Media plurality & online intermediation of news consumption

An economic assessment of potential theories of harm and proposals for evidence gathering

Prepared for

Ofcom

Riverside House
2a Southwark Bridge Road
London, SE1 9HA

Prepared by

Dr Oliver Latham, Gaber Burnik,
Dr Rafael Jiménez Durán,
Hasnain Khaki

Charles River Associates
8 Finsbury Circus
London EC2M 7EA

Date: 26 May 2022

CRA Charles River Associates

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

> *"Political thought is representative. I form an opinion by considering a given issue from different viewpoints, by making present to my mind the standpoints of those who are absent; that is, I represent them... The more people's standpoints I have present in my mind while I am pondering a given issue, and the better I can imagine how I would feel and think if I were in their place, the stronger will be my capacity for representative thinking and the more valid my final conclusions, my opinion. It is this capacity for an "enlarged mentality" that enables us to judge"*
>
> *Hannah Arendt, Truth and Politics (1967)*

A well-informed public is key to a well-functioning democracy and a pluralistic media sector has an important role to play in ensuring that voters are exposed to accurate and diverse news coverage so they can make well-informed decisions and hold government to account. This view, which is firmly supported by analysis in political science and economics, underpins the case for regulation to ensure media plurality.

However, the way that consumers access news has changed dramatically in recent years and the question is whether the consumption of news via online intermediaries, most notably through social media platforms like Facebook and Twitter or Search Engines like Google, raises new issues. We have been asked by Ofcom to distil the messages from the academic literature on the implications of increased consumption of news via online intermediaries for media plurality issues; and to identify evidence gathering initiatives to better understand these issues, taking account of gaps identified in this literature.[1]

What are the new issues raised by intermediated consumption?

The impact of intermediation on plurality issues is multi-faceted. On the one hand, exposing users to content from multiple outlets, splitting news consumption up into chunks, and helping politicians speak directly to voters might soften some of the historic concerns focused on individuals or organisations having agenda setting power or amassing sufficient control of news *output* to shape views in society.

On the other, intermediated consumption of news raises new issues, which we place into the following main categories, noting that some of them are interlinked:

- **Concerns around increased polarisation.** Online intermediaries may, deliberately or inadvertently, promote content that polarises their users (in the sense of causing consumers to be less sympathetic to ideological opponents, less likely to report being exposed to news that contradicted their beliefs and increasing the dispersion in political views). This risks a negative impact on media plurality if consumers are more entrenched and less willing or able to consider multiple viewpoints.

[1] All views expressed in the report are our own and are not necessarily shared by either CRA or Ofcom. Where we express views related to topics where we have previously done consulting work we alert the reader using footnotes. The authors bios are provided in the Appendix.

- **Concerns around "echo chambers" and/or filter bubbles.** Relatedly, consumption of news via online intermediaries may cause users to self-select news that reinforces their existing views ("echo chambers") or online intermediaries' algorithms may disproportionately promote such like-minded content ("filter bubbles"). In each case, the concern is that users are exposed to a narrower set of views potentially becoming less well informed and contributing to the polarisation effects above.
- **Concerns around the spread of misinformation.** Consumption via online intermediaries may contribute to the spread of misinformation or "fake news", especially if content can be spread by spontaneous sharing between users.
- **Concerns around algorithmic bias.** There are concerns that online intermediaries' algorithms could be biased due to financial considerations (e.g., sending more traffic to publishers in whom they have a financial interest or who make use of other revenue generating services) or in other ways (e.g., by sending more traffic to publishers of a particular ideological persuasion).
- **Concerns around negative impacts on news production.** The rise of intermediated consumption may make it more difficult for news creators to monetise their content and diminish the incentives to produce quality content. This could be due to fundamental issues with the technology (e.g., that consumption via online intermediaries reduces brand distinctiveness) or active decisions by platforms (e.g., if algorithms excessively reward "freshness" such that traffic flows to copycat articles rather than the original publisher of the content).

In each case, a recurring issue is that there may be misalignment between the incentives of online intermediaries and social welfare. For example, it might be that online intermediaries' profit maximizing incentives are to maximise consumer engagement and advertising revenue, but that this has negative social effects by contributing to one or more of the harms above. These issues may not require active decisions by platforms, but could rather be independently generated by algorithms tasked with optimising for particular objectives such as user engagement or attention. Online intermediaries may also create a new point of concentration in the value chain, even if there is low concentration amongst the publishers, which intensifies the issues stemming from then online intermediaries' incentives.

To the extent that these concerns materialise, they will have important implications media plurality and, ultimately, for the well-functioning democratic process.

What does the existing literature tell us about these concerns?

We have conducted a systematic review of academic studies of these concerns. We place particular emphasis on empirical literature in economics and political science, but also cover literature reviews and theoretical studies. We have identified 79 relevant studies which we then categorise in terms of the hypotheses they test, the methodology and data used, the platforms and geographies they cover, and their conclusions. We then take a view on the findings from the literature and the key evidence gaps. We focus attention on those studies we consider most credible and, in particular, those whose design allows them to speak to causal relationships as opposed to mere correlations.

Our key findings on the current state of the evidence are as follows.

Evidence of the impact of intermediated news consumption on polarisation:

- The literature provides evidence that consuming news via social media can increase polarisation. The majority of studies we have reviewed on this topic make this finding as do the two studies employing experimental designs which allow them to identify causal effects: Allcott et al. (2020) and Levy (2021). For example, Allcott et. al. found that paying a randomly-selected "treatment group" of Facebook users in the US to stop using Facebook resulted in these users being less polarised, but also less informed about news event, than a "control group" who did not receive such a payment.
- By contrast, there is no evidence of polarising effects of consuming news via search engines. While no equivalent experimental studies exist, studies based on observational data show search results to be less personalised and more likely to link to "mainstream" news sources. They show also that consumers using search to access news consume a more ideologically balanced portfolio of news.
- A key caveat, however, is that the studies with the most compelling research designs are US focussed and no equivalent experimental evidence exists for the UK. The UK has its own features both in terms of the dimensions of the political debate and the regulatory framework. For example, stronger regulation of broadcast news and a greater role for public service broadcasters.
- We note that there are other research studies which are more sceptical of the link between social media and polarisation. For instance, Boxell, Gentzkow, and Shapiro (2021) shows that increases in polarisation in the US predate the rise of social media, and other studies find these trends are not mirrored in Europe. Further, older citizens, who make the least use of social media, are the most polarized. Other studies note that consumption through online intermediaries is associated with users consuming a broader mix of news outlets than was the case historically. We do not see these studies as necessarily inconsistent, however. One can simultaneously accept experimental evidence that the causal effect of social media is to increase polarization while acknowledging there are other factors at work and that the extreme scenario of social media *only* showing users content that confirms their beliefs is not that widespread.
- Other unanswered questions include the precise mechanism by which polarisation occurs, the extent to which these issues are Facebook-specific or apply also to other social media platforms, and the extent to which these effects can be mitigated by targeted interventions. For example, not using Facebook at all may reduce polarisation and news knowledge, but perhaps targeted changes to platform design or algorithms could allow the benefits of increased news knowledge without increasing polarisation.

Evidence on the extent to which intermediated news consumption result in "echo chamber" or "filter bubble" outcomes. One potential mechanism for polarising effects is that intermediated news consumption, in particular via social media, creates echo chambers or filter bubbles whereby consumers do not see news that challenges their existing beliefs. Our review of the literature is that:

- **There are robust reasons to expect echo chamber and filter bubble effects to arise.** The literature establishes that consumers tend to seek out news that conforms to their prior beliefs, either because they have an intrinsic preference for such content or because they believe such news is higher quality. Experimental studies have found that consumers are more likely to engage with news aggregators that show them more partisan content that matches their prior beliefs (Bryanov et al. (2020).

