton Review." *ProPublica*, September 1, 2015. http://www.propublica.org/article / asians-nearly-twice-as-likely-to-get-higher-price-from-princeton-review.

Angwin, Julia, and Surya Mattu. 2016. "Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't." *ProPublica*, September 20, 2016. http://www.propublica.org/article/amazon-says-it-puts-customers-first-but-its -pricing-algorithm-doesn't.

Areeda, Phillip, and Herbert Hovenkamp. 2010. *Fundamentals of Antitrust Law*. New York: Aspen Law and Business, 7–72.

Authors United. 2015. "Letter to William J. Baer, Assistant Att'y Gen., Antitrust Div., Dept, of Justice." July 14, 2015. http://www.authorsunited.net/july/longdocument.html.

Bain, Joe S. 1950. "Workable Competition in Oligopoly: Theoretical Considerations and Some Empirical Evidence." *American Economic Review* 40: 35, 36–38.

Bain, Joe S. (1959) 1968. *Industrial Organization*. 2d ed. New Jersey: John Wiley & Sons

Banjo, Shelly. 2016. "Amazon Eats the Department Store." *Bloomberg: Gadfly*, September 20, 2016. http://www.bloomberg.com/gadfly/articles/2016-09-20/amazon-clothing-sales -could-soon-top-macy-s.

Banjo, Shelly, and Paul Ziobro. 2013. "After Decades of Toil, Web Services Remain Small for Many Retailers." *Wall Street Journal*, August 27, 2013. http://www.wsj.com/articles /SB10001424127887324906304579039101568397122.

Barr, Alistair. 2011. "Amazon Finds Startup Investments in the 'Cloud.'" *Reuters*, November 9, 2011. http://www.reuters.com/article/amazon-cloud-idUSN1E7A727Q20111109.

Bensinger, Greg. 2012. "Competing with Amazon on Amazon." *Wall Street Journal*, June 27, 2012. http://www.wsj.com/articles/SB10001424052702304441404577482902055882264.

Bensinger, Greg. 2014. "Amazon Raises Prime Subscription Price to $99 a Year." *Wall Street Journal*, March 13, 2014. http://www.wsj.com/articles/SB10001424052702303546204579436903309411092.

Bensinger, Greg. 2016. "Cloud Unit Pushes Amazon to Record Profit." *Wall Street Journal*, April 28, 2016. http://www.wsj.com/articles/amazon-reports-surge-in-profit-1461874333.

Bensinger, Greg, and Laura Stevens. 2016. "Amazon's Newest Ambition: Competing Directly with UPS and FedEx." *Wall Street Journal*, September 27, 2016. http://www.wsj.com/articles/amazons-newest-ambitioncompeting-directly-with-ups-and-fedex-1474994758.

Bezos, Jeffrey P. 1998. "Letter to Shareholders." Amazon.com, Inc., March 30, 1998. http:// media.corporate-ir.net/media_files/irol/97/97664/reports/ Shareholderletter97.pdf.

BI Intelligence. 2016. "The Everything Shipper: Amazon and the New Age of Delivery." June 5, 2016. http://www.businessinsider.com/the-everything-shipper-amazon-and-the-new -age-of-delivery-2016-6.

Bork, Robert H. 1978. *The Antitrust Paradox: A Policy at War with Itself*. New York: Basic Books.

Candeub, Adam. 2014. *Behavioral Economics, Internet Search, and Antitrust*. Michigan: Michigan State University College of Law.

Clark, Meagan, and Angelo Young. 2013. "Amazon: Nearly 20 Years in Business and It Still Doesn't Make Money, but Investors Don't Seem to Care." *International Business Times*, December 18, 2013. http:// www.ibtimes.com/ amazon-nearly-20-years-business-it-still-doesnt-make-money-investors -dont-seem-care-1513368.

Clifford, Stephanie, and Claire Cain Miller. 2010. "Wal-Mart Says 'Try This On': Free Shipping." *New York Times*, November 11, 2010. http://www.nytimes.com/ 2010/11/11/business/11shipping.html

CNN Money. 2002. "7 Amazon Posts a Profit." January 22, 2002. http://money.cnn. com/2002 /01/22/technology/amazon.

Cole, Paul. 2012. "Should You Use Amazon Discounted UPS Shipping?" *Seller Engine*. http://sellerengine.com/should-you-use-amazon-discounted-ups-shipping.

Cole, Robert H. 1952. "General Discussion of Vertical Integration." In *Vertical Integration in Marketing* 9: 9 (Bureau of Econ. & Bus. Research, Univ. of Ill., Bulletin No. 74, Nugent Wedding ed.).

Crane, Daniel A. 2014. "The Tempting of Antitrust: Robert Bork and the Goals of Antitrust Policy." *Antitrust Law Journal* 79: 835–47.

Del Ray, Jason. 2015. "Amazon Buys Thousands of Its Own Truck Trailers as Its Transportation Ambitions Grow." *Recode*, December 4, 2015. http://recode.net/2015/12/04/amazon -buys-thousands-of-its-own-trucks-as-its-transportation-ambitions-grow.

Del Ray, Jason. 2016. "55 Percent of Online Shoppers Start Their Product Searches on Amazon." *ReCode*, September 27, 2016. https://www.recode.net/2016/9/27/ 13078526/amazon-online-shopping-product-search-engine.

DiChristopher, Tom. 2015. "Prime Will Grow Amazon Revenue Longer Than You Think: Analyst." *CNBC*, September 11, 2015. http://www.cnbc.com/2015/09/ 11/prime-will-grow -amazon-revenue-longer-than-you-think-analyst.html.

Dini, Justin. 2000. "Amazon Losses Widen but Shares Rise After-Hours." *The Street,* February 2, 2000. http://www.thestreet.com/story/875924/1/amazon-losses-widen-but -shares-rise-after-hours.html.

Dobson, Paul W., and Roman Inderst. 2008. "The Waterbed Effect: Where Buying and Selling Power Come Together." *Wisconsin Law Review* 2008: 331, 336–37.

DOJ. 1982. "Merger Guidelines." http://www.justice.gov/sites/default/files /atr/ legacy/2007/07/11/11248.pdf.

DOJ. 1984. "Merger Guidelines." http://www.justice.gov/sites/default/files/atr/ legacy/2007/07/11/11249.pdf.

D'Onfro, Jillian. 2015. "Here Are All of Amazon's Warehouses in the US." *Business Insider*, March 24, 2015. http://www.businessinsider.com/how-many-fulfillment-centers-does-amazon -have-in-the-us-2015-3#ixzz3f3AX8zda.

Eisner, Marc Allen. 1991. *Antitrust and the Triumph of Economics: Institutions, Expertise, and Policy Change*. Chapel Hill: University of North Carolina Press, 107.

Elhauge, Einer. 2009. "Tying, Bundled Discounts, and the Death of the Single Monopoly Profit Theorem." *Harvard Law Review* 123: 397, 466–67.

Elmer-DeWitt, Philip. 2015. "This Is What Drives Apple Investors Nuts about Amazon." *Fortune*, July 24, 2015. http://fortune.com/2015/07/24 / apple-amazon-profits.

Enright, Allison. 2016. "Amazon Sales Climb 22% in Q4 and 20% in 2015." *Internet Retailer*, January 28, 2016. http://www.internetretailer.com/2016/01/28/ amazon-sales-climb-22 -q4-and-20-2015.

FCC. 2017. "Open Internet." Federal Communications Commission. Accessed September 18, 2017. http://www.fcc.gov/general/open-internet.

Ferdman, Roberto A. 2013. "Amazon Changes Its Prices More than 2.5 Million Times a Day." *Quartz*, December 14, 2013. http://qz.com/157828/amazon-changes-its-prices-more-than-2-5-million -times-a-day.

Frommer, Dan. 2015. "Half of US Households Could Have Amazon Prime by 2020." *Quartz*, February 26, 2015. http://qz.com/351726/half-of-us-households-could-have-amazon-prime-by-2020; http://perma.cc /ZW4Z-47UY.

Garcia, Tonya. 2016. "Amazon Accounted for 60% of U.S. Online Sales Growth in 2015." *Marketwatch*, May 3, 2016. http://www.marketwatch.com/story/ama zon-accounted-for-60-of-online-sales-growth-in-2015-2016-05-03.

Halpin, Nancee. 2015. "Third-Party Merchants Account for More Than Three-Quarters of Items Sold on Amazon." *Business Insider*, October 16, 2015. http://www. businessinsider.com/third -party-merchants-drive-amazon-grow-2015-10.

Hoffmann, Melissa. 2014. "Amazon Has the Best Consumer Perception of Any Brand." Adweek, July 16, 2014. http://www.adweek.com/news/advertising-branding/ amazon -has-best-consumer-perception-any-brand-158945.

Kachka, Boris. 2013. "Book Publishing's Big Gamble." *New York Times*, July 9, 2013. http://www.nytimes.com/2013/07/10/opinion/book-publishings-big-gamble. html.

Kawamoto, Dawn. 2005. "Amazon Unveils Flat-Fee Shipping." *CNET*, February 2, 2005. http://www.cnet .com/news/amazon-unveils-flat-fee-shipping.

Kennedy, Randy. 2005. "Cash Up Front." *New York Times*, June 5, 2005. http://www. nytimes.com /2005/06/05/books/review/cash-up-front.html.

Khan, Lina. 2014. "Why You Might Pay More than Your Neighbor for the Same Bottle of Salad Dressing." *Quartz*, January 19, 2014. http://qz.com/168314/why-you-might-pay-more-than-your -neighbor-for-the-same-bottle-of-salad-dressing.

Kirkwood, John. 2012. "Powerful Buyers and Merger Enforcement." *Boston University Law Review* 92: 1485–544.

Krantz, Matt. 2015. "Amazon Breaks Barrier: Now Most Costly Stock." *USA Today*, November 11, 2015. http://www.usatoday.com/story/money/markets/2015/ 11/11/amazon-pe-ratio -valuation-price/75519460.

Krugman, Paul. 2014. "Amazon's Monopsony Is Not O.K." *New York Times*, October 19, 2014. http://www.nytimes.com/2014/10/20/opinion/ paul-krugman-amazons-monopsony-is-not-ok.html.

Kucera, Daniella. 2013. "Why Amazon Is on a Building Spree." *Bloomberg*, August 29, 2013. http://www.bloomberg.com/bw/articles/2013-08-29/why-amazon-is -on-a-warehouse-building-spree.

LaVecchia, Olivia, and Stacy Mitchell. 2016. "Amazon's Stranglehold: How the Company's Tightening Grip Is Stifling Competition, Eroding Jobs, and Threatening Communities." *Institute for Local Self-Reliance*, November, 10, 2016. http://ilsr.org/wp-content/uploads/2016/11/ILSR_Amazon Report_final.pdf.

Leonard, Devin. 2016. "Will Amazon Kill FedEx?" *Bloomberg*, August 31, 2016. http://www.bloomberg.com/features/2016-amazon-delivery.

Leslie, Christopher R. 2013. "Predatory Pricing and Recoupment." *Columbia Law Review* 113: 1695–702.

Lewis, Robin. 2016. "Amazon's Shipping Ambitions Are Larger than It's Letting On." *Forbes,* April 1, 2016. http://www.forbes.com/sites/robinlewis/2016/04/01/planes-trains-trucks -and-ships/#260c3aa1408c.

Loten, Angus, and Adam Janofsky. 2015. "Sellers Need Amazon, but at What Cost?" *Wall Street Journal*, January 14, 2015. http://www.wsj.com/articles/sellers-need-amazon-but-at-what-cost-1421278220.

Mangalindan, J. P. 2012. "Amazon's Prime and Punishment." *Fortune*, February 21, 2012. http://fortune.com/2012/02/21/amazons-prime-and-punishment.

Manjoo, Farhad. 2015. "How Amazon's Long Game Yielded a Retail Juggernaut." *New York Times*, November 18, 2015. http://www.nytimes.com/2015/11/19/technology/how-amazons -long-game-yielded-a-retail-juggernaut.html.

Manjoo, Farhad. 2016. "Tech's 'Frightful 5' Will Dominate Digital Life for Foreseeable Future." *New York Times*, January 20, 2016. http://www.nytimes.com/2016/01/21/technology/techs-frightful -5-will-dominate-digital-life-for-foreseeable-future.html.

Manjoo, Farhad. 2016. "Think Amazon's Drone Idea Is a Gimmick? Think Again." *New York Times*, August 10, 2016. http://www.nytimes.com/2016/08/11/technology/think-amazons-drone-delivery -idea-is-a-gimmick-think-again.html.

Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 604 (1986) (White, J., dissenting).

Mitchell, Will. N.d. "How to Rank Your Products on Amazon—The Ultimate Guide." *StartUpBros*. http://startupbros.com/rank-amazon.

Moore, Sam. 2015. "Amazon Commands Nearly Half of Consumers' First Product Search." BloomReach, October 6, 2015. http://bloomreach.com/2015/10/amazon-commands-nearly-half-of -consumers-first-product-search.

Mulpuru, Sucharita, and Brian K. Walker. 2012. "Why Amazon Matters Now More Than Ever." *Forrester*, July 26, 2012.

O'Connor, Clare. 2015. "Walmart and Target Being Crowded Out Online by Amazon Prime." *Forbes*, April 6, 2015. http://www.forbes.com/sites/clareoconnor/2015/04/06 /walmart-and-target-being-crowded-out-online-by-amazon-prime.

Orbach, Barak. 2013. "Foreword: Antitrust's Pursuit of Purpose." *Fordham Law Review* 81: 2151–52.

Packer, George. 2014. "Cheap Words." *New Yorker*, February 17, 2014. http://www.newyorker.com/magazine/2014/02/17/cheap-words

Pasquale, Frank. 2008. "Internet Nondiscrimination Principles: Commercial Ethics for Carriers and Search Engines." *University of Chicago Legal Forum* 2008: 263.

Posner, Richard A. 1979. "The Chicago School of Antitrust Analysis." *University of Pennsylvania Law Review* 127: 925–32.

Quick Pen. 2015. "What's Driving the Amazon Stock Up Despite 188% Full Year Income Drop?" *Guru Focus,* February 8, 2015. http://www.

gurufocus.com/news/315124/whats-driving-the-amazon-stock-up
-despite-188-full-year-income-drop.

Rahman, K. Sabeel. 2015. "From Railroad to Uber: Curbing the New Corporate
Power." *Boston Review*, May 4, 2015. http://bostonreview.net/forum/k-sabeel-
rahman-curbing-new-corporate -power.

Rahman, K. Sabeel. 2016. "Private Power and Public Purpose: The Public Utility
Concept and the Future of Corporate Law in the New Gilded Age, Address at
the Association of American Law Schools' 110th Annual Meeting." January 8,
2016. Transcript on file with author.

Rajan, Raghuram, and Luigi Zingales. 2003. *Saving Capitalism from the Capitalists*.

Ray, Tiernan. 2016. "Amazon: All Retail's SKUs Are Belong to Them,
Goldman Tells CNBC," *Barrons: Tech Trader Daily*, June 16,
2016. http://blogs.barrons.com/techtrader daily/2016/06/16/
amazon-all-retails-skus-are-belong-to-them-goldman-tells-cnbc

Reiter v. Sonotone Corp., 442 U.S. 330, 343 (1979) US Supreme Court.

Rubin, Ben Fox. 2015. "As Amazon Marks 20 Years, Prime Grows to 44 Million
Members in US." *CNET*, July 15, 2015. http://www.cnet.com/news/amazon-
prime-grows-to-estimated -44-million-members-in-us.

Rubin, Chad. 2016. "The Evolution of Amazon Prime and Their Followed Success."
Skubana, March 31, 2016. http://www.skubana.com/e-commerce-trends/
evolution-of-amazon-prime.

Scherer, F. M. 2007. "Conservative Economics and Antitrust: A Variety of Influences."
In *How the Chicago School Overshot the Mark*, edited by Robert Pitofsky, 30–33.
Oxford: Oxford University Press

Seetharaman, Deepa, and Nathan Layne. 2015. "Free Delivery Creates
Holiday Boon for U.S. Consumers at High Cost." *Reuters*,
January 2, 2015. http://www.reuters.com/article/us-retail
-shipping-holidays-analysis-idUSKBN0KB0P720150102.

Slate: Moneybox. 2000. "Amazon: Ponzi Scheme or Wal-Mart of the Web?" February 8,
2000. http://www.slate.com/articles/business/moneybox/2000/02/amazon_
ponzi_scheme _or_walmart_of_the_web.html.

Sokol, D. Daniel. 2014. "The Transformation of Vertical Restraints: Per Se Illegality, the
Rule of Reason, and Per Se Legality." *Antitrust Law Journal* 79: 1003, 1014–15.

Soper, Spencer. 2015. "EBay Ends Same-Day Delivery in U.S. in Face of Amazon
Effort." *Bloomberg*, July 27, 2015. http://www.bloomberg.com/news/articles/
2015-07-27/ebay -ends-same-day-delivery-in-u-s-in-face-of-amazon-effort.

Soper, Spencer. 2015. "Inside Amazon's Warehouse." *Morning Call,* August 17,
2015. http://www.mcall.com/news /local/amazon/mc-allentown-amazon-
complaints-20110917-story.html.

Soper, Spencer. 2016. "Got a Hot Seller on Amazon? Prepare for E-Tailer to Make One
Too." *Bloomberg*, April 20, 2016. http://www.bloomberg.com/news/articles/
2016 -04-20/got-a-hot-seller-on-amazon-prepare-for-e-tailer-to-make-one-too.

Stevens, Laura. 2016. " 'Free' Shipping Crowds Out Small Retailers." *Wall Street
Journal*, April 27, 2016. http://www.wsj.com/articles/for-online-shoppers-free-
shipping-reigns -supreme-1461789381.

Stevens, Laura, and Greg Bensinger. 2015. "Amazon Seeks to Ease Ties with UPS."
Wall Street Journal, December 22, 2015. http://www.wsj.com/articles/amazon-
seeks-to-ease-ties-with-ups -1450835575.

Stewart, James B. 2014. "Booksellers Score Some Points in Amazon's Spat with
Hachette." *New York Times*, June 20, 2014. http://www.nytimes.com/2014/

06/21/business/booksellers-score-some -points-in-amazons-standoff-with-hachette.html.

Stone, Brad. 2010. "What's in Amazon's Box? Instant Gratification." *Bloomberg Businessweek*, November 24, 2010. http://www.bloomberg.com/news/articles/2010-11 -24/whats-in-amazons-box-instant-gratification;

Strauss, Karsten. 2016. "America's Most Reputable Companies, 2016: Amazon Tops the List." Forbes, March 29, 2016. http://www.forbes.com/sites/karstenstrauss/2016/03/29 /americas-most-reputable-companies-2016-amazon-tops-the-list.

Streitfeld, David. 2013. "As Competition Wanes, Amazon Cuts Back Discounts." *New York Times*, July 4, 2013. http://www.nytimes.com/2013/07/05/business/as-competition-wanes-amazon -cuts-back-its-discounts.html.

Streitfeld, David. 2013. "A New Book Portrays Amazon as Bully." *New York Times, Bits Blog*, October, 22, 2013. http://bits.blogs.nytimes.com/2013/10/22/a-new-book-portrays-amazon -as-bully.

Streitfeld, David. 2015. "Accusing Amazon of Antitrust Violations, Authors and Booksellers Demand Inquiry." *New York Times*, July 13, 2015. http://www.nytimes.com/2015/07/14/technology/accusing-amazon-of-antitrust-violations-authors-and -booksellers-demand-us-inquiry.html.

Streitfeld, David. 2015. "Amazon Reports Unexpected Profit, and Stock Soars." *New York Times*, July 23, 2015. http://www.nytimes.com/2015/07/24/technology/amazon-earnings-q2.html

Trujillo, Timothy J. 1994. "Note, Predatory Pricing Standards Under Recent Supreme Court Decisions and Their Failure to Recognize Strategic Behavior as a Barrier to Entry." *Journal of Corporation Law* 19: 809–13, 825.

Turner, Donald F., and Carl Kaysen. 1959. *Antitrust Policy: An Economic and Legal Analysis*. Cambridge: Harvard University Press.

Tuttle, Brad. 2010. "How Amazon Gets You To Stop Shopping Anywhere Else." *Time*, December 1, 2010. http://business.time.com/2010/12/01/how-amazon-gets-you-to-stop-shopping-any where-else.

Tuttle, Brad. 2016. "How Amazon Prime Is Crushing the Competition." *Time*, January 25, 2016. http:// time.com/money/4192528/amazon-prime-subscribers-spending.

Valentino-Devries, Jennifer, Jeremy Singer-Vine, and Ashkan Soltani. 2012. "Websites Vary Prices, Deals Based on User Information." *Wall Street Journal*, December 24, 2012. http://www.wsj.com/articles /SB10001424127887323777204578189391813881534.

Vara, Vauhini. 2015. "Is Amazon Creating a Cultural Monopoly?" *New Yorker*, August 23, 2015. http://www.newyorker.com/business/currency/is-amazon-creating-a-cultural-mono poly.

Wahba, Phil. 2015. "This Chart Shows Just How Dominant Amazon Is." *Fortune*, November 6, 2015. http://fortune.com/2015/11/06/amazon-retailers-ecommerce.

Weise, Elizabeth. 2015. "Amazon Prime Is Big, but How Big?" *USA Today*, February 3, 2015. http://www.usatoday.com/story/tech/2015/02/03/amazon-prime-10-years-old -anniversary/22755509.

Whitney, Lance. 2014. "Amazon Prime Members Will Renew Despite Price Hike, Survey Finds." *CNET*, July 23, 2014. http://www.cnet.com/news/amazon-prime-members-will-almost-all -renew-despite-price-increase.

Wingfield, Nick. 2016. "Amazon's Cloud Business Lifts Its Profit to a Record." *New York Times*, April 28, 2016. http://www.nytimes.com/2016/04/29/ technology/amazon -q1-earnings.html.

Wingfield, Nick, and Ravi Somaiya. 2015. "Amazon Spars with the Times over Investigative Article." *New York Times*, October 19, 2015. http://www.nytimes. com/2015/10 /20/business/amazon-spars-with-the-times-over-investigative-article.html.

Wohlsen, Marcus. 2014. "Amazon Takes a Big Step towards Finally Making Its Own Deliveries." *Wired*, September 25, 2014.http://www.wired.com/2014/09/ amazon -takes-big-step-toward-competing-directly-ups.

Woo, Stu. 2011. "Amazon 'Primes' Pump for Loyalty." *Wall Street Journal*, November 14, 2011. http://www.wsj.com/articles/SB10001424052970203503204577036 102353359784.

Wu, Tim. 2010. *The Master Switch: The Rise and Fall of Information Empires*. London: Atlantic Books

Yglesias, Matthew. 2013. "Amazon Profits Fall 45 Percent, Still the Most Amazing Company in the World." *Slate: Moneybox*, January 29, 2013. http://www.slate. com/blogs/moneybox /2013/01/29/amazon_q4_profits_fall_45 _percent.html.

Challenging Diversity—Social Media Platforms and a New Conception of Media Diversity

NATALI HELBERGER

INTRODUCTION

It is 2040.[1] The morning alarm goes, and wakes me with a careful selection of MindBook headlines. Not too negative, since the radio app knows that I am not exactly a morning person, and that bad news in the morning will negatively affect my socializability and productivity. In the bathroom, my smart mirror treats me to some compliments, and a well-balanced mix of news about health and lifestyle products, and recent tech-developments to slowly prepare me for another day at the faculty. After the first cup of coffee, MindBook considers the time ripe to present me with the more serious kinds of headlines—a new oil conflict in Antarctica, the election campaign in the United States is again in full swing, Turkey is in negotiations with Russia over Cyprus. I smile: my extra minutes on MindBook last night were well invested. Having spent half an hour clicking very purposefully on all the news about external relations, politics, and oil prices seemed to have helped to get me out of this news-about-climate-crisis-and-smart-cities loop I was stuck in for the better half of last week. Admittedly, it did help

---

1. The author would like to thank the editors for their thoughtful feedback and a stimulating discussion.

me a lot to prepare my presentation at the Ministry for Education, Culture and Science yesterday. And yet, sometimes I wish that getting the news was a little less . . .efficient. Since the decline of the general news media 30 years ago, getting the bigger picture has become more difficult.

Futuristic? A bit, but not excessively so. The way we find and receive news content is changing. Not rapidly but steadily. One key trend seems to be the fact that people access news content and media content more and more not only via traditional media but also via new information intermediaries, such as social media platforms, apps, and search engines (Reuters Institute for the Study of Journalism 2016, 2017; Pew Research Center 2016). These information intermediaries have stepped in to fill a critical gap in the news delivery chain: channeling attention and helping users to make a selection of the news that they find relevant. Information intermediaries often do not produce news themselves, neither do they see themselves as editors or as having the mission of providing citizens with a diverse set of information that we need in order to make informed choices. Rather, their business model is geared toward distributing news, connecting single articles with audiences, and realizing the advertising potential of different kinds of media content and target groups. And with the advances of data analytics and the increasing stock of data and intelligence about user preferences and interests, news has turned into a customizable product that can be carefully targeted and adjusted to individual recipients and the demands of advertisers. The presence of such data-driven, heavily targeted information intermediaries does not necessarily need to be a challenge to a diverse information environment, as long as there are alternative sources of information. But what to make of a situation in which there remain only one or a few dominant sources of information (as in the example of the fictional MindBook in the introductory scenario above)? And in the light of such a dominant player and a heavily targeted news environment, what are the prospects of still encountering diverse media content?

The focus of my chapter is on one of the central public policy objectives in media policy: media diversity.[2] I do not discuss other, equally important

2. Note that there is still considerable conceptual disagreement about the concrete meaning of the notions of "media pluralism" and "media diversity." Often, both notions are used interchangeably (see McGonagle [2011], speaking of "conceptual messiness"). McGonagle suggests a pragmatic approach in that pluralism refers to issues of media ownership and the choice of the public between different providers of services, whereas diversity refers to the range of programs and services available (2011). Along these lines, this chapter uses predominantly the notion of "diversity," and only uses "pluralism" where it is necessary to explicate the difference between issues of media ownership and the choice between different programs and services.

issues of platform dominance, such as the role of platforms in politics, their economic impact, and so forth, confident that many of those issues are covered by other chapters. Media diversity as a concept is deeply ingrained in our thinking about the role and contribution of the media in a democratic society and the idea that there shall be no one entity that can control (or dominate) the public debate. Instead, the media shall reflect the interests and needs of a heterogeneous society. In such a society, all voices have at least in principle the opportunity to make themselves heard. There is broad agreement that, as the Council of Europe has put it, "media pluralism and diversity of media content are essential for the functioning of a democratic society" and that "the demands . . . from Article 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms [right to freedom of expression] will be fully satisfied only if each person is given the possibility to form his or her own opinion from diverse sources of information" (Council of Europe 2007).

The close link between media diversity and democratic participation may also explain the vigor with which the rise of platforms and their growing influence in and on the media landscape are being met. The impact of their personalized recommendations and algorithmic filtering on users' information diet is subject to much concern and dystopian visions about filter bubbles and information bias but also targeted exclusion from news access. Depending on people's personal profile, users will get to see some kinds of information more, and others less or not at all. With the growing importance of a few large information intermediaries as sometimes the main source of information (Reuters 2017), the need to grasp the dynamics of these more centralized, data-driven (instead of editorially driven) news distribution models is ever more urgent. There is a general feeling of unease about the growing power and impact of platforms on users' media diets and yet, as Martin Moore aptly observed, it is not "[u]ntil we better understand and communicate the dilemmas they raise" that we will be able to find the effective policy responses (Moore 2016).

Regulators and policymakers across Europe are grappling with the question of what exactly the nature of these dilemmas is. Or to speak in the words of the British regulator Ofcom: "More fundamentally, the precise nature of future plurality concerns in the online news market are difficult to forecast."[3] Common to the discussions in countries such as the UK, but also Germany, France, and the Netherlands, is the difficulty of adequately conceptualizing and monitoring the impact of information intermediaries on

3. Ofcom, 2012, 27, para. 5.54.

the information landscape, or understanding where the true risks to media diversity lay.[4] The opacity of many of those platforms, and the secrecy that surrounds their algorithms and ordering mechanisms adds to this difficulty (Pasquale 2015), and requires entirely new methods of monitoring (Balazs et al. 2017). Understanding the nature of diversity concerns and potential sources of platform dominance is critical, however, to being able to identify adequate policy responses. This chapter aims to bring more conceptual clarity through developing a better understanding of platform power, how it can impact media diversity, and what the implications are for media diversity policies. In so doing, it concentrates on social media platforms. This is because of the particular role that these platforms play for news consumption but also because of the advances of at least some of these platforms into the business of distributing and aggregating news and media content. The main argument that this chapter makes is that with the arrival of information intermediaries, and social media platforms in particular, digital dominance can no longer be understood as the dominant control over content rights, outlets, or distribution channels, as used to be true with the traditional media. The true source of digital dominance is the ability to control the way people encounter and engage with information and the ability to steer their choices through the sheer knowledge about their interests and biases. More than ever media diversity has become the result of social dynamics, dynamics that are carefully orchestrated by one or few platforms. The chapter explains what implications this finding has for the way we measure and assess potential risks for media diversity on and from social platforms.

MEDIA DIVERSITY—WHAT IT IS AND WHY IT MATTERS, ALSO ONLINE

Does media diversity still matter? One could argue that in the digital information environment with its abundance of information media diversity has turned into a rather meaningless concept. Never was it possible to receive more information, not only from the national media but myriads of media companies, old and digital natives around the globe. This section will argue that "yes," media diversity still matters, but changing media consumption habits and the arrival of social media platforms requires us to further develop our conception of media diversity.

4. See, e.g., Ofcom, 2012, 25: in Germany: Die Landesmedienanstalten—ALM (2016), Digitalisierungsbericht Kreative Zerstörung oder digitale Balance: Medienplattformen zwischen Wettbewerb und Kooperation, Vistas: Berlin, 2016, in particular 14 and 74; in the Netherlands: Commissariaat voor de Media, Discussion Paper No. 877 09/2016.

