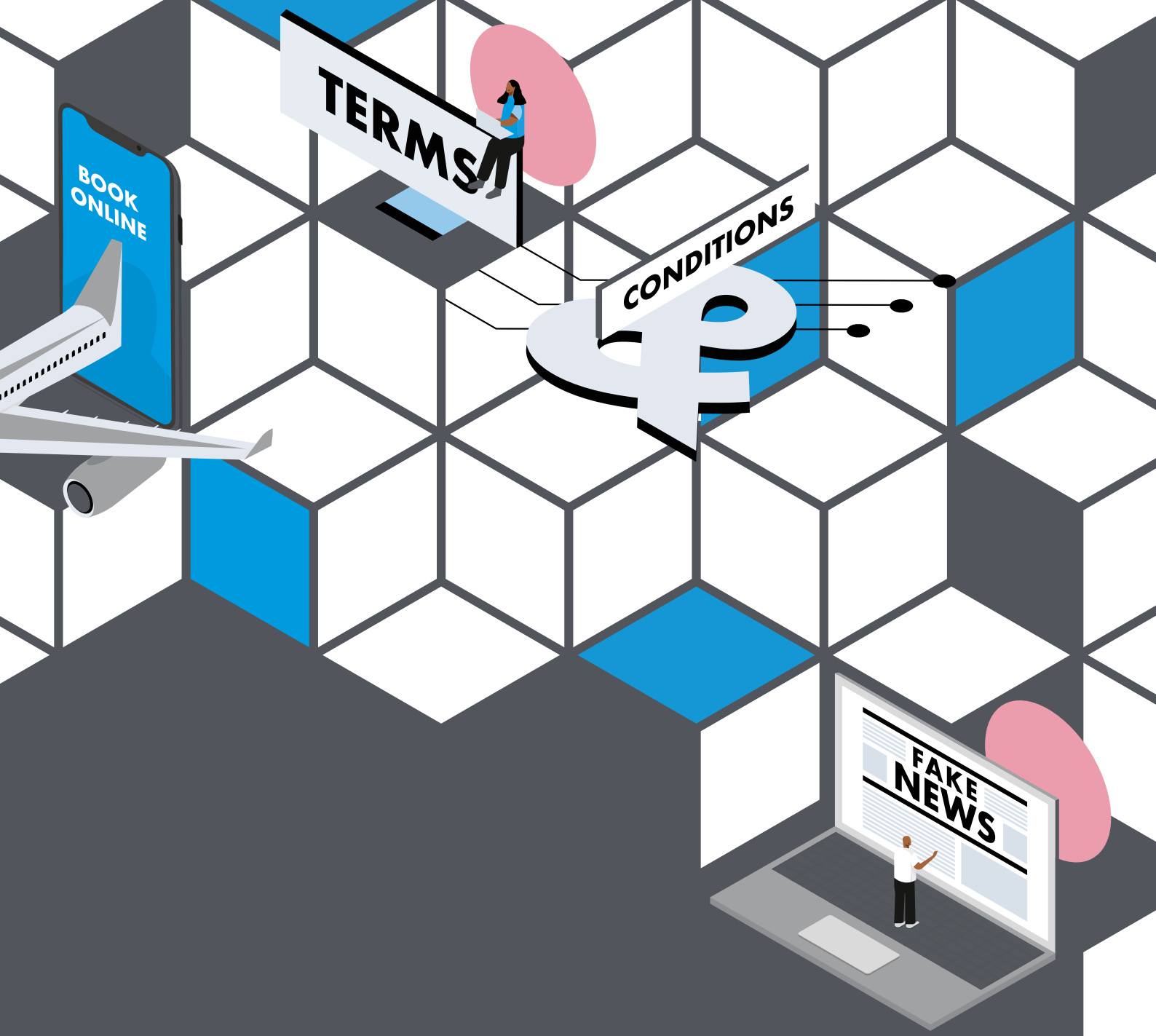


The behavioural science of online harm and manipulation, and what to do about it

An exploratory paper to spark ideas and debate

Elisabeth Costa and David Halpern

THE BEHAVIOURAL INSIGHTS TEAM

Acknowledgements

We would like to thank Lucie Martin, Ed Flahavan, Andrew Schein and Lucy Makinson for outstanding research assistance.

This paper was improved by lively discussions and useful comments from Elspeth Kirkman, Cass Sunstein, Tony Curzon-Price, Roger Taylor, Stephen Dunne, Kate Glazebrook, Louise Barber, Ross Haig, Aisling Colclough, Aisling Ni Chonaire, Hubert Wu, Nida Broughton, Ravi Dutta-Powell, Michael Kaemingk, Max Kroner Dale, Jake Appel, Matthew Davies, Flo Farghly, Toby Park, Carolin Reiner, Veronika Luptakova, Ed Fitzhugh and Pantelis Solomon.

Contents

Executive summary

This discussion paper explores the rapidly evolving landscape of how we behave and interact online, and how businesses respond. The internet has transformed how we live, work and relate to one another. We can make 'friends' around the world without ever saying 'hello', compare products from dozens of shops without leaving the house, or plan a date with a stranger without breaking a sweat. Overall it has proven a powerful force for good; delivering significant benefits to the economy and to consumers in the form of greater choice, personalisation, and incredible convenience. Yet the characteristics of online environments – the deliberate design and the ability to generate enormous quantities of data about how we behave, who we interact with and the choices we make, coupled with the potential for mass experimentation – can also leave consumers open to harm and manipulation. Many of the failures and distortions in online markets are behavioural in nature, from the deep information asymmetries that arise as a result of consumers being inattentive to online privacy notices to the erosion of civility on online platforms. This paper considers how governments, regulators and at least some businesses might seek to harness our deepening understanding of human behaviour to address these failures, and to shape and guide the evolution of digital markets and online environments that really do work for individuals and communities.[1]

The first part of this paper explores the economic and social challenges presented by the evolution of digital markets (and by increasingly spending our time online).

This first section explores how our behavioural biases manifest and are amplified in online environments, and how they can – knowingly or unwittingly – be exploited by companies seeking to manipulate or simply to capture our attention or sell more to us, more effectively.

This discussion is not seeking to be exhaustive, but rather focuses on challenges with a clear behavioural angle, and those that are emerging, where the consequences – individually and collectively – remain uncertain. The challenges we outline are well suited to being addressed with nudges and behavioural solutions, as opposed to 'harder' legislative approaches.

The specific challenges covered are the following.

The potential to exploit consumer biases online (Section 2.1.): In the same way that our behaviour in the offline world is sensitive to subtle cues from the environment, our behaviour in the online world is shaped by the design and characteristics of the websites, platforms and apps we interact with: we click on one of the first three results displayed to us when conducting an online search on our mobile phones more than 70 per cent of the time, for example. Online, behavioural biases can be magnified as we make decisions quickly and process vast amounts of information. Big data and mass experimentation open up opportunities for companies to exploit our cognitive biases and psychological weaknesses using techniques such as adding deliberate frictions (sludge) and harnessing mis-predictions and information deficits (phishing).

Understanding and accepting the 'terms of engagement' online (Section 2.2.): Paypal's terms & conditions altogether total 36,275 words – that's longer than Shakespeare's Hamlet. This is representative of consumers' online experiences: they are asked to engage with long and complex terms and conditions and notices about how companies will use and share their data. A combination of inattention and information overload means that these disclosures are largely ineffective, leaving consumers exposed and with a poor understanding of the true value exchange they are making with online companies.

[1] This paper takes a broad definition of 'online' and 'digital' markets. We discuss challenges and solutions that apply to: platforms that connect C2C, B2C, B2B; online marketplaces for goods, services and information; and social media and networking.

Trust simulations (Section 2.3.): Trust simulations are subtle and worrying techniques that lead us to infer that the seller, organisation or product can be trusted, enabling companies to steer consumers to more expensive or high margin choices, or in extreme circumstances to dupe or cheat consumers more extensively. Scam adverts, like the ones that used Martin Lewis (the famous face of MoneySavingExpert) to endorse questionable money-making schemes, spread quickly through social networks and can be challenging to identify and remove.

Attention wars (Section 2.4.): Attention is a limited resource that is fiercely contested by online players. Swipes, likes, and streaks all combine ease of use and keeping your attention. Bottomless pages keep users scrolling, removing breaks that might otherwise prompt a natural end to the task and our attention to shift elsewhere. These techniques to capture attention also have worrying spill-overs into our ability to focus and be productive in the offline world: the mere presence of a smartphone, set face down on a desk, can worsen performance on a working memory task by 11 per cent.

Predicting our preferences (Section 2.5.): We have different 'orders' of preferences. Our 'first order preference' is expressed in how we behave in the moment that the stimulus or temptation is presented. In contrast, our 'second order preference' is the choice we make for ourselves on reflection, generally separated from the immediate temptation. One concern about online markets is that they enable much more fluid expression of first order, impulsive preferences to the detriment of consumers, particularly the vulnerable. For example, online gambling environments are designed to keep people betting more, for longer: separate 'online wallets' for casino, sports, or poker gambling on betting websites encourage us to think about our money in a more fragmented way and to spend more than we might otherwise.

We also make different decisions collectively than we do as individuals. This is known as 'third order' preference: what we would choose on reflection and discussion with others. This matters because when we are online, we are often online together. As platforms are increasingly making difficult judgments about how to set the rules of their network, such as around what constitutes extremist or unacceptable content, they face a dilemma about whether this should be set by first, second, or third level preferences and whether and how to override these if they prove problematic.

More than markets: morals, ethics and social networks (Section 2.6.): There is more to markets than money. Markets are entwined with 'moral sentiments', social networks ('who knows who'), and a myriad of habits and customs that shape our lives and societies and the character of our economies. Online markets, and social media in particular, have the potential to significantly affect the character and form of our social capital. On the darker side, they can reinforce echo chambers and social and political 'bubbles' as well as propagate and potentially extend the exclusion of groups or individuals: Airbnb guests with distinctively African-American names are 16 per cent less likely to be accepted than identical guests with distinctively White names. More anonymised forms of communication also appear to licence more negative behaviour and exchanges. Yet online environments are also creating new opportunities to widen social networks; extend trust to relative strangers; and foster the 'economy of regard'. Finally, new evidence suggests that social media is impacting mental health, particularly amongst teenage girls, both via eliciting negative feelings, and via substituting time away from wellbeing-enhancing activities such as offline socialising.

Emerging problems (Section 2.7.): This final section outlines emerging and uncertain challenges such as: fake news and its impact on society and politics including the rising threat of 'deep fakes'; increasingly personalised pricing and price discrimination; the role of algorithmic bias in discrimination and injustice; and the creation of new types of 'winner-takes-most' monopolies.

The second part of this paper outlines the behavioural tools that government, regulators and industry can and should deploy to shape and guide the evolution of online markets.

Against the backdrop of rapid and genuinely unprecedented change, government and regulators need to work hard to keep up, and even more so to get ahead. The pace of change means that many of the tools of competition policy and consumer law need to be adapted and supplemented to ensure that consumers continue to benefit from technological change and firms are encouraged to continue innovating.

The discussion of what we can, and should do is organised around four themes: enhancing traditional responses; choice architecture; fostering sharper competition; and non-market design. Our focus throughout is placing a sophisticated understanding of human behaviour, including active and constructive dialogue with the public, at the heart of policy solutions. Recommendations are made to government and regulators as well as to industry, which we consider to have a central role in shaping positive online markets.

The core recommendations are highlighted below.

Section 3.1. Enhancing traditional responses, updating the traditional regulators' toolkit to provide better, clearer information about the choices consumers face online.

Smarter disclosures

There are many opportunities to use behavioural science to improve the information being provided to consumers online – from how their data is being collected to the terms on which they are using services or buying products. Encouraging smarter disclosures in these areas will help to increase consumers' level of understanding about the 'terms of engagement' online.

Recommendation: Publish evidence-based Best Practice Guides on how to improve online disclosures like Terms & Conditions and Data Privacy Notices.

Recommendation: Set the acceptable average level of understanding that can constitute informed consent, and require businesses to conduct ongoing comprehension tests to assess and improve the information they are providing to users.

Education

There remains a role for education. Here we focus on the potential to develop and disseminate useful 'rules of thumb' to help consumers navigate common challenges faced online. We also explore approaches that teach people how to recognise potentially dubious claims, and interrogate the interests of those who are making them, giving people the skills to become more resistant to a range of persuasive and potentially deceptive ploys.

Recommendation: Fund research to design and test novel approaches to build consumers' resilience against challenges like disinformation and online fraud.

Recommendation: Develop supportive apps and practices that can provide 'training wheels' to young people first using social media and engaging online.

Exhortation

Similarly, there is a clear role for the public sector to urge firms and citizens to take action, or do something differently. The UK's Chief Medical Officer, for example, recently urged parents to 'leave phones outside the bedroom when it's bedtime', and enjoy 'screen-free meal times' with their families. Yet, in some areas, governments should be ready to move beyond general exhortation and be more targeted in two ways. First, focus on urging companies rather than consumers to change their behaviour and, second, seeking to change the behaviour of specific companies as an example to industry.

Recommendation: Publically urge companies to change and improve their policies and practices, and signal willingness to intervene more strongly if change isn't satisfactory.

Recommendation: Establish an annual, consumer-led 'Sludge' award.

Section 3.2. Choice architecture, consumer-focused defaults which make choices clearer and easier to tailor to individual preferences.

Giving back individual control

One thing regulators and innovative companies can seek to do is to put as much control as possible back into the hands of the individual, allowing consumers to express their preferences and modify their online experiences to act in line with those preferences. Promising avenues include more prominent, tailored user settings, and mechanisms that allow vulnerable consumers to self-exclude or commit themselves to future actions.

Recommendation: Work with consumer bodies to identify areas where self-exclusion tools could protect consumers from online harms, and encourage the development and take up of these tools, particularly for vulnerable consumers.

Prompts, reminders and active choices

Another way of putting control back into the hands of users is through the use of prompts and reminders. These can be a powerful mechanism to elicit consumer preferences, especially where they contain a meaningful and 'active' choice – where a website or process requires the user to respond yes or no to a question before proceeding – offered at a salient moment.

Recommendation: Design and test prompts and reminders that provide consumers with active choices about how they interact with websites and platforms.

Smarter defaults

We have a strong tendency to act in line with the pre-set option, for example the time before our phone locks itself or the default privacy and notification settings of an app. The data revolution associated with online markets has moved smarter, more personalised defaults from an interesting theoretical idea into a real and widespread possibility. Guiding and prompting individuals to set smart defaults for how they use their phone or device and where they focus their limited attention is one potential avenue. These defaults could, for example, set limits on the time spent on particular activities, organise how notifications are bundled, and when and how devices are 'muted'. Another possibility is to set company-facing defaults that are firmly in favour of consumers, and that clean up sludge. For example, cancelling an online subscription, unsubscribing from mailing lists or leaving a platform should be as easy as signing up.

Recommendation: Design, test and promote cross-platform defaults that guide how consumers use their phone and other devices.

Recommendation: Make cancelling an online subscriptions, unsubscribing from mailing lists or leaving a platform as easy as signing up.

Section 3.3. Sharper competition, fostering genuine competition between market players, and encouraging new types of intermediaries.

Transparency to facilitate comparison and create accountability

We are optimistic that 'forward-looking' competition tools can give consumers greater choices over their digital services, help new companies enter and grow, and spur innovation. In particular, competition can be fostered through transparency mechanisms like user feedback systems and transparency reporting. These mechanisms are designed to facilitate meaningful comparisons and create accountability.

Recommendation: Actively encourage the wider emergence of feedback and ratings platforms, and design these systems to minimise the prevalence of fake reviews & reputation inflation.

Recommendation: Introduce transparency reports for online companies and give regulators powers to audit existing efforts and require remedial steps if they are not.

Recommendation: Conduct and publish research on the comparative performance of platforms in relation to welfare effects, online harms, data protection and other areas.

Building trust and confidence in digital comparison tools

Comparison tools are integral to well-functioning markets, but particularly digital markets where consumers are choosing between many more, and more complex products and services. Regulators should take a stronger and clearer role in catalysing these types of intermediaries and building consumer trust in them, including by helping consumers tell the difference between comparison tools. For example, highlighting the differences in the average price savings achieved; number of complaints against the site; and performance as measured by mystery shopping.

Recommendation: Publish league tables of switching sites.

Leveraging data and AI on the side of consumers

The large flows of data online tend to exacerbate the information asymmetries between suppliers and consumers. Therefore a key question is how we can rebalance this dynamic by leveraging this data on the side of consumers. We argue that this can be achieved through facilitating greater portability of user data, allowing users to take their playlists, networks and other data to rival services. The development of new intermediaries should be fostered to help users to sift through large volumes of complex information and act in line with their goals and preferences, or even allow intermediaries to act on their behalf. Data can also be used to identify vulnerable consumers and signpost them to tailored support.

Recommendation: Allow greater access and portability of data so that users – consumers, workers, and businesses – can try alternative services and switch more readily.

Recommendation: Use mechanisms like innovation funding and challenge prizes to kick start and foster intermediaries that leverage data to benefit consumers.

Next generation antitrust to support new market entrants

The scale of online markets coupled with the value of network effects and access to data means that some platforms can come to dominate the market, tipping into a 'winner-takes-most' dynamic. This section explores novel ways to use data to support new market entrants and innovation. Specifically, creating a progressive data-sharing mandate where dominant companies would be compelled to share a representative cut of anonymised data with new market entrants and competitors.

Recommendation: Actively pursue data openness, including investigating the feasibility, costs and benefits of a progressive data sharing mandate.

Section 3.4. Non-market design, fostering new norms for positive and constructive interactions online

Patterns of association

The internet and online platforms have changed the costs and benefits of trading and connecting with other people, massively expanding our 'weak ties' by connecting us with huge numbers of people we don't know, but also facilitating new forms of association and potentially discrimination that will prove challenging to current legal thinking. This raises questions about how are we to handle the emergence of overtly, or incidentally, exclusive groups that emerge, particularly on social media. A proactive approach may be possible and desirable to encourage forms of exchange and connection that foster exposure to alternative perspectives and the building of 'bridging social capital'.

Recommendation: Encourage, and potentially compel, platforms to put in place structural features to minimise discrimination. Explore how machine learning of patterns of online association, complaints mechanisms and feedback loops can identify and reduce overt and inadvertent discrimination.

Recommendation: Actively explore and encourage forms of exchange and connection that foster exposure to alternative perspectives and the building of 'bridging social capital'.

Civility

Widespread concerns about social media nurturing negative and uncivil interactions could be addressed through the evolution of a self-regulatory dynamic based on informal codes of conduct that the vast majority of users are comfortable with. These could be strengthened through feedback loops that provide feedback on why content was considered unacceptable (rather than just removing it with no explanation), and also enabling users to contest the removal of content.

Recommendation: Develop and test new prompts and defaults that encourage civility between users, and specifically encourage users to reflect before posting potentially harmful content.

Recommendation: Design systems that provide feedback – both from the platform and from other users – to users who choose to post harmful or offensive content.

Who decides?

Issues of association, civility or content have a degree of fluidity, and are strongly rooted in civil society rather than in the black and white rules of legal code. Many of the key market players are approaching market dominance, and have acquired powers and responsibilities that touch the lives of billions. There needs to be a place for users and citizens to negotiate with each other to reach a collective view on what constitutes appropriate practices and 'rules of the game'. This is especially true for those platforms that see part of their essence as nurturing a community, yet appear to lack any meaningful way for that community collectively to shape the core parameters of the platform itself. In short, there needs to be mechanisms through which the community can 'nudge the nudgers'.

Recommendation: Design new and appropriate governance mechanisms for platforms that combine expert opinion with the collective user voice to allow a platforms' community of users (and possibly all citizens) to shape the character and rules of behaviour on the platform.

Recommendation: Create a new and independent 'online' Ombudsman system to help interpret the platforms' formal and informal rules; adjudicate on contested individual cases; and advise on acceptable behaviour by users and on the evolution of the network's own rules and practices.

The Figure below maps these categories of response to the challenges we explore. It gives an indicative sense of how effective we expect each policy response to be in tackling the challenges.

Figure 1: How effective do we anticipate these responses to be?

	Section 2.1. Manipulative and exploitative practices and pricing	Section 2.2. Terms that consumers don't understand or engage with	Section 2.3. Lack of trust and trust-worthiness	Section 2.4. Attention capture	Section 2.5. Moving beyond first level preference	Section 2.6. Association: exclusion, incivility and mental health	Section 2.7. Winner-takes-most monopolies, need for innovation and disruption
Section 3.1. Enhancing traditional responses: Better, clearer information about the choices consumers face online							
Smarter disclosures	+	++	+		+		+
Education	+		+		+		+
Targeted exhortation	+++		+				+
Section 3.2. Choice architecture: Consumer-focused defaults and systems which make choices clearer and easier to tailor to individual preferences							
Giving back individual control and facilitating self-exclusion for the vulnerable	++		+	++	++	+	+
Prompts, reminders and active choices	+	++		++	++		+
Smart defaults	++		+	++	++		+
Section 3.3. Sharper competition: Fostering genuine competition between market players and encouraging new types of intermediaries							
Transparency to facilitate comparison and create accountability	+		+++		+		+++
Building trust and confidence in digital comparison tools	++		++		+		++
Leveraging data and AI on the side of consumers	++	+++	++	++	++	+	+++
Next generation antitrust to support new market entrants	+						+++
Section 3.4. Non-market design: Fostering new norms for positive and constructive interactions online							
Patterns of association: encouraging bridging social capital, reducing overt and indirect discrimination			+			+++	
Creating norms of civility			+		+	++	
Who decides? Fostering deliberative mechanisms to shape rules.	+		++	+	+++	++	+

This landscape will continue to shift under our feet. The added challenge for government and regulators is to take a proactive and experimental approach, anticipating the many areas where online markets can be harnessed for good, and where the disruptive capabilities of digital platforms has yet to hit. Active experimentation and the flexibility to adapt regulatory responses as challenges shift will also help the public sector keep pace with the rapid evolution of these markets.

This is about as important a challenge as we face in society today, and one which we need to ensure that our citizens can themselves be involved in fashioning. How we respond to, and shape, the evolving character of the digital landscape is precious not just because it is pivotal to our economies, but because it is society and the human character itself that we are shaping.

1 / Introduction

The internet has transformed how we live, work and play. Today online you can make 'friends' around the world without ever saying 'hello', compare products from dozens of shops without leaving the house, or plan a date with a stranger without breaking a sweat. The growth of the digital world has also forced a fundamental questioning of how we conceive of and run the economy: the digital economy has facilitated competition yet also concentrated market power amongst a few key players;[II] high quality products and services are available at lower prices or, in the case of many platforms, at no cost; and the connected world has delivered greater efficiency than ever before.

While the benefits of life online are significant, so too are the economic and social costs. Using the internet is increasingly simple, with many arguing that access is now a fundamental right, but using it responsibly can be far from easy. Without knowing it, we leave traces of ourselves for advertisers to target us, we can be led into sharing sensitive personal information by web design that allows us to dismiss complex terms and conditions (T&Cs) in a single click, and we spend increasing amounts of time inhabiting an online world in which anyone can throw a nasty jibe without ever looking us in the eye. You've heard of 'nudge'; these examples of the darker side of the evolution of online and digital markets give a glimpse of its manipulative cousins 'sludge' and other techniques that seek to harness our behavioural biases against us rather than for us.[1]

Of course, these phenomena are not new. For years, marketeers have used these tactics to sell us products we don't need to impress people we don't like. But as we spend more of our lives in carefully curated online spaces, it becomes harder to escape this damaging choice architecture, especially when the data we leave behind reveal things that allow for an unprecedented degree of personalised targeting. Indeed, the biggest companies reach into almost every facet of our lives: they know – perhaps more than our closest loved ones – when we go to bed, who we interact with, and the questions we have on our minds.

Against the backdrop of this rapid and unprecedented change, government and regulators need to work hard to keep up, and even more so to get ahead. Many of the tools of competition policy and consumer law are no longer fit for purpose and must be adapted to ensure that consumers continue to benefit and firms are encouraged to continue innovating. Of course, not all online problems should be addressed with a 'nudge'. The proliferation of fraud and child sexual exploitation online, for example, requires strong legislative responses, such as the recent announcements in to UK to create and enforce a duty of care for companies to keep their users safe online.[2] This discussion paper focuses on the many emerging challenges where the consequences remain uncertain and there is a clear role for nudges and behavioural solutions.

Richard Thaler, economics Nobel Laureate and longstanding adviser to the Behavioural Insights Team, often signs his books with the maxim 'nudge for good'.[III] This paper explores the ways in which governments, regulators and businesses might seek to enact this sentiment in the online world to shape and guide its evolution. It proposes ways to harness our deepening understanding of human behaviour and data to build fairer, better online markets.

[II]The FAANGs – Facebook, Apple, Amazon, Netflix and Google (Alphabet) – alone have a total market capitalisation of US$3.17tn (figures computed from CNBC, valid as of 15 March 2019). For comparison, the GDP of the UK in 2018 was US$2.81tn (figure from the International Monetary Fund).

[III]Richard Thaler uses this phrase. David Halpern, partly based on his role as What Works National Adviser – encouraging more empirical methods in policy – and as a tribute to Richard, often uses the variation 'Nudge with humility and for good'.

2 / Challenges

Let's for a moment imagine Danny: a regular guy booking a holiday online. Danny has consulted a flight aggregator and a hotel booking site and carefully weighted all of the various factors that are important to him: overall price, travel time, choice of hotel, opportunity cost and expected enjoyment of the destination. After careful deliberation and reading all of the booking conditions, he comes to an optimum decision: he books two days in Paris, mid-week, in a 3-star hotel.

Classical economics predicts we'll all make decisions in the same 'utility maximising' way as Danny, offline and with the help of Google, Skyscanner and Booking.com. Of course, it turns out that humans don't work like this. We pick the cheapest flight, forgetting to factor in the add-on cost of bringing a bag on the plane. We are influenced by the pop-up telling us that '5 people are looking at this hotel right now' and we don't read the fine print setting out the cancellation policies and what data we've shared with the various websites through the process.

Behavioural economics seeks to correct the assumptions underlying Danny's decision, instead positing a more realistic account of behaviour based on 'boundedly rational' agents.[3] [4] These agents cannot complete Danny's theoretical spreadsheet because they will often find that many aspects of the comparison are 'shrouded' – hard to estimate or observe.[5] Will I optimistically book hand luggage only and then over-pack on the day? Can I trust the restaurant recommendation of this Instagram influencer? Further, their understanding of their own behaviour, or even preferences, is not perfect. They would – like many of us – struggle to truly answer some of the most important questions: do I actually like city breaks or would I rather be in the countryside? Would I be happier at home with a good book, a bottle of Bordeaux, and the cost of the Eurostar tucked away in a savings account?