- **The empirical literature supports consumption via social media inducing echo chamber/filter bubble effects. However, the extreme case of consumers *only* seeing content that supports their views is rare.** Levy (2021) shows that inducing consumers to "like" news outlets they disagree with on Facebook, results in them seeing less polarized news content, but that these "counter-attitudinal" sources are still less likely to appear on their news feed. Fletcher, Kalogeropoulos, and Nielsen (2021) also find using UK data that, even when consumption via social media increases the breadth of content consumed, it also increases the consumption of more partisan outlets (partly explained by the fact that BBC, considered less partisan, accounts for about half of direct online news consumption). However, while there appear to be echo chamber/filter bubble effects at the margin, the proportion of consumers who see *only* news content they agree with online is limited (for instance, Fletcher, Robertson and Nielsen (2021) find that the percentage of those who consume news only from left-wing or right-wing online outlets is sub 10%) and most social media users see and consume at least some content from news outlets of a different ideological persuasion.
- **The evidence in respect of search engines is less detailed but does not support them having filter bubble effects.** As noted above, search results appear to be less individualised and more likely to link to mainstream content. This means that the echo chamber concern is less relevant than for social media platforms.
- **The literature differs as to whether the primary issue is demand-side effects (user-induced echo chambers) or supply-side effects (algorithm-induced filter bubbles), with evidence for both effects.** There is causal evidence that algorithm design impacts the pattern of consumption on Facebook. But other studies show that individual actions matter as well. Bakshy, Messing and Adamic (2015), for example, find that the network of friends is the primary driver of what content a user sees on social media.
- **Echo chamber/filter bubble effects may be moderated by consumption of news through other channels such as TV or print, but the size of these moderating effects is unclear.** The theoretical economic literature shows how access to alternative news sources and an ability to "cross check" information can act to discipline biased information transmission and most social media users still see a range of outlets. However, while the literature establishes that most users see some content that is likely to differ from their pre-existing views, it does not establish the strength of these moderating effects and the evidence from studies of polarisation indicates that it is insufficient to remove these effects entirely.
- **There remain important gaps in the evidence.** As for the evidence on polarisation, the key caveat is that most of the evidence is focussed on the US. A further key issue is that most analysis of echo chambers has focussed on the diversity of news consumption at the *outlet* level (i.e., whether users see news from both left- and right-wing outlets). This is concerning because, while intermediation may increase the number of news outlets one is exposed to, these might still present one-sided views and not necessarily increase diversity in terms of ideological *content*.

Evidence on the role of intermediated news consumption in contributing to misinformation:

- The theoretical literature explains how platform incentives in respect of echo chambers and filter bubbles can also facilitate the spread of misinformation.

- However, studies which have sought to quantify the impact of misinformation and fake news on electoral outcomes have found it to be less prevalent, and less important, than perceived in the policy debate. For example, Allcott and Gentzkow (2017) estimate that if one fake news article were about as persuasive as a single TV campaign ad, the fake news around the 2016 US presidential election would have only changed the vote shares by a few hundredths of a percentage point.
- There is evidence that platforms can reduce the spread of misinformation with evidence that misinformation on Facebook has declined since its 2016 peak. This indicates a potential role for policy as it implies that changes to algorithms and content moderation can combat misinformation.

Evidence on different sources of algorithmic bias. No studies have looked directly at whether platforms have commercial incentives to bias the visibility given to different news brands, but there are causal studies finding Twitter was biased towards right-of-centre voices[2] and incentives for bias could arise if platforms had financial stakes in news outlets either directly or because news outlets differ in their use of platforms' other products and services (e.g., ad tech) in a way that causes online intermediaries to prefer sending traffic to news outlets generating more revenue and profit than others. There may be incentives for platforms to favour other products or technologies they operate (e.g., Google's "Accelerated Mobile Pages", Facebook's "Instant Articles", or "Apple News"). We provide an overview of the literature and note that it is not yet definitive, suggesting that the issue should continue to be monitored.

What new evidence gathering initiatives would help fill the gaps in the literature?

We have identified the following key areas where additional evidence gathering would shed further light on these issues and fill the gaps left by the existing literature.

User traffic and engagement data to be gathered from platforms and publishers. One of the most useful next steps would be to gather data from platforms to answer basic measurement questions on the consumption of news via different platforms. A relatively high-level dataset tracking page views and clicks to different publishers from different platforms in different contexts (e.g., whether a link was on an article shared by a friend or from a news brand they followed directly) would shed significant light on all of the theories of harm of interest. This could then be augmented with a panel dataset for a random sample of users so one could understand what content was shown to different users on different platforms.

Requiring access to details of algorithm design and the ability to perform "A/B" tests to study the impact of algorithm design on outcomes. A complementary source of information would be for Ofcom to better understand the algorithms themselves. This could be done by gathering information on the objective functions and signals used by the core algorithms on each platform (e.g., Facebook's news feed algorithm). The information would lead to a better understanding whether basic conditions are in place for certain concerns to arise (e.g., do algorithms condition on financial considerations, to what extent are they seeking to maximise platform engagement).

Ofcom could consider conducting periodic A/B tests on platforms' algorithms by, for example, seeing how different users' news consumption patterns are if they are shown a

2 Twitter's internal study by Huszár et al. (2021).

simple chronological feed rather than an algorithmically selected one, or if changes are made to the signals that the algorithms rely upon.

Replicating key experimental studies on UK data. Because of the recurring issue that the highest-quality experimental studies are based on US data, it would be illuminating for regulators or academics to replicate the most seminal studies on UK data, most notably Allcott et al. (2020) and/or Levy (2021).

Targeted updates to the News Consumption Survey to better track issues around intermediated consumption. The NCS remains a useful tool for getting a representative picture on how individuals consume news and to get a sense for the influence of individual news outlets. However, the NCS was not designed with intermediated news consumption in mind, and it would ideally be supplemented either with additional questions or with accompanying survey initiatives.

When considering these proposals, we recognise that some of them would require the co-operation of platforms or information gathering powers. Still, we believe that the initiatives set out above would be proportionate and effective ways to measure the concerns at issue.

The rest of the report is structured as follows. Section 1 discusses relevant background and sets out the "theories of harm" at issue. Section 2 presents relevant facts on the patterns of intermediated consumption of news in the UK recorded in the News Consumption survey and on public domain information on the working of platforms' algorithms. Section 3 provides a quantitative summary of the studies reviewed when preparing this report. Sections 4-7 discusses the findings of the literature on the key theories of harm (polarisation, echo chambers/filter bubbles, misinformation, and algorithmic bias). Section 8 discusses potential evidence gathering initiatives to better understand these issues in the UK. Section 9 concludes.

1. BACKGROUND ON NEW ISSUES RAISED BY INTERMEDIATED CONSUMPTION OF NEWS

This Section provides background to motivate the analysis that follows. We summarise the historical approach to assessing media plurality in the UK before discussing the new issues and "theories of harm" raised by consumption of news through online intermediaries.

1.1. The historical approach to media plurality in the UK

Media plurality is traditionally seen as a means to an end. Citizens exposed to a plurality of views in the media are better informed on matters of local, national and international news and policy. They are then better equipped to participate in a democratic society.[3]

Ofcom has an obligation, established by the 2003 Communications Act, to review the operation of the media ownership rules which aim to promote plurality and prevent undue influence. Ofcom also has a duty to maintain a sufficient plurality of TV and radio services. Under the Act, Ofcom is also required to review the operation of the media ownership rules every three years. Alongside this review, in 2021 Ofcom launched a programme of work on the *"future of media plurality in the UK"*, reflecting changes in the market for news, including as a result of digitalisation and the rise of online intermediaries.

Ofcom has historically defined media plurality as *"ensuring that there is diversity in the viewpoints that are available and consumed, across and within media enterprises"* and *"preventing any one media owner or voice having too much influence over public opinion and the political agenda"*.[4] Its quantitative assessment has been based on an *"availability, consumption, impact"* framework where a key metric of consumption is *"share of reference"*, based on the results of the news consumption survey.

The focus of policy intervention to date has been public interest assessment of mergers, but Ofcom has flagged the potential for intervention also in the case of market exit or if one news provider were to grow organically to a size that raised plurality concerns. As well as assessing *external plurality* across news organisations, Ofcom also assesses *internal plurality* (i.e., the extent to which they present a range of views and voices).

Part of the premise for this focus on media plurality is that, in the words of the 2012 Leveson Inquiry *"It is only through [media] plurality, specifically in relation to news and current affairs,*

3 The view that media plurality is central to democratic outcomes is as old as modern democratic institutions themselves. The notion of the press as a "fourth estate" has been attributed to Edmund Burke in a 1787 parliamentary debate while Thomas Jefferson famously opined that *"The way to prevent irregular interpositions of the people is to give them full information of their affairs thro' the channel of the public papers, and to contrive that those papers should penetrate the whole mass of the people. The basis of our governments being the opinion of the people, the very first object should be to keep that right; and were it left to me to decide whether we should have a government without newspapers or newspapers without a government, I should not hesitate a moment to prefer the latter. But I should mean that every man should receive those papers and be capable of reading them."* Thomas Jefferson, 16 January 1787 Letter to Edward Carrington. (emphasis added).