Diversity policies are anchored in our ideas about functioning deliberation in a democratic society, and as such serve potentially a whole battery of goals and values, from inclusiveness, tolerance, and open-mindedness, well-informed citizens, and public deliberation, to a healthy, competitive media landscape and industry. Diversity in the media can create opportunities for users to encounter different opinions and beliefs, self-reflect on their own viewpoints (Kwon, Moon, and Stefanone 2015), enhance social and cultural inclusion (Huckfeldt, Johnson, and Sprague 2002), and stimulate political participation (Mutz 2006).

At the core of all the different values and objectives that diversity and diversity policies serve is dominance, or rather, the prevention of dominance and a situation in which one opinion, one ideology, one group or economic power dominates all others (Craufurd-Smith and Tambini 2012; Karppinen 2013; Valcke 2004). Whether one turns to the marketplace of ideas-rationale, or more deliberative or even radical conceptions of diversity—common to all of diversity's many conceptualizations (Karppinen 2013) is the ability of all voices to participate and seek an audience. The prevention of dominance as a core objective of diversity policies is also clearly reflected in the different regulatory options that have been deployed to protect and promote media diversity: the existing regulations are either concerned with avoiding and mitigating dominance or posing constraints on quasi-dominant parties so that they cannot abuse their economic and opinion power to the disadvantage of the democratic discourse and functioning media markets (Valcke 2004).

An example of the latter are the provisions that seek to promote *internal diversity of supply*, imposing more or less specific diversity requirements on one outlet. Typically that would be public service broadcasting, which, particularly in the earlier days of broadcasting, dominated the scene and was in many European countries *the* gateway to audiovisual information. Accordingly, public service broadcasting (and to a lesser extent other media services), were obliged to "enable different groups and interest in society—including linguistic, social, economic, cultural or political minorities—to express themselves" (Council of Europe 1999). Regulatory obligations to promote internal diversity or the diversity of a particular media outlet or platform include measures that guarantee a diverse composition of the programs of the public service broadcaster, provisions with the goal of protecting editorial independence, specific pluralism safeguards such as program windows, frequency sharing arrangements, provisions about the diversity of staff and program councils, list of important events, and quota rules.

Then there are measures that are directed at protecting and promoting of often referred to as *structural or external diversity*, most prominently the media-ownership rules. Ownership rules have traditionally formed the core of regulators' response to the trend toward commercialization and liberalization of the media (Karppinen 2013), with the goal of "prevent[ing] or counteract[ing] concentrations that might endanger media pluralism at the national, regional or local levels" (Council of Europe 1999, appendix, para. I). Then there are licensing requirements, the obligations to media transparency (see, extensively, Council of Europe 1994) or must-carry, due prominence rules and access obligations (Helberger, Kleinen-von Köngislöw, and Van Der Noll 2014; Council of Europe 2007).

Next to the diversity and pluralism of supply, there is also *diversity of exposure* to consider, that is, the question of how diverse the selection of content and speakers is that users are ultimately exposed to and consume. As the Council of Europe acknowledged, "pluralism is about diversity in the media that is made available to the public, which does not always coincide with what is actually consumed" (Council of Europe 1999). This is an observation confirmed by research finding that an increase in the diversity of content can under certain circumstances actually lead to a decrease in the diversity of the content consumed (Napoli 1999; Ferguson and Perse 1993; Cooper and Tang 2009; Wojcieszak and Rojas 2011). This is because people have only so much time and attention to spend on consuming media content. The greater the diversity of content, the greater the need to filter and select.

Filtering and selecting media content is an important function of information intermediaries, such as search engines and social media platforms. Their main goal is to channel audience attention and affect access to and the diverse choices people make. As such, they affect not so much the diversity of supply (social media platforms do not produce content), but rather the diversity of media content individual members of the audience are eventually exposed to (exposure diversity). And the key question is: are platforms an opportunity or threat to media diversity (and exposure diversity in particular)?

Are Social Media Platforms an Opportunity or Threat to Diversity?

The question to what extent social media platforms have added new opportunities or challenges for diversity is not easily answered. There is a growing body of research that finds evidence for a positive contribution of social media platforms to media diversity, and diversity of exposure

in particular. In its 2017 News Report, the Reuters Institute, found that users of social media were significantly more likely than non users to see sources they would not otherwise use. This finding echoes earlier research that finds that use of social media platforms can result in exposure to more diverse news (Bakshy, Messing, and Adamic 2015; incidental exposure to news: Lee, Lindsey, and Kim 2017, stressing the importance of heterogeneity of networks for this; in a similar direction Messing and Westwood 2014; or exposure dissenting opinions: Diehl, Weeks, and Gil de Zúñiga 2016). Others find evidence to the contrary, for example a lesser likelihood for exposure to cross-ideological content (Himelboim, McCreery, and Smith 2013) and the existence of echo chambers due to confirmation bias (Quattrociocchi, Scala, and Sunstein 2016). Yet others produce mixed evidence (Flaxman, Goel, and Rao 2016; Lee et al. 2014, finding that while social media platforms can increase exposure to diverse news, people who are more active in political discussions on SNSs are more likely to be polarized; Stroud 2008, on the role of exposure to particular kinds of content; Lee, Lindsey, and Kim 2017 on information overload as moderating factor, or Anspach 2017, on the importance of settings and the role that shares, likes, and comments can play for engagement).

What this research shows is that for social media, different factors than in the traditional media determine the level of diversity users are exposed to. Such factors can include the settings of the filtering and recommendation algorithms, and which kinds of content the algorithm decides to prioritize or suppress. Inasmuch, the MindBook example and the potential of its recommender to narrow down the information diet to a choice of selected topics that the algorithm considers relevant is far from being futuristic. If there is one aspect that the debates about Facebook Newsfeed have made clear, it is the impact of the recommendation mechanism on the selection of contents in users' newsfeed and the fact that the criteria that determine the selection differ strongly from the editorial criteria that matter in the traditional media (Devito 2017; Bucher 2012).

This raises the more normative question to what extent diversity and pluralism *should* still matter in the context of social media platforms. Social media platforms are not media in the traditional sense, nor is their main purpose to inform, and do so in a way that reflects the diverse topics and voices that constitute our democratic societies (Devito 2017). Still, the Council of Europe highlights the importance of diversity in the context of the criteria according to which search results are selected, ranked, or removed (Council of Europe 2012). The importance of diversity as a regulatory goal has also been highlighted by UK regulator Ofcom in its review of its diversity and plurality measurements (Ofcom 2012). In this

context, Ofcom referred explicitly to the opportunities but also challenges that come from digital platforms. According to Ofcom, "[t]here is a risk that (social media) recommendations are used in a manner that narrows citizens exposure to different points of view, by reinforcing their past habits or those of their friends" (Ofcom 2012, 25). And further: "If however they were to start exercising a greater degree of editorial control in the future, then this could raise significant plurality concerns" (Ofcom 2012, 26). In a similar vein, in the Netherlands, the Dutch regulatory authority for the media sector observed that the true risk online is not so much that the overall offer will be less diverse, but rather that the offer that is *accessible to (individual) users* may be less diverse as a result of algorithmic filtering such that users are not even aware of the size and diversity of the overall offer.[5] Seeing the growing importance of social media platforms for the way users encounter and engage information (Reuters 2017), and the impact that at least the larger platforms exercise on the overall structure of news markets and information flows (Moore 2016; Kleis Nielsen and Ganter 2017), there are strong reasons to argue that diversity should matter, in one way or other, also in the context of social media platforms. And if one follows Karppinen in ultimately conceptualizing diversity as a matter of distributing communication power (Karppinen 2013, 114), it becomes clear that leading social media sites cannot be left outside diversity considerations.

The question then is not so much whether media diversity still matters in a platform context. It does. The question is rather how and in which form. One of the reasons why policymakers find it so difficult to understand and handle the issue of diversity on social media platforms is that exposure diversity as a normative goal is still little understood and only beginning to trigger a—much needed—discussion (Craufurd-Smith and Tambini 2012; Helberger 2012; Valcke 2011). The other reason is that in order to be able to understand risks and opportunities from social media platforms potentially for diversity and pluralism (as normative goals), it is necessary to understand how exactly platform power is affecting the realization of media diversity and pluralism. The following sections therefore develop a conceptual framework to better understand the risks to, and opportunities from, social media for diversity and pluralism.

5. Commissariaat voor de Media, 15 Jaar Mediamonitor. Van Mediaconcentratie naar Mediagebruik, Commissariaat voor de Media: Amersfort, 2017, 46 (in Dutch). http://www.mediamonitor.nl/wp-content/uploads/15-Jaar-MM-pdf-LR.pdf.

UNDERSTANDING PLATFORM POWER AND DIVERSITY

In order to understand the true impact of platforms on media diversity and pluralism, it is important to view platforms in terms of their business models (and economic incentives)—the means that they use to distribute content and their role in the wider information ecology. From the point of view of diversity policies, it is also important to understand how platforms differ from the more traditional media, such as broadcasting and newspapers. This is because existing policies have been written with the more traditional media in mind, and the differences between information intermediaries and traditional media may explain why the traditional instruments are only in part suitable to address new challenges to media diversity and pluralism. Having said so, it is also important to realize that platforms are undergoing highly dynamic transitions.

Facebook is a good example. Having set out as essentially a "tech" company, for a long time news and media content were not their core business (Van Dijk, Poell, and De Waal 2016). The core business of social media platforms was providing social media services,[6] and connecting people, content producers, and advertisers. Inasmuch, social media platforms are not like traditional media, and have different business incentives. Media content has not so much the function of informing people and keeping up with an editorial mission, but rather of fueling social interactions and forming the backdrop for advertising campaigns and initiatives to keep people on the website longer. As with many user-created-content sites, however, soon the realization dawned that cat videos and vacation pictures can go only so far in arresting the attention of users—a realization that led to an increasing interest in professional content on many of these sites. Examples are Facebook's Instant articles, Trending topics, Twitter's Moments, YouTube's commissioning of professional media content, or Google's News Initiative. Common to all these initiatives is the wish to integrate professional media content into their platform—without actually producing it. As a result, the

---

6. Twitter: Connect with your friends—and other fascinating people. Get in-the-moment updates on the things that interest you. And watch events unfold, in real time, from every angle. Facebook: Connect with friends and the world around you on Facebook. Instagram: Instagram is a fun and quirky way to share your life with friends through a series of pictures. Snap a photo with your mobile phone, then choose a filter to transform the image into a memory to keep around forever. We're building Instagram to allow you to experience moments in your friends' lives through pictures as they happen. We imagine a world more connected through photos. YouTube: Geniet van je favoriete video's en muziek, upload originele content en deel alles met vrienden, familie en anderen op YouTube. Wechat: Connecting 800 million people with chats, calls and more.

relationship to professional media producers, and impact on the overall media landscape, became increasingly complicated. And it is symptomatic how M. Zuckerberg, CEO of Facebook, has moved within a relatively short time from claiming, "We're a technology company. We're not a media company,"[7] to the observation: "Facebook is a new kind of platform. It's not a traditional technology company. It's not a traditional media company. You know, we build technology and we feel responsible for how it's used. . . . We don't write the news that people read on the platform. But at the same time we also know that we do a lot more than just distribute news, and we're an important part of the public discourse."[8]

So what is it exactly that constitutes the communicative power of social media platforms such as Facebook, and that we need to be aware of when debating dominance, and the potential implications that these new players have for media pluralism and diversity? In this context, it is useful to return to the distinction between structural and internal diversity from the previous section.

Social Networks and Structural Diversity

Social media platforms do not so much affect the diversity of supply with different voices from different sources. These voices are still free to exist outside the structure of the social media platforms. Maybe the platforms' greatest structural impact on diversity is in the way that they affect diversity of exposure and media consumption and control the users' attention. This way they can affect not only the diversity of contents and plurality of sources that users encounter within the social media platform but also the vitality and diversity of the overall media landscape (since the media rely for their economic survival on access to users, and users' attention).

Social media platforms stage encounters with media content, affect the "findability" of content, order and prioritize existing content, manage and direct user attention as a scarce resource, and influence the choices users make. This happens not only through offering basic search functionality but also through algorithmic or collaborative filtering and display of personalized search results and recommendations (Schulz, Dreyer, and Hagemeier 2011; European Commission 2013; Council of Europe 2012; Van Hoboken

---

7. http://money.cnn.com/video/technology/2016/08/29/facebook-ceo-were-a-technology-company-were-not-a-media-company-.cnnmoney/.
8. https://www.theguardian.com/technology/2016/dec/22/mark-zuckerberg-appears-to-finally-admit-facebook-is-a-media-company

2012). In other words, sources of communicative power or even dominance are not, as in the traditional media, the resources to produce content, IP rights, and expertise. Instead, the source of communicative power of social platforms is the control over powerful sorting algorithms and data—data about their users, about the way users engage with content, and about the best, most effective way of pushing content under the attention of users. Thereby, social media platforms are instrumental in a more conceptual shift from mass-to personalized modes of distributing media content. This is a shift in which it is not so much ownership and control over content that matters, but knowing the users, and establishing the knowledge, relationships, and technical infrastructure to trigger the engagement of users with particular types of content. This is a shift from a situation in which the news media function as our main sources of information, to a situation in which a "MindBook" sorts our information exposure according to its own logics and users' preferences (Devito 2017).

From this it follows that the real problem with structural diversity is not so much ownership over a particular resource. The true challenge from platforms for structural diversity lies *in the relationship* between those making media content and those "owning users," their data, and the tools and technologies to distribute media content and arrest (or even monopolize) users' attention. This also means that concerns about structural diversity are no longer easily solved by counting the number of sources and diversity of content in media markets, nor will the traditional measures to protect and promote structural diversity be particularly useful in protecting and promoting structural diversity in a platform context. Instead, the key to dealing with platform power and structural diversity is to balance negotiation power, protect media independence, and ensure a fair, level playing field.

Balancing Negotiation Power

So far, the relationship between the old media and the new intermediary platforms takes the form of bilateral negotiations between traditional media outlets and the intermediaries. As Kleis Nielsen and Ganter (2017) find, these relationships can be both symbiotic and asymmetrical. "Digital intermediaries may need news in a broad sense, or at least benefit from it. But it is not at all clear that they need any one individual news media organisation, even large ones" (Kleis Nielsen and Ganter 2017). And while private ordering and the way platforms manage their relationships with users has been subject to growing attention from the perspective of contract law

and the fairness of terms of use (Loos and Luzak 2016; Wauters, Lievens, and Valcke 2014), a parallel discussion of the fairness of the terms in the agreements between media companies and publishers and broadcasters is still largely missing. Arguably, future media diversity policies need to add to their toolbox means to assess the fairness of deals in such asymmetrical relationships, as well as ways of improving the negotiation power between publishers and information intermediaries. This can include not only initiatives to promote the transparency of such deals, to stimulate the media to bundle their forces, but also ways to stimulate the openness of collaborations with third parties (similar to the way in which, e.g., telecom operators have a negotiation duty) and scrutiny of the fairness of the conditions under which media content is presented and distributed via platforms (e.g., brand visibility, client management, and data and revenue sharing). Inasmuch, the tools developed in telecommunications law (and under the European Access Directive in particular)[9] might provide an interesting route to learn from as an area in which the regulator has developed a system of assessing the fairness and openness of B2B negotiations, also and particularly from the perspective of their impact on the openness, competitiveness of, and choice within communications markets.

The Importance of Media Independence

One structural problem or danger in any asymmetrical relationship is dependence. The aspect of dependency has been also identified by Nielsen and Ganter in their study, who point in this context to "a tension between (1) short-term, operational, and often editorially led consideration and (2) more long-term strategic considerations focused on whether the organisation will become too dependent on these intermediaries for reaching audiences, and in the process will control over its editorial identify, access to user data, and central parts of its revenue model" (Kleis Nielsen, Ganter 2017). The problem of dependencies deserves to be taken seriously, particularly from the perspective of the role that the media play in the realization of the fundamental right to freedom of expression as public watchdog and fourth estate—a role that they can play only if they remain independent from states as well as from commercial power. Dommering warns that the traditional media are at risk of losing more and more of their identity

9. Directive 2002/19/EC of the European Parliament and of the Council of March 7, 2002, on access to, and interconnection of, electronic communications networks and associated facilities (Access Directive), OJ L 108, 24.4.2002, 7–20.

in their attempt to assimilate and create a functional symbiosis between themselves and intermediaries (Dommering 2016). And Van Dijk & Poell point to the risk of new dependencies as the result of a shift in the news process from "an editorial logic to an algorithmic logic," a shift whose main driver are platforms (Van Dijk, Poell, and De Waal 2016). Media law and policy in Europe have a long tradition of dealing with the independence of the media, be it the constitutional safeguards in Art. 10 ECHR against state censorship, or the extensive rules on advertising, sponsoring, and separation of editorial and commercial content in the relationship to commercial players. It is high time to revisit these rules in light of the intrinsic relationship between the media and information intermediaries.

Fair, Level Playing Field

Finally, the point about the fair, level playing field relates to the question whether it is still justified to treat offline and online media differently, and impose far stricter rules and diversity expectations/requirements on the former while maintaining a light-touch approach for the latter. For a long time, the key argument for justifying a stricter regulation of the broadcasting media vis-à-vis the online media has been the alleged persuasiveness of video (Barendt 1993). One may wonder to what extent broadcasting is still more persuasive than the communication of media content via, for example, social media platforms. Arguably, social media platforms can have an equal if not more persuasive impact, particularly if those platforms use the deep insights they have about users to refine their targeting into persuasion strategies. What is more, these platforms have the tools and power to engage users to act on information, and influence civic behavior (Moore 2016, 54). The difficulty here is understanding the true nature of editorial control/responsibility and diversity on social media platforms.

To draw a preliminary conclusion: when assessing platform power (or even dominance) over a media sector, new benchmarks need to be developed that include the amount of consumer data, characteristics of the recommendation algorithm, and number of users, activity of users, *and also* the balance in the contractual conditions between platforms and media companies, the level of independence of the media from platforms, and the existence of an equal level playing field). Doing so may also require new forms of monitoring and measuring diversity, for example, in order to be able to ascertain the level of diversity that different categories of users on different platforms are eventually exposed to.

Internal Diversity

Internal diversity considerations figure very prominently in the ongoing public policy debate about impact and responsibility of information intermediaries. These are fears about filter bubbles and echo chambers (see Pariser 2011; Sunstein 2000; High Level Expert Group on Media Freedom and Pluralism 2013). But these fears must be seen in context: as long as people have the opportunity to receive information from different sources (multisourcing), the fact that they receive a less diverse information diet on one platform can be counterbalanced by access to more diverse information on another, for example, in the public service media or the traditional press (see Zuiderveen Borgesius et al. 2016; Nguyen et al. 2014; Schmidt et al. 2017). To the contrary, in a situation in which one particular platform has become the dominant source of information (as in the MindBook example in the introduction), the internal diversity of that platform does matter. In the MindBook example, alternative sources have been crowded out of the market, and with them also the opportunity for users to learn what information they might be missing in their MindBook-only diet. Seen from this perspective it also becomes so evident why platform dominance is, or should be, of such concern for media policymakers, and why it is important to protect and promote structural diversity. In addition, and with the growing relevance of (some) platforms as important and maybe even exclusive gateway to information access (Reuters Institute for the Study of Journalism 2017), questions of internal diversity within the platforms come to the fore.

In the public policy debates so far there have been no shortage of suggestions of how to hold platforms more accountable for the diversity within their platforms and to impose internal diversity safeguards (Paal 2012; Foster 2012; Neuberger and Lobgis 2010; Schulz, Dreyer, and Hagemeier 2011). The problem with most of these suggestions, and the real challenge for future media law and policy here, is understanding how diversity works on social media platforms and what the actual contribution of platforms is to internal diversity.

Taking into account the growing number of users for whom social media platforms are the main gateway to accessing and experiencing media content, the issue of internal diversity becomes more and more pressing, and also infinitely more complex. This is because diversity on social media platforms is no longer a matter of an editor who determines what a (sufficiently) diverse mix of contents is. Diversity is increasingly also a matter of how users engage with that content, share, prioritize, like, or dislike it, and the extent to which the architecture and design of a social media platform

enables and steers such engagement. In other words, to truly understand the impact on, and power of platforms over, the internal diversity within the platform it is important to understand the impact that platforms have not only on the selection of the content itself but also on the conditions under which users encounter and can engage with content. This is essentially a user-driven perspective on diversity that corresponds to the social character of platforms.

To give but two examples: Filtering, search, and self-selected personalization are examples of activities by which users themselves actively influence the diversity of contents they wish to be exposed to (Van Hoboken 2012). And through activities such as liking, flagging, rating, and sharing, users can actively influence which contents others are exposed to (Gerlitz and Helmond 2013; Crawford and Gillespie 2016). Engagement and using (diverse) content is critical to deliberate, show different perspective, or form an opinion. On social platforms, users can actively engage with diverse content in the form of actively contributing to the deliberation (through blogs, posts, comments, etc.). They can also engage in more symbolic ways, for example, through liking, voting, rating, and so forth.

Platforms create the organizational framework and opportunities for exposure to and engagement with diverse content. Inasmuch, social media platforms not only distribute media content but also create their very own versions of "privately controlled public spheres," in which users not only encounter diverse content but also engage and deliberate, share and contest. This is where their true contribution to and power over diversity within the platform lies. And this is also where their social responsibility lies. Platforms' influence on news distribution and exposure, and ultimately diversity, can include measures and design decisions at the level of content (e.g., providing opportunities for UGC, and for user-led editing), engagement (possibilities to comment, post, or express consent or dissent), and network (through the ability to create groups, invite friends, etc.). Three examples may demonstrate my point in more detail, but also how the "diversity-by-architecture" perspective may provide new interesting avenues for diversity policies and research:

Diversity- versus Popularity-Based Recommender Design

The first and probably most obvious example is the settings of the recommender algorithms. Search, personalization, and recommendation play a rather pivotal role for both exposure to information and diverse exposure (Van Hoboken 2012). How important that role can be has been proven

once again by the fierce controversy around Facebook's Trending Topics algorithm and claims of bias. A closer look at the editorial guidelines and instructions for the human editors of Trending Topics revealed that considerations of media diversity were more or less absent in Trending Topics (meanwhile Facebook has again changed its algorithm and probably also the editorial guidelines in response to the Trending Topics criticism). Trending Topics editors were, for example, asked to get a good overview of what is trending, the Facebook Trending algorithm that notes whether topics are disproportionally often mentioned, engagement (likes, comments, and shares) and what the headlines from top news sites suggest that is trending, namely a selection of news websites that is strongly US/UK centered.[10] Arguably, the Trending Topics algorithm thereby completely failed to reflect the diversity of the media scene in Europe, local content, and so forth. More generally, many recommender systems display a certain bias toward popular recommendations or recommendations that reflect individual interests and personal relevance (DeVito 2017) (such as in the MindBook example). To the contrary, it is, at least technically, also possible to program recommendation algorithms in a way to promote more diverse exposure to contents (Adomavicius and Kwon 2011; Munson and Resnick 2010; Helberger, Karppinen, and D'Acunto 2018). More sophisticated recommendation algorithms that also take into account medium-term objectives such as diversity, or at least giving users a choice between different recommendation logics, may have a positive effect on the diversity of content users are exposed to. Also, there are more and more third-party tools and applications available whose objective it is to make people aware of their filter bubbles, to encourage them to diversify their media diet, and to stimulate their curiosity.[11] Stimulating initiatives like these, and giving prominence to such tools, could be a new and potentially far more effective approach in fostering diversity on platforms than traditional policy responses, such as the must-carry or due prominence rules suggested earlier (Foster 2012; Danckert and Mayer 2010; European Commission 2013). Arguably, dominance thereby also becomes a question of how open platforms are to alternative recommendation settings and technologies on their platforms that help users to critically question the recommendations by one party (e.g., the platform), and discover alternative recommendations by others.

---

10. BBC News, CNN, Fox News, *The Guardian*, NBC News, *New York Times, USA Today*, the *Wall Street Journal, Washington Post,* BuzzFeed News, see https://fbnewsroomus.files.wordpress.com/2016/05/full-trending-review-guidelines.pdf

11. E.g., Huffington Posts' "Flipside"; BuzzFeed's "Outside Your Bubble"; "Read across the Isle"; "Blue Feed, Red Feed"; "Escape Your Bubbles"; "Filterbubblan.se"; "AllSides."

Diversity of the Personal Social Media Platform

A growing body of research does show that the diversity and heterogeneity of the social media platform is an important aspect for the quality of the deliberative process, and the openness toward other ideas and viewpoints (Jun 2012; Huckfeldt, Johnson, and Sprague 2002; Bakshy, Messing, and Adamic 2015; Messing and Westwood 2014). And while it is true that it is primarily users who decide who will be in their social media platform, social media do exercise some influence here as well (Diehl, Weeks, and Gil de Zúñiga 2015). Facebook, for example, suggests not only certain groups and friends but also whom to (not) follow, by making recommendation for pages similar to the ones one already follows. Note that the only option offered so far is "pages similar to," and not "pages other than" or "pages likely to provide a contrasting viewpoint." Inasmuch, social platforms could learn from research that shows that the presence of dissenting minority views in a group can promote openness toward, and consideration of alternatives at a group level and enhance problem-solving (Nemeth 1986). More generally, the extent to which users encounter cross-cutting content also depends on who their friends are (Bakshy, Messing, and Adamic 2015). Accordingly, stimulating the deliberate inclusion of such minority or contrasting actors could be a way to improve the quality and diversity of engagement on social media platforms. Understanding better the dynamics of diverse engagement on social media platforms, and how personal, social, and contextual characteristics contribute to diversity of exposure may be another way of stimulating diversity online (compare, e.g., Bramoullé and Rogers 2009; Swapneel et al. 2011). Furthermore, such understanding can inform the architectural design choices that stimulate engagement with a (heterogeneous) group of friends (Anspach 2017).

Privacy and Diversity

The final example to be discussed here are the privacy settings that are offered by social media. At first sight, privacy and media diversity may not appear to have much in common, but they do. Kwon, Moon and Stefanone show, for example, that the privacy affordances that are provided by a social medium can have an effect on the way users post and engage with content, including less popular, counterattitudinal content and content reflecting minority opinions (Kwon, Moon, and Stefanone 2015). On a more fundamental level, media diversity, as a constituting factor of freedom of expression and the role of the media in a democratic society, can

only function if users enjoy a certain autonomy, that is, independence from the government or commercial forces, in making their decisions and weighting the arguments. Privacy rights, for example, can provide the necessary democratic breathing space for individuals to form their distinct and diverse identities and ideas (Richards 2008; Cohen 1996). Put differently, protecting the privacy of users, in their relationship to the media and also to advertisers and other third parties that seek to influence the way users choose and reflect on media content, is a way of protecting the very values that we hope to promote with media diversity: critical and diverse thinking. Dawes speaks in this context of a "political privacy" dimension. He argues, "[v]iewed from a civic republican perspective, therefore, the political legitimacy of the state is guaranteed by the public sphere, which in turn is dependent upon privacy" (Dawes 2014).

None of the aspects mentioned here—diversity of the recommender system, diversity of the social media platform, and level of respect for users' privacy and autonomy—are among the traditional benchmarks for assessing media diversity or dominance. And yet, as this analysis has demonstrated, these are factors that matter in the dynamic and user-driven construction of diversity online. One very concrete conclusion from this is that the assessment of diversity and the ability of particular parties to dominate the media landscape online not only must follow established criteria (such as the number of sources available, the diversity of categories of content presented, etc.) but also must be able to incorporate new criteria, including the extent to which users are (truly) free to choose between different sources and contents and enjoy both the options and the autonomy to do so.

CONCLUSIONS

This chapter has sought to sketch the contours of a new conception of media diversity, one that is able to take into account the new, deeply social dynamics of platforms in online media markets. It has argued that in order to truly understand the platforms' potential impact on media diversity and media pluralism, it is critical to look at platforms not in isolation but in their relationship toward (1) other media outlets that they distribute and (2) users.

The true impact of information intermediaries such as social media platforms on media diversity is not so much whether they are willing and able to present users with diverse packages of information in the sense that traditional media editors do. The contribution of social platforms runs

much deeper: they create the organizational and architectural framework and opportunities for exposure to and engagement with diverse content. This also means that diversity as a value or even public policy objective on social media platforms has not the same meaning as diversity in the traditional media context. Media diversity on social media platforms must be understood as a cooperative effort of the social media platform, media organizations, and users. The way users search for, engage with, like, shape their network, and so forth, has important implications for the diversity of contents, ideas, and encounters that they are exposed to. Similarly, the way the traditional media collaborate with information intermediaries to distribute content and reach viewers impacts structural diversity.

When seen from this perspective, it becomes clear that the impact of information intermediaries on media diversity is not easily understood (or monitored) with existing mainstream conceptions and measures of diversity (such as the diversity of opinions and ideas from different sources). For the same reasons existing diversity safeguards are of only limited use in protecting and promoting diversity within, and in the presence of, powerful social media platforms. Instead, future diversity policies need to turn their attention to (1) the relationship between traditional media and information intermediaries, with the goal of establishing a more equal level playing field and structural diversity; and (2) the relationship between platforms and users, with the goal of promoting the architectural and organizational measures for users to be able to encounter and engage with diverse content.

This also means that when assessing the impact of platforms on the diversity of media markets we need to include methods and factors into the media regulators' toolbox that may go beyond the traditional framework for assessing dominance. Such factors can include the balance in the contractual conditions, control over data, or sophisticated recommendation algorithms between platforms and media companies, the level of independence of the media from platforms, and the existence of an equal, level playing field. It can also include factors such as the openness toward alternative recommendation metrics and the extent to which users are truly free in choosing among different voices and opinions online.

REFERENCES

Adomavicius, G., and Y. Kwon. 2011. "Maximizing Aggregate Recommendation Diversity: A Graphtheoretic Approach." *Proceedings of the 1st International Workshop on Novelty and Diversity in Recommender Systems* (DiveRS 2011, Chicago) 2011.

Anspach, N. M. 2017. "The New Personal Influence: How Our Facebook Friends Influence the News We Read." *Political Communication*, online first, May 23, 2017. http://www.tandfonline.com/doi/abs/10.1080/10584609.2017.1316329

Bakshy, E., S. Messing, and L. A. Adamic. 2015. "Political Science: Exposure to Ideologically Diverse News and Opinion on Facebook." *Science (New York, N.Y.)* 348, no. 6239: 1130.