This is not a theoretical problem. Many of the failures and distortions in digital markets are behavioural in nature. To govern and regulate better, we must first understand how our behavioural biases manifest and are amplified in online environments, and how they might be being harnessed, manipulated and exploited.

2.1. The potential to exploit consumer biases online

In the same way that our behaviour in the offline world is sensitive to subtle cues from the environment,[6] our behaviour in the online world is shaped by the design and characteristics of the websites, platforms and apps we interact with. Online, behavioural biases can be magnified as we struggle to process vast amounts of information and tend to make decisions quickly. This leaves us open to our biases being (knowingly and unknowingly) harnessed, manipulated and exploited.

Several Nobel Laureates – Robert Shiller, George Akerlof and Richard Thaler – have written about how companies seek to manipulate and exploit our cognitive biases and psychological weaknesses.[7] Here we focus on two shades of this manipulation: the deliberate exploitation of information deficits and behavioural biases (phishing); and the deliberate addition of frictions and hassles to make it harder to make good choices (sludge).

Targeting our biases and psychological weaknesses

Companies have long used techniques to target and exploit our psychological weaknesses. Shiller and Akerlof call this 'phishing'. Offering a cash-back deal is a classic example: the vast majority of us erroneously predict that we will go to the effort of actually claiming the money. Our purchase decisions may be more sensitive to the promise of the cash than the effort required in actually claiming it, while redeeming the cash back offer is driven more by the effort involved than the cash available.[8] The seller can experiment to find the optimum level of cash-back that will tempt you, but not be so large that you will actually complete the claim. Today, a shopper may subscribe to Amazon Prime, believing that they are saving money via free shipping when in reality they are just being tempted to buy more – in fact the longer someone is a Prime member, the more they spend on the platform.[9]

These types of manipulative practices are exacerbated online. We previously had no way of knowing whether we were fast enough to be in the 'first 50 responders' to claim a free gift or discount. But now we are more likely to trust real time updates, for example, from hotel booking sites urging that there is 'only one room left at this price' or '5 people are looking at a room right now', especially if they have already shown you all the sold-out hotels in your initial search results. This is designed to harness our sense of scarcity, and our desire to see social proof that other people are making similar choices to us. But what they don't tell you is that those five people aren't necessarily looking at the same room as you.[IV] [10]

The intersection of data, imbalances of information and intelligent marketing also opens up new opportunities to exploit our biases. Conspiracy theorists might claim that our devices are listening to our conversations, when in fact the data we willingly share is more than enough to predict what we might buy, when, how much we are willing to pay and the flawed decisions we might make along the way.[V] [11] Further, the data we share online also allows companies to move away from the model of active, consumer-led searches and towards prompting us with targeted advertising and information at times when we are most vulnerable to being manipulated. The rise of 'femtech' apps to track menstrual cycles and fertility has allowed businesses to collect large and robust data sets to produce tailored and timely marketing.[12] More disturbingly, social media platforms' ability to identify users' emotional states based on their activity means that – in the alleged words of Facebook (Australia) staff – advertisers have the capacity to identify when teenagers feel 'insecure' or 'need a confidence boost'.[13]

Behavioural biases that are well documented in the offline world can be magnified and exacerbated online, leading to new and worrying opportunities to exploit information deficits and biases.

[IV] The CMA recently intervened to end these types of misleading sales tactics that put pressure on people to book more quickly.

[V] One study aiming to investigate this theory found no evidence of recorded conversations, however they did find that some companies were sending screenshots and videos of users' phone activities to third parties.

Box 1: How do our behavioural biases manifest online, and how do companies respond?

Choice and information overload

In 2018, Google processed 1.2 trillion searches; the equivalent of 170 searches by every person on the planet. Search and comparison tools can present us with many more options that we might otherwise have found, but it does not necessarily follow that we take full advantage of that extra information. We can be overwhelmed by the array of choices. This impacts whether we can make a choice at all and the quality of subsequent decisions,[14] making it more likely we'll fall back on heuristics and rules of thumb to guide our behaviour. Indeed, consumers consider on average no more than three brands when making a choice online,[15] an outcome consistent with the literature: we value options but ultimately prefer to choose from a much smaller set.[VI] [16]

Position biases and ordering effects

Besides answering a large fraction of the world's queries, Google and other search and comparison tools also determine the order and format in which search results are presented. There is evidence that suggests that details about how choices are presented online are highly consequential.[17] A review by the UK Competition and Markets Authority (CMA) concludes that 'position bias' – the tendency of those searching online to disproportionately select the top results – is, in part, driven by the order in which they are presented.[18] This trend is consistent across search engines and digital comparison tools, with the first three links accounting for 40-65 per cent of total clicks on desktop devices and more than 70 per cent of total clicks on mobile devices.[19] Online consumers therefore often compare fewer options than one might expect.[20] That means that there is significant competition between companies for top rankings, and associated costs can often be passed onto the consumer. In the US alone, companies are spending close to $80bn on Search Engine Optimisation (SEO) to improve how their websites and products rank in online search results.[21] And this position bias may lead to entrenched selling power, independent of quality, for those companies that bagged the top spot early.

Attracting attention

There is some evidence that web users have a natural tendency to process online information in certain ways, such as comparing on a horizontal rather than vertical plane.[22] Companies may use this tendency to influence how comparisons are made, for example making trade-offs between price and customer reviews, or introducing search frictions.[23] In the offline world, attributes such aesthetic appeal often influence perceptions of quality.[24] This can be accentuated online as web users make rapid assessments of the quality or trustworthiness of a site on the basis of superficial design features.[25] Alternatively, attentional biases can mean that companies are able to use 'drip' pricing, or 'bait and switch' strategies to encourage users to spend more or buy more expensive products than they intended.[26] [27]

Online disinhibition

A stronger sense of anonymity or social distance in online decision making may lead to disinhibited behaviours, whereby we make decisions that would be less likely in offline environments. This can manifest itself in various ways, such as increasing our willingness to engage in bullying or incivility, choosing unhealthier food options,[28] or spending more money on impulse purchases.[29]

[VI] Dilip Soman discusses his research on this 'overchoice' effect in his book 'The Last Mile', starting with a representative anecdote: he was once asked by the owner of a Chinese restaurant why his menu offered 155 items, yet about 80 per cent of his sales were coming from only five or six of these items.

Sludge

The conscientious nudger seeks to design systems and processes to help people make better choices for themselves. They use defaults to make it easier to save, rank deals based on best value, and tell us what others like us do. Thaler has coined the term 'sludge' to describe the deliberate use of these tactics to thwart us acting in our best interests.[30] There is plenty of sludge offline, notably in administrative paperwork burdens,[31] but online environments (and their jarring interaction with offline processes) allow sludge to thrive.

You can set up an online news subscription with one click, yet when you go to cancel you'll find yourself in a complicated loop of online and offline hurdles: calling a phone number during restricted business hours, filling out a detailed form and posting it to the company, or even being required to visit a physical store. These muddy frictions are deliberate. Companies know that they matter. Indeed, the mere friction involved in cancelling even a trial subscription is enough to discourage many people. This is strikingly illustrated in a natural experiment following a 2007 US Federal Trade Commission (FTC) decision to close down a company charging ongoing fees for worthless subscriptions. Customers enrolled for more than six months before the ruling were required to take action (by mailing a form or making a phone call) to cancel their memberships, while more recent customers were told their subscriptions would be cancelled, unless they took an action for their subscriptions to continue. Cancellations increased from 63.4 percentage points among those required to take an action, to 99.8 per cent among those who were required to do nothing.[32]

Online airline bookings are also fertile ground for sludge. An enticingly low headline flight cost and large 'book now' button is followed by a series of pages slowly adding on additional choices, information and costs. When you get to reserving your seat, you discover that you can pay online or if you want to be randomly allocated a free seat you'll need to queue at the check-in desk on the day of the flight. All designed to sell you more and discourage you from choosing free or cheaper options.

Of course, friction can also be dialled down or removed from processes to encourage behaviour that is 'self-defeating'. You can be approved for a high-cost payday loan with just a few clicks and shopping or gambling online at 3am is easy to start and difficult to stop. The practice on Netflix and other streaming platforms to automatically start the next episode in a series is a clear example of the power of removing friction, or flipping the default. People fail to press stop, and thereby watch another episode they might not have watched had they had to press play. In our own email newsletter at BIT, we found that when we switched from an opt-out model (all subscribers received the newsletter) to a new, opt-in model (subscribers were asked to confirm whether they wanted to stay on the mailing list), almost 50 per cent of subscribers opted in. This led to an increase in engagement rates with the newsletter, from about 50 per cent to almost 100 per cent, showing that removing friction can be a tool not just to keep subscriptions going but also to help align people's choices with their true preferences (more on this in Section 2.5.).

Sludge and phishing are two techniques that companies use to influence our behaviour online. The next parts of this section explore further techniques used to manipulate and exploit our behavioural biases: overloading us with information about the terms on which we are engaging with companies and each other online (Section 2.2.), simulating trust and trustworthiness (Section 2.3.), designing products, services and environments that seek to capture and hold our limited attention (Section 2.4.) and predicting our preferences and harnessing our impulsive tendencies (Section 2.5.).

The Response section proposes ideas to correct information deficits (Section 3.1.), to combat sludge through smarter defaults that make exiting a contract or service as easy as signing up (Section 3.2.), and to foster competition and squeeze out poor and exploitative practices (Section 3.3.).

2.2. Understanding and accepting the 'terms of engagement' online

Many of us have impatiently or absentmindedly ticked a box to agree to the terms of the latest iOS upgrade or the returns policy of an online retailer,[VII] [33] or clicked past the link that allows us to 'learn more' about a company's data privacy policy. Taken together, T&Cs (the contract you are agreeing to when buying a product or using a service), data privacy policies (covering how the company collects, uses, shares and manages your data), terms of use (covering rules about using a website), and cookie policies (covering how your behaviour is tracked) represent the terms on which we engage with companies and others online. That most people (91 per cent by one survey)[34] do not read the terms they 'actively' agree to is little surprise: PayPal's are longer than Hamlet.[VIII] [35] For data privacy policies, back in 2008, the average web user would have needed to set aside 244 hours, or 40 minutes a day for a year, to read through the policies of all the websites they visited.[36] Despite companies making some progress in shortening their policies, this figure is likely to be much larger today, given that time spent on digital media has more than doubled since.[37]

As well as being too long, T&Cs and privacy policies are also often incomprehensible for the average user. A review of over 200,000 privacy policies found that, on average, they require a college-level reading ability,[IX] [38] yet in the UK average reading abilities are far lower than this.[39] [X]

Figure 2: Length and complexity of typical terms of service and privacy policies

Company	Time to Read	Reading age required
Instagram terms of use and data policy	31 mins	18 years
Facebook terms of service and privacy policy	34 mins	18 years
Whatsapp terms of service and privacy policy	41 mins	21 years
Snapchat terms of service and data policy	35 mins	16 years
Twitter terms of service and privacy policy	46 mins	18 years

[VII] Web interfaces have, over many years, encouraged people to mindlessly click away dialogue boxes or agreements that stand in the way of completing their primary action. We tend to ignore these banner pop-ups because they resemble ads, or we perceive them to be of low importance because it is possible to continue by simply clicking away.

[VIII] The General Data Protection Regulation, introduced in April 2018, states that passive consent can no longer be implied through site usage and consumers need to opt in or actively agree to how their data is collected and used. GDPR Article 29: 'merely continuing the ordinary use of a website is not conduct from which one can infer an indication of wishes by the data subject to signify his or her agreement to a proposed processing operation'.

[IX] The study found that these 200,000 websites had an average Flesch Reading Ease score of 39.83. This score corresponds to a 'college' reading level.

[X] Less than 60% of 16-65 year-olds have literacy levels equivalent to a GCSE grade A*-C.

The fact that T&Cs and privacy policies are too long and too complex is not just a daily annoyance. It is problematic: we rarely understand the terms on which we are engaging with online companies, or the true value of the exchanges we are making.

In a provocative online experiment, 98 per cent of students invited by their university to join NameDrop, a fake social network, agreed to T&Cs which included the statement '...all users of this site agree to immediately assign their first-born child to NameDrop, Inc'...! [xi] [40]

Arguably the area where we least understand the exchanges we are making is on the data we agree to share, or share by default. Collectively, these data are of enormous commercial value. They allow businesses to better target advertising, improve sales conversions, and identify gaps and opportunities in the market. The value is not all one-sided: recent Salesforce research found that 82 per cent of customers are willing to share relevant information about themselves in exchange for more personalisation and convenience.[41] However, many consumers remain unaware of exactly what data they are sharing or the extent to which they are giving permission to platforms to reach into various aspects of their online and physical behaviour. This is partly because privacy policies don't say much about how they share data in the first place – a recent large-scale audit found that while third-party data collection is widespread, fewer than 15 per cent of attributed data transfers are disclosed in websites' privacy policies – and because the more obscure the third party, the less likely data transfers were to be disclosed.[42]

Box 2: Have you signed up to Facebook or Instagram? Here are some surprising data you are sharing...

- Where you are, accessed through your GPS location (depending on your device settings and permissions)
- Your phone number, even if you haven't provided it in your profile
- Access to what you see through your camera when you're using the app's camera feature regardless of whether you are recording
- Your mobile signal strength or internet connection speed
- Your contacts, who you've called and your SMS history (if you choose to sync, upload or import your phone contacts with the Instagram app, for example)
- How much battery you have left on your phone or laptop
- How much storage is available on your phone or laptop
- The other apps you have installed on your phone or laptop
- Information about other devices nearby or on your network

[xi] Note the term was arguably buried – in clause 2.3.1 of the contract – but so are many of the terms that determine what data we share, and limitations on returns, and other compensation.

Even if we were somehow able to absorb all the information given, data privacy is complicated by evidence that our privacy preferences are not stable (more on this in Section 2.3.) and contextual factors can influence perceptions of privacy, or of how our data can be used. When people are unsure about their own preferences, they may be guided by environmental cues.[43] For example, in an online experiment run by BIT to understand patient preferences for how the National Health Service (NHS) might use their data,[XII] people's data-sharing preferences varied depending on the number of options they were given. We found that 77 per cent of people said they would hypothetically share their data when choosing between two options (sharing their data or not for 'research and planning'), but only 70 per cent did so when choosing between four options (sharing for both, sharing for research only, sharing for planning only, or opting out of both).

This cocktail of friction-free agreement to T&Cs, inattention to privacy policies, widespread lack of understanding, and unstable preferences leaves consumers open to exploitation by businesses seeking to harness ever more data and share it more widely. There is a perceived inevitability – in all of these interactions, we are never given the honest option of saying 'No, I haven't read it and don't plan to'. This leaves many internet users feeling powerless in this exchange: in research conducted by the Pew Research Centre, 91 per cent of US adults surveyed agreed that consumers have lost control of how personal information is collected and used by companies.[44]

The good news is that at least some of these challenges are surmountable. In a recent set of experiments, BIT tested new and effective ways of presenting T&Cs and privacy policies to increase understanding and engagement.

The Response section outlines the results of these experiments and ideas for reform, including smarter disclosures (Section 3.1.) and more radical ideas for a system where active engagement isn't required (Section 3.3.).

2.3. Trust simulations

Alongside sludge and phishing, there are more subtle and perhaps worrying techniques that can be applied to influence our choices. Trust is an important factor in online transactions and sellers can use techniques to persuade the buyer they should be trusted. While creating trust is clearly a desirable outcome for both parties, once gained it can be exploited to sell more expensive or high margin goods, or in extreme circumstances to dupe or cheat the consumer.

In his work on persuasion, Robert Cialdini gives powerful illustrations of such practices.[45] [46] Take, for example, the restaurant waiter who starts by making a point of directing one of the guests to a 'better' but notably cheaper dish. Then later, when it comes to choosing the wine, he once again intervenes, but this time with a more expensive recommendation. The net effect is a more expensive meal – and thus a bigger tip for him – plus a grateful, generous customer. Building trust – or the appearance of trustworthiness – pays.

[XII] This experiment was run on BIT's online research platform, Predictiv (www.predictiv.co.uk). Predictiv allows organisations to make better, evidence-based decisions by rigorously testing communication, product and service design ideas in randomised control trials when field experiments are not possible or too expensive.

Box 3: Trust simulations in action

In April 2018, Martin Lewis, the famous face of consumer website MoneySavingExpert, declared he was suing Facebook for not shutting down over 1,000 scam adverts which used his image, without permission, to endorse money-making schemes. Lewis and Facebook reached a settlement in January 2019, on the conditions that Facebook donates £3 million to Citizens Advice to create a new UK Scams Action project, and launches a (new, UK-specific) reporting tool and a dedicated team to handle such complaints.[47]

As markets have evolved, consumers have got wise to some of these techniques, and regulatory responses have quashed others. However, new forms of trust simulation have emerged, varying from fake reviews or kitemarks, or misleading discount claims or reference prices, to more elaborate ploys.

Ironically, some measures intended to make consumers more careful may inadvertently be used to simulate trustworthiness. Standard legal warnings on investments, for example, with the logo of the regulator prominently displayed on the page, may make it appear the regulator has approved a product, leading consumers to be *more* trusting of it.[48] [49] Even more dangerously, field studies have suggested that getting professionals, such as medics or financial advisers, to declare conflicts of interest directly can lead their clients to be less trusting of them, yet more compliant with their advice, due to the social pressure and conflict avoidance that comes into play when we know what other people want from us.[50]

Since getting vendors to declare that they can't be trusted can backfire, it may not be surprising to discover that telling consumers when they can trust someone can also have unintended results. For example, subjects filling in information on a website that assured them their data would be held securely and not shared, were less likely to answer questions honestly than when asked on a website that offered no such assurances.[51] Trustworthiness, it seems, can be hard to reliably communicate or assess.

With respect to privacy and data security, it appears that many people don't think about their data being abused until they are reminded of the possibility. Furthermore, the cues that we use to calibrate whether someone or something can be trusted can be mimicked and manipulated.

The Response section outlines ideas to help individuals set smart defaults for how search results are ranked and ordered (Section 3.2.), and ways to foster comparison and competition and trust (Section 3.3.).

2.4. Attention wars

Our attention is a limited resource that is fiercely contested by online players. Of course, there has always been competition for our attention. Indeed our brains have evolved to enable us to filter out less important signals, and switch to the key information around us.[XIII] [52] But the evolution of online markets has kicked the attention wars up a gear. Many of the most common features of online architecture are designed to combine ease of use and attention-capture: likes or retweets capture attention and prompt frequent rechecking, bottomless pages keep users scrolling, and swipes and streaks make browsing feel like a game, removing breaks that might otherwise prompt a natural end to the task and our attention to shift elsewhere. Indeed, one recent estimate suggested that a typical Facebook news feed draws on around 25-times the computational power of that used by IBM's Deep Blue supercomputer to defeat Kasparov – the greatest human chess champion of his generation.[53] That is a lot of computational power to get the right combination of family, community and actual news, alongside advertising, to keep your attention (and the revenue flowing).

There is also a surprisingly plausible case that the attention grabbing, addictive qualities of social media and smart devices may also be spilling over into the offline world, and even denting productivity. This could help explain why productivity in the UK has essentially flat-lined since 2008, despite the digital sector growing rapidly during the same time period.[54] If workers are constantly checking their Facebook accounts and eBay bids, this is presumably displacing other activity, as well as normalising distractions and interruptions, meaning we're worse at focusing when we are actually on task.[55] In common with other addictive behaviour, social media and technology also distracts us even when we aren't directly looking at it. For example, experimental studies have found that leaving your phone in another room increased performance on a working memory task by 11 per cent compared with having it face down on your desk.[56]

At the least, we need to consider whether people have access to countervailing products and services that even up the battle for attention, and allow consumers and workers to take control.

The Response section proposes ideas to give back individual control (Section 3.2.) and leverage data on the side of consumers (Section 3.3.).

[XIII] For example, Colin Cherry researched the 'cocktail-party effect', our ability to focus on the person speaking to us while ignoring other sounds.

2.5. Predicting our preferences

We all have different 'orders' of preferences. You might say that you like Mozart best, but if you always choose to play Adele, then we – and you! – might reasonably infer that your 'first order preference' is for Adele. Perhaps you just say you like Mozart to impress other people, or perhaps you genuinely do prefer Mozart in the abstract and Adele when you are choosing which radio station to stay on.[XIV] [57] This is far from a hypothetical example, and behavioural economics is full of instances of gaps between what we say and what we do. For example, in one experiment participants were asked to choose either a healthy or an unhealthy snack to be delivered the following week and most chose the healthy option; this is different from simply saying you like healthy food since the participants took action to secure the healthier option. However, when given the chance to switch to chocolate on the day of delivery, most changed their mind and picked the less healthy option.[58] In this case, their first order preference is revealed to be chocolate and the 'second order preference' – what the participants chose for their future self – is the healthy one. Second order preferences require us to be reflective and can be easily overturned by factors such as temptation.

How online markets encourage our impulsive preferences

People, and societies, have over time developed strategies to curb our impulsive actions and encourage second order preferences to triumph. One concern about online markets is that they enable far more fluid expressions of first order, impulsive preferences. The internet removes many of the frictions around traditional purchasing, and offers more immediate gratification. Now you can order takeaway, bet on sports, and catch up with the gossip, all without leaving your bed. There are fantastic advantages to such frictionless markets, but also real dangers, particularly to the vulnerable (and most of us feel in that category some of the time).

[XIV] In economics, attempts to model such contradictions have often been framed in terms of two selves with valuation models across horizons, as in Thaler & Shefrin's 'planner vs. doer' model or Fudenberg & Levine's repeated game framework. Another common explanation is time-inconsistent preferences. These different ways of conceptualising contradictions in our behaviour are summarised in recent work on self-control interventions by Duckworth, Milkman, & Laibson.

Box 4: Design of online gambling environments

Gambling operators are well aware of the importance of providing a streamlined online experience. Below are some of the behavioural features that make online gambling so attractive, based on research we conducted for GambleAware:[59]

- **Immersion** – Online gambling provides a highly immersive environment which, coupled with evidence that gambling alters players' state of consciousness (including via dissociation), may lead to more prolonged play.
- **Anchoring** – Some operators match bets of a certain size with a free bet, thereby anchoring the player to a bet amount that is potentially larger than what they would normally spend, while also incentivising them to bet the suggested amount.
- **Defaults** – Operators may automatically suggest stakes or they may impose a minimum amount for depositing funds into users' accounts.
- **Friction** – Some activities on operators' platforms involve little friction (such as placing a bet with only a few clicks) whilst others entail significantly more (such as requiring many steps to close an account).
- **Reminders and personalisation** – Players may receive a multitude of emails, texts, and app notifications prompting them to place bets, and 'dynamic targeting' helps identify inactive players to entice them to play again. Even more severe is the potential for operators to personalise betting limits based on players' skills, for example by restricting winners from placing further bets or by giving rebates to 'big losers' to keep them gambling.[60]
- **Mental accounting** – Players may see funds in a gambling account as separate from 'real money', leading them to bet higher stakes or take more risks, and operators can even allocate funds to different 'wallets' labelled for casino, sports, or poker gambling, which might encourage spending.

Collective preferences

Finally, we make different decisions collectively than we do as individuals. If I were picking a playlist for a solo road trip, I might fill it with tracks that I've been meaning to listen to – perhaps some Mozart or new albums – but when I'm flipping through songs in the car I may reveal through what I listen to that I actually prefer Adele. If I were putting together a playlist for a party my choices would be different again, likely choosing artists that everybody knows and can sing along to like The Beatles. This is known as 'third order' preference: what we would choose on reflection and in discussion with others. Empirically, this is not the same as averaging our individual preferences, and certainly not our first order preferences. This is illustrated in experiments on group influence in juries, where the final view is often considerably different (and sometimes more extreme) than the average starting position.[XV] [61] Deliberative polls and forums also illustrate this phenomenon: a representative group of people often adopt significantly different views after hearing the views of others.[62]

In one sense, this doesn't seem very surprising. But this group shift is perhaps more significant than it seems in at least one important sense: people systematically overestimate the extent to which others hold the same view as themselves (sometimes known as the 'false consensus' effect). In other words, when people declare they prefer Mozart to Adele, they are likely overestimating the extent to which others would agree with this.

This matters because, when we are online, we are often online together. Large social networks, notably Facebook, frequently refer to their 'community' of users. As platforms are increasingly making difficult judgments about the rules of the network, such as what constitutes extremist or unacceptable content, they face a dilemma about whether this should be set by first (I listen longer to Adele), second (I choose Mozart when you ask), or third (we collectively choose The Beatles) order preferences and whether and how to override these if they prove problematic. The challenge is not in understanding these vying preferences. While pre-online businesses had to spend large chunks of their marketing budget on research to tease out even a ghostly hint of people's true first order preferences, now the questions people ask of Google, the website page they spend most time on, and the first version of the message they compose, all reveal it. Second order preferences can also be detected: what we publicly 'like' and share with others, or what we post on our carefully curated Instagram pages. And the gap between individual average behaviour and group outcomes reveals third-order preferences neatly. Now, the difficulty lies in working out how to honour these preferences and safeguard users at the same time.

The Response section proposes ideas to help individuals self-exclude, and other methods to create barriers to impulsive behaviour (Section 3.2.) and democratic mechanisms to facilitate collective deliberation and expression of third order preferences (Section 3.4.).

[XV] There is also evidence that group dynamics can sometimes lead to adopting a minority position.