4 https://www.ofcom.org.uk/__data/assets/pdf_file/0024/84174/measurement_framework_for_media_plurality_statement.pdf

that we can ensure that the public is able to be well informed on matters of local, national and international news and policy and able to play their full part in a democratic society".[5]

Modern economic analysis provides a good basis for this view. There is empirical evidence that news consumption impacts the political process (e.g., by affecting voting decisions)[6] and that access to news influences how well-informed and politically engaged citizens are.[7] Theory work has also demonstrated how a diverse media sector can play an important role in reducing the risk of political capture and better holding government to account.[8]

1.2. New theories of harm raised by intermediated consumption of news

Ofcom, and other regulators worldwide, have for some time recognised the role of online intermediaries (e.g., tech firms like Google, Facebook, Apple, and Twitter) in the consumption of news. But as the scale of these operations has grown it has become more important to assess their impact on media plurality.[9]

On the one hand, the rise of intermediated consumption of news might be expected to soften some of the concerns around plurality. First, consuming news via an online intermediary might increase the variety of news outlets users view and consume and thereby increase the range of voices they experience.[10] Second, the nature of intermediated consumption (consisting of bite-sized chunks of content recommended in a personalised way via algorithms or posts by friends and influencers) arguably reduces the agenda setting power of news outlets and owners by reducing their control over the prominence given to different stories compared to a world in which users read a publication from cover to cover. Third, and relatedly, this reduction in agenda setting power may reduce the gatekeeper role of news outlets and owners and make it more feasible for politicians and other actors to disintermediate them and speak directly to voters (e.g. as was the case of President Trump on Twitter).

On the other hand, intermediated consumption of news raises new issues. We place these concerns into four categories, noting that they are somewhat inter-linked and that the relevance of these mechanisms will differ across different intermediaries.[11]

Concerns around increased polarisation. The first category of concerns relates to the impact of intermediated news consumption on political polarisation. This can mean polarisation in terms of beliefs about underlying facts; in terms of views on the motivations

5 Leveson (2012).

6 For example, Chiang and Knight (2011) find that unanticipated newspaper endorsements influence electoral outcomes. Similarly, Martin and Yurkoglu (2017) find that random variation in exposure to different cable news stations due to positioning in the channel listing alters consumer behaviour.

7 See Snyder and Stromberg (2010), Gentzkow, Shapiro and Sinkinson (2011) and Stromberg (2015) in Anderson, Stromberg and Waldfogel (2015).

8 See Besley and Prat (2006).

9 https://www.ofcom.org.uk/consultations-and-statements/category-2/future-media-plurality-uk

10 For an example of a review of the literature taking this more optimistic view see Arguedas, AR. Robertson, CT. Fletcher, R. and Nielsen, RK (2021).

11 For example, as we discuss below, search engines' results are less personalised than the content recommended by social media platforms and social media platforms raise additional issues around the role of content sharing between users. Nevertheless, there are commonalities in the issues and potential impacts on media plurality.

of political opponents; and in the strength of feeling and willingness to compromise on political issues.[12] There is a widespread view that intermediated consumption of news, in particular via social media, exacerbates polarisation on some or all of these metrics and that this is undesirable because it "*may make democratic decision making less efficient, and may lead citizens to perceive democratic outcomes as less legitimate*".[13] Indeed, some commentators have gone so far as to argue that the rise of intermediated consumption of news via social media has threatened the democratic process itself.[14]

Multiple mechanisms have been discussed for how intermediated consumption could have these polarising effects or result in other harms such as circulation of misinformation or "fake news" (either endogenously, as a result of content "going viral" as users share it with one another, or through deliberate campaigns). In each case, there is a concern that online intermediaries do not have appropriate incentives to prevent these issues from occurring. For example, social media platforms may have a profit maximizing incentive to maximise user engagement (and hence ad revenue) without sufficient regard to the social costs that this conduct might generate.[15] The use of algorithms raises further concerns because these outcomes may emerge even with limited human input: a machine learning algorithm tasked with maximizing engagement may independently choose to promote harmful content.[16]

Echo chambers and filter bubbles. The second, but closely related, category of concerns are around "echo chambers" and "filter bubbles". These terms are often used interchangeably and the concern in both cases is that intermediated consumption of news will result in users consuming only like-minded content and that this will contribute to the polarisation issues discussed above (i.e., that different groups are basing their views on a different base of facts and will less often be confronted with news content that contradict their pre-conceived view).[17] The terms are slightly different, however, as echo chambers can simply be the result of active decisions by individuals (e.g., to choose to follow a basket of friends, celebrities, and news outlets that share their worldview) whereas filter bubbles refer specifically to the potential magnifying effects of platforms' algorithms.

12 We discuss the definitions of polarisation used in the literature in more detail below.

13 Allcott, Braghieri, Eichmeyer and Gentzkow (2020).

14 For an example see: https://www.vox.com/policy-and-politics/2019/1/22/18177076/social-media-facebook-far-right-authoritarian-populism.

15 See, for example, Frenkel and Kang (2021). The book contains various accusations that algorithms tailored to maximise engagement resulted in broader social harms and takes its title from an internal Facebook memo that concluded *"So we connect more people…That can be bad if they make it negative. Maybe it costs someone a life by exposing someone to bullies. Maybe someone dies in a terrorist attack coordinated on our tools. And still we connect people. The ugly truth is that we believe in connecting people so deeply that anything that allows us to connect more people more often is de facto good."*

16 Social media platforms can be expected to face some constraints from users and advertisers e.g., because consumers may dislike platforms with a high a proportion of partisan content or hate speech. Similarly, advertisers may avoid having their ads carried next to such content. Indeed, there are academic findings that Twitter's approach to content moderation can be rationalised based on a model of engagement maximisation. See Jimenez-Duran (2021). However, there is a legitimate concern that platforms' profit incentives to address such issues may be insufficient because they do not account for the broader social costs of such content.

17 Pariser (2011).

As above, there is a link between these policy concerns and the incentives of the online intermediaries. The same profit maximising incentive to maximise engagement discussed above in the context of polarisation, and the same reliance upon personalised recommendation algorithms, could result in online intermediaries endogenously showing consumers like-minded content.

We note that the echo chambers/filter bubbles hypothesis provides a reason to be sceptical of the optimistic view that intermediated consumption necessarily results in a broader range of news and hence reduces plurality concerns: while intermediation may increase the breadth of content *sources* accessed by users, it may narrow the range of *views* if the bundle of intermediated content is selected (either by active consumer choice or by algorithm) in a way that reduces its diversity.

However, the extent of these issues is an empirical question and one which we explore in later Sections. A related question is how intermediated consumption of news interacts with consumption of news via-non intermediated channels. For example, does watching the TV news alongside using social media prevent the more extreme concerns around filter bubbles and echo chambers from manifesting?

Concerns around algorithmic bias. The third category of concerns we consider is media plurality concerns arising from algorithmic bias. This nests a range of concerns with different underlying sources.

First, are biases due to financial incentives on the part of intermediaries. For example, platforms which are vertically integrated into news production might favour their in-house news content. There might also be other financial relationships between online intermediaries and news publishers (e.g., around their use of online advertising) which incentivise differential treatment of news outlets that may raise media plurality concerns.

Second, would be more direct concerns that platforms' algorithms are biased in favour of particular views or voices on the political spectrum.[18] Concerns could arise either from platforms pursuing overt political objectives or because, for whatever reason, their operations distort political outcomes in a particular direction. As well as design of algorithms, there is also potential for concerns to arise if platforms are making decisions around content moderation (e.g., on whether to remove certain voices from their platform).[19]

Third, would be concerns that algorithms prioritise metrics (e.g., freshness) over others (e.g., originality) which disincentivise investment in news production/investigative journalism and encourage so-called "churnalism".

18 Right-of-centre publications in the UK and US have raised this accusation against Google (see https://www.dailymail.co.uk/news/article-9491607/Mail-files-antitrust-lawsuit-against-Google.html and https://www.foxnews.com/opinion/dan-gainor-yes-google-censors-conservatives-even-liberal-journalists-now-admit-it), although others have argued that apparent bias against right-of-centre news outlets could instead reflect them having lower levels of trust (https://www.economist.com/united-states/2018/08/30/googling-the-news). A study conducted by Twitter, discussed further below, had the contrary finding that its algorithm amplified right of centre voices.

19 For example, Twitter recently decided to permanently ban the account "Politics for All" from its site and Facebook has made content moderation decisions around the treatment of discussion of the war in Ukraine.