Balazs, B., N. Helberger, K. Irion, F. Zuiderveen-Borgesius, J. Möller, B. Van Es, B. Van der Velde, and C. De Vreese. 2017. "Tackling the Algorithmic Control Crisis: The Technical, Legal, and Ethical Challenges of Research into Algorithmic Agents." *Yale Journal of Law and Technology* 19: 133–80.

Barendt, E. M. 1993. *Broadcasting Law: A Comparative Study*. Oxford; New York: Oxford: Clarendon Press.

Bucher, T. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media and Society* 14, no. 7: 1164–80.

Bramoullé, Y., and B. Rogers. 2009. "Diversity and Popularity in Social Networks." Working Paper 09-03, Centre Interuniversitaire sur le Risque, les Politiqies Économiques et l'Employ.

Cohen, J. 1996. "A Right to Read Anonymously: A Closer Look at 'Copyright Management' In Cyberspace." *Conneticut Law Review* 28: 981.

Cooper, R., and T. Tang. 2009. "Predicting Audience Exposure to Television in Today's Media Environment: An Empirical Integration of Active-Audience and Structural Theories." *Journal of Broadcasting and Electronic Media* 400–418.

Council of Europe. 1994. "Recommendation No. R(94)13 of the Committee of Ministers to Member States on Measures to Promote Media Transparency. R(94)13." Strasbourg: Council of Europe.

Council of Europe. 1999. "Recommendation No R(99)1 of the Committee of Ministers to Member States on Measures to Promote Media Pluralism." Strasbourg: Council of Europe.

Council of Europe. 2007. "Recommendation CM/Rec(2007)3 of the Committee of Ministers to Member States on the Remit of Public Service Media in the Information Society." Strasbourg: Council of Europe.

Council of Europe. 2012. "CM/Rec(2012)3 of the Committee of Ministers to Member States on the Protection of Human Rights with Regards to Search Engines." CM/Rec(2012)3. Strasbourg: Council of Europe.

Craufurd-Smith, R., and D. Tambini. 2012. "Measuring Media Plurality in the United Kingdom: Policy Choices and Regulatory Challenges." *Journal of Media Law* 4, no. 1: 35–63.

Crawford, K., and T. Gillespie. 2016. "What Is a Flag For? Social Media Reporting Tools and the Vocabulary of Complaint." *New Media and Society* 18, no. 3: 410–28.

Danckert, B., and F. J. Mayer. 2010. "Die vorherrschende Meinungsmacht von Google." *MultiMedia und Recht* 4: 219–22.

Dawes, S. 2014. Press Freedom, Privacy and The Public Sphere. *Journalism Studies* 15, no. 1: 17–32.

Devito, M. 2017. "From Editors to Algorithms. A Value-Based Approach to Understanding Story Selection in the Facebook News Feed." *Digital Journalism* 5, no. 6: 753–73.

Diehl, T., B. E. Weeks, and H. Gil de Zúñiga. 2016. "Political Persuasion on Social Media: Tracing Direct and Indirect Effects of News Use and Social Interaction." *New Media and Society* 18, no. 5: 1875–95.

Dommering, E. J. 2016. "Het Verschil van Mening: Geschiedenis van een Verkeerd Begrepen Idee." Amsterdam: Uitgeverij Bert Bakker.

European Commission. 2013. *Preparing for a Fully Converged Audiovisual World: Growth, Creation and Value*. Brussels: European Commission.

Ferguson, D., and E. Perse. 1993. "Media and Audience Influences on Channel Repertoire." *Journal of Broadcasting and Electronic Media* 37, no. 3: 31–47.

Flaxman, S., S. Goel, and J. Rao. 2016. "Filterbubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80: 298–320.

Foster, R. 2012. *News Plurality in a Digital World*. London: Reuters Institute for Journalism.

Gerlitz, C., and A. Helmond. 2013. "The Like Economy: Social Buttons and the Data-Intensive Web." *New Media and Society* 15, no. 8: 1348–65.

Helberger, Natali, Katharina Kleinen-Von Köngislöw, and Rob Van Der Noll. 2014. *Convergence, Information Intermediaries and Media Pluralism—Mapping the Legal, Social and Economic Issues at Hand*. Amsterdam: Institute for Information Law.

Helberger, N. 2012. "Exposure Diversity as a Policy Goal." *Journal of Media Law* 4, no. 1: 65–92.

Helberger, N., K. Karppinen, and L. D'Acunto. 2018. "Exposure Diversity as a Design Principle for Recommender Systems." *Information, Communication and Society* 21, no. 2: 191–207.

High Level Expert Group on Media Freedom and Pluralism. 2013. *A Free and Pluralistic Media to Sustain European Democracy*. Brussels: High Level Expert Group on Media Freedom and Pluralism.

Himelboim, I., S. McCreery, and M. Smith. 2013. "Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter." *Journal of Computer-Mediated Communication* 18, no. 2013: 154–74.

Hoboken, J. V. 2012. *Search Engine Freedom: On the Implications of the Right to Freedom of Expression for the Legal Governance of Web Search Engines*. Alphen aan den Rijn Kluwer Law International; Frederick, MD: Aspen Publishers; Biggleswade, Bedfordshire, UK: Turpin Distribution Services.

Huckfeldt, R., P. E. Johnson, and J. Sprague. 2002. "Political Environments, Political Dynamics, and the Survival of Disagreement." *Journal of Politics* 64, no. 1: 1–21.

Jun, N. 2012. "Contribution of Internet News Use to Reducing the Influence of Selective Online Exposure on Political Diversity." *Computers in Human Behavior* 28, no. 4: 1450–57.

Karppinen, K. 2013. *Rethinking Media Pluralism*. 1st ed. New York: New York: Fordham University Press.

Kleis Nielsen, R., and S. A. Ganter. 2017. "Dealing with Digital Intermediaries: A Case Study of the Relations between Publishers and Platforms." *New Media and Society*, Online first, published on April 17, 2017, 1–18.

Kwon, K., S. Moon, and M. Stefanone. 2015. "Unspeaking on Facebook? Testing Network Effects on Self-Censorship of Political Expressions in Social Network Sites." *Quality and Quantity* 49, no. 4: 1417–35.

Lee, J. K., J., Choi, C. Kim, and Y. Kim. 2014. "Social Media, Network Heterogeneity, and Opinion Polarization." *Journal of Communication* 64, no. 2014: 702–22.

Lee, S. K, N. J. Lindsey, and K. S. Kim. 2017. "The Effects of News Consumption via Social Media and News Information Overload on Perceptions of Journalistic Norms and Practices." *Computers in Human Behaviour*, online first, May 5, 2017.

Loos, M., and J. Luzak. 2016. "Wanted: A Bigger Stick. On Unfair Terms in Consumer Contracts with Online Service Providers." *Journal of Consumer Policy* 39, no. 1: 63–90.

McGonagle, T. 2011. *Minority Rights, Freedom of Expression and the Media: Dynamics and Dilemmas*. Cambridge: Intersentia.

Messing, S., and S. Westwood. 2014. "Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online." *Communication Research* 4, no. 2014: 1042–63.

Moore, M. 2016. *Tech Giants and Civic Power*. King's College London: Centre for the Study of Media, Communication & Power.

Munson, S., and P. Resnick. 2010. *Proceedings of the 28th International Conference on Human Factors in Computing Systems,* CHI'10. New York, NY: ACM, 1457–66.

Mutz, D. C. 2006. *Hearing the Other Side: Deliberative versus Participatory Democracy*. Cambridge: Cambridge University Press.

Napoli, P. 1999. "Deconstructing the Diversity Principle." *Journal of Communication* 49, no. 4: 7–34.

Nemeth, C. J. 1986. "Differential Contributions of Majority and Minority Influence." *Psychological Review* 93, no. 1: 23–32.

Neuberger, C., and F. Lobgis. 2010. *Die Bedeutung des Internets im Rahmen der Vielfaltssicherung*. Berlin: Vistas.

Nguyen, T., P. Hui, F. M. Harper, L. Terveen, and J. A. Konstan. 2014. "Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity." *Proceedings of the 23rd International Conference on World Wide Web 2014*; WWW'14, April 7–11, 2014, Seoul, Korea. ACM 978-1-4503-2744-2/14/04.

Ofcom. 2012. *Measuring Media Plurality: Ofcom's Advice to the Secretary of State for Culture, Olympics, Media and Sport*. London: OFCOM.

Paal, B. 2012. *Suchmaschinen, Marktmacht und Meinungsbildung*. Commissioned by Initiative for a Competitive Online Marketplace. Oxford: Initiative for a Competitive Online Market Place.

Pariser, E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin.

Pasquale, F. 2015. *The Black Box Society. The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.

Pew Research Center. 2016. *News Use across Social Media Platforms*. Washington, DC: Pew Research Center.

Quattrociocchi, W., A. Scala, and C. R. Sunstein. 2016. "Echo Chambers on Facebook." Discussion Paper No. 877, 09/2016, Harvard Law School.

Reuters Institute for the Study of Journalism. 2016. *Reuters Institute Digital News Report 2016*. Oxford: Reuters Institute for the Study of Journalism.

Reuters Institute for the Study of Journalism. 2017. *Reuters Institute Digital News Report 2017*. London: Reuters Institute.

Richards, N. 2008. "Intellectual Privacy." *Texas Law Review* 87, no. 2: 387–445.

Schmidt, A. L., F. Zollo, M. Del Vicario, A. Bessi, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi. 2017. "Anatomy of News Consumption on Facebook." *Proceedings of the National Academy of Sciences* 114, no. 12: 3035–39.

Schulz W., S. Dreyer, and S. Hagemeier. 2011. *Machtverschiebung in der oeffentlichen Kommunikation*. Bonn: Friedrich Ebert Stiftung.

Stroud, N. 2008. "Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure." *Political Behaviour* 30, no. 2008: 341–66.

Sunstein, C. R. 2000. "Television and the Public Interest." *California Law Review* 88, no. 2: 499–564.

Swapneel, K., J. Bell, N. Arora, and G. Kaiser. 2011. *Towards Diversity in Recommendations Using Social Networks*. Columbia University Computer Science Technical Reports. CUCS-019-11. New York: Columbia University.

Valcke, P. 2011. "Looking for the User in Media Pluralism: Unraveling the Traditional Diversity Chain and Recent Trends of User Empowerment in European Media Regulation." *Journal of Information Policy* 11, no. 1: 287–320.

Valcke, P. 2004. *Pluralism in a Digital Interactive Media Environment: Analysis of Sector Specific Regulation and Competition Law*. Brussel: Larcier.

Van Dijk, J., T. Poell, and M. De Waal. 2016. *De Platform Samenleving. Stijd om publieke Waarden in een online Wereld*. Amsterdam: Amsterdam University Press.

Wauters, E., E. Lievens, and P. Valcke. 2014. "Towards a Better Protection of Social Media Users: A Legal Perspective on the Terms of Use of Social Networking Sites." *International Journal of Law and Information Technology* 22, no. 3: 254–94.

Wojcieszak, M., and H. Rojas. 2011. "Hostile Public Effect: Communication Diversity and the Projection of Personal Opinions onto Others." *Journal of Broadcasting and Electronic Media* 55, no. 4: 543–62.

Zuiderveen Borgesius, F. J., D. Trilling, J. Möller, B. Bodó, C. H. De Vreese, and N. Helberger. 2016. "Should We Worry about Filter Bubbles?" *Internet Policy Review* 5, no. 1: 1–16.

The Power of Providence

The Role of Platforms in Leveraging the Legibility
of Users to Accentuate Inequality

ORLA LYNSKEY

INTRODUCTION

Factors such as the size, business model, and connection capacity of certain platforms, or digital intermediaries, mean they play a pivotal role in the digital ecosystem. Some platforms can be vital to the functioning of other platforms if they have assets—such as an operating system or user base—that are required by other entities to compete. Such control over access to end-users and their data may constitute a barrier to entry to certain markets or strengthen the market power of a company, leading to dominance. If a dominant platform takes advantage of this power to exploit consumers, for instance by offering unfair terms and conditions, as the German Competition Authority alleges Facebook did (Bundeskartellamt 2016), this may constitute an abuse of dominance sanctioned by competition law.

However, as a result of their pivotal position in the digital ecosystem, powerful platforms also exercise a form of power that can be distinguished from this market power, the effects of which are not captured by competition law. Attempts to define and conceptualize this power are underway. Indeed, Cohen noted in 2016 that the successful state regulation of the information economy will, among other things, require an analytically

sound conception of such platform power (Cohen 2016, 374). This chapter contributes to this endeavor by identifying some of the characteristics of this power, which shall be labeled "the power of providence." It is important to recall that "providence" can be understood as "the foreknowing and protective care of God (or nature etc.); divine direction, control, or guidance" (OED) or "an influence that is not human in origin and is thought to control people's lives" (Cambridge Dictionary). Certain platforms could be said to exercise a "power of providence" for several reasons. First, these platforms have the ability to identify users and link diverse datasets, giving them powers akin to the "all-seeing" "eye of providence." The "eye of providence" is a symbol depicting an eye, often surrounded by rays of light and usually enclosed by a triangle. It is said to represent the all-seeing eye of God watching over humanity (Wikipedia 2017). Furthermore, the triangle surrounding the eye of providence serves as a reminder of the multi-sided vantage point of platforms, which can see—and control interactions between—the users, advertisers, and goods and services providers dependent on it. Second, platforms can "control people's lives" insofar as the ability to collect and aggregate a large volume of data from a wide variety of data sources allows platforms to influence individuals in ways that have hitherto been classified by some as purely dystopian, for instance through microtargeting for political purposes. Finally, the architecture of the digital ecosystem and the terminology used to describe its processes (for instance, "machine learning") may give the impression to individuals that its influence "is not human in origin." While this is incorrect, and human input is indispensable to the functioning of digital platforms, even knowledgeable users may treat algorithmic decision-making (such as Internet search results) as "neutral" and be deterred from challenging the actions and processes of platforms as a result of the opacity and complexity of their operations.

The "power of providence" exercised by platforms raises two problems not captured by competition law. The first is that data-driven profiling can have an exacerbated impact, particularly on protected groups, when conducted by platforms with such power. This, in turn, can accentuate existing inequalities between groups in society. While profiling is now a commonplace practice, it is suggested that profiling by powerful platforms has the potential to be particularly problematic as a result of their ability to leverage their strategic position to enhance the legibility of their users. Legibility refers to attempts (previously by states but by private platforms in this context) to "arrange the population in ways that simplified the classic state functions of taxation, conscription and prevention of rebellion" through methods such as mapping or modeling that provided a "synoptic"

perspective of the population (Scott 1998, 2). Beyond this direct impact on inequality, the second concern is that this power of providence can be used to accentuate inequality indirectly. The influence of "App Stores" over the levels of privacy and data protection offered by the applications they host shall be used as an example to illustrate this point.

Data protection law is often assumed to offer a remedy to these concerns. However, as this chapter suggests, while the introduction of the EU's General Data Protection Regulation (GDPR) offers some hope in this regard, data protection law is not a panacea and some additional regulatory measures may be required in order to tackle the more systemic concerns identified in this chapter.

DIGITAL DOMINANCE AND PROFILING PRACTICES

By collecting and sorting data, dominant platforms can profile users of their platforms, and in this way individuals become visible or legible to these platforms (Taylor 2016). The techniques used to profile or categorize individuals have been clearly outlined by the UK's Competition and Markets Authority (CMA) in its report on uses of consumer data (CMA 2015), and by the US Federal Trade Commission (FTC) in its report on data brokers (FTC 2014). Data mining, and profiling in particular, can impact on the individual in two distinct ways: first, it may lead to discrimination against individuals and groups on the basis of "protected grounds," and, second, it may lead to differentiation among nonprotected groups in a way that disproportionately affects communities with certain attributes (such as lower socioeconomic status).

Data Mining and Discrimination

Discrimination occurs when an individual or group is treated in an unfavorable or prejudicial manner on the basis of a "protected characteristic." In the UK, age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion and belief, sex, and sexual orientation are all protected characteristics (Equality Act 2010). Direct discrimination occurs when an individual or entity discriminates against an individual or group by treating them in a comparably less favorable way because of a protected characteristic (Equality Act 2010, s.13). Indirect discrimination occurs when a practice, policy, or rule applies to all individuals equally yet it has a disparate impact on those with protected characteristics by placing them at a disadvantage (Equality Act 2010, s.19).

The ways in which data mining can give rise to discriminatory treatment are often simplified for descriptive purposes. Thus, for instance, Gellert et al. note, "either the difference is made between those who are included and those who are excluded (*inter*) [from the database], or the differentiation is made within the database (*intra*)" (Gellert et al. 2012, 15). Taylor similarly affirms that discrimination can operate along these lines when she notes that factors such as gender, ethnicity, and place of origin help "to determine which databases we are part of, how those systems use our data, and the kinds of influence they can have over us" (Taylor 2017, 4).

Yet a deeper understanding of how decisions regarding inclusion and exclusion are made, and differentiation within databases occurs, will be helpful in framing our subsequent discussions. Indeed, the data-mining process creates multiple opportunities for discriminatory outcomes. In order to illustrate this, Barocas and Selbst subdivide the data-mining process into component steps and identify ways in which these steps may facilitate discrimination (Barocas and Selbst 2016, 677–93). It is worth summarizing some of their findings.

The first step they suggest is to identify the "target variable"—"what data miners are looking for"—and "class labels"—the values or attributes associated with the target variable. The major concern here from a discrimination perspective is that the identification and definition of the "target variable" (e.g., "clean lifestyle") and the "labels" used to inform that (e.g., regular working patterns; low exposure to pollutants; healthy eater, etc.) might have a greater or lesser impact on groups with protected characteristics than others.

Similarly, Barocas and Selbst highlight that the data that train data-mining models (training data) may lead to discrimination if the training data itself reflects prejudice (and the model thus learns from this prejudicial example) or if the training data is based on a biased sample of the population (and thus protected groups are under- or overrepresented). For example, certain groups, which have been judged uncreditworthy in the past, and therefore offered disadvantageous terms of credit that they have struggled to meet, will have their difficulties charted and held against them in data-driven lending cycles. As Pasquale notes, "late payments will be more likely, and then will be fed into *present* credit scoring models as *neutral, objective, non-racial* indicia of reliability and creditworthiness" (Pasquale 2015, 41). In this way, the profiling actually increases the likelihood of a consumer defaulting.

The third way in which Barocas and Selbst suggest discrimination can creep into the system is during "feature selection"—the stage at which choices are made about what attributes to consider in the analysis (for
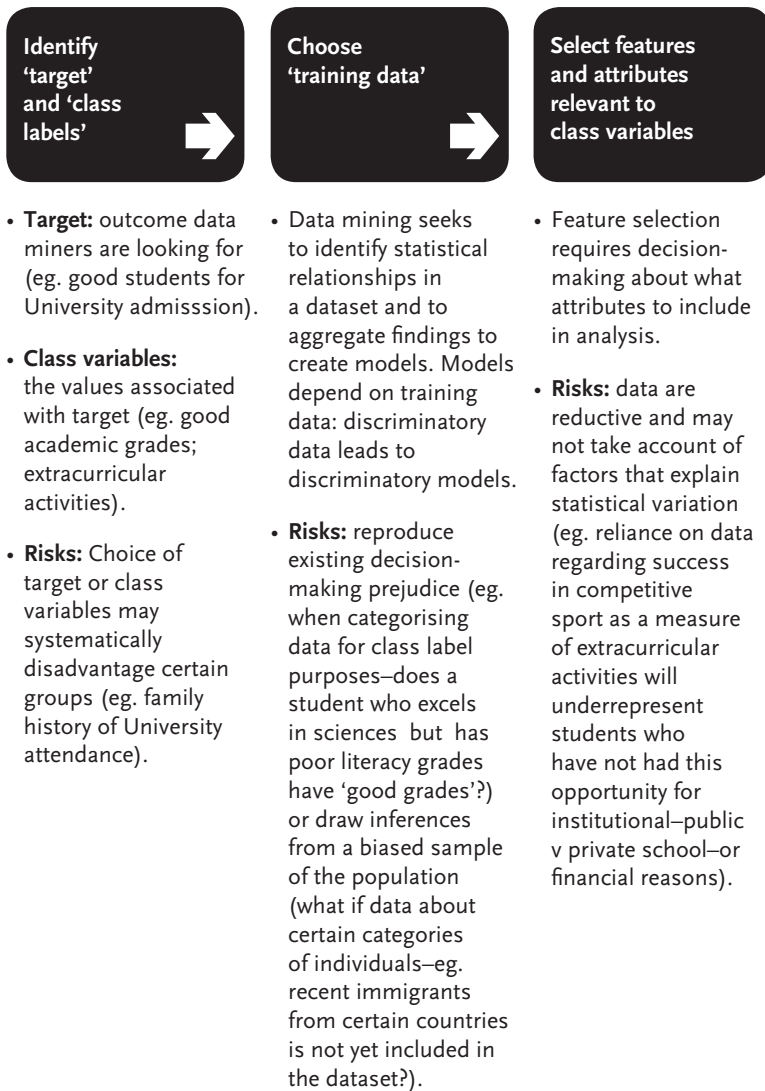
instance, in determining whether someone is a "healthy eater," one could focus on overall calorie intake, or whether they eat fresh foods). They suggest that the "reductive representations" of data may "fail to capture enough detail to allow for the discovery of crucial points of interest" (Barocas and Selbst 2016, 688). Thus, as they note, while inferences might be statistically sound, they will nevertheless be inaccurate if they are based on insufficiently granular data. For example, if participation in certain sports is used as a factor to indicate accomplishment in university admissions, it may negatively impact certain ethnic groups who are less likely to have had the opportunity to play these sports, whereas more granular data might reveal that these candidates acquired similar skills in different ways.

Indeed, statistical accuracy is not considered sufficient to justify discriminatory treatment of those in protected groups. In *Test-Achats* the European Court of Justice (ECJ) was asked to consider the legality of provisions of the EU "Gender Goods and Services Directive." Article 5(1) of that Directive prohibited the use of sex as a factor in the calculation of an individual's premiums and benefits while Article 5(2) allowed for a derogation from this prohibition when risk assessment is "based on relevant and accurate statistical data" and sex is a "determining factor." The Court held that this derogation was incompatible with the prohibition of discrimination on the grounds of sex. Advocate General Sharpston, an advisory member of the Court, had made a similar point in *Lindorfer* when she stated:

> it might be helpful to imagine a situation in which (as is perfectly plausible) statistics might show that a member of one ethnic group lived on average longer than another. To take those differences into account when determining the correlation between contributions and entitlements under the Community pension scheme would be wholly unacceptable, and I cannot see that the use of the criterion of sex rather than ethnic origin can be more acceptable.

This risk of discrimination has been recognized by policymakers. For example, former FTC Commissioner Ramirez highlighted that algorithmic data profiling can "accidentally classify people based on categories that society has decided—by law or ethics—not to use, such as race, ethnic background, gender or sexual orientation" (Ramirez 2013). Moreover, in its report on data brokers, the FTC notes that individuals are divided into segments such as "Urban Scramble" and "Mobile Mixers," which focus on minority communities and those with lower incomes, both of which incorporate a high concentration of Latino and African American consumers (FTC 2014). The FTC notes that these segments may be "more sensitive" as

a result of their reliance on characteristics such as ethnicity, income level, and education. They could therefore be labeled "redundant encodings," cases in which membership in a protected class is encoded in other data (Barocas and Selbst 2016, 691; see Figure 7.1). Indeed, data profiling can be used to conceal discriminatory treatment (Barocas and Selbst 2016, 692). Yet such attempts would constitute indirect discrimination and would therefore also be captured by the law (Gellert et al. 2012, 19).

Identify 'target' and 'class labels' →	Choose 'training data' →	Select features and attributes relevant to class variables
• **Target:** outcome data miners are looking for (eg. good students for University admisssion). • **Class variables:** the values associated with target (eg. good academic grades; extracurricular activities). • **Risks:** Choice of target or class variables may systematically disadvantage certain groups (eg. family history of University attendance).	• Data mining seeks to identify statistical relationships in a dataset and to aggregate findings to create models. Models depend on training data: discriminatory data leads to discriminatory models. • **Risks:** reproduce existing decision-making prejudice (eg. when categorising data for class label purposes–does a student who excels in sciences but has poor literacy grades have 'good grades'?) or draw inferences from a biased sample of the population (what if data about certain categories of individuals–eg. recent immigrants from certain countries is not yet included in the dataset?).	• Feature selection requires decision-making about what attributes to include in analysis. • **Risks:** data are reductive and may not take account of factors that explain statistical variation (eg. reliance on data regarding success in competitive sport as a measure of extracurricular activities will underrepresent students who have not had this opportunity for institutional–public v private school–or financial reasons).

Figure 7.1: The data mining process and the risk of discrimination.
Source: Based on categorization identified by Barocas and Selbst (2016).

Data Mining and Differentiation

In addition to discriminating against individuals with protected characteristics, data mining and profiling processes can also differentiate among individuals and groups on the basis of classifications that are not protected. In the digital context, data gathering, via tools such as first- and third-party cookies, and the subsequent mining of that data, can be used to categorize individuals in groupings according to, for instance, their perceived interests or characteristics. Tailored advertisements, goods, and services are then offered to individuals on the basis of this categorization (Article 29 Data Protection Working Party 2010, 5). Indeed, proponents of behavioral advertising regularly assert that advertisers have little interest in the actual identity of an individual: rather, they are interested in categorizing individuals into groupings (e.g., those more likely than others to be interested in folk music) (Lenard and Rubin 2010). Therefore, if the fact that an individual attends marriage counseling is a "feature" that is selected as relevant in assessing creditworthiness as a target variable, a credit card company does not need to be able to identify the individual in order to offer him more expensive credit. Thus, without necessarily identifying the individual he can be "singled out" and categorized in a way that differentiates him from others.

Like discrimination, such differentiation can exacerbate existing inequalities. An obvious example of this is differentiation on the basis of socioeconomic status, or proxies for this status. Taylor highlights that the greatest burden of dataveillance has always been borne by the poor (Taylor 2017, 2). For instance, it has been documented that data-driven law enforcement strategies have led to the overpolicing of poorer neighborhoods (see, for instance, Lum and Isaac 2016). Moreover, in addition to being subject to more surveillance and with higher stakes, Madden et al. demonstrate that poor Americans are more vulnerable to data-mining processes as a result of the devices that they use to access the Internet and the pattern of their "privacy-relevant" behavior (Madden et al. 2017, 4). For instance, they highlight that 63% of those living in households earning less than $20,000 per year mostly use their mobile phones to go online compared to just 21% of those in households earning $100,000 or more per year. This is relevant as mobile phones are less secure than other devices, such as laptops or desktops (Madden et al. 2017). The practical impact of such differentiation on the poor is recognized by the FTC when it states that big data mining "can injure the economic stability and civil rights of the poor, such as when they are targeted for predatory financial products, charged more for goods and services online, or profiled in ways that limit their employment and educational opportunities" (FTC 2016, 9–11). The

UK's CMA also recognizes that data mining can be used to differentiate among consumers based on the quality or the price of goods and services offered to them, stating that the

> collection of consumer data may enable firms to make judgments about the lowest level of quality needed by consumers/groups of similar consumers. This may enable a firm to engage in quality discrimination where quality differences are not reflected in the price of goods or services. (CMA 2015, 93)

Platforms could facilitate such practices by restricting the products that are displayed to consumers or changing the order in which they are listed to display poorer quality products first in some circumstances (Acquisti 2010, 19). According to Borgesius and Poort, despite several high-profile incidents of personalized pricing, such pricing practices seem to be used relatively rarely (Borgesius and Poort 2017, 3). The precise welfare effects of such practices are ambiguous and need to be assessed on a case-by-case basis. However, it is possible that an individual will pay more than required for goods or services, allowing the company concerned to extract more profit from their offerings and thus entailing a "transfer of wealth from the pockets of consumers to the pockets of operators" (House of Lords 2016, 75).

Borgesius and Poort highlight the factors explaining this reluctance to use personalized pricing practices, most evidently that consumers perceive such practices to be unfair and thus companies fear consumer backlash if found to be differentiating in this way. Turow's work provides vivid examples of this unfairness. For example, he explains how companies can designate individuals into categories such as "targets" or "waste" using data-mining techniques and offer discounts to them on this basis. Contrary to distributive justice intuitions, those with perceived higher spending capacity and reserve prices for products ("targets") are offered more significant discounts than price-sensitive consumers or those with lower spending capacity ("wastes") in order to entice targets to become regular consumers (Turow 2011, 108–10). As Borgesius and Poort suggest, and as discussed below, it is the surreptitious nature of such practices, when compared to signposted discounts for particular groups such as the elderly or students, that contributes to the public discomfort with these practices (Borgesius and Poort 2017, 6).

The Ability to Create Perceptions

The discriminating and differentiating impact of data mining not only compounds and exacerbates existing inequalities but also has the potential to

create further inequality by distorting perceptions. Indeed, as Helberger et al. argue, much of the concern regarding the influence of platforms, or "gatekeepers," lies in their control over access to individuals and the way in which the relationship between gatekeepers and users is shaped, rather than their control over access to information as such (Helberger, Katharina, and van der Noll 2015, 51). This echoes Zuboff's claim that if "power was once identified with the ownership of the means of production, it is now identified with the ownership of the means of behavioural modification" (Zuboff 2015, 82). A good example of this is search-ranking mechanisms based on data mining. Sweeney's research indicates that a Google search for Caucasian names presents more neutral results than a search for typically African American names (Sweeney 2013).

THE "POWER OF PROVIDENCE": THE AGGRAVATING EFFECT OF DIGITAL DOMINANCE

Data-driven discrimination and differentiation is not confined to situations of digital dominance. However, this chapter suggests that the effects of such discrimination and differentiation may be exacerbated in the presence of dominance as a result of the privileged position of dominant companies in the digital ecosystem—their "power of providence." In particular, this privileged position results in superior data-mining capacity and greater information and power asymmetries between individuals and dominant digital platforms.