2.6. More than markets: morals, ethics and social networks

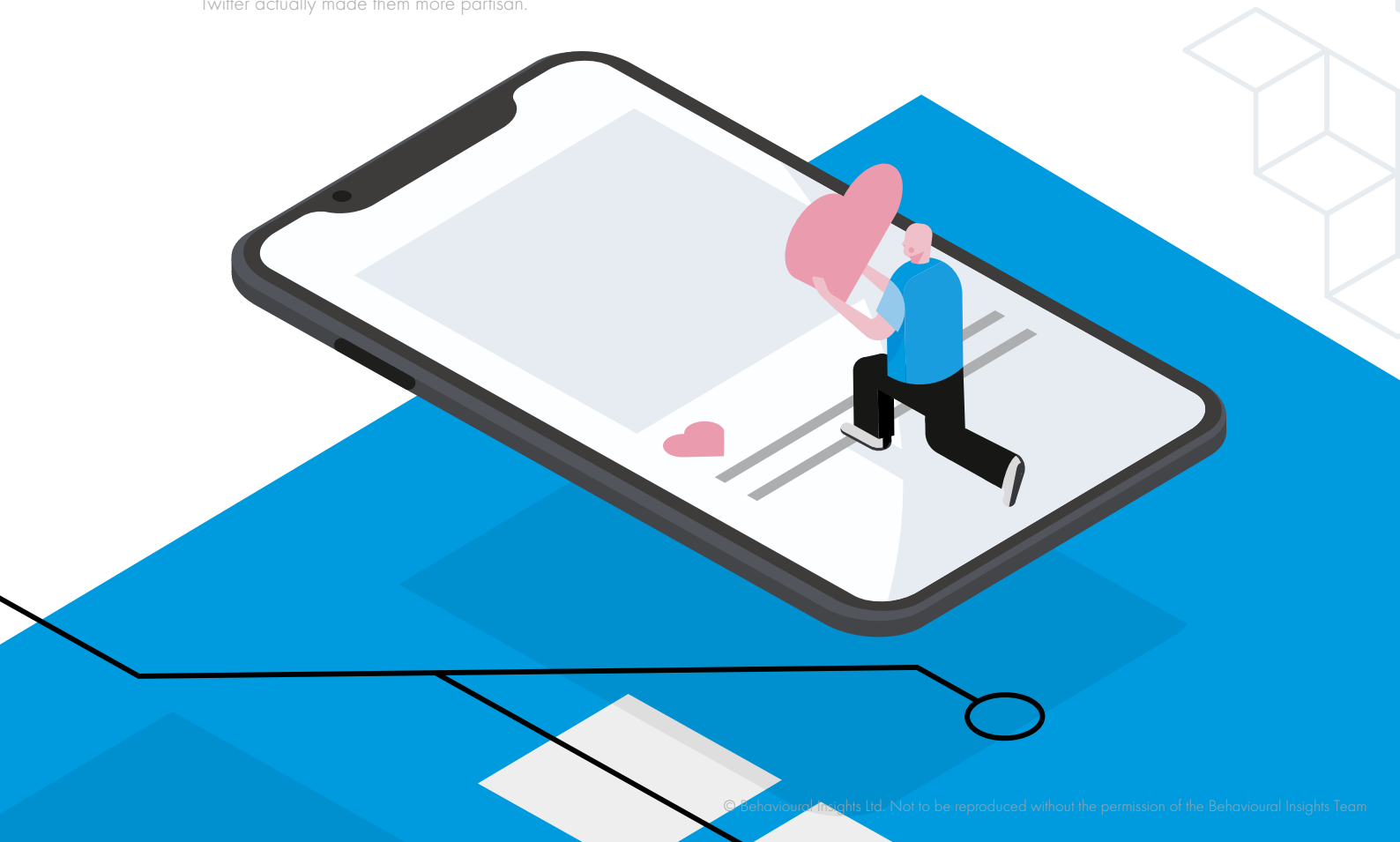
It is undeniable that digital markets are changing the nature of the economy. Yet as Adam Smith exposed, there is more to a market than money. Markets are entwined with 'moral sentiments', social networks and a myriad of habits and customs that shape our lives and societies.

Patterns of association

Social capital refers to our connections to others; the level of trust (and trustworthiness) and the informal rules and common understandings that facilitate communication, support and exchange within those networks. It helps information flow, lowers transaction costs, and drives fairer exchange.[63] At national level, it has been found to be a stronger predictor of growth than human capital (skills) and has also been shown to affect educational outcomes, crime (in both directions: stronger ties can be a double-edged sword), health and even government performance.[64]

Online markets, and social media in particular, have the potential to significantly affect the character and form of our social capital, for example, by creating 'echo chambers' that fool us into thinking our social and political 'bubbles' are representative. The scholar Cass Sunstein writes about a phenomenon that exacerbates these echo chambers: 'asymmetrical updating',[65] a strong tendency to favour evidence that confirms our beliefs and ignore or misread new evidence that doesn't.[66] This tendency has been observed in beliefs about topics such as climate change,[67] the death penalty,[68] affirmative action,[69] and sexism in male-dominated subjects and industries.[70] Social media helps asymmetrical updating thrive since many now use it as a primary news source[71] without noticing how their choices about who to follow have fundamentally altered the balance of information they receive.[72] [73] [74] Website algorithms curate what we see and prompt us where to go next based on our past usage.[XVI] [75] The powerful logic of 'people who liked this also liked that...' takes us deeper into the bubble.

[XVI] Note that despite the obvious pitfalls associated with only seeing content we agree with, exposure to diverse viewpoints is not necessarily a solution against partisanship. In fact, a 2018 study found that showing people content they disagreed with on Twitter actually made them more partisan.



And it goes further. Algorithms can create or destroy social capital more directly. They can draw on personal data to exclude people from opportunities – for example, a babysitting app that uses Facebook news feeds to rate the suitability of babysitters on the basis of whether they've discussed drugs or seem to have a 'bad attitude' online.[76] And they can propagate and potentially extend the exclusion of groups or individuals. The latter is perhaps least appreciated – a scandal waiting to break. Studies have already highlighted examples where online transactions and interactions are leading to new forms of discrimination and disadvantage. While platforms like eBay have persisted with pseudonyms, others have encouraged users to provide information about themselves to build trust and, over time, reputation. For example, Airbnb requires real names and it has been shown that guests with distinctively African-American names are 16 per cent less likely to be accepted compared to identical guests with distinctively White names and male guests in implied same-sex relationships were up to 30 per cent less likely than other couples to have their bookings accepted by Airbnb hosts.[77] [78] The discrimination was not in the form of outright rejections but more insidious, with hosts failing to respond at all.

But the impacts on social capital and cohesion are not all negative. Digital markets have been able to supplement and enhance society's stock of social capital. Platforms such as eBay have made it possible to trust relative strangers; LinkedIn can extend the 'weak ties' that play a key role in getting a job; Facebook has made it easier to stay in touch with old friends; and sites like TripAdvisor create a much wider network than word-of-mouth recommendations.

More than 15 years ago, Bob Putnam – the world's leading expert on social capital – posed a question: 'Would the internet turn into a fancy telephone (i.e. connecting us more), or a fancy TV (i.e. isolating us more)?' We now know that it has both those capabilities. Which is expressed, and for whom, appears very much up for grabs.

Economy of regard

The economic historian Avner Offer coined the phrase 'economy of regard' to capture the workings of a parallel system of exchange, far more ancient than the modern economy, but still very important today.[79] At its heart is reciprocity and social obligation. Your neighbour helps you and, at some later date, you help them. You raise your kids, and they care for you in later life (you hope). Even the so-called 'real economy' is often really driven by its more ancient sibling: our huge spending spree in the run-up to Christmas doesn't make much sense outside the economy of regard.

While it is unclear how the evolution of markets will affect the economy of regard, there are already examples that suggest it will cut both ways. Collaborative consumption and sharing platforms, such as BlaBlaCar, Streetbank, Funding Circle or Nextdoor are helping communities to share assets and skills with a powerful extension of reciprocity. Similarly, the capabilities of online matching markets could radically extend and make volunteering easier. On the flip side, it is also possible that platforms' profit motives could do the reverse, squeezing out volunteering with an enlarged 'gig economy'. Social media can also bring out the negative side of the economy of regard: unwanted or potentially costly social obligation, such as obligations to reciprocate 'likes' on Instagram, or maintain a Snapchat streak.

An uncomfortable glimpse, at least to Western tastes, into how the ability to aggregate many sources of personal data could bump into the economy of regard comes from China. In the most extreme example, young women are able to secure loans by using naked selfies as collateral; the idea being that the reputational cost of a leaked selfie is sufficient motive to guarantee repayment.[80] A more mainstream development is the Chinese 'Social Credit System'. The rating of our fellow citizens on factors such as 'social integrity' strongly echoes the logic of Offer's 'economy of regard'. It might simplify our Christmas shopping list if we can keep track of who forgot our birthday or offended our relatives. But it also reminds us of why many societies broke away from reliance on the economy of regard, and patronage, as primary methods of exchange and support.

Civility and online harassment

A related concern is how the evolution of online environments is affecting the character of our social relationships and communication with others. Most obviously, anonymised communication reduces the reputational costs usually associated with negative behaviour and exchanges.[81] This echoes long-standing findings in the psychological literature that anonymisation loosens social restraint. This can lead to more aggression (such as anonymised subjects administering more pain on others) and dishonesty, though it can also leave more room for self-expression and the ability (for some) to resist problematic social norms.[82] There is also some evidence that the characteristics that people adopt in online avatars 'leak' into their offline behaviour, for good or bad.[83]

A number of social media sites, and online marketplaces, are wrestling with this issue. Hate-speech laws pick up on some of this. But the majority of the time online incivility occupies a grey area between the law and public acceptability: hurtful comments, inaccurate reviews, bullying, and shaming. Not everyone is equally likely to be affected: gender-based online harassment is a widespread example. A survey commissioned by Amnesty International in Australia found that nearly half of women aged 18-24 had experienced online harassment, including sexist abuse, 'trolling' (posting deliberately offensive or provocative content), threats of sexual violence, the posting of intimate pictures online without consent, and 'doxxing' (the sharing of identifiable details).[84] It's clear that a self-regulatory dynamic is not enough.

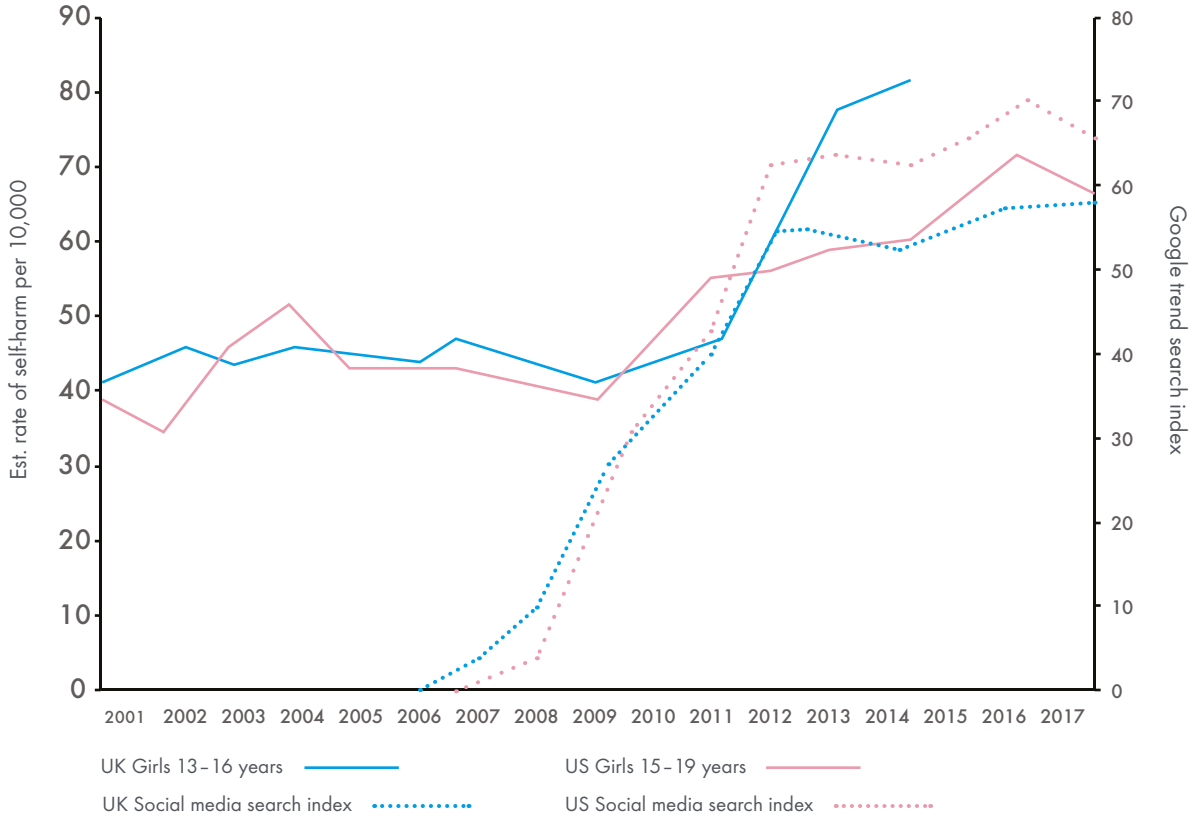
Mental health

One of the knock-on consequences of more time online is widely thought – though not proven – to be on mental health. New evidence suggests two potential factors at play: negative feelings due to hurtful interactions or negative content,[85] and substituting time away from wellbeing-enhancing activities such as offline socialising.[86]

In one recent large-scale field experiment, researchers found that Facebook users who deactivated their accounts for four weeks spent less time online, reduced political polarisation (although at the expense of factual news knowledge) and, most crucially, increased subjective wellbeing.[87] This increase was 'small but significant', in particular on self-reported happiness, life satisfaction, depression, and anxiety, and the size of the effect on overall subjective wellbeing was equivalent to about 25 to 40 percent of the effect of interventions such as therapy.

Of particular concern is the impact of social media on the mental health of younger people. A recent survey of people aged 14-25 conducted by the Royal Society of Public Health comparing the five most popular social media platforms on a range of positive and negative mental health outcomes, found that YouTube was the most positive, for example by raising awareness and self-expression, while Instagram was the most negative – via its impact on body image and fear of missing out.[88] These findings are not trivial; young adults are currently estimated to be spending an average of four hours online everyday[89], and this shift has coincided remarkably with a rise in mental health conditions in teenage girls, including depression, anxiety disorders and self-harm. The graph below shows the sharply rising trend in self-harm among teenage girls in the US and UK over the past 10 years.[90] [91] (Among boys it has remained relatively steady and at a lower base rate.) Alongside this we include a Google trends search index showing the rise in online interest in social media in both countries.[XVII] Other measures of mental health among girls such as rates of emotional disorders in the UK and depressive episodes in the US have also risen sharply over the past 10 years. [92]

Figure 3: Self harm among teenage girls and social media use



With so much cultural and technological change over the past decade it is hard to isolate causal pathways for these worrying trends. Each individual case of mental illness is complex. However, there's increasing evidence that social media is at the very least associated with these issues. One study analysed survey data from 10,904 14-year-olds in the UK to explore the relationship between social media use and depressive symptoms.[93] It found, in particular for girls, that greater social media use was associated with online harassment, poor sleep, low self-esteem and poor body image; these were in turn linked to depressive symptoms. At a minimum, this correlation serves to show that social media interacts with feelings and behaviours linked to mental health issues.

The Response section proposes ideas to help individuals self-exclude (Section 3.2.) and to use data to signpost vulnerable people to support (Section 3.3.) and discusses the 'non-market' design features that can foster social capital, civility and informed deliberation and debate (Section 3.4.).

[xvii] Google trends (https://trends.google.com/trends/) search data for Facebook, Instagram, Twitter and Snapchat in the UK and US is aggregated all platforms equally weighted) to provide an index of interest in both countries. This is a proxy for use of these platforms.

2.7. Emerging problems

We have already discussed some of the ways in which the evolution of markets, particularly towards machine learning and mass experimentation, is giving rise to significant concerns (as well as benefits), but there are more problems on the horizon. These include the impact of fake news and 'deep fakes' on society and politics; increasingly personalised pricing and price discrimination; the role of algorithmic bias in discrimination and injustice; and the creation of new types of monopolies based on the amount of data available to industry players.

Fake news and deep fakes

Another major consequence of patterns of association online (discussed in Section 2.6.) is their role in making people susceptible to misinformation and disinformation: commonly called 'fake news'.[94] Fake news is increasingly exposing people to false or polarised information that confirms their pre-existing beliefs. Beyond such asymmetrical updating, the effect of fake news is arguably enhanced by our tendency to avoid the cognitive effort required to discern between fake and real news. In fact, a 2018 study found that susceptibility to partisan fake news was better explained by lack of reasoning (low analytical thinking) than by a tendency to favour evidence that confirms pre-existing beliefs.[95]

The challenge of fake news is unlikely to resolve itself without serious intervention, as instances of fake news are only getting more difficult to spot. Firstly, while outrageously fake news receive the lion's share of media attention, real news presented through polarised perspectives is likely to be just as important, if not more so. For example, research conducted by New York University's Social Media and Political Participation Lab found that during the US 2016 presidential election, interference from Russia mainly relied on local news, and included sharing local news stories but using a highly partisan frame.[96] Secondly, current concerns about fake news may soon look quaint in the face of a slew of new 'deep fakes' brought about by software innovation such as programs that swap the face or voice of one person for another. Anyone with a PC could now have the capability to put words into the mouths, or actions into the bodies, of anyone with a reasonable digital footprint. Not surprisingly, an early application has been to create sexual content, with well-known actors having their faces imposed on pornographic videos. Insiders in the AI and digital industries warn that current software, already fairly convincing, will rapidly continue to improve.

Personalised pricing and price discrimination

Forms of subtle and overt price discrimination have long been part of market dynamics. Consumers who fail to switch providers or deals at the end of a trial period can be charged more; 10 to 50 per cent in the case of fixed-term contracts for insurance, energy and telecoms bills. As this 'loyalty penalty' affects the prices of essential services, it can also exacerbate inequality.[97] [98] [99] The CMA has published recommendations to address this, following a super-complaint from Citizens Advice.[100] [101] Its package of reforms includes measures such as moving bundled mobile customers to fairer tariffs and investigating insurance pricing practices.

While this is a worrying phenomenon offline, the additional information collected in online environments allows for increasingly precise forms of price discrimination. A summary of recent evidence by the OECD shows that retailers are using artificial intelligence (AI) to personalise customers experiences, including price discriminating.[102]

Box 5: Elements of personalised pricing

Personalised pricing is thought to include differentiation by markers such as:

- **Search history** – including cookies and browsing history, but also how you accessed the website (e.g. via a price comparison website[103]);
- **Past behaviour** – for example how often you accept surge pricing or whether you mainly purchase items on sale;
- **Device** – the Wall Street Journal reported in 2012 that a travel agency was presenting Mac users higher prices than PC users,[104] while an experimental study found different deals being offered to desktop and mobile customers;[105]
- **Location** – the same Wall Street Journal investigation found evidence of geo-pricing, which a Harvard Business Review article warns could disproportionately affect racial minorities;[106]
- **And possibly more** – such as prices differentiated on the basis of friendship networks or moods.

Data-led price discrimination raises many practical and ethical questions. For example, should health insurers be allowed to use customers' social media interactions and mail-order purchases to predict their healthcare costs?[107] Whether intentional or machine learned, we can expect to see ever more elaborate versions of precision pricing and to face tough questions about where to draw the line

Biased algorithms and AI tools

We have discussed the ways in which online players harness our behavioural biases but have not yet touched on algorithmic bias. Machine learning and AI rely on the data they receive to draw conclusions and write new rules for handling complex decisions. Doing so, they may reflect and enhance entrenched bias in pre-existing systems. 'Garbage in, garbage out', as dubbed by the World Economic Forum.[108] There are many examples of how this plays out in practice. The COMPAS algorithm, a tool created to help predict recidivism rates and assist sentencing in the US, was found to be biased against black defendants, predicting higher rates of recidivism than it would for their white counterparts.[109] [110] Amazon designed, then recalled, a recruitment algorithm which was found to be biased against women because it saw many men in the industry and inferred that 'maleness' was a desirable trait.[111] More generally, a Science study which replicated a spectrum of human biases (as exposed by the Implicit Association Test and other well-known psychological studies) showed that the human-like biases exhibited by many algorithms (e.g. reflecting gender stereotypes) result from the application of machine learning to everyday language, which is full of these stereotypes.[112]

Algorithms may reflect the biases and backgrounds of their creators. Creators and coders can inadvertently create algorithmic biases when they do not adequately represent society. In the UK, for example, 4.5 per cent of male but only 0.4 per cent of female A-level students studied computer science.[113] The top US tech companies are largely dominated by men, particularly in technical roles,[114]

and the AI industry in general suffers from a wide gender gap, with 78 per cent of the roles in the industry being filled by men.[115] The evidence is still emerging on the causal links between creators and algorithmic bias but it is clear that the early analytical choices shaping these tools matter, and therefore we should pay close attention to who is making those choices.[XVIII] For example, new research shows that the cameras and sensors on self-driving cars are less able to detect pedestrians with darker skin tones, Amazon's face recognition software is much worse at determining the gender of people with darker skin, and African American Vernacular English appears to be rated as rude or toxic on sites such as Gobo.[116]

However, we must bear in mind that the alternative to algorithms is human decision-making, which is likely on average to be more biased.[117] Algorithms can directly reduce bias in systems and a number of initiatives have focused on hunting and fixing bad algorithms.[118] For example Harvard Professor and leading academic on behavioural economics and gender equality, Iris Bohnet, runs algorithms through job ads to find and eliminate biased language.[119] Algorithms are easier to re-programme than people.

While there are still limits to how much we can remove bias from algorithms, AI offers opportunities to address inequalities – just as much as it threatens to further alienate disadvantaged groups – but only if government and regulators prove agile enough to keep up.

New monopolies

Markets are evolving to create new types of monopolies based on large user bases and streams of user data that are a powerful source of 'machine-based' innovation.[120] It is important to note that such monopolies can provide real efficiency gains and consumer benefits: interoperable calendars, contacts, email, maps and phones, for example. However, just as we've seen with previous monopolies, these advantages inevitably create opportunities for rent-seeking, suppressing competition and reducing quality.

It is worth noting that the dominant tech companies do not appear to have over-exploited the power they wield. They may have taken advantage of their position to clip out business from smaller rivals, but it is not obvious that they are secretly using their privileged access to all our data to predict (or precipitate) stock market movements to their advantage. Similarly, despite recent concerns about the use of social media for political purposes, it is not thought at this point that they have actively sought to intervene in the political systems of nations or communities for organisational gain.[XIX 121]

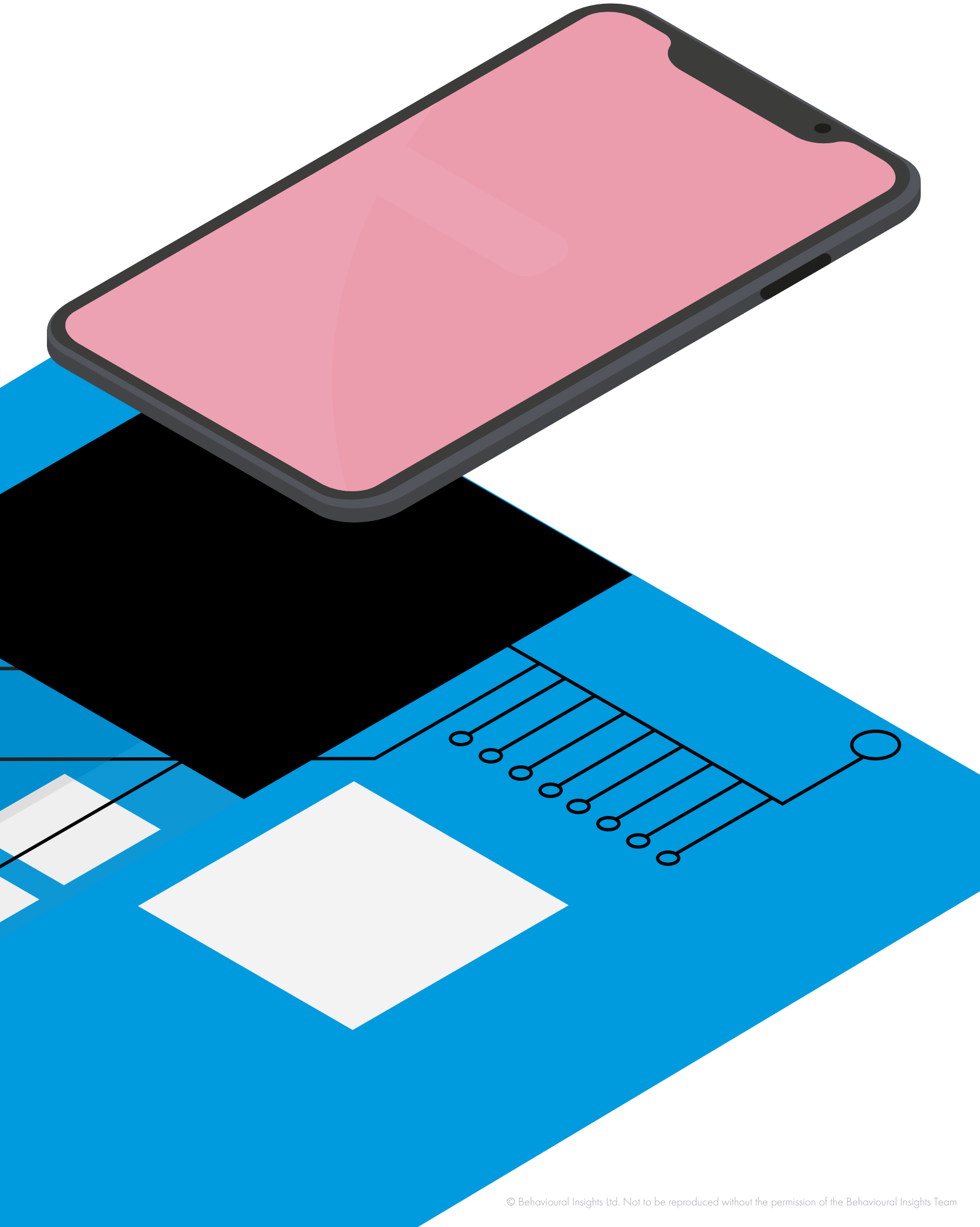
The Response section proposes ideas to leverage data and AI on the side of consumers to tackle biased algorithms, personalised pricing and fake news (Section 3.3.). It also recommends ways to sharpen competition through greater transparency and using data to support new market entrants to tackle new monopolies (Section 3.3.).

In summary, it is clear that digital markets are rapidly evolving, bringing many benefits to consumers and the economy, but also a range of economic and social challenges.

Regulators and politicians are starting to attend to the issues. Yet it doesn't seem that we have really begun to address how far our existing assumptions about how economies and societies work need to change.

[XVIII] For a fantastic and detailed discussion of how our future is being shaped by the world view of the narrow group of people designing AI, see: Broad, E. (2018) Made By Humans: The AI Condition. Melbourne, Australia: (Melbourne University Press).