Concerns around negative impact on news production. The final category of concern is that the rise of online intermediation may undermine incentives for news production and contribute to the challenges faced by the news industry.[20] This is based on the hypothesis that online intermediaries may cause lower investment into news production (e.g., because the publishers are less able to raise advertising revenue, or because of commercial links between the specific publishers and the online intermediaries). Lower investment into news production would in turn lead to lower quality reporting, with less accuracy and depth or investigative reporting, so the consumers would be less informed.[21]

There are multiple mechanisms at work.

a. First, the concerns touched on above that platform algorithms may promote more salacious news sources over more reliable ones or undermine incentives to invest in original content.

b. Second, is the concern that intermediated consumption may commodify news, reduce the scope for publishers to differentiate their brands vs. rivals, and make it harder to sustain fixed cost investments in journalism.[22]

c. Third, is the concern that the prominence of the major platforms may result in an asymmetry of bargaining power that allows them to capture a disproportionate share of the value news content brings to their platforms.

d. Fourth, is that the rise of targeted digital advertising has shifted revenue from content producers towards players with the greatest data gathering and targeting ability.

e. Fifth, is concerns that platforms are taking steps on the advertiser side of the market that exacerbate these trends and which limit publishers' ability to monetise their

20 For instance, the Lords Select Committee on Communications and Digital has raised the issue around the online ads market and suggested that platforms pay to use news https://www.parliament.uk/business/lords/media-centre/house-of-lords-media-notices/2020/november-2020/fix-dysfunctional-online-ads-market-make-platforms-pay-to-use-news-lords-committee/

21 The conclusion of David Stromberg in volume 1B of the handbook of media economics is that *"There are strong theoretical reasons to believe that the market under-supplies political news. This is because of an empirically documented positive externality from the consumption of political news on political accountability."*

22 Ofcom's 2021 report "The future of media plurality in the UK" shows that the online intermediaries' influence presents challenges for the sustainability of news publishers through unfavourable terms of service, problems with brand attribution, and through changes in the intermediaries' algorithms. Similarly in the Australian context, the 2019 ACCC Digital Platforms Inquiry finds that the platforms caused a reduction in advertising revenue for the traditional media and "the impact of this reduction in advertising revenue is most evident in relation to local and regional news providers." Commentators have argued in the context of Google's "Accelerated Mobile Pages" that "*Publishers have less control over the appearance of their work on other platforms. Because all stories look basically the same in Google [Accelerated Mobile Pages (AMP)] the platform can lend false credibility to hoaxes"*. See *"Chasing mobile search results"* by Helen Havlak, engagement editor at the Verge, published by the Nieman Journalism Lab.

content. These latter concerns include ongoing antitrust investigations globally[23] around conduct in online advertising.[24]

In the rest of this report, we focus on the first three of these issues with less attention on the impact of intermediation on incentives for news production.

2. FACTS ON CONSUMPTION OF NEWS VIA PLATFORMS

In light of the theories of harm in the previous Section, we set out some background facts on the news consumption patterns of UK consumers and public information on the algorithms used by the major online intermediaries.

2.1. Relevant findings from the news consumption survey

Ofcom's key evidence gathering tool for assessing media plurality in the UK and the nations is the "news consumption survey" (NCS). The NCS seeks to measure news consumption patterns and facilitate an assessment of the influence of different media owners and outlets, taking account of consumption across different channels (radio, TV, print, and online). A key element of this is the calculation of "share of reference" which is a means of aggregating cross-channel news consumption on a consistent basis.[25] It also tracks consumers' tendency to multi-home across news sources.[26]

The NCS's design is focussed on measuring the influence of media *outlets and owners*. The key challenge with using the NCS to understand new issues around online

23 For example, the CMA's market study into online platforms and digital advertising (https://www.gov.uk/cma-cases/online-platforms-and-digital-advertising-market-study), the European Commission antitrust case on Google Shopping (AT.39740), the Australian case on Google's dominance in ad tech (https://www.accc.gov.au/media-release/google%E2%80%99s-dominance-in-ad-tech-supply-chain-harms-businesses-and-consumers), and the US lawsuit against Google's ad tech, led by the Texas Attorney General (https://www.texasattorneygeneral.gov/news/releases/paxton-files-third-amendment-antitrust-lawsuit-against-google).

24 In the interest of full disclosure, the authors of this report are involved with said antitrust cases in particular involving Google. For an overview of the economic theory behind concerns in the specific case of ad tech see: Latham, O. Herve, M. Bizet, R. 2021. "Antitrust concerns in Ad-Tech: formalizing the combined effects of multiple conducts and behaviours", *European Competition Journal.*

25 Ofcom's measurement framework describes "share of reference" as follows: *"Share of references is calculated by asking people which news sources they use and the frequency with which they use them. Each reference is then weighted for the frequency of use and summed. The share of each source or provider can then be calculated based on their total number of references as a proportion of all references for all news sources, regardless of the platform or media."*

https://www.ofcom.org.uk/__data/assets/pdf_file/0024/84174/measurement_framework_for_media_plurality_statement.pdf

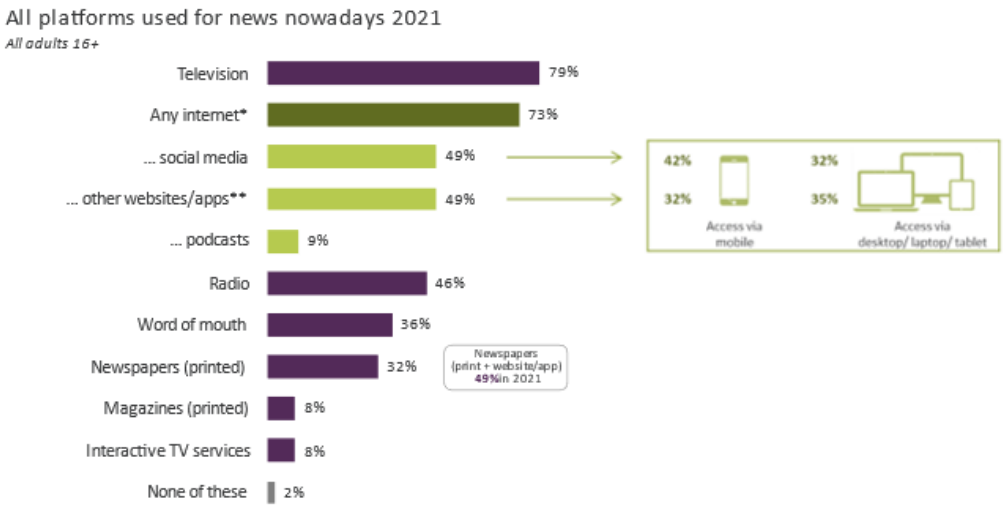
26 Analysis of multi-homing is key because academic work has underlined how the best way to think about plurality is to consider news consumption at the individual level before "aggregating up" to the society as a whole. Aggregate shares of reference are unlikely to tell the whole story because a given level of share will confer less influence if a news outlet is consumed alongside other outlets rather than in isolation. For example, one would be more worried about a situation where 25% of society each consumed only a single news outlet than a situation where everyone consumed four news outlets in equal proportions so that each had a share of 25%. See Prat (2018) and a 2014 Vox article https://voxeu.org/article/how-can-we-measure-media-power. See also Polo (2005). A more general discussion is by Eli Noam 2016. *Who Owns the World's Media?*

intermediaries is that they do not fit neatly into its measurement framework. As discussed in the documentation to the NCS, platforms and intermediaries are *"neither the provider of a news title or brand nor the producer of a news source"* and the approach to date has been to consider online intermediaries *"as a separate category distinct from the retail and wholesale classifications".*[27]

That said, while the NCS is not directly optimised for tracking consumption of news via online intermediaries it does provide some important background. The NCS has been augmented over time to capture some information on news consumption on social media, which provides some important background to understand the influence of online intermediaries.

First, it provides information on the importance of online and social media vs. more traditional forms of news consumption, finding that online news is now second only to television news in terms of penetration and that use of social media alone now exceeds radio and print news. It finds also that Facebook is the fourth most used source of news in the UK (with the only larger sources being the major TV channels). Twitter and Google have penetration rates of 24 and 28%, comparable to News Corp and Channel 4. The trends in the NCS also provide an indication that intermediated consumption of news, while drastically higher than just a few years ago, is now relatively stable, with Facebook's share actually declining slightly compared to previous years.

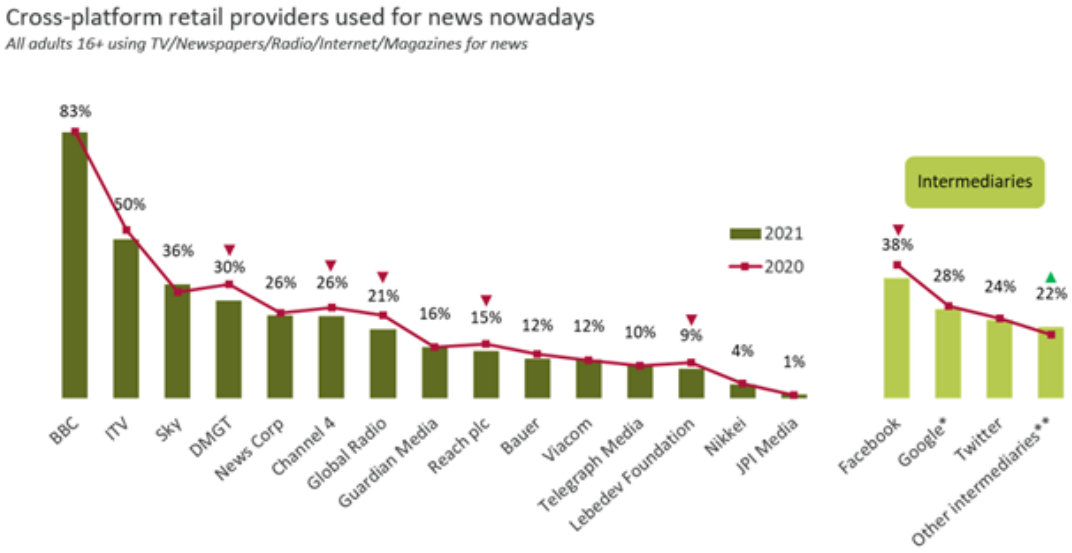
Figure 1: Penetration of online news and news consumed via social media in the UK in 2021



Source: Ofcom news consumption in the UK 2021 report, slide 19.