The "Power of Providence" of the Digitally Dominant

Specific attention ought to be paid to digital platforms that operate as "market makers—or orchestrators—in the digital ecology value chain" (Mansell 2015, 25). The size, business model, and connection capacity of these market makers mean that they play a vital role in the digital ecosystem. The EU Commission has suggested that in the near future "only a very limited part of the economy will not depend on [online platforms]" (Ansip 2015). Platforms enjoy this pivotal position when the functioning of other platforms or services is dependent on them, such as when they have assets that other entities need in order to compete (for instance, an operating system or a user base) (TNO Report 2015, 14). Indeed, in its 2017 interim report on the implementation of its Digital Single Market strategy, the European Commission notes that there is widespread concern among

businesses that some platforms may engage in practices such as favoring their own products or services over those of other businesses and restricting access to, and use of, data directly generated by a business's activities on a platform (EU Commission Communication 2017). Such an allegation of discrimination—the systematic favoring of its own service over that of other businesses—lies at the heart of the European Commission's antitrust investigation against Google, which culminated in a €2.42bn fine for the company (EU Commission—Antitrust 2017). The Commission claims that Google has systematically given prominent placement to its own comparison shopping service in its Google search engine results and demoted rival comparison shopping services results in its generic search-engine ranking. Critics have challenged the decision on the basis that it enters the unprecedented territory of establishing a principle that a company may not favor its own products over those of competitors (Lamadrid 2017).

Leaving the merits of the case aside, it does illustrate the ability—and willingness—of the EU Commission to sanction the practices of powerful digital companies. Such intervention is however ordinarily only justified when the company concerned enjoys a position of dominance, assessed in accordance with the Commission's guidance on market definition (EU Commission 2007), and when the practice concerned would lead to a decrease in "consumer welfare" (European Commission 2009). Moreover, what consumers may define as a market with a dominant player—for instance, Facebook in the market for a social networking site—often does not reflect how the market is defined for competition law purposes. Findings of market power focus solely on the economic power of the company, which may be distinct from its power over data flows or its power to influence opinions. For example, instinctively many consumers would assume that Google is dominant in the EU market for organic search engine services. However, it could be argued that no such market exists for several reasons: for instance, that organic search results compete with paid search results, or that integrated search tools in a social networking service compete with Google's search engine. If accepted, the market would be broader than a market for "organic search results" and could envisage companies such as Facebook as competitors of Google. This in turn would make a finding of market power on Google's part less likely. Equally, it could be argued that even if a market for organic search results exists, and Google has a market share of in excess of 90% of this market in Europe, it is not in a position of market power due to the low barriers to entry in the market and the fact that "competition is just a click away," a mantra in the technology sector. These empirical assessments are, of course, vigorously contested. However, they illustrate the point that although competition

law is the only legal instrument available to harness the excesses of private power and dominance, given its definition of dominance it may be of limited utility in tackling the inequalities identified earlier. This claim is further supported by the fact that the harms competition law seeks to address are primarily economic harms, and there has been significant resistance to expanding the consumer welfare paradigm to incorporate noneconomic harms (Easterbrook 1984; Odudu 2010). Therefore, while there is a lively debate about whether practices such as "price discrimination" on the basis of personalization are captured by competition law (Townley, Morrison, and Yeung 2017), many of the inequalities generated by differentiation, discrimination, and power over opinion formation are not captured by competition law as a result of these constraints (the need for "dominance" and detriment to "consumer welfare"). Yet, as discussed below, these inequalities are exacerbated by the pivotal position—the power of providence—of digital platforms that control access to infrastructure and users.

The Impact of the "Power of Providence"

The problems that data mining entails are not exclusive to the online environment. On the contrary, current business practices indicate that the line between "offline" and "online" practices is difficult to draw when it comes to the creation of profiles. Practices such as "onboarding," whereby a data broker adds offline data into a cookie to enable advertisers to use this offline activity information about consumers to determine what online advertisements to deliver to them, indicate that there may be little value in taking a distinct approach to the regulation of digital gatekeepers (CMA 2015). Indeed, in the United States, a digital rights advocacy group—the Electronic Privacy Information Center (EPIC)—alleges that Google is using credit card data to track whether the online advertisements it is offering lead to (offline) in-store purchases and that users have not been provided with adequate information about how this practice operates and how to opt-out from the practice (EPIC 2017). Given these blurred boundaries, regulators must consider whether data mining by dominant digital firms merits particular attention and, if so, why. For instance, is offering an individual a beauty product at a certain time in the day based on data mining techniques different from the practice of selling chocolate bars and snacks at supermarket checkouts? Both, it could be argued, are psychological ploys to encourage sales.

It is suggested here that dominant firms in the digital ecosystem warrant special attention because of the number of individuals with whom they have

direct contact and the extent of the data processed regarding these individuals. Therefore, while all entities—irrespective of their reach, or the extent of their data processing—might potentially have a negative impact on individuals, the actions of larger entities—with wider reach and greater data-processing capacity—has the potential to have a more significant impact on societal interests and individual rights. Indeed, the ECJ recognizes this implicitly in its *Google Spain* judgment: it highlights that both the ubiquity of Google's search engine and the quantity of personal data processed are relevant to the extent of the interference with the individual's rights when their personal data is made available through Google's search results.

In other words, the broader the reach of a service and the more personal data processed the greater the interference with individual rights. This, in turn, justifies paying particular attention to the actions of digitally dominant firms. Indeed, an analogy could be drawn here with the competition law provisions discussed earlier. Under EU competition, dominant companies are said to have a "special responsibility" such that practices, like the imposition of exclusive dealing obligations on consumers, that would be lawful for a nondominant company would be unlawful if pursued by a dominant company. This "special responsibility" is justifiable on the grounds that the actions of a dominant firm have a greater impact on competition than those of nondominant firms.

Similarly, it is argued here that the actions of dominant digital firms can also have a greater impact on the rights and interests of individuals than those of nondominant firms and that this may justify the imposition of specific duties on them that may not be appropriate for nondominant firms. For instance, Facebook has over 2 billion users as of mid-June 2017, approximately 1.2 billion of whom use their Facebook account on a daily basis (Constine 2017). Facebook's data processing potential is further enhanced as a result of its partnerships with a variety of data brokers, including some of the world's largest. For instance, Facebook partners with Acxiom, which claims to hold data on 700 million people, and Datalogix, which holds $2 trillion worth in offline purchase-based data. Facebook's extensive direct access to users as well as its data-processing capability makes it a desirable trading partner for these data brokers and gives it a superior ability to profile individuals based on what Pasquale labels a "self-reinforcing data advantage" (Pasquale 2013, 1015). While many large platforms claim that the quantity of data they process is not decisive to their success, and that it is rather their use of this data that is significant, this does raise the question of why such data-sharing partnerships are necessary and whether they are compliant with the principle of "data minimization" enshrined in many data protection laws globally.

The concern here is therefore not simply the "digital"; it is the combination of the digital with power. As suggested previously, power in this context may overlap with the concept of "market power" used in competition law and economic regulation, however it is not synonymous with market power in terms of how it is defined or measured. Indeed, one of the great challenges in this regard is that we lack even the language to describe this private power and, as a result, we fall back on the language and concepts of economic regulation and competition law. This point has not gone entirely unnoticed and, in part, explains the ongoing debates over whether competition law needs to be redesigned to remain fit for purpose in a digital era, and in particular whether a new concept of "market power" is needed (Autorité de la Concurrence and Bundeskartellamt 2016). This chapter has labeled such power the "power of providence" as part of the movement to decouple power from the limitations of antitrust "market definition" and to advocate for a reconceptualization of power in response to the quasi-regulatory role played by private platforms in society. Yet, as demonstrated in what follows, however one labels this power, its impact and effects are already tangible. Platform power both directly and indirectly exacerbates existing inequalities. It does so directly by aggravating the asymmetries between those who process personal data and those who are rendered transparent by this process to the detriment of the latter. It does so indirectly as powerful platforms, in practice, determine the terms and conditions offered by dependent service providers (such as applications in an "App store") to their users.

Exacerbating Asymmetries of Information and Power

The differential pricing practices, referred to earlier, are one example of how the asymmetry of power and information between individuals and platforms is evident.

Information asymmetries between the individual and the dominant platform enable the platform, for instance, to attempt to influence the political opinions of the individual or to engage in differential pricing practices based on an estimation of an individual's reserve price for a product or service. Individuals will perceive such practices as unfair, and they may be exploitative—for instance, in the pricing context by extracting higher rents from individuals when they are desperate or vulnerable (e.g., payday loans with excessive interest rates, or the more banal hike in taxi fares when a phone battery is dying). One of the reasons why such practices are unfair is that their operation remains opaque while the individual is simultaneously

rendered transparent. This is highlighted by Helberger et al., who—writing in the media plurality context—consider it problematic that users have no knowledge of the selection criteria on which processes of implicit personalization are based and that they are not provided with any tools to change them or "turn them off." They are therefore unable to assess or ascertain for themselves how limited their news selection is (Helberger, Katharina, and van der Noll 2015, 34; see also Helberger, this volume). Pasquale also highlights this opacity, stating that there may be "scarlet letters emblazoned on our digital dossiers" that we may not even know about (Pasquale 2015, 56). However, this lack of knowledge is not the sole problem.

Power asymmetries persist even when individuals are, for instance, given more information or the ability to view and amend the parameters that are used to generate their profiles. When individuals are co-opted into the process in this way, it does not follow that they will be able to challenge the factors influencing a particular profile (for instance, the choice of training data or "feature selection" to use the terminology above). If an individual is categorized in a manner that he or she disagrees with, for instance, "diabetic lifestyle" or "leans left" (FTC 2014, 21), a profiler may be able to argue that the inference is simply a matter of opinion rather than fact (Pasquale 2015, 32). Moreover, even if an individual knows that certain characteristics are valued, or punished, more than others when determining the terms and conditions on which goods and services are offered, this may not help them decipher how to act. For instance, according to the CMA, some grocery retailers that offer motor insurance use purchasing data from loyalty schemes to "draw inferences about household characteristics—for instance, to offer discounts to households that appeared from their shopping habits to be relatively low risk" (CMA 2015, 45). However, in the endless debates over whether butter is better for you than margarine, even if an individual were to try to conform to a profiler "ideal" this may not be possible.

De Facto Influence over the Data-Processing Practices of Service Providers

A final reason that dominant digital platforms merit particular attention is that, given the dependence of other content providers on their platform, in practice they can exercise a decisive influence over the levels of fundamental rights, such as data protection and privacy, enjoyed by individuals. For instance, the CMA acknowledges that operating systems (such as Google's Android, or the Apple OS) are responsible for the "Application Programming Interfaces (APIs) which dictate how the software and

hardware interact—including what information the app can access." These APIs control the release of information according to the privacy controls in place at the operating system level (CMA 2015, 42). In other words, it is the operating system that has the final say on the minimum level of data-processing standards offered by the applications it hosts. This means that operating systems could, in theory, exclude applications with sub-standard data use policies from their platforms. However, it would seem that platforms are doing very little to promote key data protection principles, such as data minimization (Article 6(1)(c) Directive 95/46 EC), among application providers. For example, a 2014 survey conducted by the Global Privacy Enforcement Network (GPEN) discovered that one-third of all applications requested an excessive number of permissions to access additional personal information (CMA, 2015). Moreover, the US Federal Trade Commission (FTC) has taken actions against applications such as Brightest Flashlight and Snapchat in recent years for misrepresenting how the personal data they gather is used (CMA 2015, 123–24). This is not to say that platforms are entirely inactive when it comes to promoting privacy and data protection. For instance, recent reports suggest that Google Play—the App store for Android users—culled applications from its platform on the basis of privacy and data protection concerns (Abent 2017). However, their ostensible "lowest common denominator" approach to these rights influences the extent to which these rights can be enjoyed by their users in practice. Indeed, Google Play's cull appeared only to remove egregious violators of rights from the App store, for example applications requesting sensitive permissions—such as unnecessary data from cameras or microphones—and that did not comply with the basic principles set out in the Play Store privacy policy.

Dominant platforms can, furthermore, make it difficult for individuals to take steps to defend their own rights, for instance by preventing users from using ad-blockers or by excluding privacy enhancing technologies (PETs) from their platforms. Indeed, the PET Disconnect alleges that it has been unjustly excluded from Android's Google Play application store. In its defense, Google has stated (informally) that it applies its policies consistently to all applications and that it has "long prohibited apps that interfere with other apps—such as altering their functionality, or removing their way of making money." It also notes that there are many more PETs available in the Google Play Store that comply with its policies. While the impact on rights might be minimal given the availability of competing PETs, the lack of transparency regarding Google Play Store's exclusion policy is striking. This also provides a vivid reminder of

the commercial imperatives driving Google's Play Store operations: the PET is not viewed as an application to help ensure the effectiveness of individual rights by allowing individuals to control how their personal data is processed. Rather, it is viewed as a threat to Google's bottom line as it facilitates the exercise of these rights and thus threatens the revenue streams of other apps dependent on data processing for their commercial viability.

In light of these enhanced concerns caused by platform power—the direct impact on asymmetries of power, and the indirect shaping of the de facto rights protection enjoyed by individuals—it is appropriate to query whether a "special responsibility" should be placed on powerful platforms, akin to that placed on dominant companies by competition law. At present, as a result of the limited role played by competition law to mitigate these potential harms, the primary body of rules associated with data-driven practices and their subsequent consequences, such as differentiation, are data protection rules. However, as shall now be discussed, while data protection law offers some mechanisms to mitigate the negative impact of data-mining practices, it also has its limits.

THE ROLE AND LIMITS OF DATA PROTECTION LAW

Data protection legislation applies when "personal data" are "processed" (Article 2(a) and 2(b), Directive 95/46 EC). "Personal data" and "processing" are broadly defined and, as a result, many of the potential harms of data processing by gatekeepers may be captured by data protection legislation. For instance, the concentration of data in the hands of a powerful platform may, like any other concentration of data, entail a heightened risk of data security breach (Cormack 2016, 17). Provisions in data protection legislation (theoretically) mitigate such risks by requiring those responsible for data processing to respect certain safeguards and to ensure that data-processing systems are structurally robust. Yet, to date, data protection legislation has proven to be of limited utility in regulating and curbing the excesses of data mining. Indeed, phenomena such as "Big Data" processing have emerged in spite of the ostensible tension between its operational principles and the foundational principles of data protection law. While the introduction of a new legislation framework—the General Data Protection Regulation (GDPR)—will undoubtedly improve the effectiveness of this legal regime as discussed in what follows, it would be erroneous to place too much hope in its provisions.

In the EU, a complex regulatory regime sets out a framework of checks and balances that data-processing operations must comply with to be lawful. Like its predecessor, the 1995 Data Protection Directive, the EU's GDPR applies to the automated (or systematic) processing of personal data. Once within the scope of this legal framework, the entities that determine how and why personal data are processed—the brains behind the personal data-processing operations—become "data controllers." As such, the data controllers must ensure that the data-processing operation has a relevant legal basis; that it complies with the data-processing principles and safeguards; and that the rights of individuals—"data subjects"—are respected.

Given the prevalence of commercial data mining, and the dearth of jurisprudence relating to this practice, a preliminary query is therefore whether the data-mining practices of powerful platforms fall within the scope of the data protection rules. Personal data is defined as any information that relates to someone who is identified or identifiable on the basis of that data (Article 4(1) GDPR). There are therefore three constituent elements of "personal data": it is (1) any information that (2) relates to (3) an identified or identifiable person. The Article 29 Working Party, an advisory body on data protection matters comprising representatives of national data protection agencies, has suggested that information "relates to" an individual when, among other things, the purpose of the data is to "evaluate, treat in a certain way or influence the status or behaviour of an individual" or the data is "likely to have an impact on a certain person's rights and interests" (Article 29 Working Party 2007, 9–12). Thus, the Article 29 Working Party eschews narrower understandings of the words "relate to" whereby the focus of the information is on the data subject, or the data have a clear link to the private life of the individual. It follows from the Working Party's definition that the classification or categorization of individuals through data-mining practices, for instance according to their perceived spending capacity or future interests, "relates to" individuals as this categorization determines how individuals will be treated (for example, what advertisement they will be delivered, or music suggestions they will be offered). However, the ECJ has cast doubt on such a broad interpretation of information "relating to" an individual. In *YS* the Court was asked to consider whether the data provided by an applicant for a residence permit and the legal analysis of the applicant's status in relation to that residence permit contained in a "minute" drawn up by the competent immigration officer constitute personal data. It was not contested before the Court that the data contained in the minute about the applicant (such as name, date of birth, gender, language,

etc.) constituted personal data, and the Court confirmed this finding (para 38). The Court held, however, that the legal analysis in the minute—which examined the applicant's data in light of the relevant legal provisions—did not constitute personal data (para 39). It reasoned that such legal analysis is "not information *relating to* the applicant for a residence permit, but at most, in so far as it is not limited to a purely abstract interpretation of the law, is information about the assessment and application by the competent authority of that law to the applicant's situation" (para 40; emphasis added). It justified such a finding on the basis that it was borne out by the objective and general scheme of the Data Protection Directive (para 41).

The Court's reasoning provides food for thought when transposed to the operations of dominant digital firms. It seems to suggest that the data "provided to" these firms, such as the data on an individual's browsing activities, would, like the data provided by the applicant for the residence permit, constitute personal data. However, the application of the company's algorithm to that data through data mining practices—the equivalent of the application of legal provisions to that data through legal analysis—would not constitute personal data. Indeed, Korff cautions that companies could use such reasoning in order to remove profiling from the scope of application of the data protection rules (Korff 2014). He notes that:

> After all, a profile, by definition, is also based on an abstract analysis of facts and assumptions not specifically related to the data subject—although both are of course used in relation to the data subject, and determine the way he or she is treated.

Indeed, the applicants, several Member State governments and the European Commission in *YS* had argued that the legal analysis should constitute personal data as it refers to a specific natural person and is based on that person's situation and individual characteristics (para 35).

The recent Opinion of Advocate General Kokott, an advisor to the Court, in the *Nowak* case appears to offer a counterbalance to the Court's findings in *YS*. In *Nowak* the Court is asked to consider whether an examination script constitutes personal data. The advocate general clearly opines that it does, reasoning that an examination script links the solutions it contains with the individual candidate who produces the script. As such, the "script is a documentary record that that individual has taken part in a given examination and how he performed" (para 21). In particular, she highlights that a script is intended to assess the "strictly personal and individual performance" of a candidate (para 24). She adds to this by suggesting that the comments of an examiner on a script are also personal data (para 63),

noting that "the purpose of comments is the evaluation of the examination performance and thus they relate indirectly to the examination candidate" (para 61). Despite the clear analogy between the examination corrections in *Nowak* and the legal analysis in *YS*, the advocate general does not attempt to reconcile the two. It is therefore suggested that whether a profile itself "relates to" an individual remains an open question, even following the entry into force of the GDPR.

A further bone of contention in the context of profiling is whether the personal data relates to someone who is identifiable. When determining whether someone is identifiable, recital 26 GDPR specifies that "account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person" to identify the data subject directly or indirectly. Objective factors, such as costs, the time commitment, and the current technological availability, should be taken into consideration when making this assessment. The most ardent proponents of online behavioral advertising have sought to argue that advertisers are not interested in a user's actual identity rather, they simply wish to categorize users into groups of those who are more likely than average to have certain interests or capabilities (Lenard and Rubin 2010). This is a point that has been acknowledged even by those seeking more effective regulatory responses to profiling (Barocas and Nissenbaum 2014). Indeed, Taylor highlights that data injustice tends to occur increasingly on the collective level. She notes,

> New technologies tend to sort, profile and inform action based on group rather than individual characteristics and behaviour, so that in order to operationalize any concept of data justice, it is inevitably going to be necessary to look beyond the individual level. (Taylor 2017, 14)

An example may serve to illustrate this point. Facebook categorizes its users on the basis of their user profiles and activity and then offers to connect advertisers to users with profiles that match their needs. For instance, Facebook might estimate that a user is in her mid-30s, is based in London, is interested in cycling, and works as a professional. It might therefore offer this individual advertising for spinning studios in central London on its platform. Facebook does not provide the user details to the advertiser. The advertiser may therefore argue that, even if the user clicks through to its spinning advertisement, it would not be able to identify the user on the basis of the user IP address alone. Moreover, given the broad parameters of the profile, it might argue that even if it did have the profile and an IP address it would not be able to identify an individual on that basis.

Facebook would however likely be processing personal data when connecting the advertisement to a user as, even if it categorized individuals in broad terms, it has the technical capacity to link this broad profile back to an individual. In *Breyer* the ECJ adopted a wide interpretation of identifiability. It found that a dynamic IP address could constitute personal data if the provider of a publicly available website or online media service has the legal means available to it to link that dynamic IP address with additional data to identify the individual. What is noteworthy about the Court's finding in *Breyer* is the Court's assessment of what means are "likely reasonably to be used" by a data controller to identify a data subject. In that instance, in order to identify the individual behind a dynamic IP address the website operator would need to contact a competent authority (in situ, a cyber-crime authority) who would then in turn need to contact an Internet service provider in order to obtain the additional identifying information (para 47). The availability of the mere prospect of connecting data with identifying data—even if the process for doing so is laborious—ostensibly renders that data "identifiable." This broad precedent, when coupled with the Opinion of the Advocate General in *Nowak*, therefore opens up the possibility that a profile—the application of a data-mining formula to particular personal data—constitutes personal data.

The Substantive Rights Provided by Data Protection Law

Once within the scope of the data protection regime (as there is personal data processing) the utility of the rights available to combat profiling pursuant to that framework remains hotly debated. According to the GDPR, data subjects have a right to receive specified information regarding the processing of their personal data. Among other things, the individual should be informed of the existence of automated decision-making, including profiling, and provided with "meaningful information about the logic of this automated decision-making as well as its significance and envisaged consequences for the data subject" (Articles 13(2)(f) and 14(2)(g) GDPR). Such information is also available to the data subject when exercising his or her right to access pursuant to Article 15(1)(h) GDPR.

Article 22(1) GPDR also provides that the individual shall have the "right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." There are a number of broad exceptions to this right however. It does not apply if the automated decision-making is necessary to enter into or perform a contract between the individual and

the controller, if the automated decision-making is authorized by law, or if it is based on the explicit consent of the data subject. However, where the right does not apply as the automated decision-making is based on consent or is necessary to enter into or perform a contract, the data controller must "implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision" (Article 22(3) GDPR). Whether these provisions amount to a "right to explanation" of profiling for the data subject has been the subject of vigorous doctrinal debate. On the one hand, Wachter et al. have asserted that pursuant to these GDPR provisions the individual has, at best, a right to be informed ex post about the general system functionality of an automated decision-making system as opposed to the right to receive an explanation for specific automated decisions ex ante (Wachter, Mittelstadt, and Floridi 2017). Powles and Selbst, among others, roundly contest these findings. They contend that the distinctions drawn between system functionality and specific decisions, and ex ante and ex post explanations fail to withstand scrutiny. They suggest that it is "hard to imagine a useful or meaningful explanation of system functionality that does not allow a data subject to map inputs to outputs, and to figure out what will happen in her case" (Powles and Selbst 2017). They therefore conclude that if you can explain system functionality, you can usually explain specific decisions. While this is true, it could be argued that system functionality could be compared to the relevant legal provisions applied in *YS* while a specific decision could be compared to the legal analysis, which applies a formula to particular facts. It may be in the data subject's interest to determine how a formula (whether it be a law or automated decision-making algorithm) applies to her data and not all individuals will have the skills required to work backward from the automated decision to the original personal data processed in order to decipher how a particular formula has been applied to their situation. The distinction suggested by Wachter et al. is thus not entirely irrelevant.

Nevertheless, the broader point raised by Powles and Selbst remains critical: we have yet to consider what exactly "meaningful information about the logic" of automated decision-making means and it is necessary to consider this in the wider context of the changes introduced by the GDPR. The narrow, formalistic reading of critical individual rights proposed by Wachter et al. seems to militate against the direction of travel of EU data protection law. One could point to many factors that would indicate a more generous interpretation of these GDPR rights might be preferred: for instance, one of the principal aims of data protection reform

was to enhance the effectiveness of the rights of data subjects while comparisons with the predecessor Directive remain of limited value given that the EU Charter of Fundamental Rights was not in force for much of its existence. The precise contours of these rights will therefore likely only be clarified after the new rules have been interpreted by the new EU data protection agency—the European Data Protection Board—or by the ECJ.

Rather than requiring the initiative of an individual data subject, the "right not to be subject to a decision based solely on automated data processing" ostensibly prohibits such automated decision-making in some circumstances. If true, this would differentiate this right from other data protection rights granted to the individual and potentially render it more effective in practice. As Blume has noted, other rights such as the right of access and the right to delete presuppose "the initiative of the data subject, which in practice will not often be taken" (Blume 2014, 270). Nevertheless, even the most favorable interpretation of this right for a data subject is just one small piece in the jigsaw puzzle of counterbalancing digital dominance. These provisions need to be overseen and such oversight is an onerous burden to place on individuals, particularly given that such an individual-centric system of oversight may exacerbate existing inequalities. Put bluntly, those who are poor in skills, time, or other resources, may not have the same capacity to exercise their rights as others who are, for instance, more technology literate or informed.

It is for this reason that the increased focus on more "back-end" enforcement in the GDPR is to be welcomed. The GDPR seeks to enhance the effectiveness of the rights of individuals but, in so doing, it avoids the pitfalls of the past and distributes the onus for such rights protection more evenly across stakeholders. The introduction of a general principle of accountability in the GDPR, pursuant to which the data controller must be able to demonstrate compliance with key data protection safeguards such as fair and lawful data processing, is a critical step in this regard. This principle of accountability will encourage data controllers to give adequate consideration to their data protection compliance mechanisms while the GDPR's increased emphasis on effective enforcement, through mechanisms such as enhanced sanctions, provides an added incentive to take these obligations seriously.

CONCLUSIONS

This chapter has suggested that commonplace data-mining techniques are not neutral in their application and can exacerbate existing societal

inequalities, along legally protected grounds (such as race, gender, and sexuality) as well as other grounds such as socioeconomic status. The techniques in question are not used solely by powerful platforms. However, it has been suggested that the "power of providence" of digital platforms—stemming from the volume and variety of data they process, as well as their reach—means that the effects of these data-mining techniques may be particularly pernicious in this context. This is for two reasons: first, such practices can widen existing power and information asymmetries between the individual and the data controller (widening the gulf of power between the disadvantaged individual and powerful platform even more). Second, powerful platforms have the capacity to influence the data-mining practices of providers in the broader internet ecosystem: this capacity can be used to push for higher standards of rights protection and encourage or conversely, as is presently the case, to condone or even facilitate unfair data-mining techniques. It has therefore been suggested that, just as it is appropriate to impose a "special responsibility" on firms with significant market power through competition law rules, it may be appropriate to impose additional legal responsibilities on firms with the "power of providence" as a result of the volume and variety of the personal data they process and the extent of their reach.

Data protection law is often touted as the answer to these profiling practices. However, the enforcement of this body of law has, to date, been underwhelming, in part because of the onus it places on individuals to assert rights on their own behalf. The entry into force of the EU's GDPR does offer some hope in this regard but should not be viewed as a silver bullet. Given the architecture of the digital system, it is suggested that more imaginative and holistic solutions will be required. Taylor, for instance, highlights that "markets are an important factor in establishing and amplifying power asymmetries to do with digital data, and that addressing the value chains involved in large-scale data gathering and surveillance may be a functional shortcut to controlling misuse" (Taylor 2017, 5). Getting a firmer grip and understanding of how personal data are processed by powerful platforms and networks of data brokers would be an excellent first step in the process of taming this power of providence.

REFERENCES

Abent, Eric. 2017. "Google Play Prepares to Remove Apps over Missing Privacy Policies." Slashgear.com, February 9, 2017. https://www.slashgear.com/google-play-prepares-to-remove-apps-over-missing-privacy-policies-09474464/.

Acquisti, Alessandro. 2010. "The Economics of Personal Data and the Economics of Privacy." Joint WPISP-WPIE Roundtable OECD, December 1, 2010.

Ansip, Andrus. 2015. "Productivity, Innovation and Digitalisation—Which Global Policy Challenges?" Speech at Bruegel annual meeting, September 7, 2015. https://ec.europa.eu/commission/commissioners/2014-2019/ansip/ announcements/speech-vice-president-ansip-bruegel-annual-meeting-productivity-innovation-and-digitalisation-which_en.

Article 29 Working Party. 2010. "Opinion 2/2010 on Online Behavioural Advertising." Adopted on June 22, 2010 (WP 171).

Article 29 Working Party. 2007. "Opinion 4/2007 on the Concept of Personal Data." Adopted on June 20, 2007 (WP136).

Autorité de la Concurrence and Bundeskartellamt. 2016. "Competition Law and Data." May 10, 2016. http://www.autoritedelaconcurrence.fr/doc/ reportcompetitionlawanddatafinal.pdf.

Barocas, Solon, and Helen Nissenbaum. 2014. "Big Data's End Run around Procedural Privacy Protections." *Communications of the ACM* 57, no. 11: 31–33.

Barocas, Solon, and Andrew Selbst. 2016. "Big Data's Disparate Impact." *California Law Review* 104: 671–732.

Blume, Peter. 2014. "The Myths Pertaining to the Proposed General Data Protection Regulation." *International Data Privacy Law* 4, no. 4: 269–73.

Borgesius, Frederik Zuiderveen, and Joost Poort. 2017. "Online Price Discrimination and EU Data Privacy Law." *Journal of Consumer Policy* 40: 347–66.

Bundeskartellamt. 2016. "Bundeskartellamt Initiates Proceeding against Facebook on Suspicion of Having Abused Its Market Power by Infringing Data Protection Rules." Accessed August 28, 2017. https://www.bundeskartellamt. de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_ Facebook.html.

Cohen, Julie. 2016. "The Regulatory State in the Information Age." *Theoretical Inquiries in Law* 17, no. 2: 369–414.

Competition and Markets Authority. 2015. "The Commercial Use of Consumer Data: Report on the CMA's Call for Information." CMA38, June 2015.