[XIX] Note that Facebook has sought to support increased voter turnout in general elections, for example in Iceland via a 'voter button' on users' news feeds.



3 / Response: what can, and should, we do?

It's extremely hard to buy a single piece of fruit for £1,000. You won't come across apples and oranges for sale at this price. There's no law against it: shops just know that it would lose them customers and trust: in other words, the market works.

There are, however, other important ways in which the market does not function quite as it should. Customers choose their banks, for example, based on relatively peripheral factors, such as a small gift, cash back offer, or because their cousin uses it. They then stick with that bank for years – often a lifetime – even if others offer better rates and conditions.

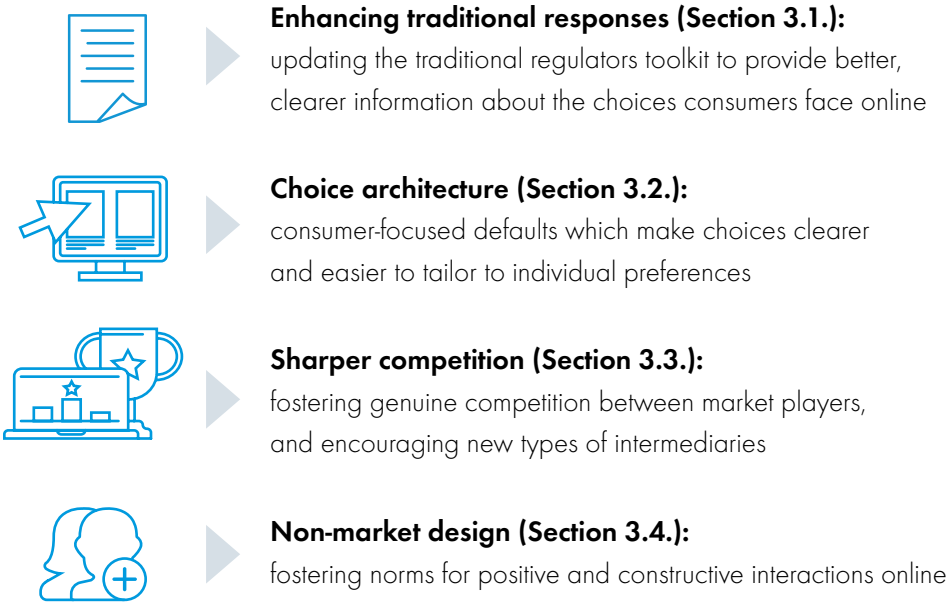
Economists David Laibson and Xavier Gabaix argue that the market sorts out fruit but not banking because of the balance between our behavioural biases, price differentiation, and the incentives between more and less sophisticated consumers.[122]

Consider the naïve consumer. They could get taken advantage of both by expensive fruit shops and bad value banks. But what about the sophisticated consumer? They will definitely avoid the expensive fruit, hence putting market pressure on shops to reduce prices, benefiting naïve consumers too. But the banks are offering differentiated pricing: charging customers a lot for unauthorised overdrafts and making a great margin on customers who don't shop around for financial products such as currency exchange. This enables banks to offer excellent value on headline services such as credit cards with no annual fees. This is a great deal for sophisticated consumer, since the naïve customers are in effect subsidising them.

So we can see why it's hard for a good, long-term value bank – offering a great product at the same price for all – to prosper in such a market, let alone 'clean up'. Even by calling out the practices of the poor value banks, the better value bank is simply alerting sophisticated customers to the subsidies available elsewhere. Both naïve and sophisticated customers are going to choose the poor value banks, though for different reasons. The unfortunate equilibrium is that the poor value banks aren't called out.

The situation we face in digital markets is similar. The collision of behavioural biases, market dynamics and differentiation between consumers have led to the many challenges that we saw earlier in this paper. An effective response to these challenges must make good use of the entire regulatory toolkit and also build on it to introduce new ways of shaping markets. And not just government and regulators: industry has a key role to play, not least to build and maintain the trust of users. We now consider some of the behavioural and data-driven tools and responses that might address the challenges outlined in Section 2.

Figure 5: An emerging framework for behavioural interventions in online markets



3.1. Enhancing traditional responses

At a simplified level, the regulators' traditional toolkit consists of 'sticks' to punish non-compliance and poor behaviour, 'carrots' to encourage companies to surpass minimum standards, and education and disclosures to better inform consumers. This section focuses on disclosures, education and exhortation, and sets out how a behavioural approach could enhance these tools.

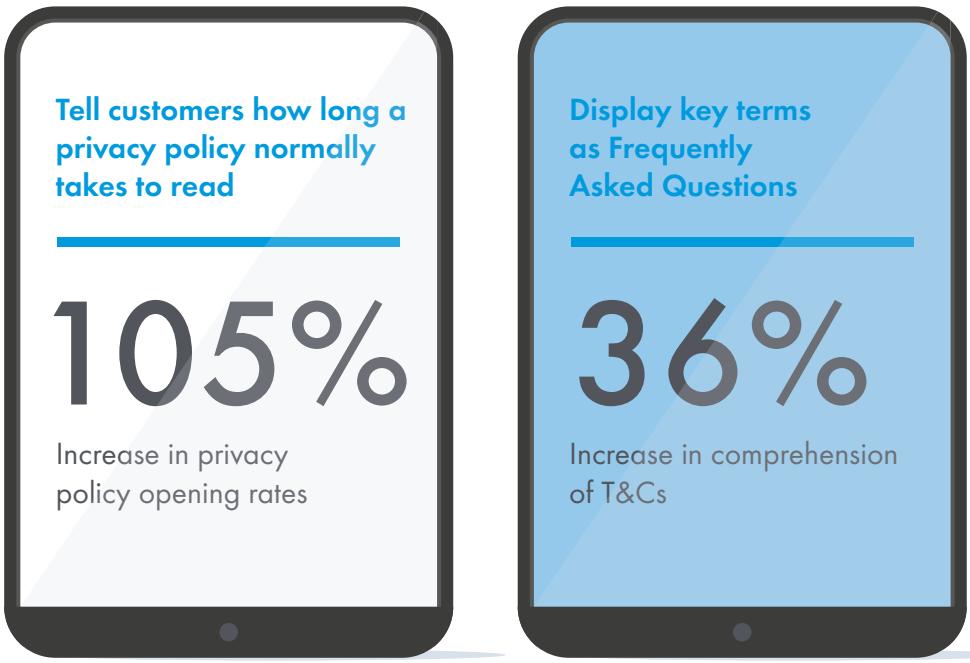
Smarter disclosures

The first line response of many regulators is to provide improved information to consumers to correct the imbalance of information between suppliers and consumers. Many behavioural studies have focused on how to improve this information by providing 'smarter' disclosures informed by behavioural science. These are often based on a simple, personalised and salient comparison. Examples include nutrient information on food; standardised Annual Percentage Rates (APR) for credit; simple information about the choices you can make with your pension The first line response of many regulators is to provide improved information to consumers to correct the imbalance of information between suppliers and consumers. Many behavioural studies have focused on how to improve this information by providing 'smarter' disclosures informed by behavioural science. These are often based on a simple, personalised and salient comparison. Examples include nutrient information on food;[123] standardised Annual Percentage Rates (APR) for credit;[124] simple information about the choices you can make with your pension pot;[125] and fuel economy measures for cars.[126] Many of the most powerful examples are from industries that are highly regulated, where the regulator has the power to compel firms to provide information in a particular format. So how do smarter disclosures translate in online markets?

In online markets, there are many opportunities to improve the information being provided to consumers – from how their data is being collected, to whether algorithms are being used to make decisions, to a website's terms and conditions (T&Cs) – and to vary when and how this information is presented. Where there is no clear regulatory power to compel companies to provide information in a particular format, these 'smarter disclosures' can be encouraged through best practice guides, codes of conduct and reporting requirements.

BIT recently concluded a series of online experiments testing ways of applying behavioural science to improve consumer comprehension of (and engagement with) online T&Cs and privacy policies. These ranged from the relatively simple, such as those typical of an online retail purchase, to more complicated agreements to engage on social media or book a room via a peer-to-peer platform. We designed and tested several techniques that increased comprehension and engagement. Telling customers how long a privacy policy normally takes to read increased privacy policy opening rates by 105 per cent. Using a question-and-answer format to present what you consider to be key terms and summarising key terms and illustrating them with explanatory icons both increased understanding of T&Cs by more than 30 per cent.

Figure 5: Improving engagement with and comprehension of terms and conditions and data privacy notices



There were also some techniques that, surprisingly, were not effective. For example, many companies present terms and privacy policies in 'layers' where customers initially only see short summaries. When they click the summary, the section expands to show more. This approach is also recommended by the Information Commissioner.[127] We tested this against a full set of T&Cs and found that customers who saw the 'layers' had a poorer understanding, possibly because of the friction involved in clicking to reveal more information.

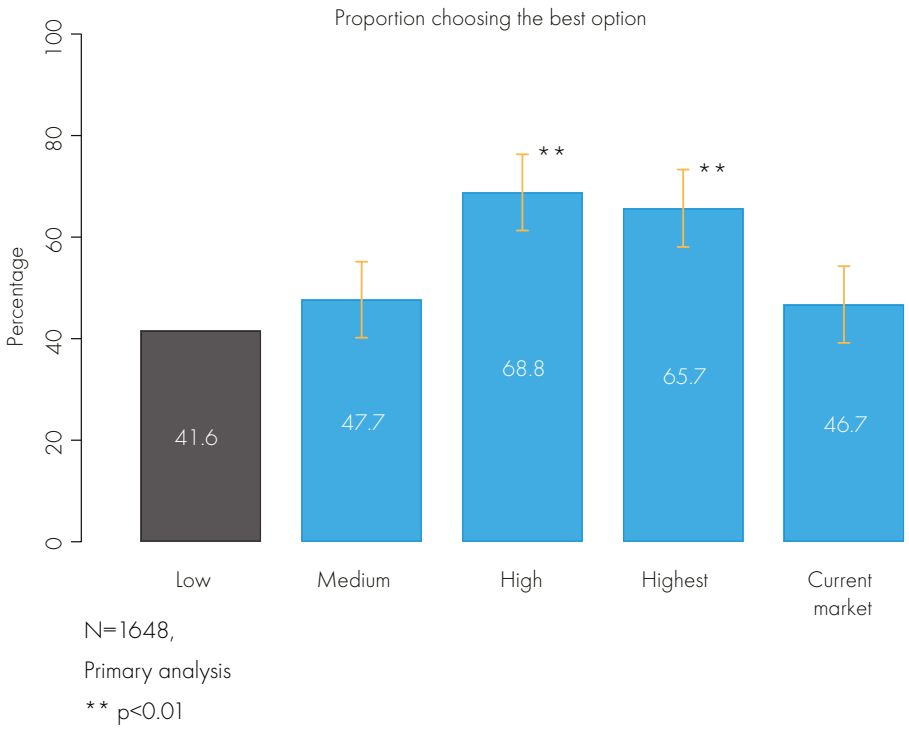
The Department for Business, Energy & Industrial Strategy (BEIS) is publishing the results of these experiments as a best practice guide for businesses.

However, smarter disclosure only gets us so far. Despite the large impact of the experiments outlined above, the baseline levels of understanding and engagement remained low (for example, our participants correctly answered on average only 45 per cent of the comprehension questions we asked them) and we suspect would be even lower if tested in the course of regular web usage alongside the many other tasks competing for our attention. Further, there are other factors that limit the effectiveness of disclosures, especially where industry goes to lengths to conceal information from consumers, where information is provided unevenly across products or providers, and where the underlying decision remains complex.

There are many ploys that market players can engage in to dilute or confuse consumers. For example, when nutrition information on products is in a green box, consumers tend to infer that the product is healthy, even if it is not. More troubling, a recent CMA investigation found that online dating sites were misrepresenting the number of active users, inflating figures by including past and present members across multiple sites.[128] Regulators should be aware of these techniques and be willing to call out poor practice.

Another more subtle effect is seen when information is provided unevenly across products or providers. A powerful illustration of this effect is from an online trial run by BIT, where consumers were asked to choose between foreign exchange options.[129] We found that adding additional, clearer information, for example including the total fee in pounds, greatly increased consumers ability to identify the best value deal. However, if only some of the providers offered this information, most of the benefits fell away (the 'current market' in the chart below). [XX]

Figure 6: The impact of transparency on consumers' ability to identify the best forex option



This is a powerful illustration of a market problem described by David Laibson and Xavier Gabaix: when faced with clear pricing on some products, and less clear or concealed pricing on others, consumers often think that the latter options are the best. The implication is not that regulators should give up on the provision of information, but that they need to attend extremely closely to how that information is contextualised, and the actual impact in a mixed market context.

Another potential pitfall is that if the underlying decision remains complex then simplifying and improving the information provided to consumers may not really help. Navigating data privacy policies is good example. As highlighted in Section 2.2., we are confronted with a bewildering range of legal agreements about how firms will collect, store, use and share our personal data. Improving these data privacy disclosures may be of some benefit, but a more promising solution is to foster the growth of intermediaries that will sort through the vast amount of information on your behalf, highlighting terms that are out of step with your pre-specified preferences (see Section 3.3.)

Recommendations

- Government, regulators and consumer groups should publish evidence-based best practice guides for businesses on how to improve online disclosures like T&Cs and data privacy policies.
- Government and regulators should set the acceptable average level of understanding that can constitute informed consent, and require businesses to conduct ongoing comprehension tests to assess and improve the information they are providing to users.
- Further research should be conducted about the most effective way to disclose how and when algorithms are being used in decisions about consumers (for example, in online loan applications), and the consequences or intentions underlying this.

See graph

[xx]The 'Low' transparency condition showed the exchange rate, flat fee and commission. 'Medium' added information about the interbank exchange rate, so that participants could see how much it differed from the exchange rate they had been offered. 'High' was the same as the 'Medium' condition, but explained how much the difference between the interbank and exchange rates meant in terms of pounds for every £1000 exchanged (e.g. 'this means you lose £20.10 for every £1000 you transfer). 'Highest' included all the information in 'High', but explained what the overall transaction would cost them in cash terms (e.g. 'You will be charged £X.XX'). The 'Current market' scenario provided participants with a mix of low and high transparency suppliers.

Education

If information by itself is not a reliable lever, what about education? It clearly has a role. Teaching people to recognise potentially dubious claims and offers, and think about whose interests are being served, can make them more savvy consumers, more resistant to persuasive and potentially deceptive ploys. For example, the 'Truth' campaign against tobacco went beyond a conventional information campaign, and encouraged young people to look behind the glossy advertising and think about the incentives of the industry.[130]

One of the most promising areas here is the development and promotion of canny 'rules of thumb' to guide consumer behaviour. This is different from traditional educational approaches that seek to build core skills to analyse and compare from first principles. For example, a traditional approach to address 'financial literacy' would involve helping people to understand compound interest and how to do basic investment appraisal. Most behavioural scientists are relatively sceptical about the efficacy of such an approach. We would instead advocate a set of simple rules of thumb akin to Harold Pollack's index card of personal finance tips like 'pay your credit card balance in full every month' and 'pay attention to fees'.[131]

In contrast to traditional education programmes, rules of thumb can be used much more widely, especially if they are developed and disseminated in conjunction with regulators or consumer bodies like Citizens Advice and draw on data on common issues that consumers face. For example, in the recent market study on Digital Comparison Tools the CMA developed simple rules of thumb for consumers searching for deals on comparison sites:

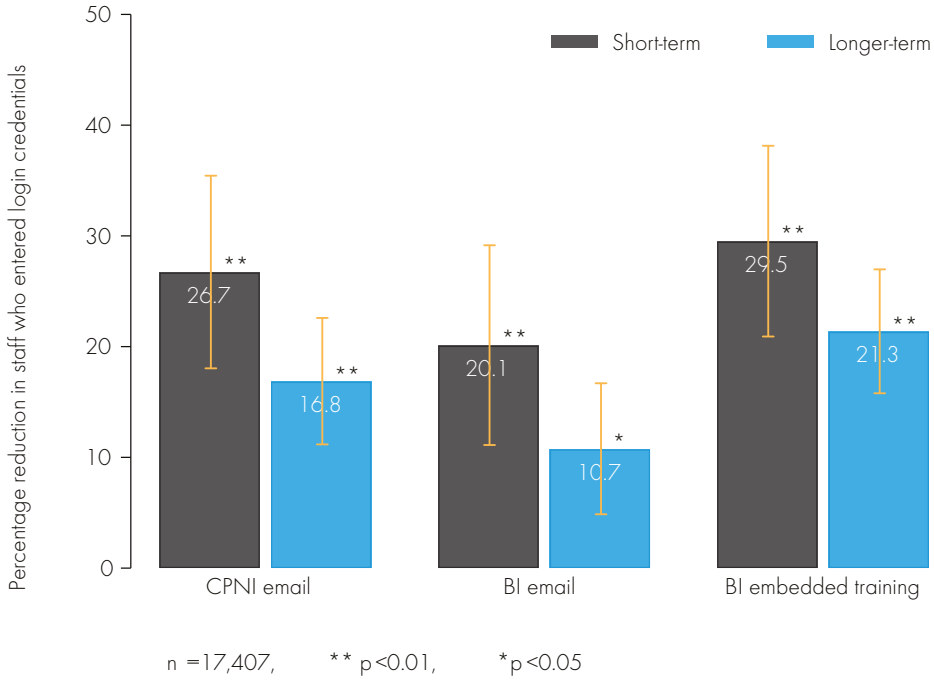
- Comparison sites can save you time and money.
- Choose carefully between comparison sites, like you would any retailer.
- Not all sites are the same, so try more than one if you can.
- Check how the site has ordered its results.[132]

There are also more specific applications of rules of thumb to guide consumers and build resilience against the challenges faced in online environments. For example, BIT recently experimented with education methods, including rules of thumb, to reduce the susceptibility of Metropolitan Police Officers to phishing emails that leave the Met vulnerable to cyber attacks[133]. The most effective preventative training was using simple rules of thumb for how officers could avoid phishing attacks[XXI]. First, a mock phishing email was sent. If officers clicked on the link in the mock phishing email and submitted their login credentials, they would then be presented with the rules of thumb, creating a 'teachable moment'. This resulted in an almost 30% reduction in staff who entered login credentials. The training was still effective three months later.

[XXI]We tested three types of preventative training. The first drew on advice from the Centre for the Protection of National Infrastructure ('CPNI email'). The second was based on BIT's own research and used simple rules of thumb for how officers could avoid phishing attacks ('BI email'). The third used the same content as the BI email, but the content was delivered following a mock (i.e. non-malicious) phishing email. If officers clicked on the link in the mock phishing email and submitted their login credentials on a mock landing page, they would then be presented with the BI anti-phishing training ('BI embedded training'), creating a 'teachable moment'."

[XXII]This experiment involved 120 schools and 11,000 children aged 10-12. Half the schools received textbooks, exercise books, and a teachers' guide, and delivered nine 80 minute lessons over the year. Teachers were trained in a 2 day introductory workshop. As part of the same project, 675 parents listened to either a public service announcement about health issues or a podcast from Informed Health Choices. Those listening to the podcast were twice as likely to pass a subsequent test assessing health claims, from 38 per cent to 71 per cent.

Figure 7: Reduction in staff entering login credentials



n =17,407, ** p<0.01, *p<0.05

Note: graph shows per cent reduction (higher = better).

In Section 2.7. we reviewed the emerging issues of misinformation, disinformation and deep fakes and the heightened challenge of discerning how much of the information we see is credible, what is rubbish, and sifting through the many variations in between. Using education-based interventions that focus on developing media literacy skills can help to address this issue by building resilience in users. For example, an experiment which asked young adults to judge the truth of (simulated) online posts found that political knowledge did not improve judgements of accuracy, but media literacy education did.[134] A novel intervention to tackle this and build resilience against misinformation is to teach children, teachers and parents to be better 'bullshit detectors'. This has been tested in an Informed Health Choices programme in primary schools across Uganda where it increased the proportion of students passing a multiple-choice test assessing their ability to understand health claims from a 27 per cent pass rate to 69 per cent.[XXII 135]

We also strongly support the development of supportive apps and practices that can provide 'training wheels' to young people first using social media and interacting online, such as the 'Own It' app being developed by the BBC.[136] These should warn about both risky online behaviour and actions that may be hurtful or harmful to others.

Recommendations

- Government and regulators should, in partnership with consumer bodies like Citizens Advice, develop and disseminate useful rules of thumb to help consumers navigate common challenges faced online.
- Government should fund research to design and test novel approaches to building consumers' resilience against challenges like disinformation and online fraud; and be willing to scale successful approaches.
- Industry and governments should develop supportive apps (or additions to the Personal, Social, Health and Economic (PSHE) curriculum) that can provide 'training wheels' to young people first using social media and interacting online.

Exhortation

A lot of policy, and certainly political activity, concerns urging citizens or firms to do something different, or differently. This is certainly true of online markets, not least where politicians and regulators are not yet confident about how best to intervene, or about their powers to do so. The UK's Chief Medical Officer, for example, recently urged parents to 'leave phones outside the bedroom when it's bedtime', and enjoy 'screen-free meal times' with their families.[137] Tech giants are urged to do more to take down content that is inappropriate or harmful, like images of self-harm or violence. Consumers are urged to be more engaged and informed about the choices they make online.

It is appropriate to allow such exhortation to run, and not to rush too readily into ill-judged or inflexible legislation. Today we smile at the thought that Victorians genuinely worried that novels were corrupting the young, and perhaps tomorrow people will smile at some of today's concerns. Exhortation is not a bad place to start where we have genuine concerns but also some doubts about direct interventions, and want to leave room for people to make up their own minds, or companies to identify the most practical way to respond.

Yet in some areas we should be ready to move beyond general exhortation, and be more targeted in two ways: first, to focus on urging companies rather than consumers to change their behaviour, and second, to look to change the behaviour of specific companies as an example to industry.

Consumers have often been exhorted to be more 'engaged' in a range of markets and, in particular, to switch more often. In energy, insurance and banking markets, economists and some regulators have bemoaned the passivity of consumers. If only more would be rational and switch, then they would get better deals and the market would work (as we saw earlier, talking about fruit). The implication is that there is nothing wrong with our model, but instead the market failures are the 'fault' of consumers. We now know enough to be wary of this interpretation. It's not just that consumers may have better things to do with their time than endlessly shop around. It's also that markets are evolving to add subtle frictions, distractions and confusions that make it hard for consumers to switch, and very easy to stick.

The recent UK Consumer Green Paper signals a significant shift away from placing the burden on consumers to act in their own best interests and make different choices.[138] While it is still important to encourage consumers to make better choices, more of our focus needs to be on nudging and prodding companies to change their behaviour to make it easier for consumers to identify better deals and to switch. More fundamentally, government and regulators need to be guiding the evolution of markets, and sometimes directly intervening, to make sure that 'good' companies and practices are the ones that are winning market share, and poor companies and practices are squeezed out. Following the suicide of a UK teenager, for example, the UK Government urged Facebook and Instagram to take more decisive action to prevent children from accessing images and videos relating to self-harm. The UK Secretary of State for Health and Social Care, Matt Hancock, indicated that if voluntary action wasn't taken then the government would consider compelling a stronger response, saying: "If we think they need to do things that they are refusing to do, then we can and we must legislate".[139]

This targeted exhortation taps into the brand and reputation of the company in question (discussed further in Section 3.3.). Instagram responded to the above pressure, stating: Over the past month we have seen that we are not where we need to be on self-harm and suicide, and that we need to do more to keep the most vulnerable people who use Instagram safe.[140] The first line of this statement – 'over the past month' – is telling: it indicates action was in direct response to pressure from UK policymakers, and by extension the media and public. The company then made commitments to ban graphic images of self-harm, limit non-graphic self-harm content in searches, hashtags and the explore tab, and provide signposts to sources of support for people.[141]

Recommendations

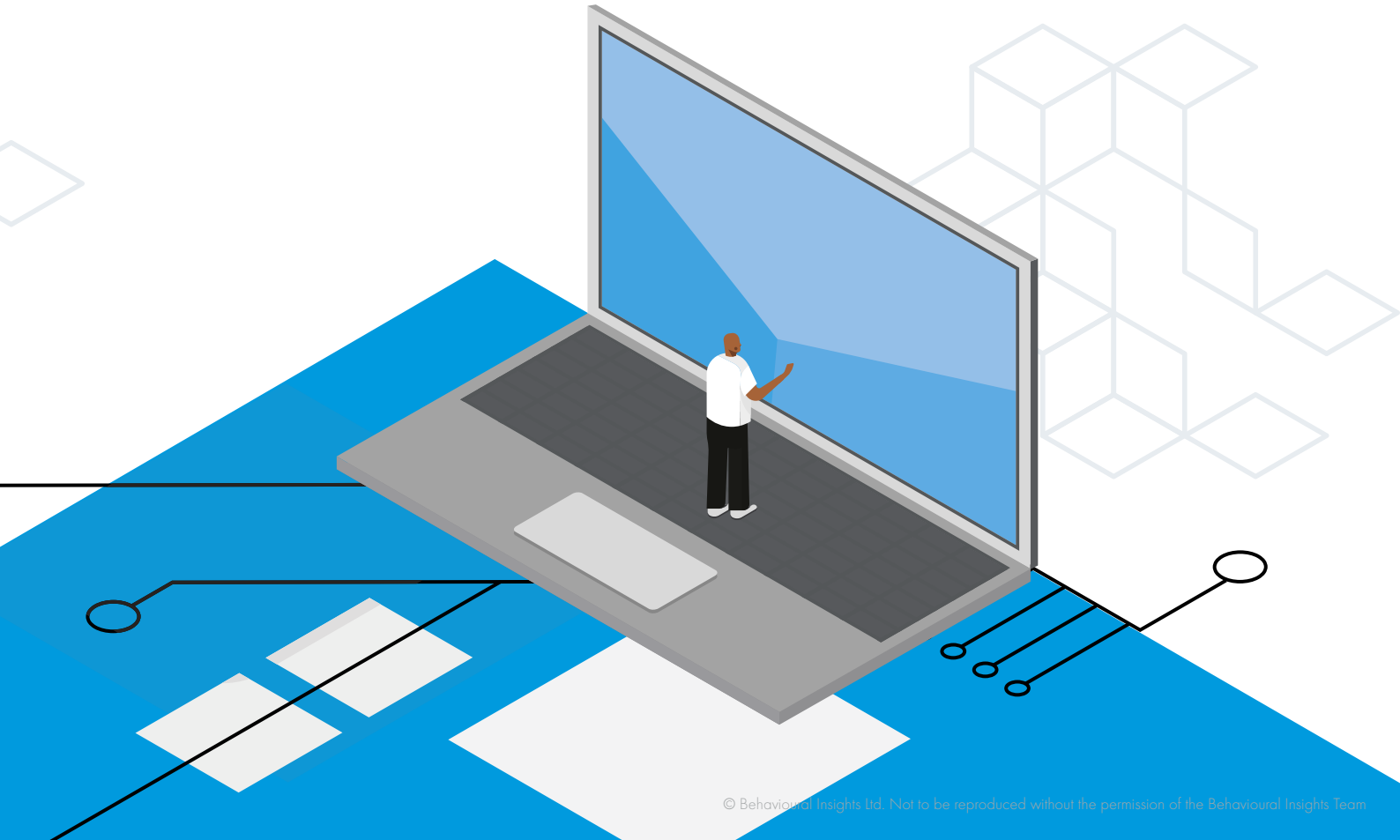
- Government and regulators should publicly urge companies to change and improve their policies and practices, and signal willingness to intervene more strongly if change isn't satisfactory.
- To help focus action, the research community should explore whether voiced concerns are correct, including through data-led investigation and exploring causal pathways.
- Consumer groups should establish an annual, consumer-led 'Sludge' award. This would expose poor behaviour and encourage consumers to recognise and call out sludge, giving industry more incentive to change.