27 See paragraph 1.23 of https://www.ofcom.org.uk/__data/assets/pdf_file/0024/84174/measurement_framework_for_media_plurality_statement.pdf

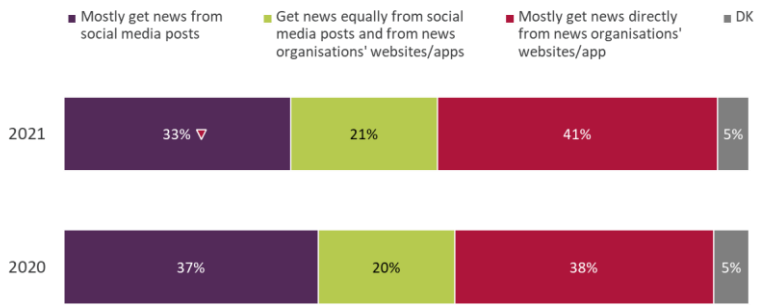
Figure 2: Summary of how the news consumption survey measures consumption across news sources and online intermediaries



Source: Ofcom news consumption in the UK 2021 report, slide 25

Second, it provides data on how consumers use online intermediaries alongside other online sources of news. It finds that around a third of consumers of online news used social media as their primary source of news online with a further 20% using a mix of social media posts and news direct from news organizations websites/apps. As we will discuss below, a key question is the extent to which consuming intermediated news alongside "traditional" news sources could have a moderating effect on the theories of harm we are examining.

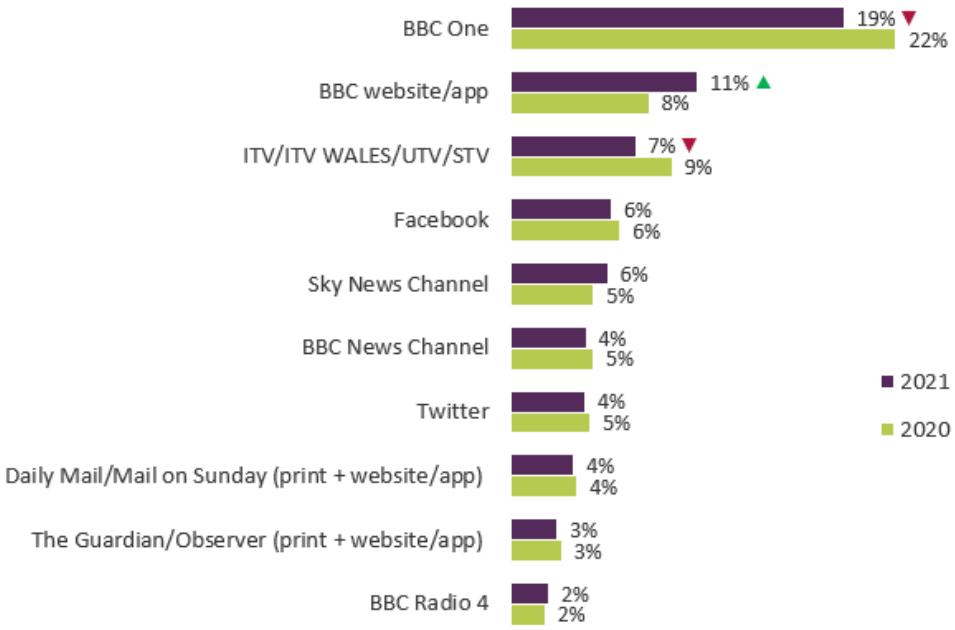
Figure 3: Breakdown of online news consumption between social media platforms and news organisations' own websites/apps



Source: Ofcom news consumption in the UK 2021 report, slide 44

Third, it directly asks users for their most important source for news finding that Facebook and Twitter are the most often identified outside of the TV news channels (albeit with the absolute number of individuals identifying them as the most important being relatively modest at 4 to 6%).

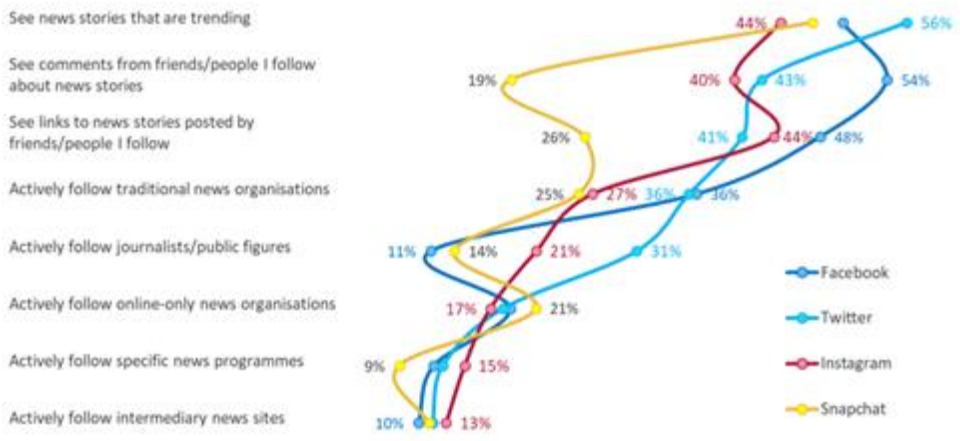
Figure 4: Proportion of respondents identifying each source as "single most important news source"



Source: Ofcom news consumption in the UK 2021 report, slide 71

Fourth, it highlights how consumption of news via social media places generally relies on recommendations made by the platform rather than by active decisions by users to consume particular outlets. As can be seen below, the most common driver of news consumption on Facebook, Twitter and Instagram is trending news stories or comments from friends or people users follow rather than active decisions to follow news outlets, journalists or public figures. This is relevant for the discussion on the distinction between "echo chambers" and "filter bubbles" discussed in the previous Section.

Figure 5: Consumption of news on social media platforms is driven by recommendations and links displayed in news feeds, rather than active decision making

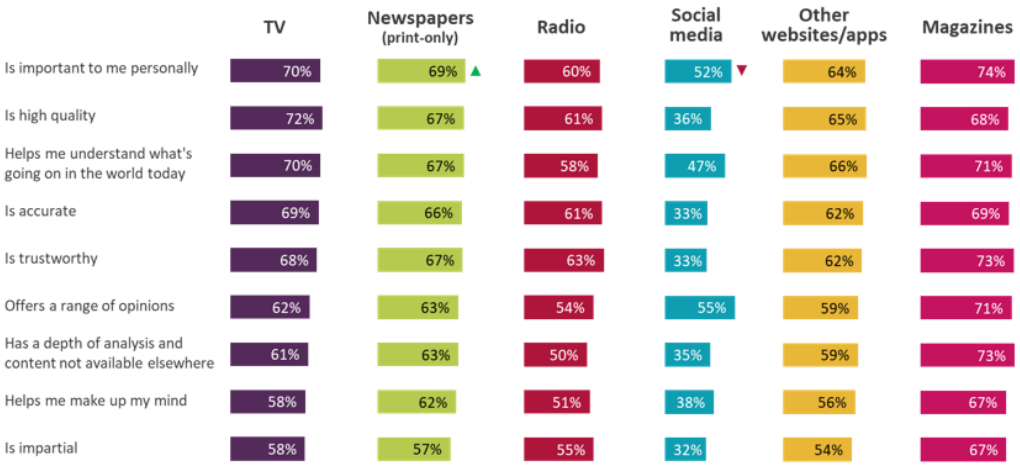


Source: Ofcom news consumption in the UK 2021 report, slide 46

Finally, the survey results show that consumers have a low opinion of the trustworthiness, accuracy, and impartiality of social media news in both absolute terms and relative to other

sources. On the one hand, this may be seen as corroborating the concerns around echo chambers and polarisation set out above, but it could alternatively suggest that consumers might discount information from social media vs. other sources, moderating these effects.