Constine, Josh. 2017. "Facebook Now Has 2 Billion Monthly Users . . . and Responsibility." Tech Crunch.com, June 27, 2017. Accessed August 27, 2017. https://techcrunch.com/2017/06/27/facebook-2-billion-users/.

Cormack, Andrew. 2016. "Is the Subject Access Request Right Now Too Great a Threat to Privacy?" *European Data Protection Law Review* 2, no. 1: 15–27.

Easterbrook, Frank. 1984. "Limits of Antitrust." *Texas Law Review* 63: 1–40.

EPIC. 2017. "EPIC Files FTC Complaint to Stop Google from Tracking In-Store Purchases." July 31, 2017. https://epic.org/2017/07/ epic-files-ftc-complaint-to-st.html.

Equality Act. 2010. https://www.legislation.gov.uk/ukpga/2010/15/contents.

EU Commission. 2009. "Guidance on the Commission's Enforcement Priorities in Applying Article 82 of the EC Treaty to Abusive Exclusionary Conduct by Dominant Undertakings." OJ [2009] C 45/7.

EU Commission—Antitrust. 2017. "Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service." Press release, June 27, 2017.

EU Commission. 2017. "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions on the Mid-Term Review on the Implementation of

the Digital Single Market Strategy: A Connected Digital Single Market for All." COM (2017) 228 final.

"European Parliament and Council Directive 46/EC of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data [1995]." (Directive 95/46) OJ L281/23.

Federal Trade Commission. 2016. "Big Data: A Tool for Inclusion or Exclusion? Understanding the Issues." June 2016. https://www.ftc.gov/system/files/ documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf.

Federal Trade Commission. 2014. "Data Brokers: A Call for Transparency and Accountability." May 2014. https://www.ftc.gov/system/files/documents/ reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

Gellert, Raphael, Katja de Vries, Paul de Hert, and Serge Gutwirth. 2012. "A Comparative Analysis of Anti-Discrimination and Data Protection Legislations." In *Discrimination and Privacy in the Information society: Data Mining and Profiling in Large Databases*, edited by Bart Custers et al., 61–89. Berlin-Heidelberg: Springer.

Helberger Natali, Kleinen-von Königslöw Katharina, and Rob van der Noll. 2015. "Regulating the New Information Intermediaries as Gatekeepers of Information Diversity." *Info* 17, no. 6: 50–71.

House of Lords—Select Committee on the EU. 2016. "Online Platforms and the Digital Single Market." HL Paper 129. April 20, 2016.

Korff, Douwe. 2014. "The Proposed General Data Protection Regulation: Suggested Amendments to the Definition of Personal Data." *EU Law Analysis*. http:// eulawanalysis.blogspot.co.uk/2014/10/the-proposed-general-data-protection. html.

Lamadrid, Alfonso. 2017. "Google Shopping Decision—First Urgent Comments." *Chilling Competition*, June 27, 2017. https://chillingcompetition.com/2017/06/ 27/google-shopping-decision-first-urgent-comments/.

Lenard, Thomas M., and Paul H. Rubin. 2010. "In Defense of Data: Information and the Costs of Privacy." *Policy and Internet* 2, no. 1: 149–83.

Lum, Kristian, and William Isaac. 2016. "To Predict and Serve?" *Significance* 13: 14–19.

Madden, Mary, Michele E. Gilman, Karen E. C. Levy, and Alice E. Marwick. 2017. "Privacy, Poverty and Big Data: A Matrix of Vulnerabilities for Poor Americans." *Washington University Law Review* 95: 53–125.

Mansell, Robin. 2015. "Platforms of Power." *Intermedia* 43, no. 1: 20–24.

Odudu, Okeoghene. 2010. "The Wider Concerns of Competition Law." *Oxford Journal of Legal Studies* 30, no. 3: 599–613.

Pasquale, Frank. 2013. "Privacy, Antitrust and Power." *George Mason Law Review* 20, no. 4: 1009–24.

Pasquale, Frank. 2015. *The Black Box Society—The Secret Algorithms That Control Money and Information*. Harvard: Harvard University Press.

Powles, Julia, and Andrew Selbst. 2017. "Meaningful Information and the Right to Explanation." *International Data Privacy Law* 4, no. 1: 233–42.

Ramirez, Edith. 2013. "The Privacy Challenges of Big Data: A View from the Lifeguard's Chair." *Keynote address—Technology Policy Institute Aspen Forum*, August 2013.

Scott, James C. 1998. *Seeing Like a State*. New Haven: Yale University Press.

Sweeney, Latanya. 2013. "Discrimination in Online Ad Delivery: Google Ads, Black Names and White Names, Racial Discrimination, and Click Advertising." *Acmqueue* 11, no. 3: 1–19.

Taylor, Linnet. 2016. "Data Subjects or Data Citizens? Addressing the Global Regulatory Challenge of Big Data." In *Information, Freedom and Property: The Philosophy of Law Meets the Philosophy of Technology,* edited by Mireille Hildebrandt and Bibi van den Berg, 81–106. Oxford: Routledge.

Taylor, Linnet. 2017. "What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally." June 26, 2017. Available at SSRN: https://ssrn.com/abstract=2918779.

TNO Report. 2015. "Digital Platforms: An Analytical Framework for Identifying and Evaluating Policy Options." TNO (2015) R11271.

Townley, Christopher, Eric Morrison, and Karen Yeung. 2017. "Big Data and Personalised Price Discrimination in EU Competition Law." King's College London Law School Research Paper No. 2017-38. Available at SSRN: https://ssrn.com/abstract=3048688 or http://dx.doi.org/10.2139/ssrn.3048688

Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth.* New Haven: Yale University Press.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. 2017. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation." *International Data Privacy Law* 7, no. 2: 76–99.

Wikipedia. 2017. "Eye of Providence." Accessed September 13, 2017. https://en.wikipedia.org/wiki/Eye_of_Providence.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30: 75–89.

Case Law

Case C-434/16 *Peter Nowak v Data Protection Commissioner*, Opinion of Advocate General Kokott delivered on July 20, 2017 EU:C:2017:582.

Case C-213/15 *European Commission v Patrick Breyer* EU:C:2017:563.

Joined Cases C-141/12 and C-372/12 *YS v Minister voor Immigratie, Integratie en Asiel and Minister voor Immigratie, Integratie en Asiel v M and S* EU:C:2014:2081.

Case C-131/12 *Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González* EU:C:2014:317.

Case C-236/09 *Association Belge des Consommateurs Test-Achats ASBL and Others v Conseil des ministres* EU:C:2011:100.

Social Media Power and Election Legitimacy

DAMIAN TAMBINI

INTRODUCTION: SOCIAL MEDIA, PLATFORM DOMINANCE, AND ELECTORAL LEGITIMACY

Debate about the Internet and democracy has evolved from starry-eyed hope (Rheingold 1995; Tambini 1998), through critical realism (Zittrain 2008; Howard 2006; Sunstein 2001), to despair (Barocas 2012; Morozov 2011; Kreiss 2012). Recent elections have called into question the promise of the Internet to provide expanding resources for information and deliberation (Tambini 2000). Growing numbers of commentators argue that the Internet agora has been displaced by the monopolized Internet of "surveillance capitalism" in which a small number of immensely powerful platform companies (Zuboff 2015) provide integrated services of targeted propaganda and misinformation undermining campaign fairness by rewarding richer campaigns and those that are increasingly able to bypass existing regulatory frameworks. In recent elections, data-driven campaigns, supported by surveillance technologies that game privacy protection to profile voters and target their weaknesses have been widely criticized. (Barocas 2012; Kreiss 2012, 2016; Howard and Kreiss 2009; Tambini et al. 2017). Some, including Epstein (this volume) go so far as to claim that powerful intermediaries such as Google and Facebook can and do influence the outcome of elections.

At the same time, the shock results of votes in the UK referendum and US elections led in 2016 to widespread questioning of the role of social media, which was seen as responsible for distributing fake news (Allcott and Gentzkow 2017; Tambini 2017), using manipulative psychometric profiling (Cadwalladr 2017), and undermining authoritative journalism (Bell, this volume; Allcott and Gentzkow 2017, 211) and ultimately the fairness and transparency of elections.

This chapter examines the charge against the social media in recent elections, with a focus on the question of dominance: whether the powerful position of a few platforms in political campaigning—and particularly Facebook—is undermining electoral legitimacy. The focus will be on the UK, which has particularly high levels of online and Facebook use, and the referendum in 2016 and general election in 2017, which offer useful contrasting examples of recent campaigns. This chapter draws on interviews conducted with campaigners on the state of the art in targeted campaigning during the referendum in 2016, and a study of online ads used in the 2017 election conducted in collaboration with the grassroots group Who Targets Me.

MEDIA AND ELECTORAL LEGITIMACY: THE FRAMEWORK

A number of national and international rules exist to prevent media and communications undermining the legitimacy and integrity of elections and referenda (Council of Europe 2003; Parliamentary Assembly of the Council of Europe 2001). On the international level, intergovernmental organizations such as the Organisation for Security and Cooperation in Europe (OSCE), the Council of Europe, the European Union, and the UN operate election-monitoring projects to ensure free and fair elections. The issue of media influence on elections, and government capture of media have become increasingly important for these monitoring missions but international organizations have done little to deal with the social media challenge.

The OSCE member states must commit to secure free and fair elections, and in particular: "[e]nsure that political campaigning can be conducted in an open and fair atmosphere without administrative action, violence, intimidation or fear of retribution against candidates, parties or voters; (and) [e]nsure unimpeded media access on a non-discriminatory basis" (OSCE 2010, 18).

These and the other commitments contained in the OSCE election guidelines and similar documents such as the Venice Commission (2010) guidelines have led to the development of sophisticated tools for monitoring mass media during elections. According to the OSCE website,

Election observation missions examine the coverage given to candidates in both state and privately owned media. Beyond parties and candidates themselves, the media are the most important source of election-related information for the public. Their ability to function freely and independently is essential to a democratic election. [. . .] An observation mission also assesses media laws, the performance of regulatory bodies, and whether media-related complaints are handled fairly and efficiently.

According to Rasto Kuzel, OSCE election media analyst, "media-monitoring projects can provide the general public with benchmarks to judge the fairness of the entire election process. This function is vital even in those countries that have a long-term tradition of freedom of speech and freedom of the media" (cited in OSCE 2017a). There have been instances in the past where elections have been scathingly criticized because of the media environment. The OSCE report on the 2015 Tajikistan elections for example, was critical of a lack of coverage of opposition parties in both state and private media (OSCE 2015, 18).

In 2017, the OSCE conducted a monitoring mission to cover UK elections, as they had done in 2015. But for the first time they added a specific media component to observe the role of key media companies in the election (OSCE 2017b).[1] A full election-monitoring mission of the OSCE according to the guidelines now includes monitoring of national media to examine evidence of systematic bias or exclusion. A key component of this is ensuring that the media are free and there is proper protection for freedom of expression, but guidance is clear that liberty is not enough: it is also necessary to ensure that media are not captured by special interests, or systematically biased against groups or interests, and that international standards such as those of the UN and the Office of Democratic Institutions and Human Rights (ODIHR) and the Council of Europe are respected.

Domestically, national election laws, media regulation, and campaign finance rules have been adapted to protect elections from the potential threat that mass media propaganda may pose, and in particular to ensure that elections are fair, clean and transparent. Election laws establish limits to spending and/or donations to election campaigns, which are defined as printing, distribution, and production of campaign messages, largely through the media. The UK for example meets its international obligations to hold free and fair elections by implementing the Representation of the People Act 1983.

1. A list of election monitoring organizations can be found on the website of the Ace Project, a UN-endorsed monitoring organization (Ace Project 2017).

In addition, media regulation provides for regulation of impartiality/ balance in broadcasting, and competition and pluralism in media systems as a whole. So for example, in addition to UK broadcasters' general requirements to broadcast news that is impartial "in matters of political and industrial controversy" they have specific duties during election periods: "Due weight must be given to the coverage of major parties during the election period. Broadcasters must also consider giving appropriate coverage to other parties and independent candidates with significant views and perspectives."[2] The UK media regulator Ofcom bases its assessment of what is a major party on previous electoral performance, but is likely in the future to delegate some of these decisions to broadcasters, who will remain bound by their general duties of impartiality.

While the overall objectives of election law and monitoring are similar in mature democracies (to make sure elections are free, fair, and transparent), means vary. Most countries control spending or donations, provide free but rationed political advertising on TV, and operate strict transparency and disclosure rules for parties and campaigns. And during the past 50 years in which broadcasting, most recently TV, has been the dominant medium, broadcasters have been subject to strict obligations to ensure that their potential to influence an election is controlled. Not only do most—at least in Europe—have balance and impartiality obligations, their role in political advertising is also regulated. For example, many democracies, including the UK, France Spain, Denmark, and Ireland operate complete bans on political advertising on TV (see Tambini et al. 2017; Holz-Bacha and Kaid 2006; Falguera, Jones, and Ohman 2014; see also Piccio 2016) and others implement partial bans. Italy for example permits it only on local TV. No such rules exist for social media.

THE SOCIAL CONSTRUCTION OF ELECTION LEGITIMACY

Despite national and international standards, "electoral legitimacy" is not a legal concept. International organizations do not inspect elections to make sure they conform to the rules, and blacklist those that do not. Rather it is a social construct (Suchmann 1995). Election monitors generally write descriptive reports on elections rather than unequivocal endorsements or condemnation. The absence of legitimacy is generally signaled not only by statements of international organizations

2. The UK Communications Regulator Ofcom operates a specific code that broadcast licensees must adhere to during election periods. See Ofcom (2017b).

and monitors but also by low turnout, protest, violence, system crisis, and the withdrawal of consent (see also Mackinnon 2012, 12). However, it is also the case that nondemocratic systems and authoritarian pseudodemocracies can also be highly legitimate in the eyes of their populations, in part because of the lack of an independent media. In systems of "competitive authoritarianism," open elections may be held, but a lack of real media independence undermines the process of open deliberation (Way and Levitsky 2002, 57–58). Therefore, the concept of legitimacy proposed for this chapter is as follows: for an election or referendum to be legitimate, results must be accepted both by international standards bodies *and* the overwhelming majority of citizens. And by contrast, where many or most citizens, and/or the majority of standards bodies and election monitors say legitimacy is lacking, we can say an election is illegitimate. Fundamentally election legitimacy is about perceived fairness. Increasingly, governance of mass media and also social media is required to guarantee such fairness.

With the rising importance of media in elections, and what some would even term the "mediatization of politics" (Garland, Couldry, and Tambini 2017; Esser 2013; Kunelius and Reunanen 2016; Hepp 2013) monitors are increasingly taking notice of media system requirements in their assessments. International standards bodies have outlined standards for the media. The obvious next point is whether those standards need to be updated for a period in which social media are increasingly displacing print and broadcasting.

CAMPAIGNS MOVE ONLINE

A growing number of researchers and commentators are concerned about data-driven political campaigning and message targeting on social media. The concerns include privacy (Howard 2006; Kreiss and Howard 2010; Cohen 2012; Barocas 2012); transparency (Kreiss and Howard 2010); campaign finance (Butrymowicz 2009); and the (in)ability of existing electoral laws to maintain a level playing field and thus election legitimacy (Pack 2015; Barocas 2012; Ewing and Rowbottom 2011; Tambini 2017). Researchers have raised longer-term concerns with the undermining of the quality of deliberation; since 2016 the concern has been with the proliferation of messages that were either inconsistent with, or contradictory to, other communications from a campaign. Or third-party messages that were deliberately designed to mislead or provoke. There is also a longer term worry about "political redlining," that is, the ability to

target messaging on a narrow segment of the electorate (Barocas 2012) and exclude others, because they are less likely to vote or do not belong to key swing demographics; and with the overall transparency of political deliberation (Ewing and Rowbottom 2011). One area of concern that links these various claims is the notion that effective targeting may undermine voter autonomy: voters for whom social media is the dominant source of news and information could be inundated with a constant stream of skewed, politically interested messaging that would drown out opposing views; a new form of targeted propaganda.

Following the shock results of the 2016 Brexit referendum and the US election, a wide range of concerns were raised about social media campaigning by a wider range of public commentators. The influence of deliberately targeted "fake news" messages, and the potential for foreign intervention in domestic campaigns, including spooky "psychometric profiling" have been raised by journalists such as Carole Cadwalladr of the *Observer* newspaper (Cadwalladr 2017).

At the time of writing, several investigations into the use of targeting were ongoing: In addition to the US Special Prosecutor's investigation of Russian involvement in the 2016 elections, The Information Commissioner's office (the UK regulator for freedom of information and data protection) was examining the use of data for campaign purposes (Denham 2017); and an investigation by the UK electoral supervisor the Electoral Commission examined potential breaches of campaign funding reporting obligations relating to provision of database and targeting services by Leave.EU (Electoral Commission 2017). While the international agencies such as the OSCE that are responsible for electoral supervision and monitoring have been relatively slow to respond to the challenge of social media, the Council of Europe has carried out a feasibility study for a new recommendation on how democracies might regulate the new practices (Council of Europe 2017).

Despite this gathering storm of debate, there has been a lack of robust and disinterested information on how the campaigns actually work. Research into data-driven campaigning has tended to rely on interviews (Moore 2016; Anstead 2017), ethnography (Nielsen 2012), or legal analysis (Butrymowicz 2009). There is surprisingly little analysis of the messages themselves, or of the validity of some of the more worrying claims about new forms of propaganda. A partial exception is Allcott and Gentzkow (2017). The key proposal of the theoretical literature, namely that the legitimacy of elections and referenda is undermined by these new campaigning tools, has not been tested, and there remains a rather large gap between hype (generally of the dystopian variety) and understanding of how targeted campaigning on social media has in fact been deployed.

The UK referendum of 2016, like the US election of the same year, led to a shock outcome.[3] The discussion following the referendum predictably focused on why there was such a contrast with previous votes, and a tendency to "blame" unwelcome political changes on the Internet. In particular, concerns were expressed about misinformation and "fake news" being distributed online without the skeptical filter of journalism, and about targeted messaging online (Allcott and Gentzkow 2017). Commentators, who themselves had been sidelined by new opinion leaders online, looked for someone to blame, and Facebook was convenient.

In May 2017, after a series of shorter stories, Carole Cadwalladr published a detailed "exposé" relating to opaque links, data sharing, and cross-funding between the UK referendum and the US Trump campaign. Cadwalladr closed the article arguing that "Britain in 2017 . . . increasingly looks like a 'managed' democracy. Paid for by a US billionaire. Using military-style technology. Delivered by Facebook. . . . the first step into a brave, new, increasingly undemocratic world" (Cadwalladr 2017).

In the article she alleged not only that both campaigns were using sophisticated data-driven social media targeting campaigns but also that there was a degree of cross-funding (through provision of benefits in kind such as data services), coordination of campaign data, and learning between the two campaigns. For the politically displaced, the story was attractive, as it offered support to the claim that the result was illegitimate.

In comparison with other advanced democracies, the UK has a very active online population, and users are particularly engaged on social media. More than 82% of British adults used the Internet daily or almost daily in 2016 according to the Office for National Statistics (ONS 2016), and 27% of online adults reported using Facebook on a daily basis. The Internet was according to Ofcom the only news platform with a growing number of users since 2013: 48% of UK adults say they use the Internet to get their news (Ofcom 2017a). According to the same report, 27% of UK adults say they get news from Facebook.

Social media, according to the data from a 2017 report, are the fastest growing news source sector: "overall, 47% of those who use social media for news say they mostly get news stories through social media posts, compared to 30% in 2015." This survey evidence is self-reported, and different surveys vary to an extent. According to the Reuters Institute *Digital News*

---

3. The author acknowledges the excellent research assistance of Sharif Labo for this section.

Report 2017 (reported in this volume), 41% say they use social media for "news" in the UK.

This shift online, and to social media, is reflected also in advertising spending, though estimates of spend vary. Between 2008 and 2016 the "digital" (online) share of US political ad spend rose from 0% in 2008 to an estimated 10% in 2016 (E-Marketer 2016). Given evidence from interviews with campaign leaders (Tambini 2016), and spending returns to the UK Electoral Commission,[4] much more than 10% of election marketing budgets is now spent on digital. In 2015, the first year in which digital spending was reported separately by the Electoral Commission, around 23% of the total spend was digital, with the majority of this being spent on Facebook (Electoral Commission 2016). In the United States, which remains dominated by TV spend, almost a billion dollars, or 10% of spend on political ads was forecast to be spent on online advertising in the 2016 election (E-Marketer 2016).

The reason for this rapid shift of campaign activity online is simple. Social media advertising appears to be more cost-effective than other less "smart" forms of advertising. Of particular interest to political strategists and campaigners is the fact that data-driven campaigns offer superior targeting and audience-segmentation capabilities. Campaigns can get the messages they think will be most persuasive to people who are undecided but likely to vote, in the constituencies that might swing the election, or key voters in a referendum. What is attractive to advertisers is that they can target those key strategic voters with the messages that are most likely to swing those voters on the basis of demographic, political, and even potentially psychometric profiling. According to campaign leaders, strategists are following audiences online, and developing more sophisticated approaches to online advertising. This is generally combined with an attempt to develop shocking and resonant "shareable" messages to harness the organic sharing of propaganda online. According to Andy Wigmore, the campaign director of Leave.EU:

> It didn't matter what was said in the press. The more critical they were of us when we published these articles to our social media, the more numbers we got. So it occurred to us that actually Trump was onto something because the more outrageous he was the more air time he got, the more air time he got the more outrageous he was. . . . The more outrageous we were the more air time we got in the normal media and the more airtime—which was always critical—, the more support we got. . . . The more outrageous we were, we knew that the

4. Researchers examined spending returns as they were published by the Electoral Commission and categorized the payees according to their basic function, in order to identify social media and other forms of spend.

press were going to attack us, which is what they did. We are now anti-establishment full throttle. The more outrageous we were the more attention we got. The more attention we got, the bigger the numbers. (Andy Wigmore, interview, September 2016)

How a Data-Driven Social Media Campaign Works

In order to gain a rich understanding of data-driven campaigning on social media we interviewed referendum campaign leaders.[5] This builds on the work of Anstead (2017) and others. Seven semistructured interviews were conducted with a common template of questions designed to enable the campaigners to outline their approaches to data-driven campaigning, voter profiling, and social media messaging. The interviews were conducted in London August–November 2016, following the referendum to exit the EU. Three were conducted on the phone/Skype, and the others were conducted in person.

In practice, it is impossible to separate the mass media campaign from the social media campaign, and it is impossible to separate the "organic" social media campaign driven by "voluntary" sharing and liking and the use by campaigns of the commercial advertising services offered by social media. Effective campaigns use those three elements together. But in what follows the focus is on the paid element, which has particular implications for election legitimacy, and which often fuels and primes the organic social media campaign, which in turn feeds mass media with stories.

On the basis of the literature review and expert interviews carried out following the 2015 general election and the 2016 Brexit referendum, it is
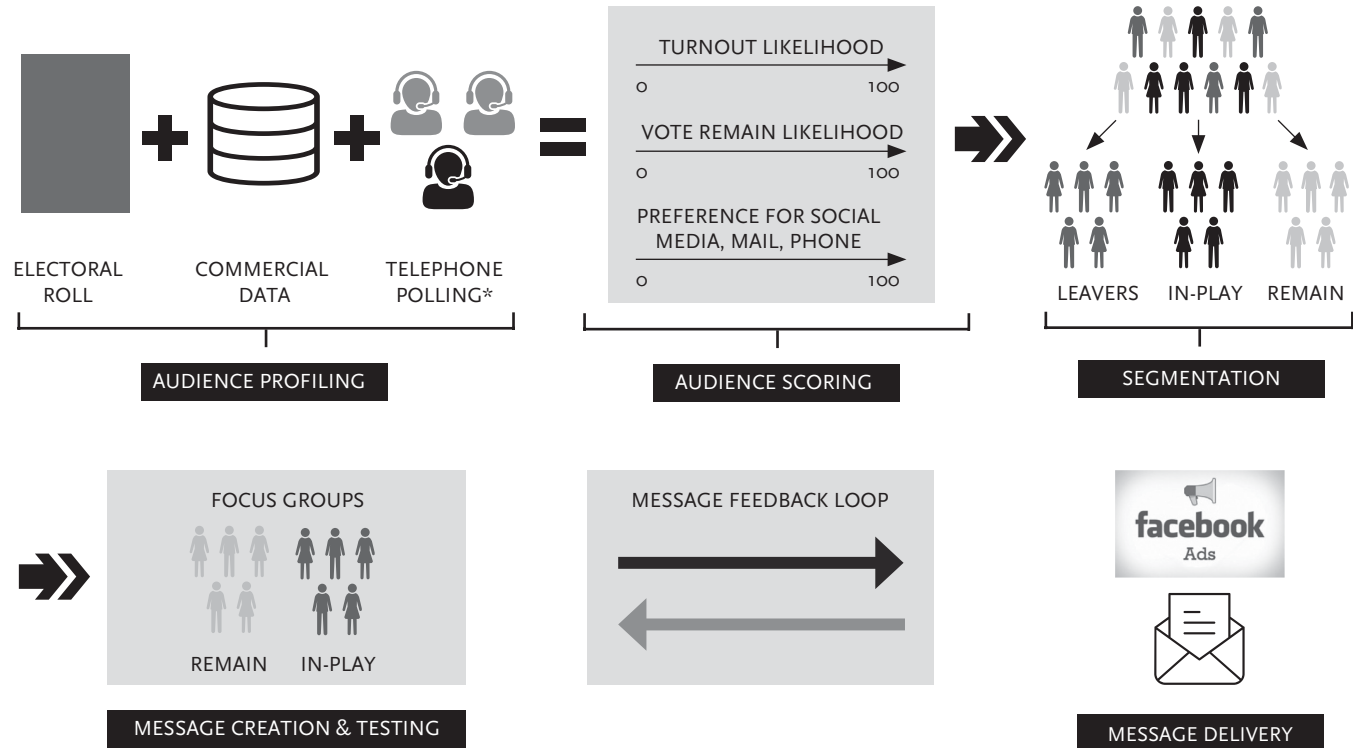
5. To gain an insight into the message-targeting and communications strategy of a modern political campaign we interviewed the key participants from the two officially designated sides: Stronger In and Vote Leave. We were interested in speaking with people who had close operational detail of the campaign strategy; how the key messages were decided on, message sign-off and audience segmentation. We anticipated this would require authorization from senior figures in the campaign and so chose to approach these senior figures first and asked them to suggest people to speak with throughout the campaign organization. We e-mailed interview requests to the heads, deputy heads, and campaign managers. We secured interviews with Jack Straw and Lucy Thomas, the director and deputy director of Stronger In, and Matthew Elliott, the CEO of Vote Leave. These interviews provided the names of other individuals, consultancies, and agencies involved in the campaigns that we subsequently approached as well as providing useful operational detail of the campaigns, especially on the Stronger In side. We also interviewed Andy Wigmore of Leave.EU. All interviews were transcribed and analyzed according to a meaning-condensation process with a focus on ascertaining expert views on processes of segmentation and profiling. Respondents were asked to go on the record and did so. The following section is based on a thematic analysis of their responses.

possible to outline the following generic stages in building a social media campaign (Figure 11.1 and 11.2):

1. Building the audience. Using a wide range of database-building techniques, campaigners build databases of potential supporters, link these to various forms of purchased and freely available data, such as the electoral register, existing party membership, and canvassing lists, cold-calling records, and "opt in" data-harvesting techniques such as surveys, competitions, and petitions, which are increasingly carried out online.

2. Audience segmentation. There are various approaches to audience segmentation, which combine the following types of criteria: (1) marginality: Is the voter situated in a constituency that is possible to win, that is, a target constituency? Is the voter undecided?; (2) the basic demographic information attached to this voter (e.g., gender, age, income, education); (3) previous voting record (including likelihood of actually voting); (4) evidence of current opinions and "hot-button" issues; and (5) social media activity and degree of its influence. The different campaigns in 2016 each had a slightly different approach to profiling, but each attached a score and a profile to each potential voter using data from the electoral role. In elections, parties are able to learn between elections, but in referenda regulation requires them to "start from scratch" (Matthew Elliott, interview, September 2016) and destroy data on completion.

4. Message creation and testing. The process of finding messages that are effective and resonate with potential voters has in recent years involved extensive "focus group" testing, and repetition of a narrow range of messages that have been vetted and signed off by senior politicians. The social media campaign, by contrast tends to be more dynamic, with messages devised and tested online throughout each day of the campaign using processes of "A/B" testing, whereby messages are selected on the basis of their resonance rather than ideological or political selection.

5. Message targeting and delivery. Many campaigners report that they are focusing more of their advertising spend on digital, and they are doing this because they have a clear sense that social media platforms in particular are much more cost-effective than for example, press, display, or direct mail marketing techniques. The question of whether specific messages are targeted on the basis of the segmentation and profiling techniques described at (2) is the black box of research on social media and campaign targeting. Campaigners frequently claim that they are able to target messages on an individual basis, and serve individually targeted messages that are designed to appeal to particular demographic, education level, psychological, or geographic groups.