3.2. Choice architecture

Whether regulators (and companies) like it or not, online markets are drawing them into the business of choice architecture. How overt should a choice be? When should it be made? Is it a once-off or, if not, when and how should it be re-prompted? To what extent should choices be different, or presented differently, for different groups?

Giving back individual control

One thing regulators and innovative companies can seek to do is to put as much control as possible back into the hands of the individual. This is not the same as just encouraging people to switch product or service, not least since in many contexts there isn't a good alternative. Rather it is about allowing consumers to be able to express their preferences and to modify their online experiences in line with those preferences.



User settings and controls

Adding in specific and obvious user controls, particularly on quasi-monopolistic platforms, could be a tool for regulators – and a market advantage for challengers. Just as now a user can easily adjust the font size on a screen to suit their eyesight, they might also be able to – or be prompted to – adjust settings that:

- **Control the level, type or source of advertising** (or other) material they are seeing (noting the potential commercial implications of this);
- **Control how their data is collected and shared;** or
- **Choose or weight the criteria by which things are ranked and presented,** for example by independent retailer, female commentator, lowest total price, most reliable news source or healthiest option.

Currently, it's hard to change privacy settings and other controls buried way down in settings menus and a key challenge is how to prompt users to engage with and adjust these settings and controls. Google Chrome content settings, for example, are hidden deep down in the 'advanced' settings, where many users are unlikely to find them. Sometimes this is sludge; done to add deliberate friction, to make it harder for users to restrict commercial re-use of their data or even leave the platform. But sometimes it's done for good reason – to avoid filling screens with rarely used choices or because a choice that might seem important to one user may not be so for another.

There may be a role for regulators here, in encouraging or requiring companies to make user settings more prominent, either during ongoing use or in the set-up phases. However, this activity is more likely to be driven by fostering the development of new intermediaries (discussed in more detail in Section 3.3.). For example, the MIT Media Lab's Center for Civic Media is building a tool called Gobo to aggregate content from large platforms and then enable users to customise that content.[XXIII] It gives users a series of sliders to curate what news they see and what is hidden. For example, you can express a preference to see wider news sources or more female commentators. Gobo and other intermediaries could offer more control and customisation to users without regulators requiring large platforms to change their own user controls.

Self-exclusion

For vulnerable people and certain categories of behaviour, just adjusting user settings won't be enough. Industry and intermediaries need to design tools that allow people to self-exclude or commit themselves to future actions that give them the power to protect themselves from harmful behaviour.

Research conducted by the Money and Mental Health Policy Institute (MMHPI) found that nearly half of the 12 million UK consumers with mental health problems stated that they would like to set controls like spending limits in online shops.[142] People with addictions to online gambling or pornography could benefit from similar options to self-exclude or to limit their access to particular websites or certain types of payments.

[XXIII] https://gobo.social/

Monzo, the UK challenger bank, has led the way for the financial sector by introducing a gambling block.[XXIV] Customers who are problem gamblers can block gambling transactions on their Monzo card. The individual's freedom to choose to gamble is preserved, but there is an appropriate amount of 'positive friction' involved if they want to reverse the block: a 48-hour delay and requirement to speak with customer service.[143] Other retail banks – Starling, Barclays, Santander, Lloyds and the Royal Bank of Scotland – have followed suit.[144]

There is huge potential for industry and intermediaries to build targeted self-exclusion tools. A few promising areas:

- **Category-based blocks.** Akin to the gambling block, where people choose to block certain sites, payments or behaviours all of the time;
- **Time-based blocks.** For example, plug-ins like the MMHPI's Shopper Stopper that block spending in certain online stores, or at certain times of the day (like after 10pm when people are more likely to spend or behave compulsively); and
- **Blocks that leverage social networks.** For any form of self-exclusion, the delay on lifting the block could be combined with a requirement to speak with a trusted friend or support person to discuss whether you really want the block lifted.

While industry is moving ahead on some of these areas, there is a clear role for government and consumer bodies. They need first to identify the areas where self-exclusion tools will be useful for consumers – the Monzo gamble block was designed following recommendations by MMHPI. Citizens Advice data on consumer advice trends is an excellent place to start. And while innovation by 'disruptive' firms benefits consumers and is to be encouraged, many of the harmful online behaviours need a more coordinated approach. You may decide to turn on the gamble block with Monzo, but what if you have two other current accounts without it? These tools are most effective where they are either delivered through a cross-cutting intermediary, or the action of industry is coordinated.

Recommendations

- Government and regulators should foster the growth of intermediaries that can give individuals more control and ability to customise their online experiences.
- Government and regulators should work with consumer bodies to identify areas where self-exclusion tools could protect consumers from online harms, and encourage the development & take up of these tools, particularly for vulnerable consumers.

[XXIV]BIT is also currently working with Monzo to better understand how gambling behaviour interacts with other financial behaviour

Prompts, reminders and active choices

Another way of putting control back into the hands of users is through the use of prompts and reminders. These have been used to great effect in the offline world, to encourage people to switch their energy provider;[145] [146] [147] attend, cancel or rebook their medical appointments;[148] [149] attend careers appointments;[150] pay their court fines;[151] and save money or repay their credit cards.[152] [153]

Online, these prompts and reminders can be a powerful mechanism to elicit consumer preferences (discussed in Section 2.5.). This is especially the case if they contain a meaningful and 'active' choice – requirements to respond 'yes' or 'no' to a question before proceeding – offered at a salient moment. An illustration of the power of an active choice is that researchers found that presenting consumers with an active choice on whether they would pick up their prescription from the pharmacy or have it home-delivered moved the take-up rate for home delivery from only 6 per cent to 42 per cent. It was not that they couldn't have chosen home-delivery before: it just was not the default option.[154]

On Instagram you can (if you can find where) 'manage your time' by setting a daily reminder of how long you've spent on the platform. When you reach the allotted time a pop-up says 'You've spent 10 minutes on Instagram today'. You can, however, click straight past this (perhaps in exasperation and revealing your first order preference to continue using the app). A more effective prompt designed to reduce the effects of attentional capture and elicit second order preferences might remind you that you've reached the time you set yourself, but then require you to actively choose to continue: 'Thanks, but I don't want to stick to the limit I set myself today' or 'Yes, I'd like to close to app now'. Ideally it would also impose some (small) effort if you did want to continue, like clicking through another screen or logging back in.

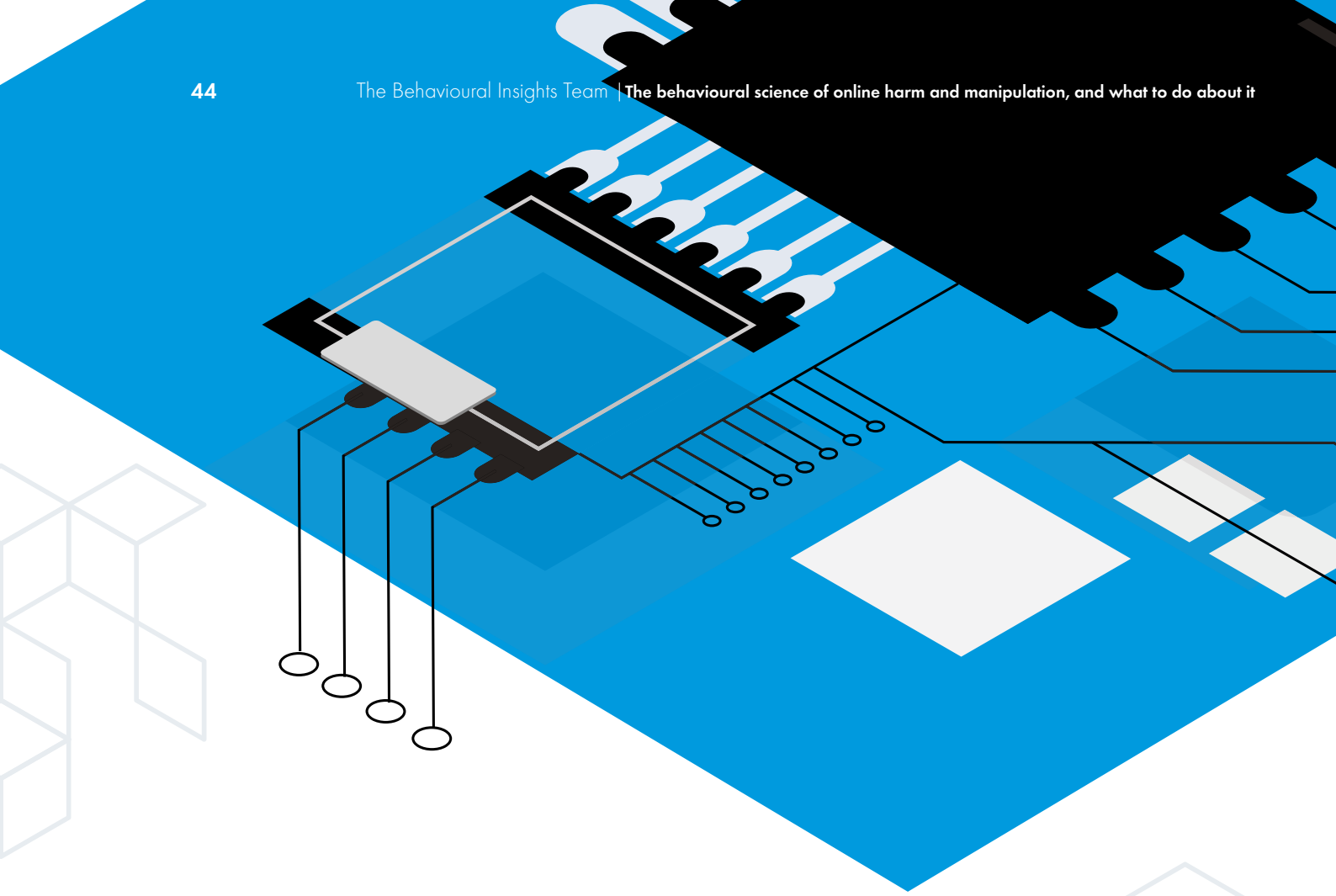
Adding in these types of prompts, reminder and active choices can also be positive for companies. It helps them understand more about users' preferences and sends a strong signal that the company is taking steps to be overtly on the users' side. In particular, helping users to enjoy interacting with platforms and services can only be positive. Our hypothesis is that building and actively encouraging these types of tools will avoid a tipping point where users choose to switch or self-exclude.

Recommendations

- Industry should design and test prompts and reminders that provide consumers with active choices about how they interact with websites and platforms.
- Legal processes, such as around possible litigation or complaints, should start to differentiate between choice architectures that enable the consumer to more actively consent or shape their choices and those that don't.

Smart defaults

User settings and prompts are a powerful way of putting control back in the hands of users, but there is also the risk that these could erode the convenience and advantages we now expect. The large tech companies and others have obsessively removed frictions to make using their platforms as easy and intuitive as possible – from one click purchases to watching an entire series of House of Cards by just pressing play. On the other hand, toggles and prompts work because they involve friction costs that interrupt the mindless flow, waking up our more deliberative mode of thinking.



One way through this dilemma is the use of 'smart' or personalised defaults. We have a strong tendency to act in line with the pre-set option – whether it is the preset length of time before our phone locks itself, or the percentage of our salary paid into our pension. Public policy has already made great gains from harnessing defaults more effectively, and the data revolution associated with the growth of the digital economy has moved these smarter, more personalised defaults from an interesting theoretical idea to a real and widespread possibility.

Smart defaults for device users

The relatively new Apple Screen Time feature (released in September 2018) sends you a weekly report about how much time you spend on your iPhone or other devices, how many notifications you are sent and how often you pick up your phone, as well as allowing you to 'set limits for what you want to manage'. Far from welcoming this feedback, the initial reactions to Screen Time reports were of widespread horror.[155] Anecdotally, realizing just how much time they spent on their phones, people quickly started avoiding the reports. Despite this, a powerful cross-platform default could move beyond information and feedback to guide and prompt individuals to set smart defaults for how they use their devices and where they focus their limited attention (see Section 2.4.). These defaults might focus on:

- **The amount of time spent on your phone or device** in total, or on specific apps or categories (social networking vs productivity) before you are 'locked out' or your device switches to black and white or slows down
- **When and how you are sent notifications, to minimise distractions and cognitive disruption** – in a 'bundle' twice a day rather than every hour or every time there is new activity[xxv]

[xxv] Forthcoming research by Nick Fitz (currently under review) suggests that such bundling of smartphone notifications may also improve wellbeing. Slides presenting the results of this research are available at: https://osf.io/fbk7w/

- **When and how you 'mute' your phone or device,** for example during dinner time or from an hour before bedtime until you wake up, or when you are driving
- **Highlighting dynamic social norms** about how other people are adjusting their device use in response to personalised feedback

Such defaults also have the potential to reduce unintended consequences and associated harmful effects of attention catching, habit-forming design.

Symmetry by default

As discussed in Section 2.1., the tactics that make cancelling online subscriptions difficult are classic sludge. While it's often possible to subscribe with a single click, opting-out or cancelling can take much more effort. Inspired by the GDPR's philosophy that 'it shall be as easy to withdraw as to give consent', a powerful default to curb this kind of exploitative behaviour is that cancelling an online subscription, unsubscribing from a mailing list or leaving a platform should be as easy as signing up. This would powerfully rebalance market practices in favour of consumers, and facilitate switching and competition in digital markets. The burden of effort should be placed on businesses to simplify the exit process – they invest heavily in designing easy sign-up processes, and it should be equally possible to invest in removing frictions at the other end.

Recommendations

- Government should work with industry to design, test and promote cross-platform defaults that guide how consumers use their phones and other devices.
- Government should pass new legislation requiring industry to make cancelling an online subscription, unsubscribing from a mailing list or leaving a platform as easy as signing up.

3.3. Fostering sharper competition

The contrast we painted between fruit and banks at the start of this chapter helps us to see, albeit in a simplified form, how even competitive markets can lead to situations with widespread and stable detriment to consumers. As we saw in Section 2., these dynamics can be exacerbated online, especially as digital markets – by virtue of scale and network effects – are subject to tipping to the point where one market player dominates. Greater competition, on its own, is unlikely to solve this issue where there are behavioural biases in play (leading to systematic errors that can be exploited), heterogeneity in the strength of biases (naive and sophisticated customers), and the ability to price differentiate (or facilitate that differentiation through targeted advertising).

In line with the findings of the recent UK Digital Competition Expert Panel, we are optimistic that 'forward-looking' competition tools can play a large role in giving consumers greater choices, helping new companies enter and grow, and spurring innovation.[156] There is a role here for updated antitrust interventions, like more flexible mergers policy and enforcement. However, this section focuses on more nimble mechanisms to foster greater competition: transparency, facilitating comparisons and smarter use of data.

Transparency to facilitate comparison and create accountability

Online feedback mechanisms

Even in the age of big data, sometimes the most useful information lies in the heads, and experiences, of the users. Unfortunately, while companies can now generally easily marshal their knowledge about their customers (and suppliers), it is much harder for consumers or users to do the same. Platforms like TripAdvisor, consumer experience systems such as Reevoo, or the feedback mechanisms built into platforms such as eBay and Amazon fundamentally change this dynamic by shining a light on aspects of the products (e.g. quality of service, or communication speed) that would otherwise be, as Xavier Gabaix and David Laibson put it, 'shrouded'.[157]

Such 'deshrouding' – revealing key information about quality and service – enables consumers to find better, more personally suited products, and helps good firms and products grow and take market share while poor quality products and firms get squeezed out or forced to improve. Harvard Professor Mike Luca has studied the effects of ratings and review platforms in some detail. He has documented how a one-star increase in a restaurant's rating on Yelp led to a 5-9 per cent increase in revenue.[158] Interestingly, the benefits of better ratings disproportionately fall to smaller businesses, not to generic chains. This makes sense given that most consumers already know what to expect from a burger at McDonalds, but know less about an independent burger restaurant so may be much more influenced by the ratings of others.

Given the power of aggregate consumer feedback to 'deshroud' markets and empower consumers, it raises the question of why TripAdvisor-style mechanisms aren't seen everywhere. One answer is that they thrive in certain conditions. These sites work in markets where consumers have reasonably frequent transactions (consumer goods vs. arranging a funeral); where the consumer can relatively easily judge quality (restaurant food vs. the performance of a pension) and where a critical mass of experiences can be meaningfully aggregated (quality of a given hotel vs. a personal carer).

However, our view is that ratings and reviews could have broader applications and that there are markets that could be made to work far better through a strong dose of consumer feedback. For example, a key finding of the Taylor Review was that the UK labour market needs greater transparency and understanding between workers and employers.[159] While a prospective employee can easily find out what they might get paid, more qualitative elements of a workplace, such as the character of management practices or the opportunities for progression, are harder to discover. The UK Government has agreed to five principles that drive the quality of work: overall worker satisfaction, good pay, participation and progression, wellbeing, safety and security, and voice and autonomy.[160] Transparent measures in these areas – developed and promoted by both government and online job boards – could help employees choose workplaces with higher wellbeing, higher pay and good progression. It would also foster more competition between employers on these metrics.

Of course, there are potential issues with feedback systems too, notably fake reviews – highlighted by the alarming but amusing case of a backyard shed becoming the top rated restaurant in London on TripAdvisor[161] – and the conundrum of reputation inflation. A recent paper argues that the effectiveness of rating systems can deteriorate over time.[162] This is because reviewers feel pressure to leave 'above average' ratings, so ratings become clustered at the top rather than distributed in a way that allows users to differentiate between suppliers. Reputation inflation seems to be most acute in peer-to-peer platforms like online labour markets and the sharing economy: feedback scores are often the only sign of quality and the 'reflected costs' for reviewers can be high. A reviewer rating an Uber driver may worry about the driver's return review of them, or feel guilty about harming an under-performing driver's future prospects of work.

Left unchecked, fake reviews and reputation inflation can undermine consumer trust.[163] But there are things we could do. For example, systems could be designed to:

- **Raise the benefit of being truthful** by providing incentives for users who generate feedback, emphasising reviews as a service to fellow consumers. For example, Google reviews tell people how many others have found their review helpful;
- **Allow other users to rate reviews** for accuracy and usefulness (or even humour!);
- **Impose sanctions for untruthful or fraudulent reviews.** For example, in the US, the Federal Trade Commission recently took its first major action against fake online reviews by bringing a successful case against a seller buying fake reviews for its product (a purported weight loss supplement which had no scientific basis) and posting these fake reviews on Amazon.[164] Some online platforms, like videogame marketplace Steam, are also using algorithms to detect patterns of reviews that are unlikely to have been made by a human, such as where there are an excessive numbers of reviews, and then deleting or down-rating these reviews;[165]
- **Lower the personal cost for the reviewer of leaving a poor review**, for example, simultaneously revealing consumer and supplier ratings or anonymising ratings through aggregation;
- **Only display feedback once a threshold number of reviews has been collected** and make the distribution of feedback scores publically available alongside the average score and number of reviews;
- **Supplement and validate feedback against other forms of information** such as complaints data, 'friends and family' or 'net recommender' prompts or surveys.

Both the issues associated with online feedback mechanisms, and the solutions that can be designed to help tackle these issues, are currently an open area of research.[XXVI] Rapidly designing flexible solutions will only become more important as fraudsters adapt to the changing online environment by using new strategies, such as the empty box scam on Amazon,[XXVII] [166] the creation of 'click farms' where workers are tasked with continuously reviewing specific products to boost their ratings,[167] or even the 'weaponisation' of online platforms' fraud detection systems by small businesses who, in an attempt to shut down competitors, post large numbers of positive reviews on their competitors' sites to make it look like they are buying fake reviews.

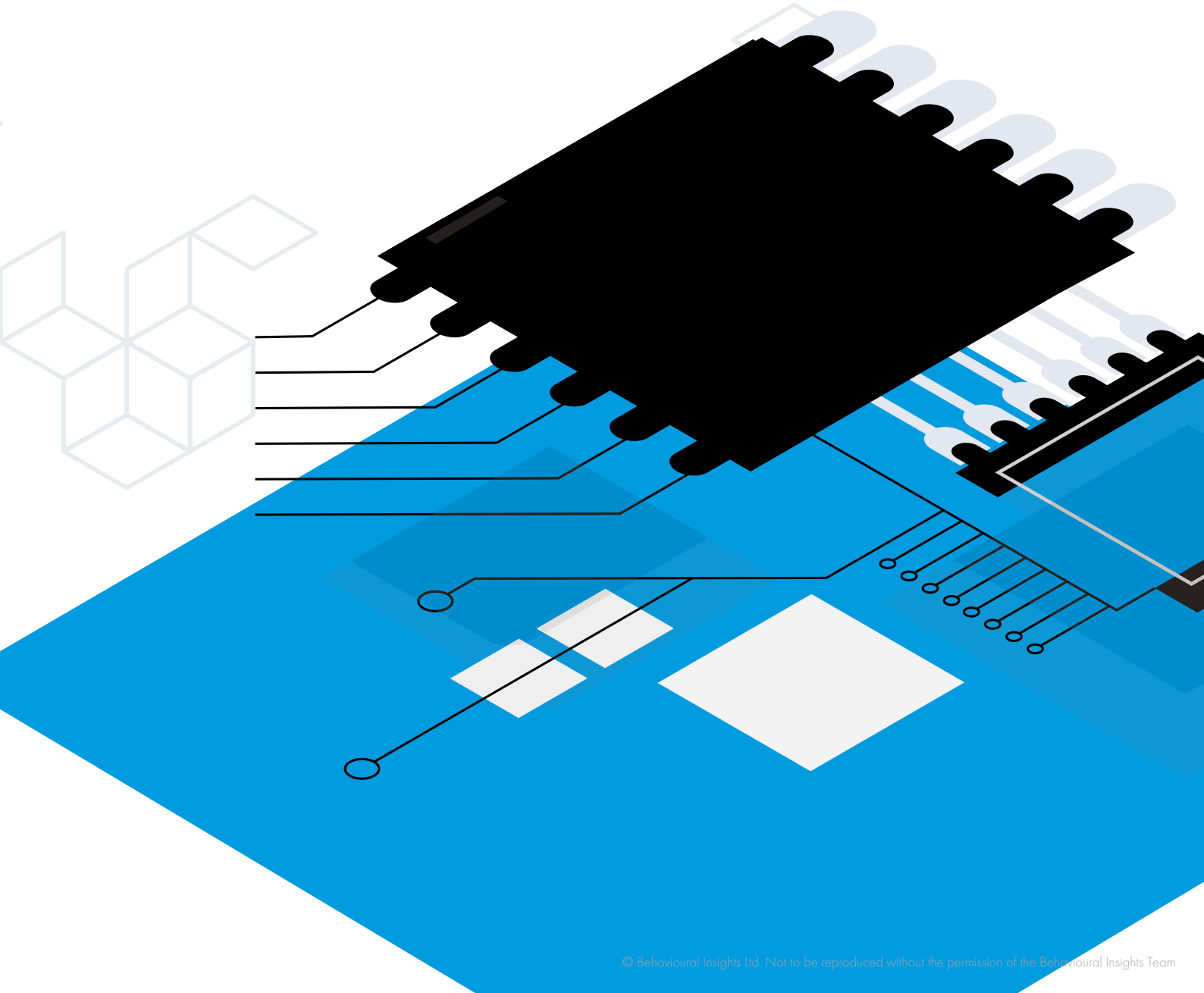
[XXVI] For a detailed discussion of these issues see: Watt, M & Wu, H. (2018) Trust mechanisms and online platforms: a regulatory response, Harvard Kennedy School Policy Analysis Exercise for the Federal Trade Commission.

[XXVII] This is an increasingly common scam in which buyers order an item, and upon receipt of the item, contact Amazon's customer service, claiming that they received an empty box or that the item never arrived, so that they can claim a refund or a replacement item.

Transparency reporting

Another way to promote transparency and public accountability is to compel industry, particularly social media and news platforms, to publish regular reports on key metrics: complaints; how they are protecting user data and privacy; the type, volume and results of experiments they are running; and how they are addressing online harms like misinformation and disinformation on their sites.

Well-designed transparency reports that can be used to rank company performance in a salient and meaningful way provide a framework to drive excellence and best practice, at the same time as clipping the poorest practices and businesses. This is especially the case if the reporting system is designed to target advertisers and intermediaries who interact with companies and have reputational incentives to hold them to account on issues like dubious experiments, fake news and hateful content.



Of course not all of this activity need be driven by regulatory or voluntary requirements on industry. There is a clear role for civil society to undertake cross-platform research into issues like the welfare effects of social media and the prevalence of fake news and communicate the results to consumers. This research can also be used to check or calibrate companies' own transparency reports. Section 2.6. mentioned the RSPH's recently published survey results about how almost 1,500 young people feel the most popular social media platforms impact their health and wellbeing. This report is not just for policy wonks; rather, its findings can help young people and parents make meaningful comparisons about social media use by ranking the net impact of different platforms on wellbeing. Given the pace of change, we suggest funding an annual survey, including a larger range of age groups and increasing the sample to break down the results for young people by gender. The existing report could also be enhanced by supplementing and calibrating the survey data with experimental evidence on the welfare effects of different platforms.