Figure 6: Rating of reliability of social media vs. other platforms



Source: Ofcom news consumption in the UK 2021 report, slide 74

In Section 8 we discuss how the NCS measurement framework and questions could be adjusted to shed further light on issues around the intermediated consumption of news.

2.2. Background on platforms' recommendation algorithms

Many of the plurality concerns around intermediated consumption derive from the design of the algorithms used to recommend news content to consumers. As we discuss in Section 8, better understanding and auditing the workings of these algorithms should be a key priority for future evidence gathering, but here we provide a brief summary of what is known about the major platforms' algorithms based on information in the public domain.

A key challenge is that all of these algorithms can be expected to rely on machine learning and artificial intelligence (AI) techniques. These techniques are difficult to audit as they are based on analysing a large number of variables ("signals") and weighting these variables to achieve an objective (e.g. maximizing engagement or more accurately predicting an outcome) in an automated manner, rather than in a way specified by the designer of the algorithm.[28] As such, they are to some extent "black box" and it may not be possible to perfectly understand why an algorithm outputted a particular result. Still, understanding what signals these algorithms use and what objectives they seek to maximize can provide insight into how they work and what they do.

28 Typically, the approach is to "train" algorithms on one dataset and then verify their performance on a new, unfamiliar dataset. The intention is to avoid "overfitting" (whereby an algorithm performs well on a given dataset by capturing all of the peculiarities of this dataset but does not generalise to new settings).

Facebook's News Feed algorithm. Facebook's public description is that its algorithm seeks to display the most relevant content based on a range of signals ("likes", who else is tagged, format (e.g., text or video), and so on). It assigns a score to each potential piece of content that could be displayed based on these signals and then downweighs the score of content triggering certain flags (e.g., for clickbait and disinformation). The weights on different signals are personalised (e.g., a consumer who comments more on posts may have more weight placed on the number of comments on a piece of content) and the algorithm is also designed to show a mix of content on the page. The objective function of the algorithm is not entirely clear, but it is reported that the algorithm historically prioritised "time on site", but now also gives weight to encouraging interaction (e.g., posting or likes).[29] The exact signals used are also unclear, but the number is reportedly in excess of 10,000. Concerns have been raised that some choices around the signals to include (e.g., basing prominence on more emotive reactions like angry emojis instead of simple likes) could result in more polarised news feeds.[30]

Twitter's timeline algorithm.[31] Historically, Twitter's home screen simply showed all of the tweets by people the user was following in chronological order. Since 2016[32], however, Twitter users by default see Tweets and other users recommended by an algorithm. Twitter reports that the *"algorithmic Home timeline displays a stream of Tweets from accounts you have chosen to follow on Twitter, as well as recommendations of other content we think you might be interested in based on accounts you interact with frequently, Tweets you engage with, and more."* Twitter's ranking signals take into account recency, relevance (e.g., based on location), engagement (how popular it is and how many users in a person's network are interacting with it), and media type.[33] Twitter users also have the option of influencing the Tweets they see in their timeline and the priority in which they appear via Twitter Lists.[34] As Twitter's algorithm relies on machine learning, the content it shows is somewhat unbeknown to the platform, prompting them to introduce a Responsible Machine Learning initiative to analyse its algorithm.[35]

As we discuss further below, Twitter found that its algorithm amplified tweets about political content from elected officials compared to political content on the reverse chronological feed.[36] They also found greater algorithmic amplification of tweets from the political right than the political left in six out of seven countries studied, and that right-leaning news outlets see greater algorithmic amplification than left-leaning news outlets.

29 https://www.washingtonpost.com/technology/interactive/2021/how-facebook-algorithm-works/

30 https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/

31 We also note that Twitter uses algorithms for other parts of their platform, such as to suggest Topics they think a user will like, to determine what topics are trending, and to suggest accounts to follow.

32 https://blog.twitter.com/en_us/topics/company/2021/rml-politicalcontent

33 https://blog.hootsuite.com/twitter-algorithm/#How_the_Twitter_algorithm_works_in_2022

34 https://help.twitter.com/en/using-twitter/twitter-lists

35 https://blog.twitter.com/en_us/topics/company/2021/introducing-responsible-machine-learning-initiative

36 https://blog.twitter.com/en_us/topics/company/2021/rml-politicalcontent

Instagram. Instagram reports that they *"use a variety of algorithms, classifiers, and processes, each with its own purpose."*[37] Instagram had been using a chronologically ordered feed since launch, but in 2016 they decided to change this as *"people were missing 70% of all their posts in Feed, including almost half of posts from their close connections"*. Currently all parts of the platform (Feed, Explore and Reels) use their own algorithm.

Instagram also discuss how they rank posts for Feed and Stories[38], mentioning that they consider thousands of signals. They note that the most important signals across Feed and Stories, in rough order of importance, are information about the post (such as how popular it is, location, media type and recency), information about the person who posted (such as how many times users have interacted with that individual in the previous few weeks), a user's activity and a user's history of interacting with someone. Instagram then predicts how likely a user is to interact with a post.

Instagram also report that, for Feed, the five interactions Instagram's algorithm most closely looks at are a user's likelihood to spend a few seconds on a post, comment on it, like it, save it, and their likelihood to tap on the profile photo. The more likely one is to take an action; the more weight Instagram gives to that action and the higher up that post will be in the Feed.

Instagram also try to avoid showing too many posts from the same person in a row and apply labels to posts with misinformation such that they are shown lower in Feed and Stories. Users are able to affect what they see by selecting 'close friends', muting people and marking recommended posts as 'not interested'.

We also note that in January 2022, Instagram announced they are testing three different views on the home screen (two of which allow for posts to be ordered chronologically).[39]

37 https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works

38 The majority of content shown on Explore and Reels is from accounts a user doesn't follow, as opposed to Feed and Stories. On Explore, they look at signals to find photos and videos a user might be interested in, noting one might see posts about related topics that they have signalled interest in. They then rank the content they think a user will be interested in based on how likely a user is to interact with the post (with most important interactions being likes, saves, shares). The most important signals they consider in doing this are (in rough order of importance) information about the post, user history of interacting with the user who posted, user activity, and information about the user who posted. For Reels, they pay specific attention to what might entertain users and they survey people to better understand this, *"with an eye towards smaller creators".* The most important predictions they make are how likely a user is to watch the full reel, like it, say it was entertaining or funny or go to the audio page. The most important signals they use to ascertain this are (in rough order of importance) user activity, user history of interacting with the user who posted, information about the reel, and information about the user who posted.

39 https://twitter.com/mosseri/status/1478767105444966401

Snapchat. There is less information about how Snapchat's algorithm works in comparison to the other platforms detailed above. Snapchat deploys algorithms so that content from friends that a user interacts with a lot shows up higher in the Friends section of the app.[40] Snapchat also has a Discover section (which has stories from publishers, creators, the community and more[41]) where it is reported that the algorithm shows content that a user is likely to be interested in[42] – however, the exact mechanism it uses is unknown. Similarly, they also introduced 'Spotlight' which shows videos that are up to 60 seconds long.[43] It is not clear how the algorithm determines what content is shown to users though it is thought to be influenced by content a user has previously watched and how long for.[44]

Google Search. Google Search uses a *"whole series of algorithms"* to find relevant search results.[45] Their algorithms look at a variety of factors, including the words used in a user's search query, relevance and usability of pages, expertise of sources and user location and settings. They weight each factor depending on the nature of the query, for example, recency is a larger factor for queries about current news topics than it is for a word definition. Google reports also subjecting its search results to live tests and evaluation by external Search Raters.[46] The key factors they look at when determining search results are:

- Meaning of query – to establish what information a user is searching for they build language models to work out what strings of words they should look up in the index. They also consider what category of information a user wants (e.g., in terms of recency, or language).
- Relevance of webpages – their algorithms analyse the contents of webpages to determine if they contain information relevant to the search query. To do this, the algorithms consider keyword matching to the search query (as well as whether the pages contain other relevant content) in addition to using aggregated and anonymised interaction data.
- Ranking useful pages – the algorithms analyse *"hundreds of different factors"* to rank the best pages first. This includes considering recency, the number of times a user's search terms appears, whether the page has a good user experience, whether the site seems to be valued for similar queries by other users and if other prominent websites on the query matter link to that page. They also use algorithms to identify spam and remove sites that violate Google's webmaster guidelines.

40 https://martech.org/snapchats-major-redesign-algorithmic-feed-means-brands/#:~:text=Algorithmic%20control&text=These%20algorithms%20are%20intended%20to,go%20for%20the%20Discover%20tab

41 https://support.snapchat.com/en-US/a/discover

42 https://martech.org/snapchats-major-redesign-algorithmic-feed-means-brands/#:~:text=Algorithmic%20control&text=These%20algorithms%20are%20intended%20to,go%20for%20the%20Discover%20tab.