FOCUS GROUPS   PHONE SURVEY

ATTITIDUNAL + DEMOGRAPHIC FACTORS

SEGMENTATION

100

LEAVE

ELECTORAL
ROLL

facebook
Ads

BBC
LOCAL RADIO

Commercial   Individual
Proprietary   Propensity
Data   Models

HEADS vs HEARTS   DISENGAGED MIDDLE   MESSAGE DELIVERY
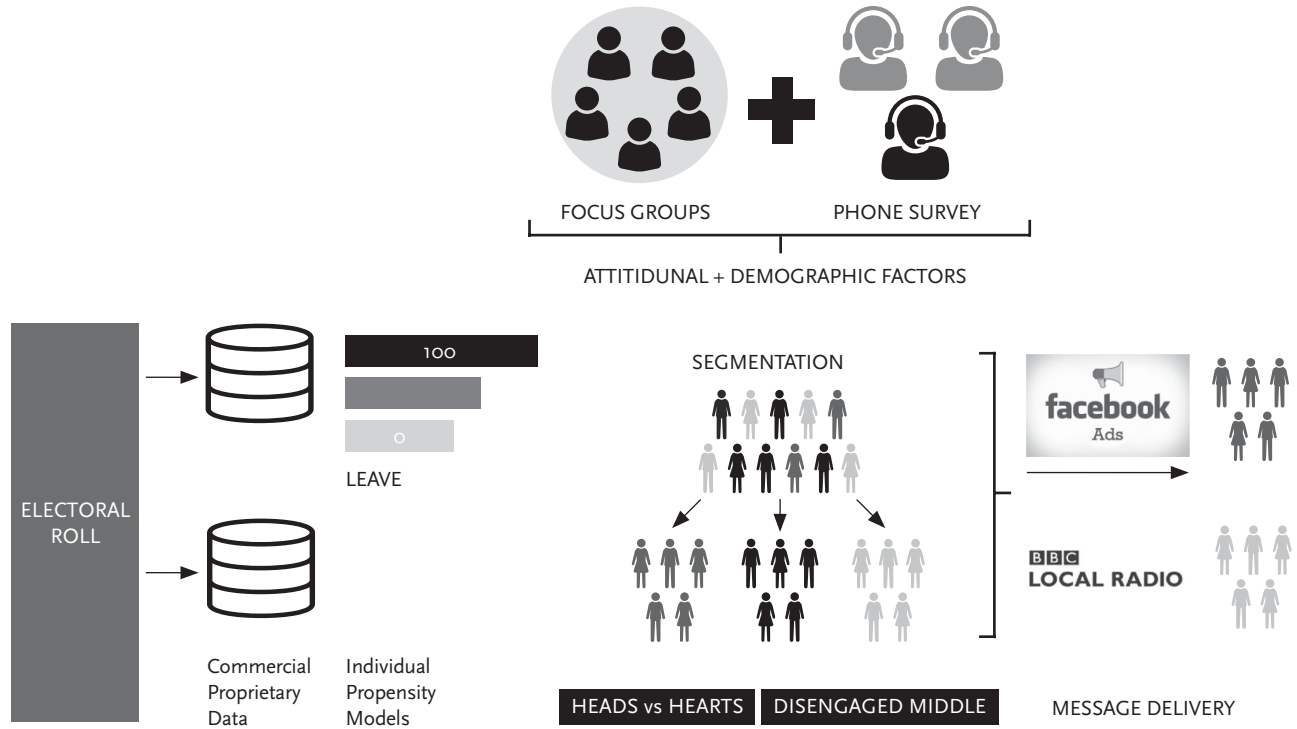
Figure 11.1: Data-driven profiling in online political advertising: Interview findings on the Vote Leave Campaign.

ELECTORAL ROLL + **COMMERCIAL DATA** + **TELEPHONE POLLING*** =

TURNOUT LIKELIHOOD
0   100

VOTE REMAIN LIKELIHOOD
0   100

PREFERENCE FOR SOCIAL MEDIA, MAIL, PHONE
0   100

LEAVERS   IN-PLAY   REMAIN

AUDIENCE PROFILING   **AUDIENCE SCORING**   **SEGMENTATION**

FOCUS GROUPS

REMAIN   IN-PLAY

MESSAGE FEEDBACK LOOP

facebook Ads

MESSAGE CREATION & TESTING   **MESSAGE DELIVERY**

*National Rep. Sample

Figure 11.2: Data-driven profiling in online political advertising: Interview findings on the Stronger In Campaign.

THE REFERENDUM: VOTER PROFILING AND SEGMENTATION IS GETTING SMARTER

All the campaigns used a wide combination of techniques to build the audience and profile and segment it. This involved complex modeling of relationships among demographic characteristics, previous expressions of opinion, and stated voting intentions. Such profiling can involve hundreds of data points from dozens of sources. As Will Straw, CEO of Stronger explained,

> These were opinion groups with demographic characteristics. So for the segmentation— . . . they identified common traits based on how people answer specific questions. Such demographic characteristics as well, but mainly based on their answers to questions that have been asked. What that threw up was some really interesting characteristics of these different groups. So you could say that the average person in this segment would be better or worse off than average, would be overall younger than average, would get their media from the BBC versus newspapers versus online. Would have these attitudes to the EU. These other issues would be of interest to them. Whether they are members of particular groups and so on. So some quite good general information. Then throughout the campaign we used that sub-segmentation to drive our focus group work. So when we had focus groups, I think we had close to thirty focus groups over the course of the campaign, we would get—You might have four to eight different tables up the focus group depending on the size, but it would be a male heads versus heart and a female strong sceptic group [. . .] Then we would have monthly depth polls which went back through the segmentation and we could see how the segments were shifting, both in their total numbers but also in their views of the Referendum. Then we would underneath that be able to track how people responded to different questions, certainly immigration question or the economy. How were we best able to get our messages across to those different groups.

Given that this process of segmentation and profiling is subsequently used in order to determine to whom messages are addressed and which messages are addressed to those voters the cumulative effect of this data-driven profiling is of interest: it is likely, for example, that this profiling procedure may inadvertently result in different messages being targeted on the basis of protected characteristics, such as ethnic or religious grouping. Profiling and segmentation has always taken place to an extent on a geographical basis; these new techniques merely offer a much cheaper and effective way of doing so and thus may raise new concerns (see Lynskey,

this volume). Profiling and segmentation has always taken place but rapid innovation makes individual level targeting much more efficient and sophisticated.

MESSAGE TARGETING AND DELIVERY

One of the striking things about all the major campaigns to leave the EU is that they both took the strategic decision to focus the majority of their resources and energy on Facebook. There was strong agreement that it was simply the most effective form of political advertising. All the leaders said that Facebook was crucial, and particularly the two Leave campaigns. Andy Wigmore claimed that his team made a strategic decision early in the campaign to put the entire ad budget into Facebook. And this was true also of his counterparts in the other (official) Leave campaign, such as Matthew Elliott.

> **Elliott:** . . . almost nothing went in traditional advertising. Maybe one or two things which were more aimed at the press and getting coverage, but almost nothing went on traditional advertising.
> **DT:** A lot on social media and—
> **Elliott:** A load on social media, a lot of it geared towards the end of the campaign.
> **DT:** So increasingly that social media spend is Facebook?
> **Elliott:** Facebook yes.

EU REFERENDUM CAMPAIGN FOCUS: EXPENSES FOR INDIVIDUALS WITH LESS THAN £250K SPEND

In order to further understand how the campaigns were approaching social media, and test some of the claims made by our interviewees, we also examined the Electoral Commission on returns on the referendum. Taking one illustrative example, the returns released in early 2017 show that social media now account for most of the spending of the major parties. While the overall sums are relatively small, due to the Electoral Commission spending caps, social media have become the largest recipient of advertising spending, with most of this going to Facebook (Figure 11.3, Table 11.1).

The data covers those campaigners that reported spend of between £10,000 and £250,000 at the EU Referendum. Any individual or organization that intended to spend more than £10,000 was required to register as a
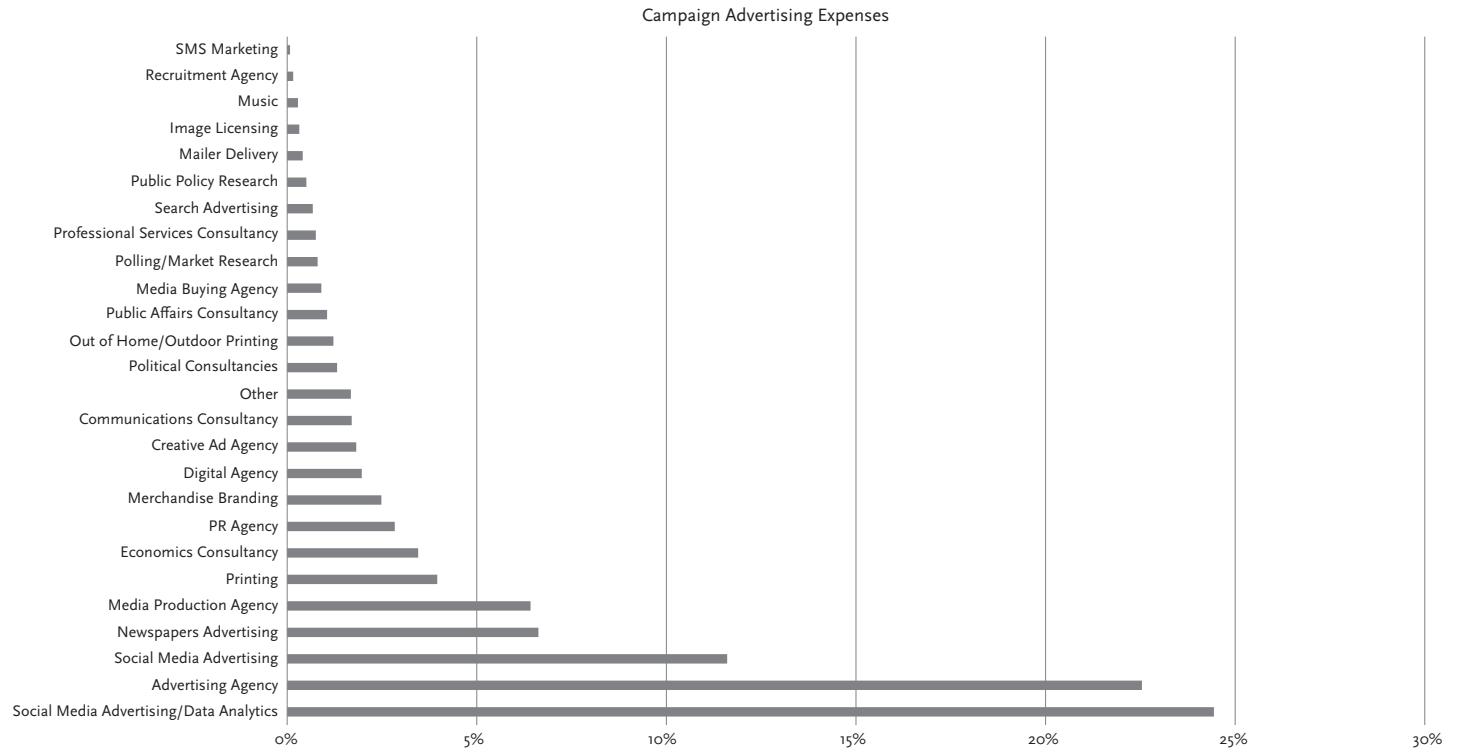
Figure 11.3: Campaign ad spend: breakdown.

Source: Analysis of Electoral Commission spending returns.

Table 11.1. MARKETING, MEDIA, AND MARKET RESEARCH SPENDING TOTALS

Category	Spend	Percentage
TOTAL	**£3,172,565.83**	
Social Media Advertising/ Data Analytics	£775,315.18	24%
Advertising Agency	£715,059.35	23%
Social Media Advertising	£368,085.52	12%
Newspapers Advertising	£210,169.50	7%
Media Production Agency	£203,565.10	6%
Printing	£125,554.95	4%
Economics Consultancy	£109,594.80	3%
PR Agency	£90,006.22	3%
Merchandise Branding	£78,805.80	2%
Digital Agency	£62,371.99	2%
Creative Ad Agency	£57,792.58	2%
Communications Consultancy	£54,000.00	2%
Other	£53,318.45	2%
Political Consultancies	£41,730.00	1%
Out of Home/Outdoor Printing	£38,723.16	1%
Public Affairs Consultancy	£33,382.80	1%
Media Buying Agency	£28,583.80	1%
Polling/Market Research	£25,489.60	1%
Professional Services Consultancy	£24,000.00	1%
Search Advertising	£21,400.00	1%
Public Policy Research	£16,034.10	1%
Mailer Delivery	£13,034	0%
Image Licensing	£10,133.00	0%
Music	£9,000.00	0%
Recruitment Agency	£5,016.00	0%
SMS Marketing	£2,400.00	0%

Source: Analysis of Electoral Commission spending returns.

"permitted participant" and submit expenses to the Electoral Commission earlier than groups spending more than £250,000. A few parties who spent in excess of £250,000 submitted their expenses earlier. The expenses analyzed are in the categories of marketing, media, and market research. They make up 66% of the total expenses of £4.8 million reported. Expenses outside the campaign period are not included.

One difficulty we encountered analyzing this data is that a great deal of the advertising spend is channeled through intermediaries such as advertising agencies. Advertising agencies tend to be active across different media. That said, the highest spend was in social media both through

agencies and directly. By examining the spending returns we found that most social spend went to Facebook. An important implication of this is that social media spend, which is growing to become a disproportionate size of the pie, is hardly broken down at all. It therefore becomes an obscure black box in regulated campaigns.

## IS FACEBOOK BECOMING A ONE-STOP-SHOP FOR ELECTION CAMPAIGNING? SOME FINDINGS FROM THE LSE/WHO TARGETS ME PROJECT

During the 2017 UK General Election the social enterprise Who Targets Me persuaded approximately 11,000 volunteers to download a browser plugin. The plugin scraped political advertising from their Facebook feeds and created a large database that contained the almost 4.5 million records of exposure to Facebook ads (Figures 11.4–11.7). Voters continued to volunteer during the election campaign, and this, together with obvious self-selection biases, means that the data is not a representative record of all the ad exposures, but it is a valuable record of a large sample of advertisements that can provide some general indications of the kinds of activities of party political advertisers and of Facebook users.[6]

These initial results from the LSE/Who Targets Me research collaboration offer significant evidence that Facebook is not only an important part of the message delivery machinery for targeted advertising services but also is emerging into a one-stop-shop for fundraising, recruitment, profiling, segmentation, message targeting, and delivery. This vertical integration of campaign services, and its operation by a company that in most of the globe is foreign, will have serious implications for future election legitimacy if it is to continue unchecked.

SOME IMPLICATIONS OF SOCIAL MEDIA CAMPAIGNING

The shift to social media therefore poses some serious potential concerns for election legitimacy but, partly because of the lack of transparency of the

6. The dataset is a collection of 1,341,004 impressions of 162,064 unique Facebook advertisements. The data was gathered between May 27, 2017, and June 18, 2017, via a Chrome plugin installed by volunteers taking part in the Who Targets Me project (https://whotargets.me/). The project is intended to capture and save the content of political Facebook ads served to participating volunteers, and more information on the
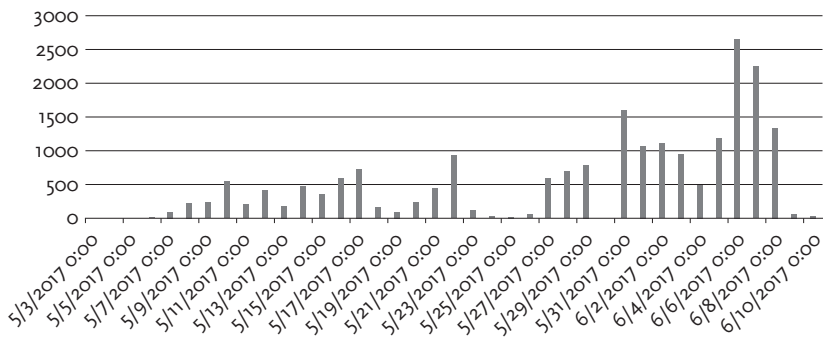
Figure 11.4: Total impressions of political ad per day.
Total number of ads served to our sample on Facebook during the election campaign. Note that the sample grew during the campaign, so this should not be seen as an indication of numbers of ads viewed.
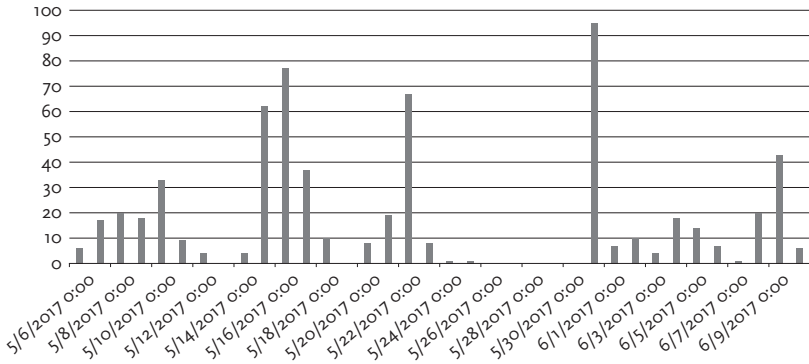
Note: This is a count of the total number of ads served (unique ad impressions) per day to users with the Who Targets Me plugin installed. Data from May 30 has been removed from this graph, due to an error in the plugin on that day which caused an unknown number of duplicate ad impressions to be recorded on that day. The data on which this graph is based is a database of 1,341,004 adverts captured by the Who Targets Me project, of which 20,958 were judged to be political in nature on the basis of a filter applied to the names of the advertisers named as responsible for the ads. The filter sought to detect the main political parties by searching for text matches to *labour*, *conservative*, *liberal democrat*, *ukip*/*independence party*, *momentum* (where * is a wildcard and the search was case-insensitive) in the names of the advertisers.

platform, and of the process of campaigning, claimsare difficult to assess through research. This fuels the conspiracy theories.

In addition to what seems to be a process of consolidation and vertical integration of campaign activity in one platform, namely Facebook, allegations have been made of various forms of foreign involvement, biases in distribution of key messages, bias against small parties, bias against new entrants, bias against parties with socially diverse supporters, bias against certain campaign messages/issues, and bias against certain groups of voters—so-called redlining (Kreiss 2012).

Such biases may be unintentional or deliberate. As a hypothetical example, if a party or campaign emerged that was standing on a platform of breaking up social media companies, there would be a strong incentive for social media companies to undermine the visibility of that party. This example may, or may not be far-fetched, but parties already exist that propose radical, sometimes statist solutions that would be hostile to the economic model of the platform companies. Electoral supervision exists
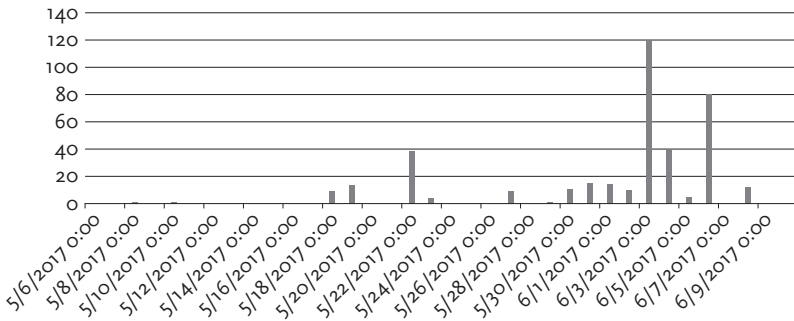
plugin and the team that developed it can be found at https://whotargets.me/about. Volunteers agreed that data could be scraped from their Internet browser when they viewed Facebook. This enabled researchers to monitor the different types of messages that were viewed. Graphs presented here outline the basic content of messages during the GE2017. Future research will analyze targetting strategies, content, and profiling.

Figure 11.5: Political posts containing the word "join."

Posts containing the word "join" or "joining" were more evenly spread throughout the campaign. The high volumes indicate that parties were active in using Facebook as a recruitment campaign—to build their databases.
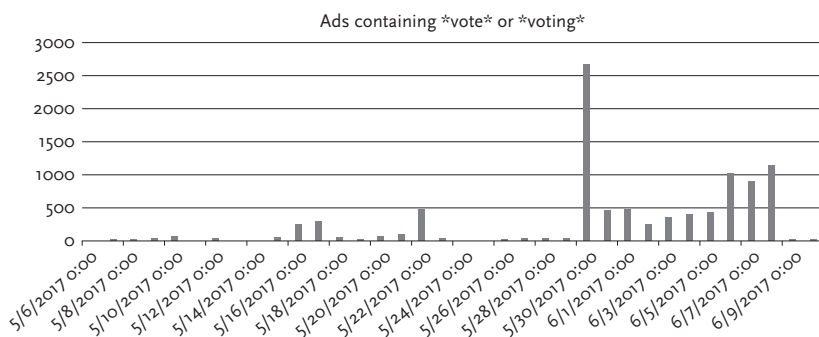
Note: This is a count of the total number of ads served (unique ad impressions) per day to users with the Who Targets Me plugin installed, filtered to include only ads (conservatively) run by political parties or allies (Labour, Momentum, Liberal Democrat, Conservative, UKIP) and containing particular keyphrases. Data from May 30 has been removed from this graph, due to an error in the plugin on that day which caused an unknown number of duplicate ad impressions to be recorded on that day. The data on which this graph is based is a database of 1,341,004 adverts captured by the Who Targets Me project, of which 20,958 were judged to be political in nature on the basis of a filter applied to the names of the advertisers named as responsible for the ads.

Figure 11.6: Ads containing the word *donat*.

The relatively high volume of ads containing the words "donate" or "donation" confirms that FB was a significant fundraising platform for parties throughout the campaign and even after it.

Note: This is a count of the total number of ads served (unique ad impressions) per day to users with the Who Targets Me plugin installed, filtered to include only ads (conservatively) run by political parties or allies (Labour, Momentum, Liberal Democrat, Conservative, UKIP) and containing particular keyphrases. Data from May 30 has been removed from this graph, due to an error in the plugin on that day which caused an unknown number of duplicate ad impressions to be recorded on that day. The data on which this graph is based is a database of 1,341,004 adverts captured by the Who Targets Me project, of which 20,958 were judged to be political in nature on the basis of a filter applied to the names of the advertisers named as responsible for the ads.

Figure 11.7: Ads containing the word "vote" or "voting."
Adverts from all parties containing the words "vote" or "voting." These instructional posts cluster at the end of the campaign period.

Note: This is a count of the total number of ads served (unique ad impressions) per day to users with the Who Targets Me plugin installed, filtered to include only ads (conservatively) run by political parties or allies (Labour, Momentum, Liberal Democrat, Conservative, UKIP) and containing particular keyphrases. Data from May 30 has been removed from this graph, due to an error in the plugin on that day which caused an unknown number of duplicate ad impressions to be recorded on that day. The data on which this graph is based is a database of 1,341,004 adverts captured by the Who Targets Me project, of which 20,958 were judged to be political in nature on the basis of a filter applied to the names of the advertisers named as responsible for the ads.

to ensure that elections—and the deliberative processes that surround them—are seen to be fair. They are increasingly powerless to do so in the face of opaque platforms.

In order for elections to be legitimate, voter choices should be demonstrably free and not constrained by propaganda or subject to any form of control or deceit. This is another reason why targeting has been an issue, and "filter bubble" (Sunstein 2001; Pariser 2011) concerns have arisen. While the "jury is out" on the extent to which intermediaries narrow or broaden access to sources of information (see Newman and Fletcher this volume; Ofcom 2017a; Helberger this volume) the danger of social media targeting offers new opportunities in election campaigns for those wishing to shift opinion and votes with scant regard for the truth.

There have thus been important concerns about voter autonomy and new forms of manipulation and propaganda. According to the UK election lawyer Gavin Millar,

> Section 115 of the 1983 Act creates an offence of "undue influence." Amongst other things this [. . .] prohibits impeding or preventing the free exercise of the franchise by duress or any fraudulent device or contrivance. In its long history it has been used against priests and imams preaching politics to the faithful, as well as those who circulated a bogus election leaflet pretending to be from another party [. . .] To me the most concerning is the impact of the targeted

messaging on the mind of the individual voter. A "persuadable" voter is one thing. A vulnerable or deceived voter is quite another. (Millar 2017)

Foreign intervention has been a feature of much of the public debate, particularly links between the Trump campaign and Russia and the Brexit campaign and the United States, and involvement of Russia in various elections in France and Germany. In the UK this has led to the Electoral Commission enquiring about the funding of the various leave campaigns for example.

It will be pointed out that allegations about social media bias and control are speculation. But speculation and conspiracy theory is what undermines trust in democracy. One of the basic premises of free and fair elections is that the contest is free and fair, *and perceived as such*. This is why simplicity and transparency are so important. While media system capture and bias is inevitable in a mass media system, whether that is one dominated by private media, public media, or some variant (Hallin and Mancini 2004), those biases are by their nature transparent and obvious for everyone to see.

WAS IT FACEBOOK WOT WON IT?

If an election is swung by a private company it is more likely to lose legitimacy in the eyes of citizens and the international community.[7] The evidence from the UK is mixed: on the one hand, the mere fact that there has been a loud debate on these issues since the 2016 referendum suggests that data-driven campaigning has had a negative impact on election legitimacy. But others claim that this is simply sour grapes—losers questioning the process. There is something in both arguments and they are not mutually exclusive. Empirical data on the role of Facebook in the overall information ecology is ambivalent, in part because Facebook data is difficult to access.

Facebook is market-dominant as a social media company (particularly if we include Instagram and WhatsApp) but not as a media company. In terms of time spent, and survey reports on where people get their news, it is certainly not dominant. But in terms of deliberation and information gathering related to elections, it is becoming the crucial platform in some countries, which is reflected in the shift of UK political advertising onto

7. The title of this section is a reference to an infamous front page headline in the British tabloid the *Sun*, which gleefully claimed the day after the 1997 election victory of Tony Blair that "it was the Sun Wot Won it."

the platform over the past five years. Facebook, in particular, is emerging as a vertically integrated one-stop-shop for fundraising, recruitment, database building, segmentation, targeting, and message delivery. As a result, there is a paradox: the complex process of deliberation and debate during an election cycle, the flow of ideas, memes, reversals of public opinion, and fluctuations of fortune of individual politicians is now more knowable than ever before. The problem, for most democracies, is that it is knowable by a company based in California that has no intention of sharing that knowledge with anyone, apart from those able to pay for it, without asking too many questions about what they will do with this data or where they are based.

This is not Facebook's fault, but it is a fact, and in the history of elections it is a novel one. There are multiple sensitivities about foreign involvement in media systems. Most countries have maintained rules preventing foreign ownership of media companies under pressure from trade liberalization (this after all is why Rupert Murdoch had to take US citizenship) and the United States, the UK, and most other mature democracies have specific laws that prohibit foreign involvement in campaign funding. So the mere fact of a private, foreign company having this position cuts across the spirit of these previously existing laws.

WHY DOMINANCE MATTERS

Until now, this chapter has focused on the implications for democratic legitimacy of data-driven social media–based election campaigns. The question that follows is to what extent this is a problem of dominance—or conversely whether increasing choice, plurality, and switching between social media platforms could mitigate any of these concerns.

The short answer is that dominance matters. A good deal of the concerns we have discussed would be allayed, to an extent by more competition and pluralism in social media platforms.

Censorship Effects

If a nondominant platform takes down a post, that could be described as editorial discretion. If a dominant platform takes down or blocks a post, a person or a topic, that is censorship. It is of little import whether the material is taken down by a human due to a rule violation, or by an algorithm for reasons that are not understood. Dominant platforms censor.

Prominence Effects

Platforms can also use their dominant position to promote messages. This has been most evident when Google and others took positions in relation to intellectual property and net neutrality discussions in the US Congress, and platforms have also lobbied on gay rights issues. This is of course what is traditionally done by newspapers, which is why they are subject to sector specific merger and competition rules that limit market concentration.

Propaganda Bubbles

If one company holds data on you, and one profile is sold on to advertisers and fed into the relevance algorithm that determines what you are exposed to, there is the danger that this single profile will determine the "filter bubble" (Pariser 2011) effect of what you are exposed to. These are complex processes, and as yet little understood (Helberger this volume; Newman and Fletcher this volume). In the context of elections, the "propaganda bubble" effect could undermine legitimacy if there is a genuine lack of exposure pluralism (Helberger 2012) such that individual autonomy and free will is undermined, and deliberation undermined. In other words, each citizen might be better served by living within multiple "filter bubbles" operated with separate data ecologies.

Lack of Competitive Discipline

Where there are high switching costs and consumer lock-in (Barwise and Watkins this volume) users may be less able to exert "democratic discipline" on platforms—for example by demanding greater control over personal data, more transparency about relevance and prominence, and due process and "put back" rights in relation to takedowns and blocking. There is increasing evidence that Facebook is becoming a "one-stop-shop" for political campaigns that need to gather, profile, segment, and target, and that consumer lock in due to a lack of data portability compounds the effects of this.

Separation of Powers and Checks and Balances

Like branches of government, social media companies should be balanced by countervailing power; which can be provided by other social media companies.

A dominant company like Facebook, particularly one that is offering a vertically integrated "one-stop-shop" for election services, is in a historically unique position, and as a foreign company it is a position that if left unchecked will be corrosive of trust and democratic legitimacy.

Some of this is speculation. Some of this, we will be told by Facebook and others, could be wrong. But that is, at least in part, the point: because of a lack of transparency, speculation is necessary. Because of opacity and speculation, electoral legitimacy and democratic legitimacy more widely, suffers. Plurality of platforms would provide an important safeguard to democratic legitimacy.

Social media are not transparent, and the shift of campaigns online undermines the principle of transparency. To a certain extent this directly undermines existing regulation. The Political Parties, Elections and Referendums Act in the UK places a number of requirements on parties to be open about the funding and governance of campaigns. These exist so that citizens can be clear on who is behind any party of campaign. For example, campaigns are obliged to label their leaflets and other materials that. In 2016, the Electoral Commission admitted that these transparency requirements were not possible to enforce effectively online (Electoral Commission 2016). In a world of leaflets, campaigns could simply provide "imprint" information in small print on each leaflet which specified which campaign was behind the leaflet, and voters (and journalists and other campaigns) could find detailed information about the funding of that campaign on the Electoral Commission website. Social media advertising, where ad messages take a simpler format and do not include imprint information, was undermining that key tenet of transparency.

UNDUE INFLUENCE: THE CRISIS OF ELECTORAL LEGITIMACY

An election in the UK shares many of the features of a village fête. People gather in their local village hall or primary school and are met by volunteers puffed up with civic pride. Votes, like raffle tickets, are carried in battered steel boxes to bigger local secondary schools and counted by more local volunteers. The politicians wear retro rosettes, and tears are shed in the great climax of civic participation, when the teller, often in ceremonial garb, announces the count.

Part of the reason for the fusty process and archaic technology, in the era of big data and instant AI-driven feedback is ritual, and part of it is about trust. The two go together, and they are both important factors in the social construction of legitimacy.

But the crucial factor in the legitimacy of elections is fairness. Profound political change and party realignment always involves contestation of legitimacy, and the events of 2016 and 2017 have been no exception. Both losers and winners have raised concerns about recent elections and referenda, but there have been some themes that link them, and also concern social media: foreign interference, message targeting, and database-driven campaigning that subverts existing election supervision law.