On the basis of these transparency reports, consumer groups or regulators might consider developing regulated descriptors, similar to film classifications, to clearly signal the reliability and 'safety' of different platforms. This could greatly help understanding, for example of which platforms or environments are safe for children, more appropriate for adults, or those in which many adults might feel uncomfortable or want to exercise a greater degree of caution.

Recommendations

- Governments, regulators, consumer organisations and trade bodies should actively encourage the wider emergence of feedback and ratings platforms.
- Industry should pay particular attention to the design of feedback systems to minimise fake reviews and reputation inflation.
- Governments should introduce transparency reports for online companies and give regulators powers to audit existing efforts and require remedial steps if they are not adequate.
- Civil society (and academia) should continue to conduct and publish research on the comparative performance of platforms in relation to welfare effects, online harms and data protection.

Building trust and confidence in digital comparison tools

Comparison tools are integral to well-functioning markets, but particularly digital markets where consumers are choosing between many more, and more complex, products and services. In principle, it doesn't matter how many alternative broadband plans or credit card options exist, as long as there are comparison tools that reliably sift through them on consumers' behalf.

In their recent market study, the CMA argued that realising the benefits of digital comparison tools (DCTs) depended on consumers having 'sufficient trust and confidence to use DCTs in the first place, and enough understanding to choose and use them effectively'.[168] When the CMA asked UK consumers about their experiences, they found that people were mainly positive about DCTs but also had concerns especially regarding transparency about their use of personal information. Given broader issues about economic 'hollowing out' and 'most favoured nation' agreements,[XXVIII] [XXVIX] it is vital that consumers are able to tell the difference between a good comparison tool and a bad one.

[XXVIII] Hollowing out', as defined by the CMA in the same DCTs market study, refers to the decrease in quality (for example, a worse insurance cover) that can result from an undue focus on price.

[XXIX] Most favoured nation' agreements, as defined by the CMA in the same DCTs market study, are a type of contract between suppliers and DCTs which prevent suppliers from offering better prices on one DCT than on another and can therefore reduce competition between DCTs.

Regulators should take a stronger and clearer role in catalysing these types of intermediaries, and building consumer trust in them, including by helping consumers tell the difference between comparison tools. For example, they could publish league tables of switching sites on their own sites, so that consumers can see the rankings and, ideally, sort by best price, best service or other criteria. The power of this form of transparency would lie in its deterrent effect on poor comparison sites by creating a risk to their brand and reputation.

Recommendation

- Regulators should publish league tables of switching sites covering metrics like: the average price savings achieved; number of complaints against the site; and performance as measured by mystery shopping.

Leveraging data and AI on the side of consumers

At the core of many market problems is a deep information asymmetry between suppliers and consumers, even the sophisticated ones. Online markets widen this gap because the sheer amount of data available to the supplier gives a huge advantage. And so a key question is how we can rebalance this dynamic by enabling consumers to use it too?

Giving consumers more access to, and control of, their data

Since 2011, the UK Government has pursued the Midata initiative in regulated markets, working with businesses to increase access to personalised consumer data in electronic formats. Midata is designed to redress the imbalance between consumers and firms by providing consumers with information about their consumption over the previous year. But, while there have been some issues with the implementation of Midata as currently configured, it is also unlikely to be flexible enough to address these imbalances in online markets.

The data that consumers generate online can help platforms enhance the user experience, but at the same time it can increase the cost to consumers of switching platform or even trying out new services. Once you've invested in building playlists and establishing a network on Spotify, it is very costly to move to a new music streaming service and start again. If users were able to take the data that they generate to other services, it would make trying alternatives and switching far less costly – an idea also recommended by the Digital Competition Expert Panel.[169]

An expansion of Midata and similar programmes that facilitate open standards could improve access to data for consumers, workers, businesses and suppliers alike. For consumers, that could mean being able to join a new service and upload their historical purchasing data – books and music for example – resulting in a better customer experience. For self-employed gig economy workers this would mean they could take their existing customer review data to new platforms; for example, a Uber driver should be able to take their 5-star rating as a starting point to a rival car sharing service.

This type of access to and portability of user-generated data is likely to be key to fostering competition between companies and shifting the balance of power between platform and user. Of course, the channel is just part of the puzzle. Once access and portability are secured, behavioural insights can be harnessed to help consumers use this information to actually make the switch.

Fostering new types of intermediaries

The challenges presented by the evolution of online markets create opportunities for new types of intermediaries to help users to sift through large volumes of complex information (and even potentially profit from their data) and help redress the balance of power between big tech and consumers. These intermediaries are only set to become more important over time as more data is created. A few areas ripe for innovation include:

- **Intermediaries to decipher terms and conditions** and privacy notices on your behalf.[xxx] Useful features might include: the ability to set your privacy preferences upfront and then have any terms that don't align with your preferences highlighted to you; highlighting terms that differ from industry standard; and a simple rating of T&Cs based on how well they protect your interests compared to those of the supplier;
- **Intermediaries to give consumers control of their data** and allow them to easily exercise their rights. Yo-Da (short for 'your data', it is) is a start-up allowing consumers to understand what companies have access to their personal data and to restrict the use of it should they wish.[xxxi] Essentially this allows consumers to exercise their rights under GDPR with a few clicks;
- **Intermediaries that generate or aggregate trustworthiness ratings,** for example, a plug-in that rates the factual accuracy or trustworthiness of news sources. Such a tool could use a combined rating of the accuracy of given sources based on previous accuracy levels as measured by independent fact checker sites. Or a browser plug-in that aggregates the various trust ratings for online suppliers from Trustpilot, Trusted Trader, Yelp etc.;
- **Automated switching services** allowing consumers to be regularly and automatically switched onto the best deal, based on pre-set preferences, across multiple markets.[xxxii] These intermediaries could also use banking transaction data to highlight areas where individuals could get a better deal and then facilitate switching;
- **Online 'advocates' powered by AI.** A more radical suggestion comes from UK economist Tony Curzon-Price in considering how to marshal computational power on the side of the consumer.[170] Curzon-Price argues that the AI on the side of the consumer should be equivalent to – or at least a fair match against – the AI being leveraged by platforms like Facebook in deciding how to order and curate your online experiences for maximum attentional capture and thus maximum revenue. He suggests that every consumer ought to have the right to be represented online by an intelligent, robotic avatar with a fiduciary responsibility to the consumer. Such an avatar – let's think of it as the true 'econ' imagined by classical economists but never seen in real life – could, for example, process, re-order and curate an individual's Facebook newsfeed.

[xxx]For example, 'Terms and Conditions; Didn't Read' is a website and browser add-on that will give you info on key good/bad aspects of a website's T&Cs and a A-E rating. Timely, salient, short and simple! https://tosdr.org/

[xxxi]For information see: https://www.yo-da.co/

[xxxii]In the UK there are existing powers which allow regulators to push through APIs that make it much easier for consumers to compare products, including enabling consumers to get easy access to their own consumption data (as discussed above). Exercising these powers effectively would enable intermediaries (and suppliers) to give consumers more personalised, low cost advice across a range of markets – from choice of annuity products to energy supplier or credit card. The recent Consumer Green Paper and BEIS Energy Midata consultation start to work through the issues of cost, compliance and security that regulators will have to tackle to take automatic switching to the next level.

Some of these intermediaries are already emerging, but our strong view is that their development should be accelerated by government support through innovation funding or challenge prizes. Government should intervene in a market if there is sufficient evidence to establish that there is market failure that leads to an inefficient economic outcome and that government intervention will make a cost-effective improvement to market functioning. We consider that these conditions exist in areas of online markets market, particularly where there are deep information asymmetries between companies and consumers (particularly as a result of data collection), public and mixed good characteristics, and behavioural barriers that prevent consumers from acting in their own best interests (as set out in Section 2.).

Further, Government can kick start these markets by working with new intermediaries in areas where government is a data controller or has significant influence over how data is collected, stored and used, for example, in the health and transport sectors.

Using data to identify vulnerable consumers and signpost support

If an individual is consistently shopping online between midnight and 4.00am, or spending large sums on online gambling sites, there are several parties who can see this pattern of behaviour. The person's bank. Online retailers. Gambling sites. Google. The list goes on. Yet, none of these parties have obligations to identify these patterns as possible warning signs of vulnerability, nor to signpost sources of support to these people.

In regulated markets, at least, there is a live debate about regulators using data science to identify vulnerable consumers based on their behaviours and to share this information securely across sectors and firms. The UK Energy and Water regulators have made strides in pooling their data to identify vulnerability.[171] This will ultimately allow firms to offer tailored support to vulnerable consumers – for example access to priority services registers – without them having to self-identify. Your bank, for instance, might suggest that you set up a gambling block or spending limits. Google might promote search results that signpost charities that specialise in problem gambling or a plug-in to self-limit the hours you can shop online.

Of course, this raises issues of data privacy and ethics that are not to be taken lightly. Should governments hold this data or do we need new institutions to govern the uses of this data and foster trust and trustworthiness?[xxxiii] These issues could be constructively explored using the types of deliberative mechanisms we will discuss in Section 3.4.

Recommendations

- Allow greater access and portability of data so that users – consumers, workers and businesses – can try alternative services and switch more readily.
- Government and regulators should use innovation funding and/or run Challenge Prizes to kick start and foster intermediaries that leverage data to benefit consumers.
- Regulators should investigate the case for a new consumer right to be represented online by an advocate powered by AI.
- Explore the case for extending existing work and obligations to use data to identify vulnerable customers to non-regulated markets.

[xxxiii] See for example the discussion by Geoff Mulgan and Vincent Straub on how data trusts, collaboratives and coops can help govern data for the maximum public benefit: https://www.nesta.org.uk/blog/new-ecosystem-trust/

Next generation antitrust to support new market entrants

As seen in Section 2.7., the scale of online markets, coupled with the value of network effects and access to data, mean that some platforms can come to dominate the market, tipping into a 'winner-takes-most' dynamic. Even if a new entrant developed a better, more innovative product or service, it would be difficult for it to compete against dominant players. And even if they did, many start-ups actively set out to be acquired by the incumbents at a premium.

Various ideas are being debated to lower the barriers to entry into data-rich markets. Should companies be required to share their algorithms with competitors, regulators or the public? Perhaps, but it's difficult to follow a recipe if you don't have the ingredients. And few algorithms are static. Machine learning algorithms adapt as they learn from new data: sharing an algorithm with a competitor merely provides a snapshot in time – a small piece of a giant puzzle. So what about sharing the data itself? The Digital Competition Expert Panel, recognising the importance of data as a driver of market power and concentration, recommends opening up data held by digital businesses and providing reasonable access to competitors.[172]

The question is, how do we pursue data openness in a proportionate way while protecting privacy? One novel idea proposed by several legal scholars and economists is to create a progressive data-sharing mandate.[173] Dominant companies would be compelled to share a representative cut of anonymised data with new market entrants and competitors. This requirement would kick in when a company's market share hits a certain threshold and ratchet up so that the larger its market share, the more or more detailed data to be shared. Companies sharing the data won't lose the benefits of it, nor the intellectual property and investment they have made in turning that data into usable insights and business strategy.

A step further would be to establish regulated arrangements to access actual data. For example, the UK Energy Regulator recently established a database to hold the data of almost 10 million UK households on expensive standard variable energy tariffs: their energy consumption, personal details, and current energy deal, amongst other things.[174] The database is designed to allow the regulator to work with energy companies and intermediaries to intervene in the market, for example to experiment with ways to encourage these 'sticky' customers to switch to better deals or better providers. Similar databases could be established in online markets, with data trusts or other intermediaries controlling access as well as the terms on which new entrants could access this data or experiment with new product or business ideas.

In both incarnations, the principle of data openness spreads the value of big data around, allowing new entrants to innovate and compete and keeping the dominant players on their toes.

Recommendation

- Government and regulators should actively pursue data openness, including investigating the feasibility, costs and benefits of a progressive data sharing mandate.

3.4. Non-market design

As we saw in Section 2.6., markets have a deep symbiotic relationship with non-market factors – Adam Smith's 'sentiments' or values; habits of trustworthiness and common understandings; and social networks. The ways we interact and transact online, social media in particular, are entwined especially deeply in these non-market factors, and have the potential both to enhance and to degrade them.

Patterns of association

The internet and online platforms have changed the costs and benefits of trading and connecting with other people. They can massively expand our 'weak ties', connecting us with huge numbers of people we don't know well, or at all, but whom we can at least adequately trust to honour a transaction (especially when we can see feedback on reputation or performance).

Technology can also be applied, at least in principle, to transactions or exchanges that are traditionally considered non or only partially marketised, notably care-based exchanges. Just as eBay ratings enable us to trust someone we have never met to send us a product in exchange for advance payment, online matching markets are starting to offer ways to find a carer for our children or elders. Technology can solve both complex issues about time and availability (who can look after my children, grandmother, pet tomorrow?), and trustworthiness (can I entrust those near and dear to me to this person?).

The 'sharing economy', too, is growing, but has yet to reach anything like its full potential. Trust is a major barrier. It is likely that for these platforms to reach scale, user authentication – and reputation – will need to be more securely cracked. Users really need to know the person is who they say they are, and that they can be trusted. This may require blended authentication – official verification blended with private or third sector identification. Users may have to be required to use their real names and identities, or have the ability for these to be confirmed by a third party or mediator.[xxxiv]

It is also hard for such platforms to reach critical mass, especially where many of the exchanges are care-based or in the form of extended reciprocation (e.g. time banks). Given the strong public good aspects of such platforms, it may require extensive action, and possibly funding, by state or third-sector actors to build or extend such platforms, though it is also possible that some existing platforms could build this kind of functionality into their current offers.

One issue discussed in Section 2.6. concerns the rules around discrimination and association that are to apply in evolving online markets.

Some forms of discrimination may be adequately picked up by existing legislation. Where a platform is found to be facilitating or encouraging discrimination by protected characteristic, such as race or gender, on price or access to a service, legal action may require it to change its practice. These platforms should be required to alter their choice architecture to minimise discrimination, for example by reducing salience of photographs or user names that may provide clues about ethnicity, gender or sexual orientation (at least until after a booking is accepted) and encouraging complaints as part of feedback mechanisms. Data analytics could make it much easier to detect certain of these forms of discrimination more accurately than in 'traditional' markets.

However, it is likely that some aspects of online markets may facilitate new forms of association and potentially discrimination that will prove challenging to current legal thinking. Social media makes it easier for individuals to enter special interest groups, and exchange and trade internally within these

[xxxiv] For example, Nextdoor (nextdoor.co.uk) is an online platform where people in a neighbourhood verify their names and addresses and can then chat privately to their neighbour for everything from finding a babysitter to planning a local event.

groups. We know, for example, that many job opportunities are identified through social networks.[xxxv] [175] This raises questions about how we handle the emergence of overtly, or incidentally, exclusive groups. How should we handle the emergence of groups that form around, say, white men who went to a particular college, or specific ethnic groups within a given profession? These issues have been powerfully illustrated by recent controversies in France around the Facebook group of male journalists, the Ligue du LOL, which actively organised to harass female journalists.[176] In this case the behaviour that was revealed caused outrage, but there will be many such groups that refrain from such overt behaviour, but nonetheless provide benefits and contacts along the lines of an 'old boys club'.

Box 6: Reducing bias in the hiring process with Applied[xxxvi]

Applied, BIT's first spin out, is a people platform focused on using behavioural and data science to improve hiring decisions. Applied uses (and carries out) research to design products that help teams to hire in ways that are smarter, fairer and more diverse including through anonymising candidate applications to reduce the risk of unconscious bias. Since its launch in 2016, more than 90,000 candidates have applied for jobs through the platform, which involves being tested on job-relevant tasks rather than CVs or pedigree.

In the past year, Applied has expanded to service over 60 organisations – including government departments in the UK and abroad, start-ups, and large corporates like Hilton, Penguin Random House and GroupM – and recruited for an array of jobs ranging from product managers to ministerial speechwriters, interns to chief marketing officers. In addition to scaling the platform and successfully securing venture capital backing, 2018 saw the Applied team expand its product into better interviews and tools to improve the inclusivity of the language used in job descriptions.

The issue may be partly handled by using techniques to reduce bias in recruitment, purchasing and other areas. These should definitely be used to their full potential, but may not entirely address the subtler forms of discrimination and disadvantage that identity or interest-based groups may be associated with. The answer is likely to be in the emergence of a new body of practice or law that sets out the extent to which groups may be allowed to exclude others in online markets.

A proactive approach may also be possible and desirable – in other words, actively encouraging forms of exchange and connection that tend to foster exposure to alternative perspectives and the building of 'bridging social capital'. Specific examples of practices or policies that could be built on include:

- **News feeds and other tools that prompt you to read alternative, opposing or perhaps even evidence-based views.** This is an area that is well-suited to experimentation;
- **Highlighting comparison points that are not compounded with protected group characteristics,** for example user reviews and recommendations could highlight 'people with similar interests' liked X rather than 'similar people' liked Y; and
- **Active encouragement of bridging or interlocking forms of social capital,** such as work by BIT using our 'Networky' platform[xxxvii] to securely introduce young people to others from different backgrounds with whom they will be going to a new school or doing National Citizens Service.

[xxxv] Despite the rise in online job sites an estimated 80 per cent of jobs are not advertised and some companies aim for up to 50 per cent of their hires to be referrals from existing staff (and provide rewards to incentivise this).

[xxxvi] https://www.beapplied.com/

[xxxvii] https://networky.co.uk/

Recommendations

- Regulators should encourage, and potentially compel, platforms to put in place structural features to minimise discrimination.
- Regulators should explore how machine learning of patterns of online association, complaints mechanisms and feedback loops can identify and reduce overt and inadvertent discrimination.
- Policymakers, third sector and industry should actively explore and encourage forms of exchange and connection that foster exposure to alternative perspectives and the building of 'bridging social capital'.

Civility

As we noted in Section 2.6., there are widespread concerns that social media, and to some extent other online platforms (such as in hostile reviews) can nurture uncivil comments and tone. Much of this is well below what would be defined as 'hate speech', and regulators are appropriately wary about getting drawn into the territory of Stalin's 'Ministry of Everyday Life'.

Much better would be to see the evolution of self-regulatory dynamics more akin to 'politeness' or informal codes of conduct that the majority of users are comfortable with (combined with a sufficiently dynamic means for users to exit or temper exchanges that they find unacceptable).[XXXVIII 177]

Such evolution needs to be guided by users, and builds the case for stronger mechanisms for user influence over online platforms and social media (discussed below). But we can start to see glimpses of the kinds of practice and policy that this might lead to:

- **Defaults that buffer potentially hurtful or offensive posts and responses for a short time,** enabling users to change their minds;
- **Nudging users to consider the likely impact of hurtful content.** This could increasingly be done with live detection and flagging, using machine learning trained on recipient reactions or third-party readers or viewers (or even to originators when they are in a 'cool' state). These should be able to be overridden but would first ask users to reflect: 'Are you sure this won't cause offense?' or 'This doesn't sound like something you would normally say, are you sure you want to post it?'; and
- **Feedback to users whose content is felt to be hurtful or offensive,** to enhance the constructive impact on both sender and recipient. This is an area well-suited to further experimentation.

Larger market players are also starting to explore different ways in which users can be given feedback on why content was considered unacceptable (rather than just removing it with no explanation), and enabling users to contest content. We believe these feedback loops can provide a very important channel through which social norms or semi-formal rules can develop, in effect creating a largely self-regulatory dynamic. The logical corollary, at least for larger platforms, will be the development of some form of Ombudsman, ideally independent, to adjudicate difficult cases.

[XXXVIII] See for example the battle over the character of Reddit, including the interesting battle over how to fill in a million pixel image.

Recommendations

- Platforms should develop and test new prompts and defaults that encourage civility between users, and specifically encourage users to reflect before posting potentially harmful content.
- Platforms should design systems that provide feedback – both from the platform and from other users – to users who choose to post harmful or offensive content. These should explain why the content was considered unacceptable. Disagreements that go to fundamental points of principle should be referable to an Ombudsman (see below).

Who decides?

Issues of association, civility and content have a degree of fluidity, and are strongly rooted in civil society rather than in the black and white rules of legal code (or market maximisation).

Many of the key market players are approaching market dominance, and have acquired powers and responsibilities that touch the lives of billions. As discussed in some detail in Section 2.5., it is not enough to construct policy or practice for such entities solely on the basis of the first level 'expressed preference' of users – i.e. if they click it, it's right. Fostering a choice architecture that enables users to express their more reflective, or second order, preferences helps (see Section 3.2.). But there also needs to be a place for people to negotiate with each other to reach a collective view on what constitutes appropriate practices and the 'rules of the game'. This is especially true for those platforms that view themselves as nurturing a community, yet appear to lack any meaningful way for that community collectively to shape the core parameters of the platform itself. In short, we need ways for the community to 'nudge the nudgers' – and particularly when exit is difficult.

Box 7: Engaging young people to create a more positive online environment

In partnership with the Vincent Fairfax Family Foundation, BIT held a two-day citizens' jury and design sprint mash-up called the #NoFilter Forum to discuss teenagers' ethical use of technology. Over two days, we brought together 61 young people aged between 12 and 16 to debate and deliberate with one another, hear from experts and co-create solutions to create a safer and more secure online environment.

The ideas ranged from tech solutions to moderate the type of content they see, to ways to inform and spark conversations between parents and children about navigating the online environment.

Perhaps even more profound than the ideas generated, was what we learned by spending two days with a diverse range of young people. We found that young people do think about their impact on others and are concerned about showing empathy. But they often struggle to act in line with their values and wished they had more support and understanding from parents, other adults and their peers to find practical ways to resolve these conflicts. They generally were not aware of the impact that industry and platforms have on their actions; rather, they felt solely responsible for their behaviour. They expressed a desire to be an authentic part of developing solutions and were critical of adults who involve them in tokenistic ways. These lessons became guiding principles for the interventions developed out of the citizens' jury.

For this reason, we suggest that there is a key governance gap with respect to online platforms and markets. There are a number of models that could be used to address this gap. National (and supra-national) governments can intervene more directly to set the rules on behalf of their citizens – something that we are already seeing in areas such as hate speech and data portability. Nobel Laureate Jean Tirole proposes a more flexible option of encouraging industry and other parties to propose possible solutions, with competition regulators formally responding to give some direction but without creating new or binding rules. Tirole calls this 'participative antitrust'.[178]

We believe there is a case for going further still, with governments working closely with companies and citizens to get the framework right. In the spirit of debate, we suggest that key elements should include: a high level, independent appellate system; a cross-national ethics body; and the collective user voice. These are described below.

A high level, independent appellate system. Building on companies' own internal complaints mechanisms, fundamental questions that relate to matters of principle (such as classes of unacceptable content, or use of data) should be escalated to an independent body. Mark Zuckerberg has proposed a form of independent 'Supreme Court' for Facebook that content rules can be escalated to. Legal academics on both sides of the Atlantic are exploring related ideas. A key element of such systems is that they need to hear contrasting views; to publicly wrestle with issues of principle; to have independent but fully paid 'judges'; and to generate a body of principles and rules to guide the practices of both companies and citizens. It is likely any single 'supreme' body will need to be built on a number of independent review bodies, given differences in cultural sentiments around the world. It is also likely that each of the major platforms will need their own internal review systems, but it would be better if they shared a common top-level appellate system.

A cross-national ethics body. It is possible that this could be integrated into a top-level appellate system, or even built into existing international governance systems such as the United Nations. Prototypes of such a body are already emerging. Within the UK, a group of independent experts that Google's Deep Mind invited in to set the parameters of their work offers a possible model. The UK Centre for Data Ethics and Innovation potentially expands this approach across companies, academia and governments, and could evolve into an international body. The Turing Institute[XXXIX] and Tegmark-Musk[XL] research bodies have elements of what is needed in a top-level ethics body – the capacity to proactively (i.e. without cases being brought) explore difficult ethical and operational issues, and to express a view. However, an effective ethics body will need to go further, establishing principles and guidelines and having some powers to enforce these parameters.

Collective user voice. Expert opinion and voice are necessary but not sufficient to create a legitimate and effective governance framework. The practices in online markets are not just an add-on to existing market and non-market practices – they increasingly are the market and the substance of social interaction. It is important that citizens are able, collectively and reflexively, to shape this landscape: it is their lives and relationships that are the matter being discussed. Realistically – and methodologically – everyone can't spend their lives debating the rules for this new world. But it is possible to build in deliberative mechanisms that take representative samples of citizen-users, immerse them in the issues for a couple of days, and ask them to give a view on behalf of the community. These mechanisms range from Fishkin-style 'deliberative polls',[179] where a couple of hundred people are exposed to a range of perspectives and then surveyed on their newly informed views, to deliberative forums within which a random sample of citizens are given the freedom to formulate their own proposals and are asked to reach a collective view (e.g. where 80 per cent of them agree).[XLI 180] As with the appellate system, it is likely that such forums would need to run in parallel in different regions of the world. They should be held regularly, at a frequency to be established – possibly around twice a year – with a substantially or entirely new sample for each one.[XLII]

[XXXIX] Alan Turing Institute Data Ethics Group. See: https://www.turing.ac.uk/research/data-ethics

[XL] Elon Musk has donated to the AI Safety Research Program at the Future of Life Institute co-founded by Max Tegmark. See: https://futureoflife.org/ai-safety-research/

[XLI] For example see the deliberative forum held in Iceland in the wake of the 2008 financial crisis.

[XLII] It is important that each forum is substantially 'refreshed' so that it remains representative of the general population. There are some variations that are possible, such as that citizens are recruited into a panel that serves across more than one issue (for efficiency grounds), or that a proportion of people from a previous forum are retained to 'seed' or bring expertise into the next one. The advantage of the latter is that you have a group of relatively expert citizens who are familiar with the process able to help others who are new. The (significant) downside is that the voices of these relatively expert citizens may hold disproportionate sway, distorting the sample and process away from representative. This may be attenuated through skillful mediation.

Ideally, these mechanisms will be used in combination, to enable citizens to genuinely shape the character of online environments and the way that we interact with each other. An appellant-style system helps deliver individual and community level justice, and helps a body of practice to evolve. A cross-national ethics body can anticipate issues and proactively shape a framework that maximises good for all. The overt incorporation of a collective and informed citizen voice brings both legitimacy and a powerful check and balance.