43 https://invideo.io/blog/snapchat-spotlight/#:~:text=Spotlight%20is%20a%20TikTok%2Dlike,users%20in%20a%20dedicated%20section

44 https://www.pocket-lint.com/apps/news/snapchat/154739-what-is-snapchat-spotlight-how-to-create-send-spotlight-snap

45 https://www.google.com/search/howsearchworks/algorithms/

46 https://www.google.com/search/howsearchworks/mission/users/

- Usability of webpages – Google Search uses algorithms to promote more usable pages over less usable ones (holding everything else constant).
- Context and settings – they consider information such as user location, previous search history and settings.

Google has been reported to change its algorithm over 600 times a year and aspects of its algorithm changes have been investigated by regulators (e.g., in the long-running antitrust cases around its promotion of its comparison shopping business and demotion of rival services).[47]

3. OVERVIEW OF THE ACADEMIC LITERATURE ON THESE THEORIES OF HARM

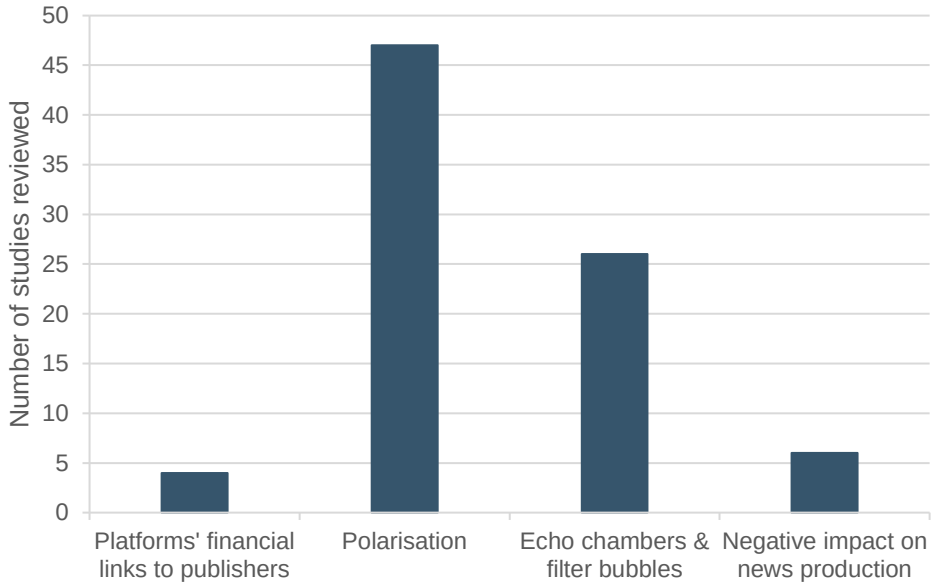
In this Section we summarize the results of our review of the academic literature. We focus on academic articles (both peer-reviewed publications and working papers) addressing the theories of harm of interest with particular attention to those with a significant quantitative component. We have identified 79 studies meeting these criteria, although we cite some additional studies on related topics as part of our discussion.

Figure 7 below summarises the number of studies relevant for each theory of harm.[48] We take a relatively expansive definition which captures studies which speak to the economic mechanism at issue. For example, a study might look at whether social media users' consumption patterns are consistent with echo chambers but might not look at whether echo chambers are propagated by the use of social media or the extent to which this outcome is driven by social media platforms' decisions. The most commonly reviewed topics are around misinformation, polarisation and filter bubbles/echo chambers. Very few studies have directly considered financial incentives for platforms to favour in-house news outlets and so we have conducted a separate review (discussed below) of the broader literature on self-preferencing by platforms in other contexts.

47 See: https://ec.europa.eu/competition/elojade/isef/case_details.cfm?proc_code=1_39740. In the interest of full disclosure, we note that the authors have advised complaints in respect of Google's algorithm design in Shopping Search and Local Search as well as on other issues including around ad tech.

48 We note that some studies speak to multiple theories and hence the sum of the four bars exceeds the total number of studies reviewed. We note also that some studies may speak to a theory of harm indirectly (e.g., they may consider the extent to which filter bubbles or echo chambers exist rather than necessarily the extent to which decisions by platforms exacerbate these issues).

Figure 7: Breakdown of studies by theory of harm



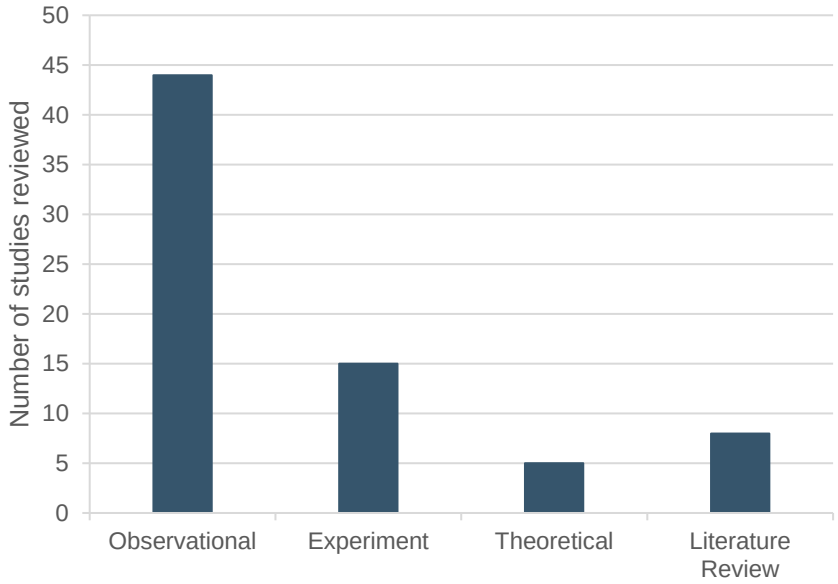
Source: CRA analysis. Note: Some studies are relevant for more than one theory of harm and thus have been counted twice.

We have also categorised the studies on dimensions relevant for considering their policy implications. We consider questions of "internal validity" (the empirical design used to try and measure the effects of interest and whether they seek to determine causal effects rather than mere correlation); and "external validity" (whether the results can be expected to generalize to other settings or if they are specific to particular geographies or platforms).

Issues relating to internal validity and measurement

Observational vs. experimental studies. We classify studies as observational, theoretical, experimental or literature reviews. Observational studies typically rely on surveys and/or tracking of web browsing behaviour. These studies are useful for gathering factual and descriptive information (e.g., on the pattern of consumption online and the extent to which intermediated news is consumed alongside traditional online news outlets), but are less likely to provide information on causal relationships. Under "experimental", we include studies which use techniques to identify causal relationships such as a randomised control trials, A/B testing, or natural experiments. The other categories cover articles that are primarily a literature review but offer additional insights, and those that present a theoretical model to explain the mechanisms behind the issues of concern.

Figure 8: Breakdown of studies by empirical design



Source: CRA analysis.

Note: Some studies include two types of data (e.g., observational and experiment data) and have been counted for both categories.

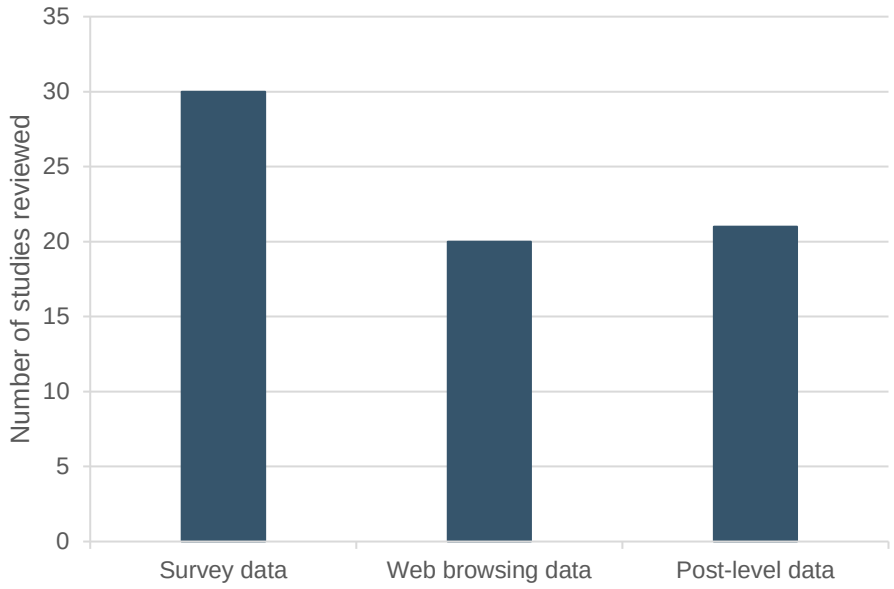
As can be seen from the Figure, the majority of studies rely on observational data. This is not a reason for them to be dismissed: in these early stages of understanding the issues, careful factual work to, for example, understand the patterns of news consumption on platforms, are of significant interest. However, we place greater weight on studies which use experimental or quasi-experimental designs to establish causal relationships.