While election designs can be complex, the principle and process of counting Xs on papers could not be more intuitive and widely understood. Transparency has extended to the process of information and to the campaigns itself. While it is clearly the case that in free media systems private media exercise significant influence on the outcomes; the bias and selectivity of those media are there for everyone to see, and newspapers in particular have been freely selected by readers in part for the biases they represent in competitive markets regulated for competition, media plurality and diversity.

According to the tests set out earlier in this chapter, electoral legitimacy in the UK is still intact: international organizations and British subjects still view electoral processes as legitimate. But, particularly with regard to the UK Referendum, cracks are beginning to show. This chapter has examined how data-driven campaigning—and Facebook dominance—can undermine legitimacy. The wider issue here may be that while social media still *in theory* offer new opportunities for democracy, the increasingly commercial and increasingly smart, data-driven social media may in the long term be on a collision course with the open, voluntary, equal public deliberation required by democracy. Some of the corrosive effects of social media can be mitigated if citizens are provided with the appropriate information and the tools needed to switch platforms and exert some competitive pressure. Continuing dominance and monopoly positions, particularly by opaque foreign companies, are likely to be particularly corrosive of trust, fairness, and legitimacy.

Many of the issues raised in this chapter are features of social media per se, not any one platform or the fact of dominance. But, and here is the central point, dominance exacerbates the problem. Put in another way, an increased plurality of social platforms would go a long way to addressing many of them.

to Louis Knight-Webb and the Who Targets Me volunteers for providing access to their Facebook data.

REFERENCES

Ace Project. 2017. "Election Observation Portal." Accessed September 7, 2017. http://aceproject.org/electoral-advice/dop.

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31, no. 2: 211–36.

Anstead, Nick. 2017. "Data-Driven Campaigning in the 2015 UK General Election." *The International Journal of Press/Politics* 22, no. 3: 294–313.

Barocas, Solon. 2012. "The Price of Precision: Voter Microtargeting and Its Potential Harms to the Democratic Process." In *Proceedings of the First Edition Workshop on Politics, Elections and Data*, 31–36. PLEAD '12. New York: ACM. doi:10.1145/2389661.2389671.

Butrymowicz, Daniel W. 2009. "Loophole.com: How the FEC's Failure to Fully Regulate the Internet Undermines Campaign Finance Law." *Columbia Law Review* 109: 1708–51.

Cadwalladr, Carole. 2017. "The Great British Brexit Robbery: How Our Democracy Was Hijacked." *Observer*, May 7, 2017. https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy.

Cohen, Julie. E. 2012. "What Privacy Is For." *Harvard Law Review* 126, no. 7: 1904–33.

Council of Europe. 2003. "Recommendation Rec(2003)4 of the Committee of Ministers to Member States on Common Rules against Corruption in the Funding of Political Parties and Electoral Campaigns." Adopted by the Committee of Ministers on April 8, 2003, at the 835th meeting of the Ministers' Deputies. https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805e02b1.

Council of Europe. 2017. "Feasibility Study on the Use of Internet in Elections, MSI-MED (2016)10rev." Strasbourg: Council of Europe. https://rm.coe.int/16806fd666.

Denham, Elizabeth. 2017. "The Information Commissioner Opens a Formal Investigation into the Use of Data Analytics for Political Purposes." Information Commissioner's Office blog, May 17, 2017. https://iconewsblog.org.uk/2017/05/17/information-commissioner-elizabeth-denham-opens-a-formal-investigation-into-the-use-of-data-analytics-for-political-purposes/.

E-Marketer. 2016. "US Political Ad Spending, by Format, 2008–2016." Last Modified April 21, 2016. http://www.emarketer.com/Chart/US-Political-Ad-Spending-by-Format-2008-2016-billions-of-total/189566.

Electoral Commission. 2016. "UK Parliamentary General Election 2015: Campaign Spending Report." February 2016. http://www.electoralcommission.org.uk/__data/assets/pdf_file/0006/197907/UKPGE-Spending-Report-2015.pdf.

Electoral Commission. 2017. "Electoral Commission Statement on Investigation into Leave.EU." Electoralcomission.org, April 21. 2017. https://www.electoralcommission.org.uk/i-am-a/journalist/electoral-commission-media-centre/news-releases-Referendums/electoral-commission-statement-on-investigation-into-leave.eu.

Esser, Frank. 2013. "Mediatization as a Challenge: Media Logic versus Political Logic."
In *Democracy in the Age of Globalization and Mediatization*, edited by Hanspeter
Kriesi et al., 155–76. Basingstoke: Palgrave Macmillan.

Ewing, Keith D., and Jacob Rowbottom. 2011. "The Role of Spending Controls: New
Electoral Actors and New Campaign Techniques." In *The Funding of Political
Parties: Where Now?*, edited by Keith D. Ewing, Jacob Rowbottom, and Joo-
Cheong Tham, 77–91. Abingdon: Routledge.

Falguera, Elin, Samuel Jones, and Magnus Ohman, ed. 2014. *Funding of
Political Parties and Election Campaigns: A Handbook on Political Finance*.
Stockholm: International Institute for Democracy and Electoral Assistance.
http://www.idea.int/sites/default/files/publications/funding-of-political-
parties-and-election-campaigns.pdf.

Garland, Ruth, Damian Tambini, and Nick Couldry. (2017). "Has Government Been
Mediatized? A UK Perspective." *Media, Culture and Society*. In Press. Available at
http://eprints.lse.ac.uk/70662/

Hallin, Daniel C., and Paolo Mancini. 2004. *Comparing Media Systems: Three Models of
Media and Politics*. Cambridge: Cambridge University Press.

Helberger, Natali. 2012. "Exposure Plurality as a Policy Goal." *Journal of Media Law* 4,
no. 1: 65–92.

Hepp, Andreas. 2013. *Cultures of Mediatization*. Cambridge: Polity.

Holz-Bacha, Christina, and Lynda Lee Kaid. 2006. "Political Advertising in
International Comparison." In *The Sage Handbook of Political Advertising*, edited
by Lynda Lee Kaid and Christina Holz-Bacha, 3–13. Thousand Oaks: Sage
Publications.

Howard, Philip N., and Daniel Kreiss. 2009. "Political Parties & Voter
Privacy: Australia, Canada, the United Kingdom, and United States in
Comparative Perspective." World Information Access Project Working
Paper #2009.1. Seattle: University of Washington. https://ssrn.com/
abstract=2595120.

Howard, Philip. 2006. *New Media Campaigns and the Managed Citizen*.
Cambridge: Cambridge University Press.

Kreiss, Daniel. 2012. "Yes We Can (Profile You): A Brief Primer on Campaigns and
Political Data." *Stanford Law Review Online*, 64, no. 70 (February): 70–74.

Kreiss, Daniel, and Philip N. Howard. 2010. "New Challenges to Political
Privacy: Lessons from the First US Presidential Race in the Web 2.0 Era."
International Journal of Communication 4: 1032–50.

Kunelius, Risto, and Esa Reunanen. 2016. "Changing Power of Journalism: The Two
Phases of Mediatization." *Communication Theory* 26, no. 4: 369–88. doi:10.1111/
comt.12098.

Levitsky, Steven, and Lucan A. Way. 2002. "The Rise of Competitive Authoritarianism."
Journal of Democracy 13, no. 2: 51–65.

MacKinnon, Rebecca. 2012. *Consent of the Networked: The Worldwide Struggle for
Internet Freedom*. New York: Basic Books.

Millar, Gavin. 2017. "Undue Influence." *The New European,* May 12, 2017.

Moore, Martin. 2016. "Facebook, the Conservatives and the Risk to Fair and Open
Elections in the UK." *Political Quarterly* 87, no. 3: 424–30.

Morozov, Evgeny. 2011. *The Net Delusion: The Dark Side of Internet Freedom*.
New York: Public Affairs.

Nielsen, Rasmus Kleis. 2012. *Ground Wars: Personalized Communication in Political
Campaigns*. Princeton: Princeton University Press.

Ofcom. 2017a. *News Consumption in the UK: 2016*. London: Ofcom. https://www.ofcom.org.uk/__data/assets/pdf_file/0017/103625/news-consumption-uk-2016.pdf.

Ofcom. 2017b. "Elections and Referendums." In *The Ofcom Broadcasting Code (with the Cross-promotion Code and the On Demand Programme Service Rules)*. April 3, 2017. London: Ofcom. https://www.ofcom.org.uk/__data/assets/pdf_file/0009/100116/broadcast-code-april-2017-section-6.pdf.

ONS. 2016. *Internet Access—Households and Individuals: 2016. What the Internet Is Used for and Types of Purchases Made, by Adults (Aged 16 or Over)*. London: Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2016.

OSCE. 2010. *Election Observation Handbook*. 6th ed. Warsaw: OSCE Office for Democratic Institutions and Human Rights.

OSCE. 2015. "Republic of Tajikistan Parliamentary Elections 1 March 2015: OSCE/ODIHR Election Observation Mission Final Report from May 15, 2015." Warsaw: OSCE. http://www.osce.org/odihr/elections/tajikistan/158081?download=true.

OSCE. 2017a. "Media Analyst." Osce.org. Accessed September 7, 2017. http://www.osce.org/odihr/elections/44233.

OSCE. 2017b. "United Kingdom Early General Election 8 June 2017. Final Report." Accessed September 14, 2017. http://www.osce.org/odihr/elections/uk/336906?download=true?.

Pack, Mark. 2015. "Constituency Expense Limits Are Dying off in the UK, but Neither Politicians nor the Regulator Will Act." Markpack.org (blog), March 20, 2015. http://www.markpack.org.uk/130283/internet-speeds-up-the-killing-off-of-expense-controls-in-marginal-seats/.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. London: Viking.

Parliamentary Assembly of the Council of Europe. 2001. "Recommendation 1516 (2001), Financing of Political Parties." Adopted by the Standing Committee, acting on behalf of the Assembly, on 22 May 2001. http://www.assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=16907&lang=en.

Piccio, Daniela R. 2016. "The state of political finance regulations in Western Europe." International IDEA Discussion Paper 13/2016. Stockholm: International Institute for Democracy and Electoral Assistance. http://www.idea.int/sites/default/files/publications/the-state-of-political-finance-regulations-in-western-europe.pdf.

Rheingold, Howard. 1995. *The Virtual Community*. London: Minerva.

Suchmann, Mark. 1995. "Managing Legitimacy: Strategic and Institutional Approaches." *Academy of Management Review* 20, no. 3: 571–610.

Sunstein, Cass R. 2001. *Republic.com*. Princeton: Princeton University Press.

Tambini, Damian. 1998. "Civic Networking and Universal Rights to Connectivity: Bologna." In *Cyberdemocracy*, edited by Roza Tsagarousianou, Damian Tambini, and Cathy Bryan. London: Routledge.

Tambini, Damian. 2000. "The Civic Networking Movement: The Internet as a New Democratic Public Sphere?" In *Citizenship, Markets and the State*, edited by Colin Crouch, Damian Tambini, and Klaus Eder, 238–60. New York: Oxford University Press. Available online at http://eprints.lse.ac.uk/2895/.

Tambini, Damian. 2016. "In the New Robopolitics, Social Media Has Left Newspapers for Dead." *Guardian*, November 18, 2016. https://www.theguardian.com/commentisfree/2016/nov/18/robopolitics-social-media-traditional-media-dead-brexit-trump.

Tambini, Damian. 2017. "Fake News: Public Policy Responses." Media Policy Brief 20. London: Media Policy Project, London School of Economics. http://eprints.lse.ac.uk/73015/1/LSE%20MPP%20Policy%20Brief%2020%20-%20Fake%20news_final.pdf.

Tambini, Damian, Sharif Labo, Emma Goodman, and Martin Moore. 2017. "The New Political Campaigning." Media Policy Brief 19. London: Media Policy Project, London School of Economics. http://eprints.lse.ac.uk/71945/7/LSE%20MPP%20Policy%20Brief%2019%20-%20The%20new%20political%20campaigning_final.pdf.

Venice Commission. 2010. "Guidelines on Political Party Regulation." Warsaw: OSCE Office for Democratic Institutions and Human Rights. https://www.osce.org/odihr/77812?download=true.

Zittrain, Jonathan L. 2008. *The Future of the Internet and How to Stop It*. New Haven: Yale University Press.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30, no. 1: 75–89.

CHAPTER 13

I Vote For—How Search Informs Our Choice of Candidate

NICHOLAS DIAKOPOULOS, DANIEL TRIELLI,
JENNIFER STARK, AND SEAN MUSSENDEN

Search engines such as Google mediate a substantial amount of human attention, acting as algorithmic gatekeepers and curators as people seek and navigate news and information online. A 2014 survey found that 51% of American respondents got news in the last week via a search engine (Media Insight Project 2014), most likely with either Google or Bing, which together represent a practical duopoly, dominating about 86% of the US market for desktop search ("Rankings—comScore, Inc" 2017). Yet, as proprietary systems, little is known about search engine information biases and how they may be shaping the salience and quality of users' information exposure. This is of particular importance in considering how voters become informed during elections. The primary focus of this chapter is to begin to illuminate the multifaceted roles that Google's search technologies have played in algorithmic information curation in the 2016 US elections.

The concern over the power of search engines in mediating access to information is not a new one. As early as 2000 Introna and Nissenbaum expounded on issues of accountability and transparency (Introna and Nissenbaum 2000), and these concerns have been echoed in more recent work (Laidlaw 2009; Granka 2010). Research on the potential for search engines to impact people has enumerated a number of possible repercussions, including the potential to affect attitudes (Knobloch-Westerwick, Johnson, and Westerwick 2015), to alter perceptions based on image

presentations (Kay, Matuszek, and Munson 2015), and to lead to anticompetitive tendencies by privileging preferred products or pages (Hazan 2013; Edelman 2010; Edelman and Lockwood 2011; Alam and Downey 2014).

The fundamental reason why search rankings are so powerful is that the order of information has a substantial impact on human attention and reliance: not only do people click on top results more than lower results (Agichtein et al. 2006) but also they believe something is more relevant if it is in a higher position (Pan et al. 2007). In the media environment, search engines have the power to "grant visibility and certify meanings" as Gillespie (2017) has written. Because search engines act as gatekeepers to information, the way they are designed and wielded can yield subtle or overt power. For instance, the prevailing design value baked into (and optimized for in) Google search is that of "relevance," but this may come at the cost of considering other possible design values such as fairness or information diversity (Van Couvering 2007). In China the state exercises its will to censor and block politically inexpedient information via the predominating nationalized search engine, Baidu (Jiang 2013, 2014).

In politics the most startling evidence for the power of search engines to impact election processes comes from Epstein and Robertson (2015). In their laboratory study, they showed that by manipulating the ordering of supporting information for candidates in a mock search engine, they could shift the voting preferences of undecided voters. Other work has explored differences between search engines in terms of the types of sources surfaced for election-related queries. Muddiman (2013) found that Google ranked campaign-controlled and advocacy pages higher than Yahoo, for instance. Research has shown that users search for candidate information during key campaign events such as debates and gaffes (Trevisan et al. 2016), however query strategies used to seek information about candidates are an underresearched topic.

Legal scholars have explored the complexities of applying regulation to search engines (Grimmelmann 2014; Bracha and Pasquale 2008) in order to quell their potential for abuse. One of the key barriers to regulating search engines in the US jurisdiction is that their output is currently interpreted as free speech by courts (Laidlaw 2009). While there may be other regulatory options for ensuring search engine accountability, such as setting up nonpublic limited-access courts (Bracha and Pasquale 2008), these only partially address accountability of such systems due to their nonpublic nature. Given the challenges to the regulation of search engines, an alternative to achieving algorithmic accountability (Diakopoulos 2015) of search engines, and the approach we have taken here, is to audit their results by gathering data around specific candidate-related queries.

Yet there are a number of confounds to studying search engine results using an auditing approach. Factors such as the query terms used, language setting, geolocation, search history, and logged-in status, not to mention randomization or A/B testing that the search engine may be undertaking make it difficult to see the outputs of the system as stable (Ørmen 2015; Xing et al. 2014; Kliman-Silver et al. 2015). Moreover, the unavoidable constitution of the search engine as a sociotechnical system (Gillespie 2017) that embeds feedback loops with user inputs make it difficult to disambiguate the role of the algorithm versus the role of the human user. And this is further compounded by the knowledge that different demographics and users may exhibit different types of information-seeking behavior (Weber and Jaimes 2011). The choice of query terms can itself lead to politically biased results, as can the "universe" or input bias of the pages indexed by the search engine. Teasing apart differences in results arising from query term selection, input bias, and actual algorithmic ranking bias is a methodologically challenging proposition (Magno et al. 2016; Kulshrestha et al. 2017).

Taking these methodological challenges into account to the extent possible, this chapter details four distinct case studies/audits. Each case illustrates how Google mediates candidate information in different ways: (1) *Search Results*: the set of sites ranked on the first page of search results, including their support or opposition to the candidate; (2) *Issue Guide*: the presentation of an issue guide integrated into search results on candidates constituted from algorithmically curated quotations from each candidate gleaned from news coverage; (3) *In the News*: the presentation of news information about each candidate as framed in the "In the News" section, which occupies privileged screen real-estate at the top of search results; and (4) *Visual Framing*: candidates are visually framed in search results as a consequence of the image selections that Google makes.

The following subsections detail each of these case studies, examining why and how each facet of the search engine is important to consider with respect to how voters gather information about candidates. We then conclude the chapter with a comparison and contrast of the cases, including methodological challenges and an elaboration on where additional work is needed to better understand how search engines inform voters.

CASE 1: SEARCH RESULTS

Previous research has indicated that "biased search rankings can shift the voting preferences of undecided voters by 20% or more" (Epstein and Robertson 2015). The impact of that bias may be felt more if users assume

that the results are neutral and, therefore, trustworthy. To begin to articulate such biases in a real search engine a crowdsourced analysis was used to explore differences in how primary candidates in the 2016 US presidential election were presented by Google.

Methods

Search result links were collected on December 1, 2015, from nonpersonalized Google searches for each candidate's complete name (i.e., "first-name last-name"). The top 10 ranked results for each of the 16 candidates (13 Republicans and 3 Democrats) were collected. The focus on the first page was grounded in knowledge of search users' behavior: users end up clicking on the first 10 results 70% of the time (Loiz 2014).

The process to determine whether the linked webpages returned by the search were positive or negative for each candidate was crowdsourced using the Amazon Mechanical Turk (AMT) microtask market. Each link was presented to three separate AMT workers, who determined whether the linked webpages returned by the search were favorable or oppositional toward each candidate.

Several iterations of the instructions for workers were tested with pilot data. As a basis for that process, an initial sample of the websites was rated by the researchers and measured against the ratings of AMT respondents. Language and instructions for the task were adjusted until there was a good level of agreement between the researcher rating and the respondents' ratings. The intraclass correlation coefficient between ratings done by AMT and the researcher's baseline rating was 0.912.

The final version of the AMT task included simple instructions on how to read the websites. The respondents were instructed to only evaluate the linked webpage, to consider only the primary content of the webpage, and to decide whether the webpage favored or opposed a particular candidate. Workers were directed not to consider source reputation, but instead to focus on the content of the specific article.

We incorporated strategies into our task design to help ensure quality responses from crowdsource workers (Kittur, Chi, and Suh 2008) including asking respondents to answer a series of four questions: (1) "Please summarize the content of the website in one sentence"; (2) "To what extent does the linked webpage favor (candidate name)?"; (3) "To what extent does the linked webpage oppose (candidate name)?"; and (4) "Please explain the reasoning behind your ratings in a few sentences. Be sure to include relevant excerpts from the page that support your ratings."

Questions 1 and 4, which were answered via free text, were used to control the quality of the information. Question 1 was designed so that respondents had to explicitly demonstrate that they, in fact, had read and understood the page. Question 4 was designed to ensure that respondents had fully considered the page; it also helped explain the respondent's reasoning. Any response that did not include answers to Questions 1 and 4 were rejected, and another worker was then asked.

In questions 2 and 3, each responder had to rate each website on scales of "favorability" and "opposition" to a particular candidate, on a scale of 1 ("not at all") to 5 ("a lot"). This rating was not done using a single scale because some websites were found to be more complex than merely favoring or supporting; many of them included different angles and opinions on the same candidate. This was the case, for instance, for news articles that had quotes from both Democrats and Republicans about the same candidate, or that recounted both controversies and accomplishments by the same person.

At the end of the crowdsourcing process, each page had three scores for favorability and three scores for opposition, which were averaged. The averages of the favorability and opposition scores were subtracted from each other to calculate an average net positivity score. Those average net positivity scores were then averaged by candidate and party.

Analyzing Net Positivity

The average net positivity score for all websites returned by the top 10 search results was 1.2 on a scale of -4 (very opposed) to 4 (very favorable). The Democrats had higher average scores (1.9) than the Republicans (1.0). The candidate with the highest average was Democrat Bernie Sanders, with 3.2. The lowest was Republican George Pataki, with 0.3.

Analysis included not only the average net positivity score but also the total number of positive and negative web pages for each search result. Any result with an average net positivity score between -0.5 and 0.5 was considered neutral. Below and above those thresholds, they were considered, respectively, negative and positive. Democrats had, on average, 7 positive search results in those top 10, whereas GOP candidates had an average of 5.9. Sanders was also the only candidate with no negative pages surfaced in his first 10 web results; the average candidate had 2.1. Sanders had 9 positive websites compared to the average of 6.1, and one neutral, whereas the average number of neutral results per candidate was 1.8. Republican Ted Cruz had the most negative results, with five, compared with five positive results.

The results for candidates tended to decline in positivity as you move down the rankings on the first page. On average, the net positivity score of the first result was 2.2. By the 10th ranked result, the average score dropped to 0.9 (Figure 13.1).

Looking closer at the individual results, the negative websites tended to appear in the bottom of the first page of Google results. Campaign websites—which included candidate webpages and social media profiles—were clustered at the top. The negative websites tended to be news articles that were mostly critical of the candidates.

Google presented a higher proportion of negative news articles about Republican candidates than their Democratic counterparts. Democrats tended to have more official sites, social platforms, and, to a lesser degree, more positive coverage on the first page. Six of the top 10 results for Bernie Sanders were run by the campaign. They were a Facebook profile, a Twitter profile, a YouTube account, and three campaign website pages. The other links were to a Wikipedia page and three news stories, all favorable to Sanders. Ted Cruz, on the other hand, had only two websites and a Facebook profile. The other seven links were news articles, and for Cruz, those were mostly negative.

The way Google ranks and selects websites, therefore, opens the way for digitally savvy campaigns to take hold of the top 10 results. Official sources of information appear to be privileged and come to dominate the most coveted and attention-getting positions on the results page.
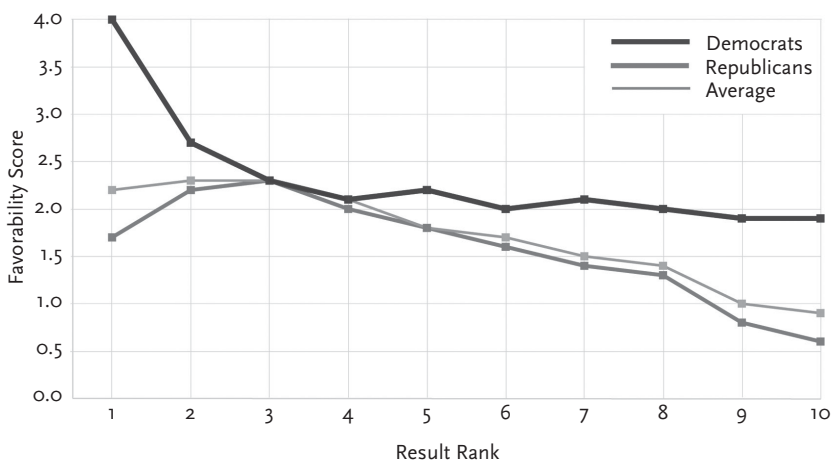
Figure 13.1: The decay in the average positivity score on the first page of web results for presidential candidates in the US presidential election of 2016.
Source: Original chart produced by Daniel Trielli.

On February 1, 2016, Google deployed a new feature for users searching for US political information regarding the presidential election candidates (Schonberg 2016). Whenever a user searched the name of a candidate in the presidential primaries taking place at the time, the results page displayed a list of 16 political issues (later expanded to 17) each containing several position statements by the candidate. These statements were sourced from news articles that were linked alongside the quotes and, according to our contact with Google public relations the guide, would "only show issue topics from candidates when they have submitted a statement or if there is a sufficient amount of news articles that include relevant quotes from them that our algorithm has found" (see Figure 13.2).

Google's ability to edit together and inject information that it deems relevant to candidates is another instance of how their dominance may affect the public's ability to get fair and balanced information. This case study was motivated by a need to understand how the design of the new feature might bias how people perceive the candidates.

The analysis presented relies on data collected by observing the information box constituting the issue guide. The data collected includes quotes by the candidate, the name of the candidate, the broad issue or topic the quote is referring to, the rank position of the quote within its topic, the link to the article from which the quote was extracted for the infobox, the website that is the source for that article, and the date the article was published.

Data was collected three times, at three-week intervals, on April 1, April 22, and on May 13, 2016. In the first two collections, the tool presented quotes from five candidates: Democrats Hillary Clinton and Bernie Sanders, and Republicans Ted Cruz, John Kasich, and Donald Trump. By the time the third collection was performed, both Kasich and Cruz were removed from the list, since they had withdrawn from the campaign.

Quantifying Statements

The initial analysis involved counting the number of quotes by each candidate presented by the information box and then looking at the change across each data collection. The first discovery was that the number of quotes increased over time. No quotes were removed between each measurement. New quotes were added as the candidates made new statements
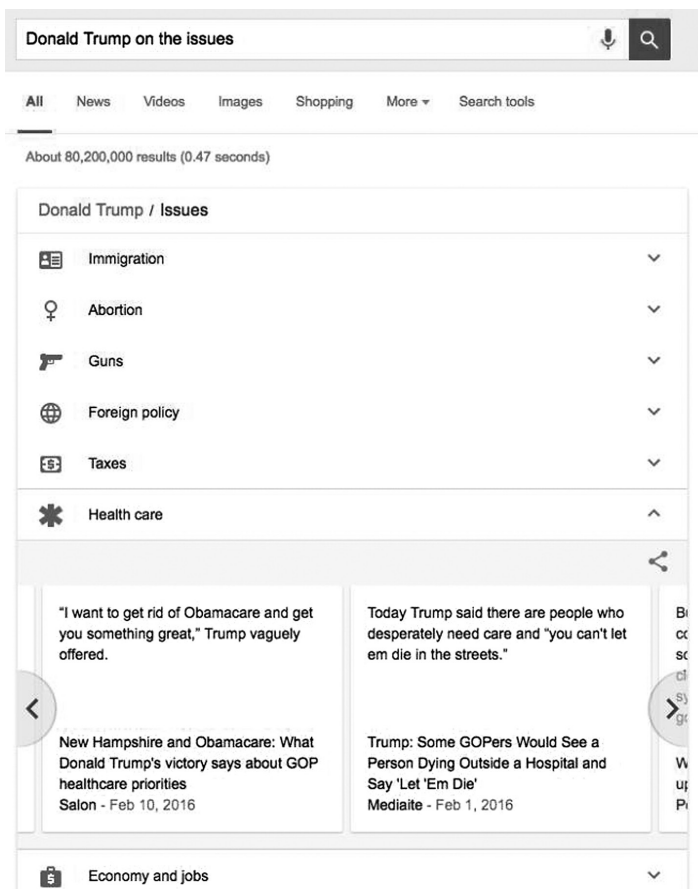
Figure 13.2: The Google campaign issue guide showing positions for Donald Trump, collected June 2, 2016.
Source: Screenshot of webpage taken and cropped by authors.

that were covered by new articles. This increase over time happened with all candidates.

However, over time a disparity between the totals from candidate to candidate was detected. On average, candidates had 191 statements listed in the tool on April 1. But after separating by ideology, the Democratic candidates had an average of 258, while the Republicans had an average of 147. This continued through the second collection (average of 287 statements for Democrats versus 166 for Republicans) and the third, after two Republicans dropped out (300 and 160). In this final collection, Hillary Clinton had more statements than Donald Trump in 13 of the 17 topics; Donald Trump had more statements in three topics; and there was an equal number of statements in one topic.

The challenge, then, was to try to determine whether that disparity could be the result of an imbalanced input of news articles that preceded the aggregation by the information box. If the volume of coverage of the candidates was divergent, there would be fewer articles about some candidates, which would lead to fewer statements extracted and presented. Fewer news articles about Donald Trump would, therefore, indicate bias in the coverage indexed, not necessarily the Google information box curation algorithm. However, searches for each of the three remaining candidates conducted on Google News showed the opposite. In May of 2016, Google News had 941 indexed articles about Trump, 710 about Clinton, and 630 about Sanders. Data from LexisNexis corroborated that Trump was the focus of a larger number of news stories during the election ("U.S. Presidential Campaign Tracker" 2016).

News Sources

We also considered the distribution of sources for the articles from which statements were extracted, in order to determine whether a source or group of sources was dominant in the tool. In the May 13 collection, 326 different news organizations served as the sources for the statements of the candidates. Some sources had more articles represented than others. Out of a total of 735 statements from the three remaining candidates, 63% came from sources that, individually, had a small participation—1 to 5 articles listed in the information box. Another 24% came from sources that were in the middle tier—6 to 20 statements. And 14% of the statements came from four news sources, that is, 1.2% of the total of 326 sources: *The Hill, Politico, The Washington Post*, and *The Huffington Post*.

Additionally, Google itself was the source for 17 statements each by Bernie Sanders and Hillary Clinton. Unlike Donald Trump, Sanders and Clinton decided to use a specific Google-based solicitation for self-presentation on issues, and those were also included in the information box. This difference in how the candidates provided (or did not provide) information accounts for only about 12% of the disparity in totals we refer to in the previous section.