Recommendations

- Governments should work with industry and citizens to design new and appropriate governance mechanisms for platforms that combine expert opinion with the collective user voice to allow a platform's community of users (and possibly all citizens) to shape its character and rules of behaviour.
- Governments should work with platforms to create a new and independent 'online' Ombudsman system to help interpret the network's formal and informal rules; adjudicate on contested individual cases; and advise on acceptable behaviour by users and on the evolution of the network's own rules and practices.

4 / Conclusion and what's next

Just over a decade ago, President Obama was celebrated for using social media to engage the nation in debate about aspirations for the future and mobilize millions of voters to exercise their democratic rights. Today the world is reeling from the Cambridge Analytica scandal and speculation about electoral interference in the 2016 US Presidential elections. Understandably, the tone and focus of the public debate has shifted dramatically: from hesitation and caution about regulatory interventions for fear of curbing free speech, creativity and innovation in the tech sector, to ministerial declarations that the era of self-regulation of the internet is over, culminating in the creation in the UK of a new independent regulator to enforce a statutory duty of care for companies to keep their users safe online.

Despite the enormous benefits created by online markets – innovations that have led to greater choice and incredible convenience – governments and the public have rightly begun to sharpen their focus on the market power and dominance of the large tech companies, questioning many aspects of their conduct including the way they collect and use personal data to the boundaries of their liability.

It is time for industry and government to address the real and appropriate public concerns that exist around issues ranging from algorithmic bias, through disinformation, to the mental health of children and young people online. As we've set out in this paper, a sophisticated understanding of human behaviour, including active and constructive dialogue with the public, should be at the heart of designing successful policy solutions. And in addressing these issues, governments should take an active role in catalysing and supporting new market entrants, and shaping and designing markets that really work for everybody.

This landscape will continue to shift under our feet. And so the added challenge for government and regulators is to harness the positive opportunities to build more effective and efficient markets, while also anticipating the many areas where the disruptive capabilities of digital platforms has yet to hit. Overall, government and regulators should be prepared to experiment to bridge the growing mismatch between rapid experimentation and product innovation (measured in weeks and months) and the evolution of regulatory regimes and even more so the regulatory enforcement tools we have at our disposal (measured in years). A behavioural and experimental approach will help the public sector keep pace with the rapid evolution of these markets.

This is about as precious and important a challenge as we face in society today, and one which we must ensure that our citizens can themselves be involved in fashioning. Our governments and regulators stand at the vibrant intersection between civil society and market functioning. How we respond to, and shape, the evolving character of the digital landscape is precious not just because it is pivotal to our economies, but because it is society and the human character itself that we are shaping.

Authors

Elisabeth Costa

Senior Director, Consumers, Economic Growth and Energy

Elisabeth Costa is a Senior Director at the Behavioural Insights Team, where she leads a portfolio of teams working on economic policy. Elisabeth's portfolio covers consumer markets, energy and sustainability, financial decision-making, economic growth and encouraging businesses to be more productive, as well as more diverse and inclusive. Elisabeth was the lead author of BIT's report on applying behavioural insights to regulated markets and developed the team's work on building better, fairer consumer markets, running trials with regulators, consumer bodies and industry. In particular, Elisabeth's work focuses on encouraging consumers to switch to better deals in energy markets; helping people save, repay debts and plan for retirement; and improving consumers' understanding of online contractual terms, financial products and economic information. Elisabeth regularly contributes to policy discussions and debate on digital markets and online decision-making, financial behaviour and consumer vulnerability. Prior to joining BIT, Elisabeth held senior roles at the Australian Treasury and completed her postgraduate studies at Harvard Law School.

David Halpern

Chief Executive

David Halpern is the Chief Executive of the Behavioural Insights Team. He has led the team since its inception in 2010. Prior to that, David was the founding Director of the Institute for Government and between 2001 and 2007 was the Chief Analyst at the Prime Minister's Strategy Unit. David was also appointed as the What Works National Advisor in July 2013. He supports the What Works Network and leads efforts to improve the use of evidence across government. Before entering government, David held tenure at Cambridge and posts at Oxford and Harvard. He has written several books and papers on areas relating to behavioural insights and well-being, including Social Capital (2005), the Hidden Wealth of Nations (2010), and co-author of the MINDSPACE report. In 2015 David wrote a book about the team entitled Inside the Nudge Unit: How Small Changes Can Make a Big Difference.

Appendix: Full list of recommendations

3.1. Enhancing traditional responses

Smarter disclosures

- Government, regulators and consumer groups should publish evidence-based best practice guides for businesses on how to improve online disclosures like T&Cs and data privacy policies.
- Government and regulators should set the acceptable average level of understanding that can constitute informed consent, and require businesses to conduct ongoing comprehension tests to assess and improve the information they are providing to users.
- Further research should be conducted about the most effective way to disclose how and when algorithms are being used in decisions about consumers (for example, in online loan applications), and the consequences or intentions underlying this.

Education

- Government and regulators should, in partnership with consumer bodies like Citizens Advice, develop and disseminate useful rules of thumb to help consumers navigate common challenges faced online.
- Government should fund research to design and test novel approaches to building consumers' resilience against challenges like disinformation and online fraud; and be willing to scale successful approaches.
- Industry and governments should develop supportive apps (or additions to the Personal, Social, Health and Economic (PSHE) curriculum) that can provide 'training wheels' to young people first using social media and interacting online.

Exhortation

- Government and regulators should publicly urge companies to change and improve their policies and practices, and signal willingness to intervene more strongly if change isn't satisfactory.
- To help focus action, the research community should explore whether voiced concerns are correct, including through data-led investigation and exploring causal pathways.
- Consumer groups should establish an annual, consumer-led 'Sludge' award. This would expose poor behaviour and encourage consumers to recognise and call out sludge, giving industry more incentive to change.

3.2. Choice architecture

Giving back individual control

- Foster the growth of intermediaries that can give individuals more control and ability to customise their online experiences.
- Government and regulators should work with consumer bodies to identify areas where self-exclusion tools could protect consumers from online harms, and encourage the development & take up of these tools, particularly for vulnerable consumers.

Prompts, reminders and active choices

- Industry should design and test prompts and reminders that provide consumers with active choices about how they interact with websites and platforms.
- Legal processes, such as around possible litigation or complaints, should start to differentiate between choice architectures that enable the consumer to more actively consent or shape their choices and those that don't.

Smart defaults

- Government and regulators should publicly urge companies to change and improve their policies and practices, and signal willingness to intervene more strongly if change isn't satisfactory.
- To help focus action, the research community should explore whether voiced concerns are correct, including through data-led investigation and exploring causal pathways.
- Consumer groups should establish an annual, consumer-led 'Sludge' award. This would expose poor behaviour and encourage consumers to recognise and call out sludge, giving industry more incentive to change.

3.3. Fostering sharper competition

Transparency to facilitate comparison and create accountability

- Governments, regulators, consumer organisations and trade bodies should actively encourage the wider emergence of feedback and ratings platforms.
- Industry should pay particular attention to the design of feedback systems to minimise fake reviews and reputation inflation.
- Governments should introduce transparency reports for online companies and give regulators powers to audit existing efforts and require remedial steps if they are not adequate.
- Civil society (and academia) should continue to conduct and publish research on the comparative performance of platforms in relation to welfare effects, online harms and data protection.

Building trust and confidence in digital comparison tools

- Regulators should publish league tables of switching sites covering metrics like: the average price savings achieved; number of complaints against the site; and performance as measured by mystery shopping.

Leveraging data and AI on the side of consumers

- Allow greater access and portability of data so that users – consumers, workers and businesses – can try alternative services and switch more readily.
- Government and regulators should use innovation funding and/or run Challenge Prizes to kick start and foster intermediaries that leverage data to benefit consumers.
- Regulators should investigate the case for a new consumer right to be represented online by an advocate powered by AI.
- Explore the case for extending existing work and obligations to use data to identify vulnerable customers to non-regulated markets

Next generation antitrust to support new market entrants

- Government and regulators should actively pursue data openness, including investigating the feasibility, costs and benefits of a progressive data sharing mandate.

3.4. Non-market design

Patterns of association

- Regulators should encourage, and potentially compel, platforms to put in place structural features to minimise discrimination.
- Regulators should explore how machine learning of patterns of online association, complaints mechanisms and feedback loops can identify and reduce overt and inadvertent discrimination.
- Policymakers, third sector and industry should actively explore and encourage forms of exchange and connection that foster exposure to alternative perspectives and the building of 'bridging social capital'.

Civility

- Platforms should develop and test new prompts and defaults that encourage civility between users, and specifically encourage users to reflect before posting potentially harmful content.
- Platforms should design systems that provide feedback – both from the platform and from other users – to users who choose to post harmful or offensive content. These should explain why the content was considered unacceptable. Disagreements that go to fundamental points of principle should be referable to an Ombudsman.

Who decides?

- Governments should work with industry and citizens to design new and appropriate governance mechanisms for platforms that combine expert opinion with the collective user voice to allow a platform's community of users (and possibly all citizens) to shape its character and rules of behaviour.
- Governments should work with platforms to create a new and independent 'online' Ombudsman system to help interpret the network's formal and informal rules; adjudicate on contested individual cases; and advise on acceptable behaviour by users and on the evolution of the network's own rules and practices.

1 Thaler, R. H. (2018). 'Nudge, not sludge'. *Science, 361*(6401), 431.

2 Department for Digital, Culture, Media and Sport. (2019). Online Harms White Paper. Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf*

3 Simon, H. (1986). Rationality in Psychology and Economics. *Journal of Business, 59,* S209-S224.

4 Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioural Economics. *American Economic Review, 93*(5), 1449-1475.

5 Gabaix, X. & Laibson, D. (2006). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics, 121*(2), 505-540.

6 Dijksterhuis, A., Smith, P. K., Van Baaren, R. B., & Wigboldus, D. H. (2005). The unconscious consumer: Effects of environment on consumer behavior. *Journal of Consumer Psychology, 15*(3), 193-202.

7 Akerlof, G. A. & Shiller, R. J. (2016). *Phishing for Phools: The Economics of Manipulation & Deception.* Princeton, NJ, USA: Princeton University Press; Thaler, R. H. (2018). Nudge, not sludge. *Science,* 361(6401), 431.

8 Soman, D. (1998). The illusion of delayed incentives: Evaluating future effort-money transactions. *Journal of Marketing Research, 35*(4), 427-437.

9 Kim, E. (2015). These numbers explain why Amazon wants to give so much free stuff to Prime members. *Business Insider.* 20 October. Retrieved from: *https://www.businessinsider.com/survey-shows-amazon-prime-members-are-more-loyal-2015-10?r=US&IR=T*

10 Competition and Markets Authority. (2019). Hotel booking sites to make major changes after CMA probe. Press release. 6 February. Retrieved from: https://www.gov.uk/government/news/hotel-booking-sites-to-make-major-changes-after-cma-probe

11 Pan, E., Ren, J., Lindorfer, M., Wilson, C., & Choffnes, D. (2018). Panoptispy: Characterizing Audio and Video Exfiltration from Android Applications. *Proceedings on Privacy Enhancing Technologies, 18*(4), 1-18.

12 Tiffany, K. (2018). Period-tracking apps are not for women. *Vox.* 16 November. Retrieved from: *https://www.vox.com/the-goods/2018/11/13/18079458/menstrual-tracking-surveillance-glow-clue-apple-health*

13 Levin, S. (2017). Facebook told advertisers it can identify teens feeling 'insecure' and 'worthless'. *The Guardian.* 1 May. Retrieved from: *https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens*

14 Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology, 25*(2), 333-358.

15 Competition and Markets Authority (2017). *Online Search: Consumer and firm behavior – A review of the existing literature.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/607077/online-search-literature-review-7-april-2017.pdf*

16 Soman, D. (2015). *The Last Mile: Creating Social and Economic Value from Behavioural Insights.* Toronto, Canada: University of Toronto Press. (pp. 54-61)

17 Morgan, J., Ong, D., & Zhong, Z. (2018). Location still matters: Evidence from an online shopping field experiment. *Journal of Economic Behaviour & Organisation, 146,* 43-54.

18 Competition and Markets Authority (2017). *Online Search: Consumer and firm behavior – A review of the existing literature.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/607077/online-search-literature-review-7-april-2017.pdf*

19 Competition and Markets Authority (2017). *Online search: Consumer and firm behaviour.* Report. Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/607077/online-search-literature-review-7-april-2017.pdf*

20 Competition and Markets Authority (2017). *Online Search: Consumer and firm behavior – A review of the existing literature.* Retrieved from *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/607077/online-search-literature-review-7-april-2017.pdf*

21 McCue, TJ (2018, July 30). SEO Industry Approaching $80 Billion But All You Want Is More Web Traffic. *Forbes.* 30 July. Retrieved from: *https://www.forbes.com/sites/tjmccue/2018/07/30/seo-industry-approaching-80-billion-but-all-you-want-is-more-web-traffic/#12ecbfe27337*

22 Shi, S. W., Wedel, M., & Pieters, F. G. M. (2013). 'Information acquisition during online decision making: A model-based exploration using eye-tracking data'. *Management Science, 59*(5), 1009-1026.

23 Ellison, G., & Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica, 77*(2), 427-452.

24 Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

25 Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction (TOCHI), 18*(1), 1.

26 Competition and Markets Authority (2017) *Online Search: Consumer and firm behavior – A review of the existing literature.* Retrieved from *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/607077/online-search-literature-review-7-april-2017.pdf*

27 Huck, S., & Zhou, J. (2011). *Consumer behavioural biases in competition: A survey.* Retrieved from *http://londoneconomics.co.uk/wp-content/uploads/2012/06/Consumer-behavioural-biases-in-competition-OFT1.pdf*

28 Goldfarb, A., McDevitt, R. C., Samila, S., & Silverman, B. S. (2015). The effect of social interaction on economic transactions: Evidence from changes in two retail formats. *Management Science, 61*(12), 2963-2981.

29 Jahshan, E. (2017, May 11). Impulse spending online continues to grow. Blog post. Retrieved from *https://www.retailgazette.co.uk/blog/2017/05/impulse-spending-online-grow/*

30 Richard Thaler defines 'Sludge' as 'essentially nudging for evil', either by discouraging behaviour that is in a person's best interest, or by encouraging behaviour that is 'self-defeating'. See Thaler, R. H. (2018). Nudge, not sludge. *Science, 361*(6401), 431.

31 Sunstein, Cass R., Sludge and Ordeals (November 20, 2018). Duke Law Journal, Forthcoming.

32 Letzler, R., Sandler, R., Jaroszewicz, A., Knowles, I. T., & Olson, L. (2017). Knowing When to Quit: Default Choices, Demographics and Fraud. *The Economic Journal, 127*(607), 2617-2640.

33 Böhme, R., & Köpsell, S. (2010, April). Trained to accept?: a field experiment on consent dialogs. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2403-2406). ACM; Pernice, K. (2018, April 22). *Banner Blindness Revisited: Users Dodge Ads on Mobile and Desktop.* Retrieved from: *https://www.nngroup.com/articles/banner-blindness-old-and-new-findings/*

34 Deloitte. (2017). 2017 Global Mobile Consumer Survey: US edition. The dawn of the next era in mobile (p. 12). Retrieved from: *https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-2017-global-mobile-consumer-survey-executive-summary.pdf*

35 Parris, R. (2012, March 23). *Online T&Cs longer than Shakespeare plays – who reads them?.* Blog post. Retrieved from *https://conversation.which.co.uk/technology/length-of-website-terms-and-conditions/*

36 McDonald, A. M. & Cranor, L. F. (2008). The Cost of Reading Privacy Policies. *I/S: A Journal of Law and Policy for the Information Society, 4*(3), 543-568.

37 Meeker, M. (2018, May). *Internet Trends 2018.* Report presented at the Code 2018 Conference, Rancho Palos Verdes, CA. Retrieved from *https://www.kleinerperkins.com/files/INTERNET_TRENDS_REPORT_2018.pdf*

38 Libert, T. (2018). An Automated Approach to Auditing Disclosure of Third-Party Data Collection in Website Privacy Policies. *The 2018 Web Conference,* April 23-27, 2018, Lyon, France. doi>10.1145/3178876.3186087; and Dubay, W. H. (2004). *The principles of readability.* Costa Mesa: Impact Information.

39 Department for Business, Innovation, & Skills. (2012). The 2011 Skills for Life Survey: A Survey of Literacy, Numeracy and ICT Levels in England. *BIS Research Paper Number 81.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/36000/12-p168-2011-skills-for-life-survey.pdf*

40 Obar, J. A., & Oeldorf-Hirsch, A. (2016). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society,* 1-20.

41 Salesforce. (2018). *State of the Connected Customer.* Second Edition: Retrieved from *https://www.salesforce.com/blog/2018/06/digital-customers-research.html*

42 Libert, T. (2018). An Automated Approach to Auditing Disclosure of Third-Party Data Collection in Website Privacy Policies. *The 2018 Web Conference,* April 23-27, 2018, Lyon, France. Retrieved from: *https://doi.org/10.1145/3178876.3186087*

43 Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509-514.

44 Pew Research Centre. (2016). The state of privacy in post-Snowden America. *Pew Research Centre Fact Tank*. 21 September. Retrieved from: *https://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/*

45 Cialdini, R. (1984). *Influence: The psychology of persuasion.* New York, USA: Quill.

46 Cialdini, R. (2016). *Pre-suasion: A revolutionary way to influence and persuade.* London, UK: Penguin Random House.

47 Mason, C. (2019). Martin Lewis settles lawsuit as Facebook agrees to donate £3m to anti-scam charity and launch new scam ads reporting button. *Money Savings Expert.* 23 January. Retrieved from: *https://www.moneysavingexpert.com/news/2019/01/martin-lewis-drops-lawsuit-as-facebook-agreed-to-donate-p3m-to-a/*

48 Cox, R. & de Goeij, P. (2016). Regulatory Certification, Risk Factor Disclosure and Investor Behavior. Available at SSRN: *https://ssrn.com/abstract=2822212* or *http://dx.doi.org/10.2139/ssrn.2822212*

49 Smart, L. (2018). Don't look here: do risk warnings really work?. *Insight: Opinion and Analysis from the FCA.* Retrieved from: *https://www.fca.org.uk/insight/dont-look-here-do-risk-warnings-really-work*

50 Loewenstein, G., Sunstein, C. R., & Golman, R. (2014). Disclosure: Psychology Changes Everything. *Annual Review of Economics, 6,* 391-419.

51 John, L. K., Acquisti, A., & Loewenstein, G. (2011). Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information. *Journal of Consumer Research, 37*(5), 858-873.

52 Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975-979.

53 Curzon-Price, T., Paper given to the Association of European Competition Judges, Vienna, 9 June 2017.

54 Haldane, A. G. (2017). *Productivity Puzzles.* Speech given at the London School of Economics, London, UK. 20 March. Retrieved from: *https://www.bankofengland.co.uk/-/media/boe/files/speech/2017/productivity-puzzles.pdf?la=en&hash=708C7CFD5E8417000655BA4AA0E0E873D-98A18DE*; Haldane, A. G. (2018). *The UK's Productivity Problem: Hub No Spokes.* Speech given at the Academy of Social Sciences Annual Lecture, London. 28 June. Retrieved from: *https://www.bankofengland.co.uk/-/media/boe/files/speech/2018/the-uks-productivity-problem-hub-no-spokes-speech-by-andy-haldane*; House of Commons Exiting the EU Committee. (2017). *Sectoral Report: Technology.* Retrieved from: *https://www.parliament.uk/documents/commons-committees/Exiting-the-European-Union/17-19/Sectoral%20Analyses/36-Technology-ICT-Report%20FINAL.pdf*

55 Nixon, D. (2017). Is the economy suffering from the crisis of attention?. *Bank Underground.* 24 November. Retrieved from: *https://bankunderground.co.uk/2017/11/24/is-the-economy-suffering-from-the-crisis-of-attention/*

56 Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain Drain: The mere presence of 'one's own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research, 2,* 140-154.

57 Thaler, R. H. & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy, 89*(2), 392-406; Fudenberg, D. & Levine, D. K. (2006). A dual-self model of impulse control. *American Economic Review, 96*(5), 1449-1476; Duckworth, A. L., Milkman, K. L., & Laibson, D. (2019). Beyond Willpower: Strategies for Reducing Failures of Self-Control.

58 Read, D., & van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes, 76*(2), 189-205.

59 Behavioural Insights Team. (2018). *Can behavioural insights be used to reduce risky play in online environments?* Report. Retrieved from: *https://about.gambleaware.org/media/1869/gambleaware-phase-iii-report_updated-v1.pdf*

60 Davies, R. (2018). Ladbrokes wooed problem gambler – then paid victims £1m. *The Guardian.* 17 December. Retrieved from: *https://www.theguardian.com/society/2018/dec/17/ladbrokes-wooed-problem-gambler-then-paid-victims-1m;* Wood, G. (2016). Bet365 faces legal action over delay in paying winning punter £54,000. *The Guardian.* 28 June. Retrieved from: *https://www.theguardian.com/sport/2016/jun/28/bet365-legal-action-delay-paying-punter-54000;* and Anonymous (2015). All bets are off: why bookmakers aren't playing fair. *The Guardian.* 2 August. Retrieved from: *https://www.theguardian.com/global/2015/aug/02/betting-horses-gambling-bookmakers-accounts-closed*

61 Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.* Oxford, England: Houghton Mifflin; Sunstein, C. R. (2009). *Going to extremes: How like minds unite and divide.* New York, USA: Oxford University Press; Moscovici, S., & Personnaz, B. (1980). Studies in social influence: V. Minority influence and conversion behavior in a perceptual task. *Journal of Experimental Social Psychology, 16*(3), 270-282.

62 Fishkin, J. S. (1991). *Democracy and deliberation: New directions for democratic reform* (Vol. 217). New Haven, CT: Yale University Press.

63 Halpern, D. (2005). *Social Capital. Malden,* MA, USA: Polity Press. (pp. 43-72)

64 Halpern, D. (2005). *Social Capital.* Malden, MA, USA: Polity Press. (pp. 43-195)

65 Sunstein, C. R. (2009). *Going to Extremes: How Like Minds Unite and Divide.* New York, NY, USA: Oxford University Press; and Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media.* Princeton, NJ, USA, and Oxford, UK: Princeton University Press.

66 Ross, L. & Anderson, C. A. (1982). Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments. In Kahneman, D. et al. (eds). *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge, UK: Cambridge University Press. (pp. 129-152)

67 Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2017). How People Update Beliefs about Climate Change: Good News and Bad News. *Cornell Law Review,* 102(6), 1431-1443.

68 Lord, C. G. et al. (1979). Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

69 Taber, C. S. & Lodge, M. (2006). Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science, 50*(3), 755-769.

70 Handley, I. M. et al. (2015). Quality of evidence revealing subtle gender biases in science is in the eye of the beholder. *PNAS, 112*(43), 13201-13206.

71 Matsa, K. E. & Shearer, E. (2018). News Use Across Social Media Platforms 2018. Pew Research Centre Journalism & Media. Report. 10 September. Retrieved from: *http://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/*

72 Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit Debate. *Social Networks, 50,* 6-16.

73 Quattrociocchi, W., Scala, A., & Sunstein, C. (2016) Echo chambers on Facebook. *Harvard Law School Discussion Paper No. 877.* Available at: *http://www.law.harvard.edu/programs/olin_center/papers/pdf/Sunstein_877.pdf*

74 Barbera, P., Jost, J., Nagler, J., Tucker, J., & Bonneau, R. (2015). Tweeting from Left to Right: Is online political communication more than an echo chamber? Psychological Science, 26(10), 1531-1542.

75 Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Fallin Hunzaker, M. B., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. PNAS, 115(37), 9216-9221.

76 Harwell, D. (2018). Wanted: The 'perfect babysitter.' Must pass AI scan for respect and attitude. *The Washington Post.* 23 November. Retrieved from: *https://www.washingtonpost.com/technology/2018/11/16/wanted-perfect-babysitter-must-pass-ai-scan-respect-attitude/?amp;utm_term=.efea589e7bb4&noredirect=on&utm_term=.f0e637579e9e;* Lee, D. (2018). Predictim babysitter app: Facebook and Twitter take action. *BBC News.* 27 November. Retrieved from: *https://www.bbc.co.uk/news/technology-46354276*

77 Edelman, B., Luca, M., & Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics,* 9(2), 1-22.

78 Ahuja, R. & Lyons, R. C. (2019). The Silent Treatment: LGBT Discrimination in the Sharing Economy. *Oxford Economic Papers,* gpz025.

79 Offer, A. (1997). Between the Gift and the Market: *The Economy of Regard. The Economic History Review, 50*(3), 450-476.

80 Leavenworth, S. (2016). China's 'naked loans' force female students to bare all in return for more cash. *The Guardian.* 15 June. Retrieved from: *https://www.theguardian.com/world/2016/jun/15/chinas-naked-loans-force-female-students-to-bare-all-in-return-for-more-cash*

81 McAuliffe, W. H. B., Forster, D. E., Pedersen, E. J., & McCullough, M. E. (2018). Experience with anonymous interactions reduces intuitive cooperation. *Nature Human Behaviour, 2,* 909-914.

82 Zimbardo, P. (1969). The human choice: Individuation, reason, and order vs. deindividuation, impulse, and chaos. In Arnold, W. J. & Levine, D. (eds). *Nebraska Symposium on Motivation.* Lincoln, NE, USA: University of Nebraska Press. (pp. 237-307); Diener, E., Fraser, S. C., Beaman, A. L., & Kelem, R. T. (1976). Effects of deindividuation variables on stealing among Halloween trick-or-treaters. *Journal of Personality & Social Psychology, 33,* 178-18.

83 Yee, N. & Bailenson, J. (2007) The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research, 33*(3), 271-290.

84 Amnesty International. (2016). Australia: Poll reveals alarming impact of online abuse against women. News Release. Retrieved from: *https://www.amnesty.org.au/australia-poll-reveals-alarming-impact-online-abuse-women/*

85 Verduyn, P., Ybarra, O., Résibois, M., Jonides, J., & Kross, E. (2017). Do social network sites enhance or undermine subjective wellbeing? A critical review. Social Issues and Policy Review, 11(1), 274-302.

86 Kelly, Y., Zilanawala, A., Booker, C., & Sacker, A. (2018). Social media use and adolescent mental health: findings from the UK Millennium Cohort Study. EClinicalMedicine, 6, 59-68.