The next step was to determine whether there was an ideological imbalance in the sources from which the statements came. A database of ideological leanings of websites from prior research (Bakshy, Messing, and Adamic 2015) was used to determine whether there was a political bias in the news sources used in the information box. The database scores websites from -1 to 1, the negative values meaning they are preferred (by Facebook shares

and likes) by users who are left-wing and positive values meaning they are preferred by users who are right-wing.

Applying the bias scores from the study to the 83 overlapping sources used by the Google issues guide, provided an average bias score of -0.16, meaning that overall there is a slight liberal bias in news sources. However, those 83 sources cover only 50% of all the articles in the information box. When considering only the top 10 sources for the Issues tool, all of which were covered by the political bias database, the bias score was more liberal at −0.31.

Editorial Choices

Editorial choices for the Google information box were also analyzed. The first element in that inquiry was the order in which the topics were listed. Previous research has indicated that items listed higher on the page will get more attention from searchers (Agichtein et al. 2006). In the first data collection, the order of issues was: Immigration, Abortion, Guns, Foreign Policy, Taxes, Health Care, Economy and Jobs, Civil Liberties, Crime and Safety, Environment, Education, Budget and Spending, National Security, Medicare and Social Security, Veterans, and Energy. In the following data collections, it was mostly unchanged, except for the addition of Gay Marriage, in the sixth position between Taxes and Health Care. It is also noteworthy that when the results page was initially presented to the user, the list was trimmed and expandable; only the top four topics were displayed, lending even more significance to the prioritization of the first topics.

The order of topics is not alphabetical, indicating some other prioritization of the issues. A similarity with polls conducted by Gallup (Newport 2016) and the Pew Research Center (Doherty, Kiley, and Jameson 2015) and with Google Trends was investigated to see whether the order was determined by public interest or search trends, but no correlation was found. Additionally, the addition of Gay Marriage indicates a human editing process for two reasons: first, the statements shown in this category were previously listed under "Civil Liberties"; and second, Google Trends showed no spike in searches related to same-sex marriages that could explain its surge in the electorate's interest.

In conclusion, a variety of strategies used in this inquiry surfaced the presence of biases in the Google information box, with respect to representation of candidates in their proportion of statements, dominance of sources, political bias of sources, and editorial choices. Still, the absence of algorithmic transparency (Diakopoulos and Koliska 2016) regarding how

the tool was designed and how it curated information limited our ability to establish why we observed these patterns.

CASE 3: IN THE NEWS

At the top of the result page for searches about newsworthy topics during the 2016 US elections Google featured a box with related news articles labeled "In the News" (later relabeled as "Top Stories" in November 2016 after the election). Because of its prominence, the Google "In the News" box had the potential to lead users to particular news sources, and to direct information seekers to particular narratives and frames around candidates, thus shading their impressions. With so much potential for guiding attention, questions of ideological diversity and news-value of the articles "In the News" become critical and motivate the current case study (Figure 13.3).

For this analysis article links were collected from Google using a non-personalized browser every hour between May 31 and July 8, 2016, resulting in a total of 5,604 links. Along with the URL for the articles, metadata was collected, including the name of the source for the article (usually a news organization), the text of the link (usually the title of the article), the string of text that shows how long the story has been posted or updated ("X hours ago"), and the exact time the link was scraped. The 5,604 links were nonunique, meaning that they could point to the same article in multiple instances. In total, there were 972 unique articles collected.
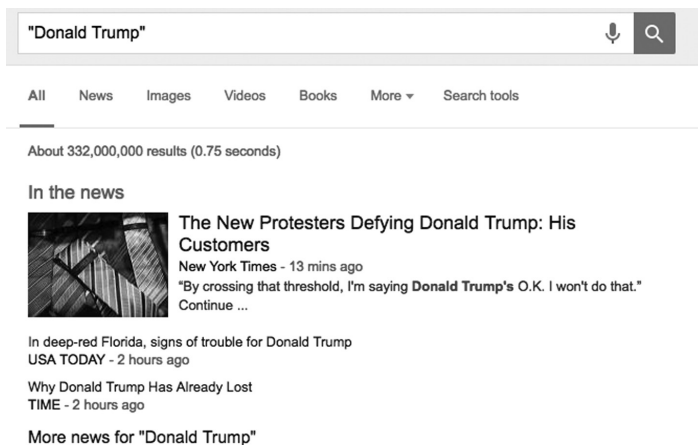
Figure 13.3: Google "In the news" box for Donald Trump, collected October 17, 2016.
Source: Screenshot of webpage taken and cropped by authors.

Privileged Sources

To determine whether particular sources dominated the Google "In the News" box, the total number of links collected (including repetition) from each website (i.e., top level domain) was measured against the total number of links displayed. Out of the 113 news sources, 60 had articles linked nine or fewer times during the time frame of collection. Another 35 sources had between 10 and 50 links displayed in the Google "In the News" box. Sixteen sources had between 50 and 500 links showcased. The remaining two sources combined, CNN and New York Times, account for 2,476 links (44.2%) listed. Of the 5,604 links listed in the Google "In the News" box, 1,276 came from CNN and 1,200 originated from the New York Times domain. For example, between June 2 and June 6, an article titled "Hillary Clinton's Evisceration of Donald Trump" appeared in 90 different hourly measurements, indicating a strong staying power and emphasis in Google's results.

However, the sources that are most frequent in the Google "In the News" boxes do not rely only on single articles that are repeated multiple times. In fact, when only unique links are considered, CNN and the *New York Times* also have high prevalence in the ranking. Out of the 972 unique articles displayed, 152 belonged to the *New York Times* and 142 to CNN, meaning that, combined, they amounted to 30% of the unique articles. In third place, NBCNews.com had 69 articles, or 7% of the total unique links.

The Google "In the News" box also prioritizes the three links that it presents. There are always three links that are displayed in the tool, but the first in that list has more prominence due to additional information displayed. All of the first listed links include a short summary, or snippet of the article, and 99.4% of them also include an image. The distribution of sources in those first links listed was also measured, to investigate whether or not some sources also had more dominance in the premium spot. Again, CNN and the *New York Times* dominate the first listed link, with 1,211 out of 1,868, or 64.4% of them being from these two sources. This is a startling concentration of attention oriented toward only two publishers of information.

When evaluated on a weekly basis, some variation in the ranking of the sources for articles was detected. The most clear example is the *Washington Post*, a news organization that was in the top five sources of links for the week of June 12 to June 18, when it rose to third place.

One hypothesis is that the *Washington Post* had a temporarily higher profile than usual, since it became the focus of a controversy concerning the presidential election coverage. The increase of the newspaper in the Google "In the News" box rankings coincided with the then-candidate

Donald Trump criticizing the *Post* on Twitter and revoking their credentials for campaign events. However, more research is necessary to conclude that becoming the focus of news itself could itself propel a news organization to higher prominence in the box.

Freshness

Using the metadata relating to time, we found that "freshness" of articles also appears to be a relevant factor for prevalence in the "In the News" box. When all articles are considered, 30.5% of all links are marked as being less than three hours old. A smaller proportion, 4.8%, of links listed were more than one day old, and occurred mostly on weekends. Additional research is needed to determine the specifics of how Google's algorithms considers article timestamps and update times. For example, the article "Hillary Clinton's Evisceration of Donald Trump," may have had staying power because it was continually updated and considered "new" by Google because of an updated timestamp.

Non-News Organizations

Not all linked articles came from news organizations. Of the 5,604 total links, 214 came from Twitter (the official accounts of Hillary Clinton, Donald Trump, and Chelsea Clinton were the originators of those links), 58 from YouTube, and 24 from the Federal Bureau of Investigation. Also included in the list of sources are websites that are controversial sources of news, such as Breitbart News, Infowars, and Russia Today.

It is important to note that the Google "In the News" box shifted from a strict news organization background to a more open aggregator of news-related links in 2014. Until that year, the box was called "news for . . ."; then, it changed to "in the news." That was the year in which users started noticing non-news links appearing in the box, such as press releases from companies and content from Reddit and YouTube.

The results presented earlier can lead to different threads of inquiry. The analysis sheds light on the disparity in attention that some sources of news achieve via the box, but it still remains unclear *why* such disparities came about. There is a clear predilection toward fresher and newer content, privileging the daily news cycle or at least the appearance of it via the updated publication date of articles. The inclusion of non-news links is a complicating factor, raising questions about Google's very definition of a news source.

CASE 4: VISUAL FRAMING

The first three case studies presented focus on textual results from search engines such as links to articles and websites. However, Google also presents images on the main results page. Google's image box is located in the top right section of the main results page, and typically contains five to seven images. The box draws attention to specific images of the candidates, and to the articles or information linked via these images. In this way, Google may contribute to shaping users' attitudes and perceptions by visually framing candidates in particular ways.

Previous work suggests that the visual portrayal of candidates in election campaigns may have an impact on electoral outcomes. For instance, the competence of a candidate is inferred from attractiveness (Todorov et al. 2005), and beauty itself is positively associated with votes (Berggren, Jordahl, and Poutvaara 2010). Furthermore, people are unlikely to change their candidate pick even after they are provided additional information regarding candidates' competencies (Todorov et al. 2005).

In this case we examine how Hillary Clinton and Donald Trump were framed in the images surfaced on search results pages by Google including analyses of emotional and gestural content, the sources of images and their political leanings, and image rank positions. Images, their sources, and the ranking position (i.e., first, second, etc.) were collected for the queries "Hillary Clinton" and "Donald Trump" once per day from September 3 until October 28, 2016. As a baseline, we also collected images from Google Image Search for each candidate for the same time period, resulting in 353 images of Clinton and 298 of Trump after removing images that did not contain the candidate, or that contained multiple faces.

In our sampling period there were nine unique images from the image box of Clinton, and 11 of Trump. Unique images were identified using a difference hash algorithm (Buchner 2017) which converts the images to grayscale, reduces them to 72 pixels, and then computes the difference between neighboring pixels producing an alphanumeric hash unique to that image. To analyze the emotional content of the images, we used an API (Microsoft Azure 2017), which provided a confidence score for each of eight emotions, comprising anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise.

Clinton was deemed happy in seven images, neutral in one, and surprised in another, whereas Trump was happy in three images, neutral in seven, and angry in one. This distribution of emotions was somewhat different from the distribution observed in Google Image search baseline images, where Clinton exhibited a neutral expression in the majority of

images. In terms of gestural content, Clinton was gesturing in only one of nine image box images (11%) versus 40% of baseline images, while Trump was gesturing in four of 11 image box images (36%) versus 57% of baseline images (Figure 13.4).

Altogether, these images frame the candidates in a manner typical of representation of women and men in news media (Rodgers, Kenix, and Thorson 2007; Kwak and An 2016), with Clinton portrayed as happy and more passive (fewer gestures), and Trump depicted as serious, and more energetic (more gestures). Energetic gesturing is associated with dominance and power (Burgoon and Dunbar 2006), and effective leadership is typically associated with masculine traits including dominance (Koenig et al. 2011). Therefore, these images may reinforce the gender stereotype that men are strong leaders, while women, who display characteristics incongruent with typical leadership traits, are ineffective.

Because images are hyperlinked, Google's choice of images also serves to direct attention to specific sources of those images. Research shows that people show more interest in unconventional news photos when presented outside the context of the article (Mendelson 2001). Such atypical images in Google's image box would be particularly powerful in directing attention to the related articles as they are also presented outside the context of the source article. This could have an impact on voters, as outlets may portray candidates differently, and may impart partisan bias within the images that are published. Sources for Clinton's baseline images were mostly left-leaning, but left-leaning sources of image box images were even
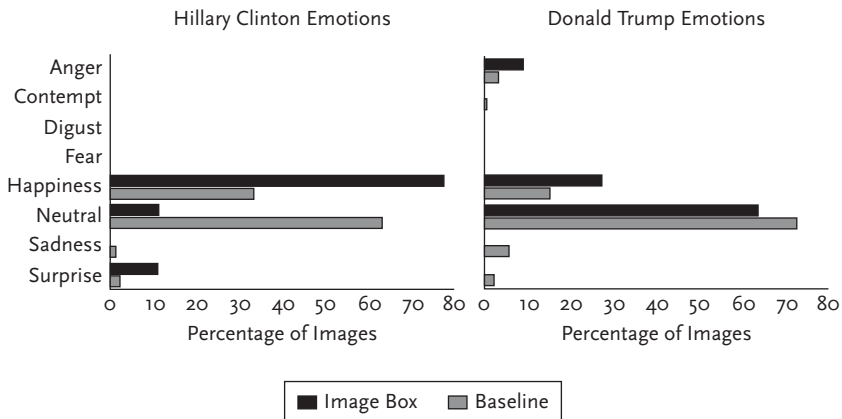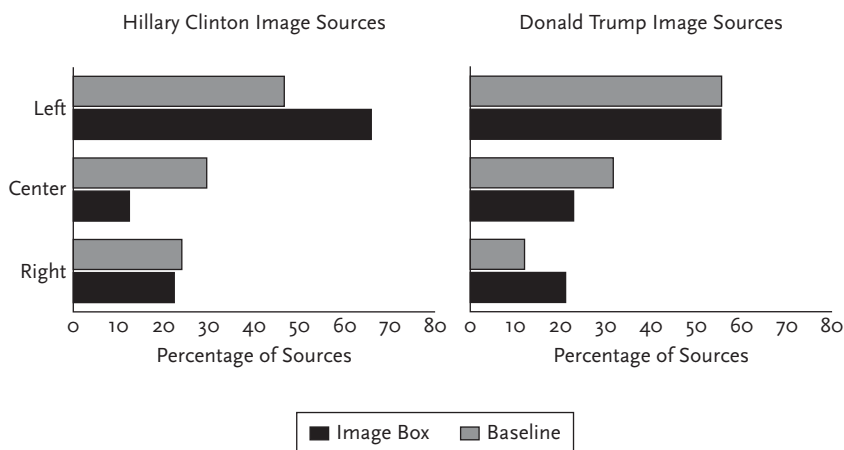

Figure 13.4: Percentage of Clinton and Trump images from the image box or from Google Images (Baseline) that show listed emotions.
Source: Original chart produced by Jennifer Stark.

more highly represented (Stark and Diakopoulos 2017). While sources of Trump's baseline and image box images were also mostly left-leaning, centrist sources were reduced and right-leaning sources were augmented in the image box compared with the baseline (Figure 13.5).

In terms of preferential position of the images within the image box, Google+ was always in the first ranking position, which affords the largest image, meaning that while Google is privileging their own social media platform, knowing this, the campaigns themselves can control the visual in the most dominant position. Clinton's Wikipedia image was in second position, suggesting it is second in importance. Trump's Wikipedia page was second also, but after several edits to the Wikipedia page's main picture, this image was dropped altogether. The remaining rank positions of images for both candidates were less stable. Notably for Clinton, many images changed rank between the 3rd and 24th of September. During this time Clinton was reported as suffering from pneumonia, apologized for calling Trump supporters "deplorables," and suggested Russia was intervening in the election process (Figure 13.6). The news cycle does appear to impart some variability in the visual portrayal of the candidates after the top image or two, but does not seem to be the only driver, given that the primary rank positions were independent from news. Altogether, Google's images are selected from sources across the ideological spectrum, and prioritize images from its own social media site Google+.

There are several limitations with this case study. Our baseline consists of images collected from Google Images search results. Although we

Figure 13.5: Percentage of sources for Clinton and Trump images from the image box or Google Images baseline identified from each listed ideology.
Source: Original chart produced by Jennifer Stark.
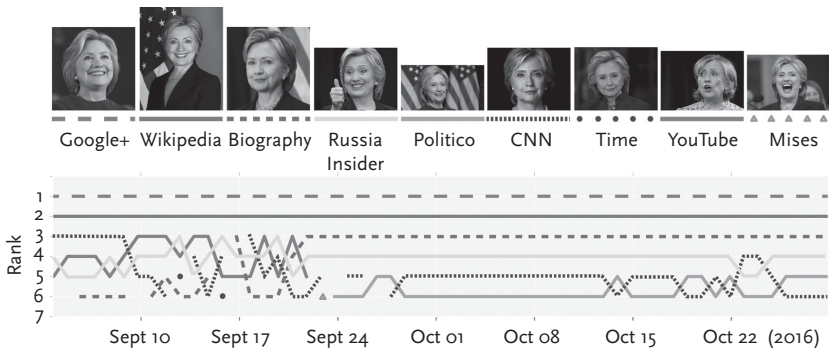
Figure 13.6: Image box rank positions of Clinton across time.
Source: Original illustration produced by Jennifer Stark.

expected this baseline to represent a universe of Google images from which Google selects images for the image box, not all images found in the image box were also in the baseline. Alternative baselines may be more representative of all images that exist for each candidate, for example, using a longer search time period, or images collected from additional search engines like Bing and Yahoo! Moreover, the bias of photographers and editors in selecting images that feed indices like Google Images is not addressed. We also relied on computer vision algorithms for emotional content information, which in turn are trained on data with its own unspecified sampling bias.

Open questions include how the images are selected for the image box, and the role that algorithms or human editors may play in those selections as well as whether source bias would be the same for a different time period, or for politicians in different countries. Our results suggest that Google's image results reflect visual gender differences with respect to leadership roles present in society: that women are happy and passive and men are serious and active. An open question is whether such gendered visual frames, from search engines in particular, can impact perception of presidential competency.

DISCUSSION

Through a series of case studies this chapter has begun to characterize the range of influences that Google's search technologies were having on information curation during the 2016 US elections. Results clearly show the myriad editorial choices that Google's algorithms make in shaping the information environment, including a focus on official sources in the main rankings, ordering in the issue guide that may have unduly privileged

certain issues, the dominance of a small set of news sources highlighted "In the News," and differences in the visual framing of a female versus a male candidate. These results constitute a set of observations that raise important questions for future work, and suggest political conversations that need to be undertaken about the editorial role of search engines in political life: What *should be* the responsibilities of Google as an important intermediary of information for voters?

Studying search engine results is a methodologically challenging undertaking. Though we were able to observe several instances of how results may shift attention toward candidates we lacked any form of scientific control, or ability to run experiments, and found it difficult to explain *why* we were seeing many of the results patterns that we saw. Still, we believe that it is a valuable public interest endeavor to report on the observed patterns of information that dominant entities like Google mediate, given that they can have a substantial effect on exposure to political information and bias.

For the Issue Guide case we found it valuable to consider the *dynamics* of the results from Google and we believe this should be an important aspect that informs the methodologies of algorithm audits of search engines in future work. While taking individual snapshots of data allows for concrete analyses, the dynamics of how the results change over time (and how the search giant may be responding to public pressure) is important to track for accountability purposes. Writing "stories" about results may be augmented by building dashboards to track and reflect search results over time.

In two of the cases, for the Issue Guide and for Visual Framing, the concept of defining *baselines* emerged. Defining an appropriate baseline dictates what the expectation for an algorithm should be, thus informing what is perceived as interesting and newsworthy for an audit. But what is the "right" baseline? In our work we made logical arguments about the expected input data sets that Google algorithms would use to curate quotes or images, however other reasonable baselines could be considered. Some may consider results less compelling if the baseline does not come from a sample independent of the platform under study. We believe that the public impact of these types of audits hinges on making a strong and well-reasoned claim for the expected information that a search engine should provide, and then showing that the search engine meets that expectation or violates that expectation. Additional algorithmic transparency information could inform the baseline definition process, and more generally may be considered by regulators exploring how targeted transparency policies could balance the dominance of information intermediaries.

Several of the case studies presented focused on characterizing the sources that Google surfaces in results. We think diversity is an important

frame for considering the information that voters have about candidates. For instance, although Google predominantly operates on a logic of *relevance*, the idea of information diversity could help to mitigate issues of political polarization that challenge society. In these cases we have considered the ideological diversity present in sources, however other definitions of diversity are also possible, such as by looking at ownership diversity (i.e., who owns a particular source), and even topical diversity (i.e., the topical frames that are apparent in information surfaced by search engines). Future work might be usefully oriented toward building community resources that can reliably tag sources according to these dimensions of diversity so that audits can easily incorporate various diversity measures.

ACKNOWLEDGMENTS

REFERENCES

Agichtein, Eugene, Eric Brill, Susan Dumais, and Ragno Ragno. 2006. "Learning User Interaction Models for Predicting Web Search Result Preferences." In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3. doi:10.1145/1148170.1148175.

Alam, Mohammed A., and Doug Downey. 2014. "Analyzing the Content Emphasis of Web Search Engines." In *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1083–86. doi:10.1145/2600428.2609515.

Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348, no. 6239: 1130–32. doi:10.1126/science.aaa1160.

Berggren, Niclas, Henrik Jordahl, and Panu Poutvaara. 2010. "The Looks of a Winner: Beauty and Electoral Success." *Journal of Public Economics* 94, no. 1–2: 8–15. doi:10.1016/j.jpubeco.2009.11.002.

Bracha, Oren, and Frank Pasquale. 2008. "Federal Search Commission? Access, Fairness, and Accountability in the Law of Search." *Cornell Law Review* 93, no. 1149: 1149–210. doi:10.2139/ssrn.1002453.

Buchner, Johannes. 2017. "ImageHash 3.4." *Python Software Foundation*. https://pypi.python.org/pypi/ImageHash.

Burgoon, Judee K., and Norah E. Dunbar. 2006. "Nonverbal Expressions of Dominance and Power in Human Relationships." In *The SAGE Handbook of Nonverbal Communication*, edited by Valerie Manusov and Miles L. Patterson, 279–78. Thousand Oaks, CA. doi:10.4135/9781412976152.n15.

Diakopoulos, Nicholas. 2015. "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures." *Digital Journalism* 3, no. 3: 398–415. doi:10.1080/21670811.2014.976411.

Diakopoulos, Nicholas, and Michael Koliska. 2016. "Algorithmic Transparency in the News Media." *Digital Journalism*, forthcoming. doi:https://doi.org/10.1080/21670811.2016.1208053.

Doherty, Carroll, Jocelyn Kiley, and Bridget Jameson. 2015. "Contrasting Partisan Perspectives on Campaign 2016." http://www.people-press.org/2015/10/02/contrasting-partisan-perspectives-on-campaign-2016/#.

Edelman, Benjamin. 2010. "Hard-Coding Bias in Google 'Algorithmic' Search Results." http://www.benedelman.org/hardcoding/.

Edelman, Benjamin, and Benjamin Lockwood. 2011. "Measuring Bias in 'Organic' Web Search." http://www.benedelman.org/searchbias/.

Epstein, Robert, and Ronald E. Robertson. 2015. "The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections." *Proceedings of the National Academy of Sciences of the United States of America* 112, no. 33: E4512–E4521. doi:10.1073/pnas.1419828112.

Gillespie, Tarleton. 2017. "Algorithmically Recognizable: Santorum's Google Problem, and Google's Santorum Problem." *Information, Communication and Society* 20, no. 1: 63–80. doi:10.1080/1369118X.2016.1199721.

Granka, Laura A. 2010. "The Politics of Search: A Decade Retrospective." *Information Society* 26, no. 5: 364–74. doi:10.1080/01972243.2010.511560.

Grimmelmann, James. 2014. "Speech Engines." *Minnesota Law Review* 98: 868–952.

Hazan, Joshua G. 2013. "Stop Being Evil: A Proposal for Unbiased Google Search." *Michigan Law Review* 111, no. 5: 789–820.

Introna, Lucas D., and Helen Nissenbaum. 2000. "Shaping the Web: Why the Politics of Search Engines Matters." *Information Society* 16, no. 3: 169–85. doi:10.1080/01972240050133634.

Jiang, Min. 2013. "The Business and Politics of Search Engines: A Comparative Study of Baidu and Google's Search Results of Internet Events in China." *New Media and Society* 16, no. 2: 212–33. doi:10.1177/1461444813481196.

Jiang, Min. 2014. "Search Concentration, Bias, and Parochialism: A Comparative Study of Google, Baidu, and Jike's Search Results From China." *Journal of Communication* 64, no. 6: 1088–110. doi:10.1111/jcom.12126.

Kay, Matthew, Cynthia Matuszek, and Sean A. Munson. 2015. "Unequal Representation and Gender Stereotypes in Image Search Results for Occupations." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI '15*, 3819–28. doi:10.1145/2702123.2702520.

Kittur, A., E. H. Chi, and B. Suh. 2008. "Crowdsourcing User Studies with Mechanical Turk." In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*, 453–56. http://dl.acm.org/citation.cfm?id=1357127.

Kliman-Silver, C., A. Hannak, D. Lazer, C. Wilson, and A. Mislove. 2015. "Location, Location, Location: The Impact of Geolocation on Web Search Personalization." In *Proceedings of the 2015 Internet Measurement Conference*, 121–27. doi:10.1145/2815675.2815714.

Knobloch-Westerwick, Silvia, Benjamin K. Johnson, and Axel Westerwick. 2015. "Confirmation Bias in Online Searches: Impacts of Selective Exposure before an Election on Political Attitude Strength and Shifts." *Journal of Computer-Mediated Communication* 20, no. 2: 171–87. doi:10.1111/jcc4.12105.

Koenig, Anne M., Alice H. Eagly, Abigail A. Mitchell, and Tiina Ristikari. 2011. "Are Leader Stereotypes Masculine? A Meta-Analysis of Three Research Paradigms." *Psychological Bulletin* 137, no. 4: 616–42. doi:10.1037/a0023557.

Kulshrestha, Juhi, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P. Gummadi, and Karrie Karahalios. 2017. "Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media." In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*, 417–32. doi:10.1145/2998181.2998321.

Kwak, Haewoon, and Jisun An. 2016. "Revealing the Hidden Patterns of News Photos: Analysis of Millions of News Photos Using GDELT and Deep Learning-Based Vision APIs." The Workshops of the Tenth International AAAI Conference on Web and Social Media News and Public Opinion: Technical Report WS-16-18.

Laidlaw, Emily B. 2009. "Private Power, Public Interest: An Examination of Search Engine Accountability." *International Journal of Law and Information Technology* 17, no. 1: 113–45. doi:10.1093/ijlit/ean018.

Loiz, Dana. 2014. "How Many Clicks Does Each SERP Get?—Google Organic Click-Through Rate Study 2014." *Advanced Web Ranking Blog*. https://www.advancedwebranking.com/blog/google-organic-click-through-rates-2014/.

Magno, Gabriel, Belo Horizonte, Camila Souza Araújo, and Virgilio Almeida. 2016. "Stereotypes in Search Engine Results: Understanding the Role of Local and Global Factors." Workshop on Data and Algorithmic Transparency (DAT'16). http://datworkshop.org/papers/dat16-final35.pdf.

Media Insight Project. 2014. "The Personal News Cycle: How Americans Choose to Get Their News." https://www.americanpressinstitute.org/publications/reports/survey-research/how-americans-get-news/.

Mendelson, Andrew. 2001. "Effects of Novelty in News Photographs on Attention and Memory." *Media Psychology* 3, no. 2: 119–57. doi:10.1207/S1532785XMEP0302.

Microsoft Azure. 2017. "Emotion API." https://azure.microsoft.com/en-gb/services/cognitive-services/emotion/.

Muddiman, Ashley. 2013. "Searching for the Next U.S. President: Differences in Search Engine Results for the 2008 U.S. Presidential Candidates." *Journal of Information Technology and Politics* 10, no. 2: 138–57. doi:10.1080/19331681.2012.707440.

Newport, Frank. 2016. "Democrats, Republicans Agree on Four Top Issues for Campaign." http://www.gallup.com/poll/188918/democrats-republicans-agree-four-top-issues-campaign.aspx.

Ørmen, Jacob. 2015. "Googling the News: Opportunities and Challenges in Studying News Events through Google Search." *Digital Journalism* 4, no. 1: 1–18. doi:10.1080/21670811.2015.1093272.

Pan, Bing, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. "In Google We Trust: Users' Decisions on Rank, Position, and Relevance." *Journal of Computer-Mediated Communication* 12, no. 3: 801–23. doi:10.1111/j.1083-6101.2007.00351.x.

"Rankings—comScore, Inc." 2017. Accessed June 27. http://www.comscore.com/
Insights/Rankings.

Rodgers, Shelly, Linda Jean Kenix, and Esther Thorson. 2007. "Stereotypical
Portrayals of Emotionality in News Photos." *Mass Communication and Society*
10, no. 1: 119–38. doi:10.1080/15205430701229824.

Schonberg, Jacob. 2016. "On the Road to the 2016 Elections with Google Search."
Google Blog. https://www.blog.google/products/search/on-road-to-2016-
elections-with-google/.

Stark, Jennifer A., and Nicholas Diakopoulos. 2017. "Using Baselines for Algorithm
Audits." In *Proceedings of the 2017 European Data and Computational Journalism
Conference*, 3. http://datajconf.com/#schedule.

Todorov, Alexander, Anesu N. Mandisodza, Amir Goren, and Crystal C. Hall. 2005.
"Inferences of Competence from Faces Predict Election Outcomes." *Science* 308,
no. 5728: 1623–26. doi:10.1126/science.1110589.

Trevisan, Filippo, Andrew Hoskins, Sarah Oates, and Dounia Mahlouly. 2016.
"The Google Voter: Search Engines and Elections in the New Media
Ecology." *Information, Communication and Society*, 1–18. doi:10.1080/
1369118X.2016.1261171.

"U.S. Presidential Campaign Tracker." 2016. *Lexisnexis*. https://www.lexisnexis.com/
en-us/products/newsdesk/presidential-tracker.page.

Van Couvering, Elizabeth. 2007. "Is Relevance Relevant? Market, Science, and
War: Discourses of Search Engine Quality." *Journal of Computer-Mediated
Communication* 12, no. 3: 866–87. doi:10.1111/j.1083-6101.2007.00354.x.

Weber, Ingmar, and Alejandro Jaimes. 2011. "Who Uses Web Search for What: And
How." In *Proceedings of the Fourth ACM International Conference on Web Search
and Data Mining*, 15–24. doi:10.1145/1935826.1935839.

Xing, Xinyu, Wei Meng, Dan Doozan, Nick Feamster, Wenke Lee, and Alex C. Snoeren.
2014. "Exposing Inconsistent Web Search Results with Bobble." In *Passive
and Active Measurement. PAM 2014. Lecture Notes in Computer Science*, edited
by M. Faloutsos and A. Kuzmanovic, 8362: 131–40. Switzerland: Springer.
doi:10.1007/978-3-319-04918-2_13.