87 Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2019). The Welfare Effects of Social Media. NBER Working Paper No. 25514. Retrieved from: *https://www.nber.org/papers/w25514*

88 Royal Society for Public Health. (2017). #StatusOfMind: Social media and young people's mental health and wellbeing. Report. Retrieved from: *https://www.rsph.org.uk/uploads/assets/uploaded/62be270a-a55f-4719-ad668c2ec7a74c2a.pdf*

89 Behavioural Insights Team. (2018). We spoke to teenagers about tech this is what they told us. Blog post. Retrieved from: *https://www.bi.team/blogs/we-spoke-to-teenagers-about-tech-this-is-what-they-told-us/*

90 US data based on emergency department visits data from; Mercado, M. C., Holland, K., Leemis, R. W., Stone, D. M., & Wang, J. (2017). Trends in emergency department visits for nonfatal self-inflicted injuries among youth aged 10 to 24 years in the United States, 2001-2015. *Jama, 318*(19), 1931-1933.

91 UK data based on GP primary care records from; Morgan, C., Webb, R. T., Carr, M. J., Kontopantelis, E., Green, J., Chew-Graham, C. A., ... & Ashcroft, D. M. (2017). Incidence, clinical management, and mortality risk following self harm among children and adolescents: cohort study in primary care. *bmj, 359*, j4351.

92 UK data on emotional disorders from NHS Digital (2017) Mental Health of Children and Young People in England. Available at: *https://files.digital.nhs.uk/F6/A5706C/MHCYP%202017%20Summary.pdf;* US data on major depressive episodes from the National Survey on Drug Use and Health. Available at: *https://pdas.samhsa.gov/#/survey/NSDUH-2017-DS0001*

93 Kelly, Y., Zilanawala, A., Booker, C., & Sacker, A. (2018). Social media use and adolescent mental health: findings from the UK Millennium Cohort Study. *EClinicalMedicine, 6,* 59-68.

94 House of Commons Digital, Culture, Media and Sport Committee. (2019). *Disinformation and 'fake news': Final Report.* HC 1791. Available at: *https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf*

95 Pennycook, G. & Rand, D. G. (2018). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition,* in press. Retrieved from: *https://www.sciencedirect.com/science/article/pii/S001002771830163X*

96 Yin, L., Roscher, F., Bonneau, R., Nagler, J., & Tucker, J. A. (2018). Your Friendly Neighborhood Troll: The Internet Research Agency's Use of Local and Fake News in the 2016 US Presidential Campaign. *NYU SMaPP Data Report 2018.1.* Retrieved from: *https://s18798.pcdn.co/smapp/wp-content/uploads/sites/1693/2018/11/SMaPP_Data_Report_2018_01_IRA_Links.pdf*

97 Behavioural Insights Team. (2018). The Loyalty Penalty: why do disengaged customers pay more for essential services? Blog post. Retrieved from: *https://www.bi.team/blogs/the-loyalty-penalty-why-do-disengaged-customers-pay-more-for-essential-services/*

98 Behavioural Insights Team. (2018). How can we end the loyalty penalty? Blog post. Retrieved from: *https://www.bi.team/blogs/how-can-we-end-the-loyalty-penalty/*

99 Citizens Advice. (2018). *A price of one's own: An investigation into personalised pricing in essential markets.* Retrieved from: *https://www.citizensadvice.org.uk/Global/CitizensAdvice/Consumer%20publications/Personalised%20Pricing%20Report%202018.pdf*

100 Competition & Markets Authority. (2018). *Tackling the loyalty penalty: Response to a super-complaint made by Citizens Advice on 28 September 2018.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/766795/Loyalty_Penalty.pdf*

101 Citizens Advice. (2018). *Excessive prices for disengaged consumers: A super-complaint to the Competition & Markets Authority.* Retrieved from: *https://www.citizensadvice.org.uk/Global/CitizensAdvice/Consumer%20publications/Super-complaint%20-%20Excessive%20prices%20for%20disengaged%20consumers%20(1).pdf*

102 OECD. (2018). *Personalised Pricing in the Digital Era.* DAF/COMP(2018)13. Retrieved from: *http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP(2018)13&docLanguage=En*

103 European Commission. (2018). *Consumer market study on online market segmentation through personalised pricing/offers in the European Union.* Final Report. Retrieved from: *https://ec.europa.eu/info/sites/info/files/aid_development_cooperation_fundamental_rights/aid_and_development_by_topic/documents/synthesis_report_online_personalisation_study_final_0.pdf*

104 Mattioli, D. (2012). On Orbitz, Mac Users Steered to Pricier Hotels. *The Wall Street Journal.* 23 August. Retrieved from: *https://www.wsj.com/articles/SB10001424052702304458604577488822667325882*

105 Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). Measuring Price Discrimination and Steering on E-commerce Web Sites. *Proceedings of the 2014 Conference on Internet Measurement Conference,* Vancouver, BC, Canada, 5-7 November. Retrieved from: *https://dl.acm.org/citation.cfm?doid=2663716.2663744*

106 Mohammed, R. (2017). How Retailers Use Personalized Prices to Test What You're Willing to Pay. Harvard Business Review. 20 October. Retrieved from: *https://hbr.org/2017/10/how-retailers-use-personalized-prices-to-test-what-youre-willing-to-pay*

107 Allen, M. (2018). Health Insurers Are Vacuuming Up Details About You – And It Could Raise Your Rates. *NPR.* 17 July. Retrieved from: *https://www.npr.org/sections/health-shots/2018/07/17/629441555/health-insurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates*

108 Cairns, A. (2019). Why AI is failing the next generation of women. *World Economic Forum.* 18 January. Retrieved from: *https://www.weforum.org/agenda/2019/01/ai-artificial-intelligence-failing-next-generation-women-bias/*

109 Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016). How we analysed the COMPAS recidivism algorithm. *ProPublica.* 23 May. Retrieved from: *https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm;* Dressel, J. & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances, 4*(1), eaao5580;

110 Corbett-Davies, S., Pierson, E., Feller, A., & Goel, S. (2016). A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. *The Washington Post.* 17 October. Retrieved from: *https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/?utm_term=.cdc75481a70a*; Yong, E. (2018). A popular algorithm is no better at predicting crimes than random people. *The Atlantic.* 17 January. Retrieved from: *https://www.theatlantic.com/technology/archive/2018/01/equivant-compas-algorithm/550646/*

111 McGregor, J. (2018). Why robots aren't likely to make the call on hiring you anytime soon. *The Washington Post.* 11 October. Retrieved from: *https://www.washingtonpost.com/business/2018/10/11/why-robots-arent-likely-make-call-hiring-you-anytime-soon/?noredirect=on&utm_term=.a0cc89375b4a;* Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters.* 10 October. Retrieved from: *https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G*

112 Caliskan, A., Bryson, J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183-186.

113 Department for Education. (2018). *Revised A level and other 16-18 results in England, 2016/2017.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/676389/SFR03_2018_Main_text.pdf*

114 Richter, F. (2019). Female workers in the tech industry: The tech world is still a man's world. *Statista.* 8 March. Retrieved from: *https://www.statista.com/chart/4467/female-employees-at-tech-companies/*

115 World Economic Forum. (2018). Global Gender Gap Report 2018: Assessing Gender Gaps in Artificial Intelligence. Retrieved from: *http://reports.weforum.org/global-gender-gap-report-2018/assessing-gender-gaps-in-artificial-intelligence/*

116 Chung, A. (2019). How Automated Tools Discriminate Against Black Language. *OneZero.* 28 February. Retrieved from: *https://onezero.medium.com/how-automated-tools-discriminate-against-black-language-2ac8eab8d6db*

117 Miller, A. P. (2018). Want Less-Biased Decisions? Use Algorithms. *Harvard Business Review.* 26 July. Retrieved from: *https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms*

118 Sharma, K. (2018). Can We Keep Our Biases From Creeping Into AI?. *Harvard Business Review.* 9 February. Retrieved from: *https://hbr.org/2018/02/can-we-keep-our-biases-from-creeping-into-ai*; Courtland, R. (2018). Bias detectives: the researchers striving to make algorithms fair. *Nature, 558,* 357-360. Retrieved from: *https://www.nature.com/articles/d41586-018-05469-3*

119 Staley, O. (2017). Harvard economist Iris Bohnet on how to design a hiring process that's fair for everyone. *Quartz at Work*. 17 October. Retrieved from: *https://qz.com/work/1080530/harvard-economist-iris-bohnet-says-to-eliminate-bias-its-easier-to-change-systems-than-change-people/*

120 Mayer-Schonberger, V. & Ramge, T. (2018). Reinventing Capitalism in the Age of Big Data. New York, USA: (Basic Books).

121 Grassegger, H. (2018). Facebook says its 'voter button' is good for turnout. But should the tech giant be nudging us at all? The Guardian. 15 April. Retrieved from: *https://www.theguardian.com/technology/2018/apr/15/facebook-says-it-voter-button-is-good-for-turn-but-should-the-tech-giant-be-nudging-us-at-all*

122 Gabaix, X. & Laibson, D. (2006). Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *The Quarterly Journal of Economics, 121*(2), 505-540.

123 Ducrot, P., Julia, C., Méjean, C., Kesse-Guyot, E., Touvier, M., Fezeu, L. K., ... & Péneau, S. (2016). Impact of different front-of-pack nutrition labels on consumer purchasing intentions: a randomized controlled trial. *American Journal of Preventive Medicine, 50*(5), 627-636.

124 Money Advice Service, Behavioural Insights Team, & Ipsos MORI. (2018). *A behavioural approach to managing money: Ideas and results from the Financial Capability Lab*. Report. Available at: *https://www.fincap.org.uk/en/articles/the-financial-capability-lab*

125 Behavioural Insights Team. (2017). Improving engagement with pension decisions: The results from three randomised controlled trials. Report for the UK Government's Pension Wise service. Available at: *https://www.bi.team/publications/improving-engagement-with-pension-decisions-the-results-from-three-randomised-controlled-trials/*

126 Ungemach, C., Camilleri, A. R., Johnson, E. J., Larrick, R. P., & Weber, E. U. (2018). Translated Attributes as Choice Architecture: Aligning Objectives and Choices Through Decision Signposts. *Management Science, 64*(5), 1975-2471.

127 Information Commissioner's Office. (n.d). *The right to be informed: What methods can we use to provide privacy information?* Retrieved from: *https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/the-right-to-be-informed/what-methods-can-we-use-to-provide-privacy-information/*

128 Competition and Markets Authority. (2018). Online dating giant vows clearer path to love. Press Release. 13 June. Retrieved from: *https://www.gov.uk/government/news/online-dating-giant-vows-clearer-path-to-love*

129 Behavioural Insights Team. (2018). The impact of improved transparency of foreign money transfers for consumers and SMEs. Final report. Retrieved from: *http://www.behaviouralinsights.co.uk/publications/the-impact-of-improved-transparency-of-foreign-money-transfers-for-consumers-and-smes/*

130 Hicks, J. J. (2001). The strategy behind Florida's 'truth' campaign. *Tobacco Control, 10,* 3-5. Retrieved from: *https://tobaccocontrol.bmj.com/content/10/1/3*

131 Friedman, Z. (2017). 9 Simple Money Rules All On 1 Index Card. *Forbes*. 9 March. Retrieved from: *https://www.forbes.com/sites/zackfriedman/2017/03/09/9-money-rules-index-card/#63481d2b2c09*

132 Competition and Markets Authority. (2017). Digital comparison tools market study. Final report. Retrieved from: *https://assets.publishing.service.gov.uk/media/59c93546e5274a77468120d6/digital-comparison-tools-market-study-final-report.pdf*

133 Behavioural Insights Team. (2019). Annual Report 2017-2018. Retrieved from: *https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf*

134 Kahne, J. & Bowyer, B. (2017). Educating for Democracy in a Partisan Age: Confronting the Challenges of Motivated Reasoning and Misinformation. *American Educational Research Journal, 54(1),* 3-34.

135 Nsangi, A., Semakula, D., Oxman, A. D., Austvoll-Dahlgren, A., Oxman, M., Rosenbaum, S., Morelli, A., Glenton, C., Lewin, S., Kaseje, M., Chalmers, I., Fretheim, A., Ding, Y., & Sewankambo, N. K. (2017). Effects of the Informed Health Choices primary school intervention on the ability of children in Uganda to assess the reliability of claims about treatment effects: a cluster-randomised controlled trial. *Lancet,* 390(10092), 374-388.

136 *https://www.bbc.com/ownit*

137 Office of the Chief Medical Officer. (2019). United Kingdom Chief Medical Officers' commentary on 'Screen-based activities and children and young people's mental health and psychosocial wellbeing: a systematic map of reviews'. Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/777026/UK_CMO_commentary_on_screentime_and_social_media_map_of_reviews.pdf*

138 Department for Business, Energy & Industrial Strategy. (2018). *Modernising Consumer Markets: Consumer Green Paper.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/699937/modernising-consumer-markets-green-paper.pdf*

139 BBC. (2019). Mental health: UK could ban social media over suicide images, minister warns. 27 January. Retrieved from: *https://www.bbc.co.uk/news/uk-47019912*

140 Mosseri, A. (2019). Changes We're Making to Do More to Support and Protect the Most Vulnerable People who Use Instagram. Instagram Information Centre. Press Release. Retrieved from: *https://instagram-press.com/blog/2019/02/07/changes-were-making-to-do-more-to-support-and-protect-the-most-vulnerable-people-who-use-instagram/*

141 Mosseri, A. (2019). Changes We're Making to Do More to Support and Protect the Most Vulnerable People who Use Instagram. Instagram Information Centre. Press Release. Retrieved from: *https://instagram-press.com/blog/2019/02/07/changes-were-making-to-do-more-to-support-and-protect-the-most-vulnerable-people-who-use-instagram/*

142 Murray, N. (2017). *The Shopper Stopper: An evaluation.* Report. Retrieved from: *https://www.moneyandmentalhealth.org/wp-content/uploads/2017/10/Shopper-Stopper-Evaluation-Report.pdf*

143 Ledward, N. (2018). Block gambling transactions from your Monzo account. *Monzo.* 19 June. Retrieved from: *https://monzo.com/blog/2018/06/19/gambling-block-self-exclusion/*

144 Schraer, N. (2019). Bank of Scotland, Halifax, Lloyds and Santander planning to let you block gambling transactions. *Money Savings Expert.* 4 January. Retrieved from: *https://www.moneysavingexpert.com/news/2019/01/more-banks-plan-gambling-restrictions/*

145 Tyers, R., Sweeney, M., & Moon, B. (2019). Harnessing behavioural insights to encourage consumer engagement in the British energy market: results from a field trial. *Journal of Behavioral and Experimental Economics.* In press, retrieved from: *https://doi.org/10.1016/j.socec.2019.03.001*

146 Ofgem. (2017). Results from the Cheaper Market Offers Letter Trial. Retrieved from: *https://www.ofgem.gov.uk/publications-and-updates/results-cheaper-market-offers-letter-trial*

147 Ofgem. (2018). Open letter: Active choice collective switch trial – early findings. Retrieved from: *https://www.ofgem.gov.uk/publications-and-updates/open-letter-active-choice-collective-switch-trial-early-findings*

148 Perron, N. J., Dao, M. D., Kossovsky, M. P., Miserez, V., Chuard, C., Calmy, A., & Gaspoz, J. M. (2010). Reduction of missed appointments at an urban primary care clinic: a randomised controlled study. *BMC Family Practice, 11*(1), 79.

149 Hallsworth, M., Berry, D., Sanders, M., Sallis, A., King, D., Vlaev, I., et al. (2015). Stating Appointment Costs in SMS Reminders Reduces Missed Hospital Appointments: Findings from Two Randomised Controlled Trials. *PLoS ONE,* 10(9): e0137306. Retrieved from: *https://doi.org/10.1371/journal.pone.0137306*

150 Behavioural Insights Team. (2019). Annual Report 2017-2018. Retrieved from: *https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf*

151 Behavioural Insights Team. (2019). Annual Report 2017-2018. Retrieved from: *https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf*

152 Behavioural Insights Team. (2019). Annual Report 2017-2018. Retrieved from: *https://www.bi.team/wp-content/uploads/2019/01/Annual-update-report-BIT-2017-2018.pdf*

153 Behavioural Insights Team. (2018). Financial Capability Lab – detailed findings. Retrieved from: *http://www.behaviouralinsights.co.uk/wp-content/uploads/2018/05/Financial-Capability-Lab%E2%80%93Detailed-Findings%E2%80%93May18-1.pdf*

154 Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2019). Active Choice, Implicit Defaults, and the Incentive to Choose. *NBER Working Paper No. 25473.* Retrieved from: *https://www.nber.org/papers/w25473*

155 Harris, J. (2018). Screen Time is here and people are shocked by how long they spend on their phone. *The Independent.* 25 September. Retrieved from: *https://www.independent.ie/world-news/and-finally/screen-time-is-here-and-people-are-shocked-by-how-long-they-spend-on-their-phone-37352684.html*

156 Digital Competition Expert Panel. (2019). *Unlocking digital competition.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf*

157 Gabaix, X. & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics, 121*(2), 505-540.

158 Luca, M. (2016). Reviews, reputation, and revenue: The case of Yelp. com. *Harvard Business School Working Papers 12-016,* Retrieved from: *https://www.hbs.edu/faculty/Publication%20Files/12-016_a7e4a5a2-03f9-490d-b093-8f951238dba2.pdf*

159 Taylor, M. (2017). *Good Work: The Taylor Review of Modern Working Practices.* Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/627671/good-work-taylor-review-modern-working-practices-rg.pdf*

160 HM Government. (2018). *Good work: A response to the Taylor Review of Modern Working Practices.* Retrieved from: *https://www.gov.uk/government/publications/good-work-the-taylor-review-of-modern-working-practices*

161 Butler, O. (2017). I Made My Shed the Top Rated Restaurant On TripAdvisor. *Vice.* 6 December. Retrieved from: *https://www.vice.com/en_uk/article/434gqw/i-made-my-shed-the-top-rated-restaurant-on-tripadvisor;* Low, V. (2017). Prankster Oobah Butler tricks TripAdvisor into thinking his shed is London's best restaurant. *The Times.* 7 December. Retrieved from: *https://www.thetimes.co.uk/article/prankster-oo-bah-butler-cooks-up-top-restaurant-the-shed-at-dulwich-fx6hfwj89*

162 Filippas, A., Horton, J. J., & Golden, J. (2018). Reputation Inflation. *Proceedings of the 2018 ACM Conference on Economics and Computation,* Ithaca, NY, USA, June 18 - 22, Retrieved from: *https://dl.acm.org/citation.cfm?id=3219222*

163 Filippas, A., Horton, J. J., & Golden, J. (2018). Reputation Inflation. *Proceedings of the 2018 ACM Conference on Economics and Computation,* Ithaca, NY, USA, June 18 - 22, Retrieved from: *https://dl.acm.org/citation.cfm?id=3219222*

164 Federal Trade Commission. (2019). FTC Brings First Case Challenging Fake Paid Reviews on an Independent Retail Website. Press Release. 26 February. Retrieved from: *https://www.ftc.gov/news-events/press-releases/2019/02/ftc-brings-first-case-challenging-fake-paid-reviews-independent*

165 Hall, C. (2017). Steam has a new roadblock to combat so-called 'review bombs'. Polygon, 21 November. Retrieved from: *https://www.polygon.com/2017/11/21/16687498/steam-review-bombs-helpful-system-trolls*

166 Alvarado, F. (2018). The Hot New Millennial Trend Is Scamming Amazon for Big Money. *Vice.* 11 October. Retrieved from: *https://www.vice.com/en_us/article/8xjbng/the-hot-new-millennial-trend-is-scam-ming-amazon-for-big-money*

167 Arthur, C. (2013). How low-paid workers at 'click farms' create appearance of online popularity. *The Guardian.* 2 August. Retrieved from: *https://www.theguardian.com/technology/2013/aug/02/click-farms-appearance-online-popularity*

168 Competition and Markets Authority. (2017). *Digital comparison tools market study.* Final report. Retrieved from: *https://assets.publishing.service.gov.uk/media/59c93546e5274a77468120d6/digi-tal-comparison-tools-market-study-final-report.pdf*

169 Digital Competition Expert Panel. (2019). Unlocking digital competition. Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf*

170 Curzon-Price, T., Paper given to the Association of European Competition Judges, Vienna, 9 June 2017.

171 Richings, S. (2017). Making Better Use Of Data: Identifying Customers In Vulnerable Situations. *UK Regulators Network.* Blog post. Retrieved from: *https://www.ukrn.org.uk/making-better-use-of-data-identi-fying-customers-in-vulnerable-situations/*

172 Digital Competition Expert Panel. (2019). Unlocking digital competition. Retrieved from: *https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf*

173 Mayer-Schonberger, V. & Ramge, T. (2018). *Reinventing Capitalism in the Age of Big Data.* London, UK: John Murray (Publishers).

174 *https://www.ofgem.gov.uk/consumers/household-gas-and-electricity-guide/how-switch-energy-supplier-and-shop-better-deal/ofgem-disengaged-customer-database*

175 See Belli, G. (2017). How Many Jobs Are Found Through Networking, Really?. *PayScale.* 6 April. Retrieved from: *https://www.payscale.com/career-news/2017/04/many-jobs-found-networking*; and Fisher, C. (n.d.) Top 5 Ways People Find Their Jobs. *Career Trend.* Retrieved from: *https://careertrend.com/top-5-ways-people-jobs-7797.html*

176 Breeden, A. (2019). Facebook Group of French Journalists Harassed Women for Years. *The New York Times.* 12 February. Retrieved from: *https://www.nytimes.com/2019/02/12/world/europe/la-ligue-du-lol-sexual-harassment.html*

177 Marantz, A. (2018). Reddit and the struggle to detoxify the internet: How do we fix life online without limiting free speech? *The New Yorker.* 19 March. Retrieved from: *https://www.newyorker.com/magazine/2018/03/19/reddit-and-the-struggle-to-detoxify-the-internet*

178 Tirole, J. (2017). *Economics for the Common Good.* Princeton, NJ, USA: Princeton University Press.

179 Fishkin, J. S., & Luskin, R. C. (2005). Experimenting with a democratic ideal: Deliberative polling and public opinion. *Acta politica, 40*(3), 284-298.

180 Fillmore-Patrick, H. (2013). The Iceland experiment (2009-2013): a participatory approach to constitutional reform. DPC Policy Note, 2.

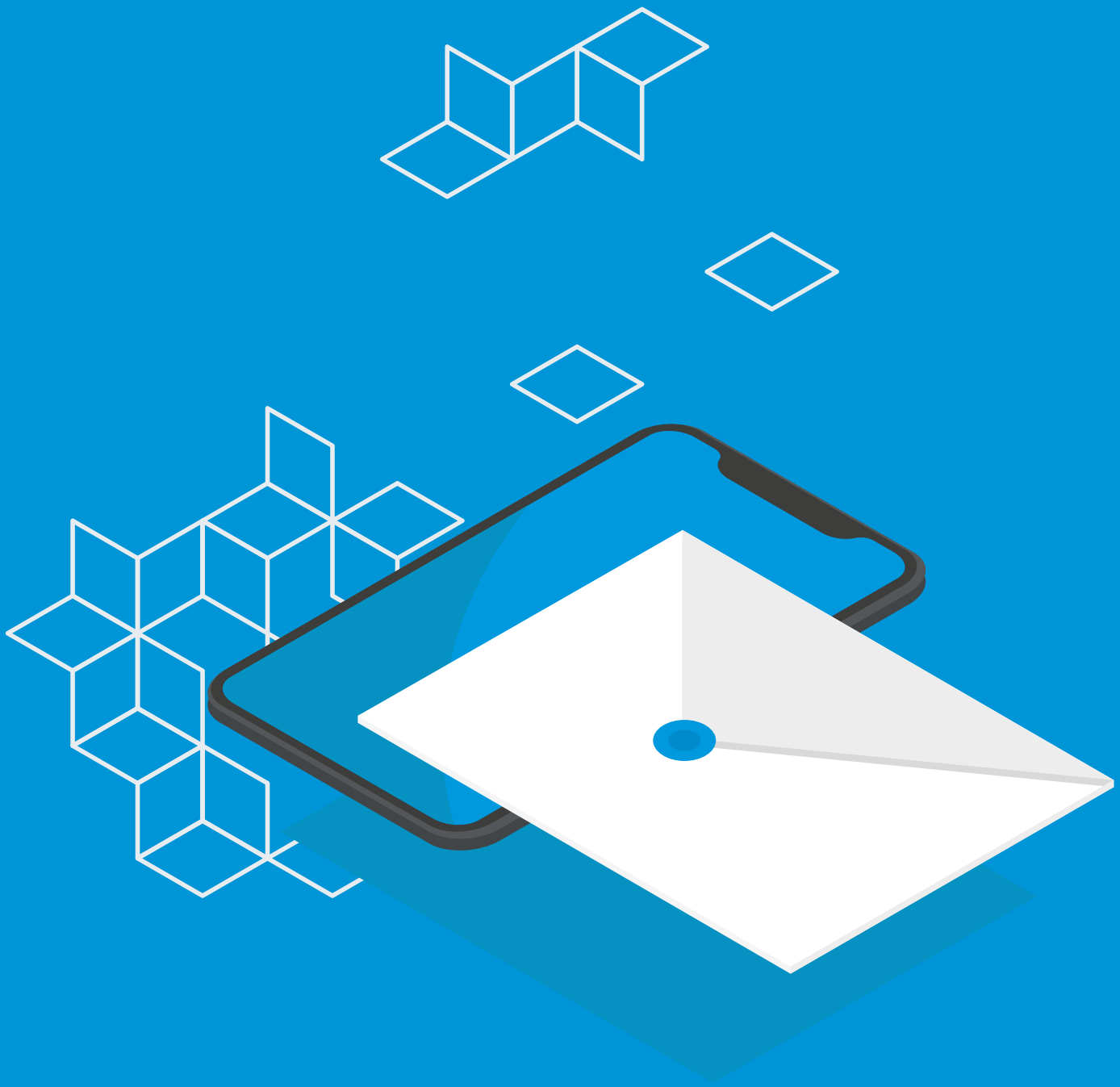
BEHAVIOURAL EXCHANGE

London 2019

Stay in touch

Contact the team at **www.bi.team/contact**

Stay up to date with our latest work, findings and job vacancies. Subscribe for email updates at **www.bi.team/subscribe**



The Behavioural Insights Team

www.bi.team

4 Matthew Parker Street
London SW1H 9NP

THE
BEHAVIOURAL
INSIGHTS
TEAM