Approach to measuring consumer activity and consumption. Understanding each of the theories of harm requires tracking consumer behaviour and news consumption on platforms. The following are the main categories of data collection approaches:

- Surveys of participants' news consumption behaviour. For example, this might involve asking the participants to identify with an ideological slant and then study their behaviour such as sharing stories and the types of news articles they are being shown by the online intermediaries.
- Web browsing data, based on tracking visits to websites and/or observing the navigation process either via tracking services or browser plugins.
- Post-level data involves the study of how social media posts are shared, the links they include, hashtags, etc, typically in the context of studying the spread of misinformation.

The Figure below summarizes the approaches used across studies:

Figure 9: Breakdown of studies by approach used to measure consumer interaction with and consumption of news



Source: CRA analysis. Note: Miscellaneous data gathering methods (such as emails sent with links to measure click-through rates) have not been included. Survey data refers to survey and questionnaires. Web browsing data refers largely to tracking visits to websites and/or observing the navigation process.

Focussing on web browsing data specifically, we have seen the following third-party sources used in the literature:

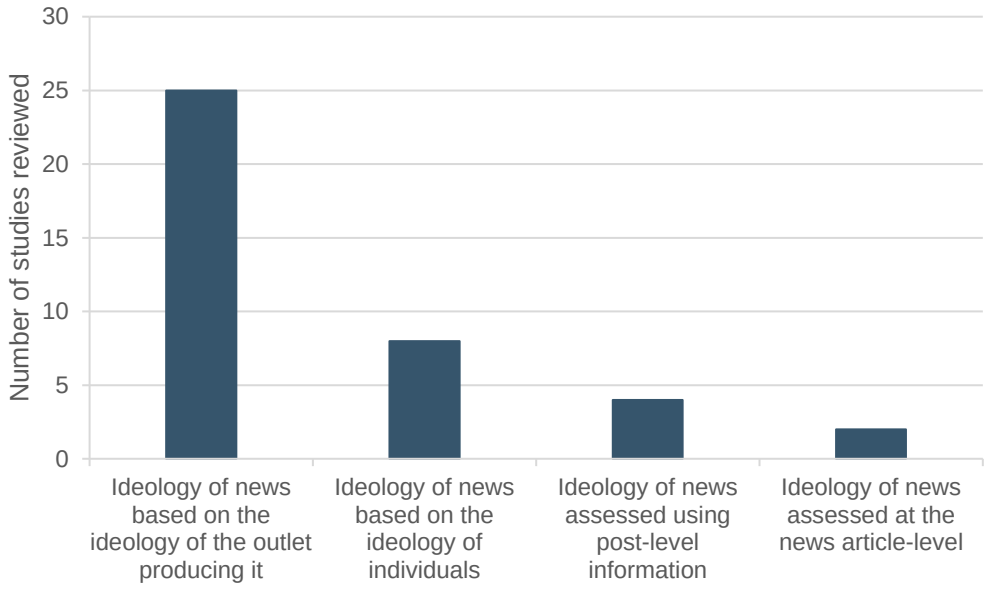
- Alexa – indexing database measuring daily reach to different websites, gathering data from browser extensions installed on people's computers as well as from measurement services offered to websites. These data exclude mobile browsing and do not capture news viewed directly on social media sites, for example, when people read headlines within Facebook or Twitter news feeds.
- BuzzSumo – a commercial content database that tracks the volume of user interactions with internet content on Facebook, Twitter, and other social media platforms. This provides a measure on the volume of Facebook engagements and Twitter shares for news stories.
- comScore – data provider from its panel of over one million US resident internet users (panellists voluntarily install software on their computers to permit monitoring of their browsing behaviour), as well as providing a measure of ideology of the news sites visited.
- Reuters Institute Digital News Report data – reports typically relying on commissioning surveys (e.g., via YouGov) on online news consumption
- Web Historian – an open-source extension for Google Chrome, that accesses respondents' browser history stored on their computers

In addition to this list, we are aware of other third-party sources such as "Reality Mine" and Ipsos's "IRIS" panel. Several studies also rely on proprietary data from the platforms themselves (most notably Facebook and Twitter). For example, Bakshy, Messing, and Adamic (2015) used a dataset from Facebook and Ferrara, Jiang, and Ren (2021) used a Twitter data set collected by Chen, Lerman, Ferrara (2020).

Approach to measuring ideology of news content. For many of the theories of harm, an important component is measuring the ideology of the content consumers are exposed to. This is necessary to, for example, assess whether consumers are consuming content that mirrors their existing views as predicted by the echo chambers and/or filter bubbles hypothesis. Figure 10 breaks down studies according to how ideology of content is measured.

We see three main approaches in the literature: measuring ideology at the outlet level (e.g., designating a newspaper as left- or right-leaning); at the individual level (e.g., referring to a survey where individuals would identify themselves with a political leaning); or at post-level (e.g., by using tweet hashtags to designate the ideology of each post).

Figure 10: Breakdown of studies by approach used to measure ideology of content



Source: CRA analysis.

Note: Some studies measure the ideology of content through two different approaches and have been counted for both approaches.

As shown in the Figure, the most common approach is to measure the ideology of news consumption using the ideology of the *outlets* producing the content. This is understandable given data constraints, but a fuller analysis would ideally look at the ideology of individual pieces of content. While the literature has identified empirical techniques, such as textual analysis, to classify news content (more detail is provided in Appendix D) we are not aware of any studies applying these techniques to understand the impact of consuming news via online intermediaries. This is an important gap for future research.

Issues relating to external validity

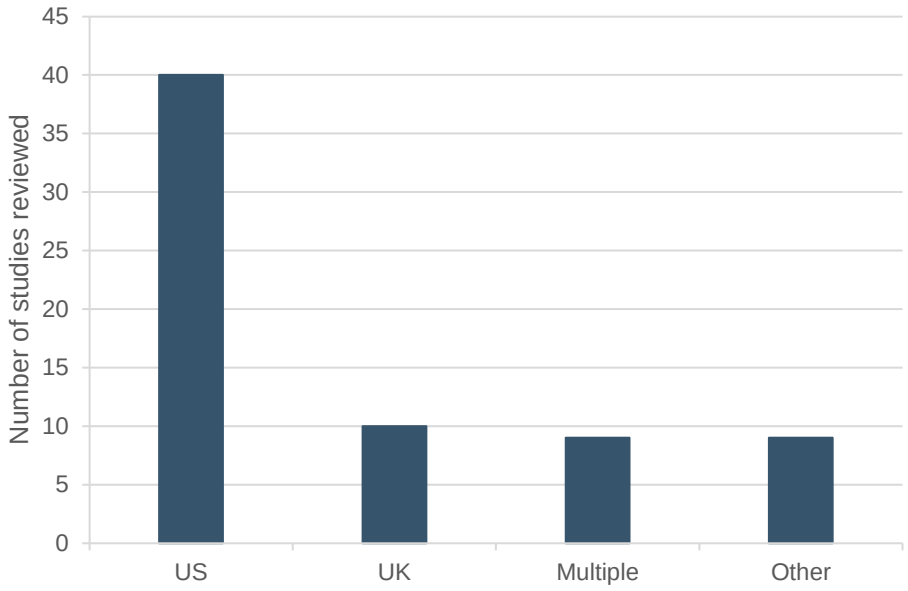
Geographies analysed. The Figure below shows that a disproportionate share of studies focus on the US with four times as many US studies as UK ones.[49] Further, as we discuss

49 The studies typically involve cross-sections of the US internet users, rather than the specific subgroups (e.g., students). The public sources of data are discussed above in this section, while Appendix C lists the geographies of each study considered.

in subsequent Sections, this discrepancy is even more pronounced for the methodologically sophisticated and credible studies.

This points towards a significant evidence gap. While the fundamental economic mechanisms might be reasonably expected to be common across countries (e.g., that consumers find content that mirrors their prior belief more engaging and that algorithms' emphasis on maximizing engagement could lead to filter bubbles), different countries clearly display differences in the dimensions of polarisation and hence the potential harms that might be at play. US politics exhibits some political differences which are less common to the UK (around the Trump presidency, abortion, etc.) as well as not exhibiting other UK-specific political differences (e.g., discussion around the UK's exit from the European Union). The UK also has a different regulatory framework, in particular around public service broadcasting, which could alter the impact of online intermediaries.

Figure 11: Geographies covered by the studies



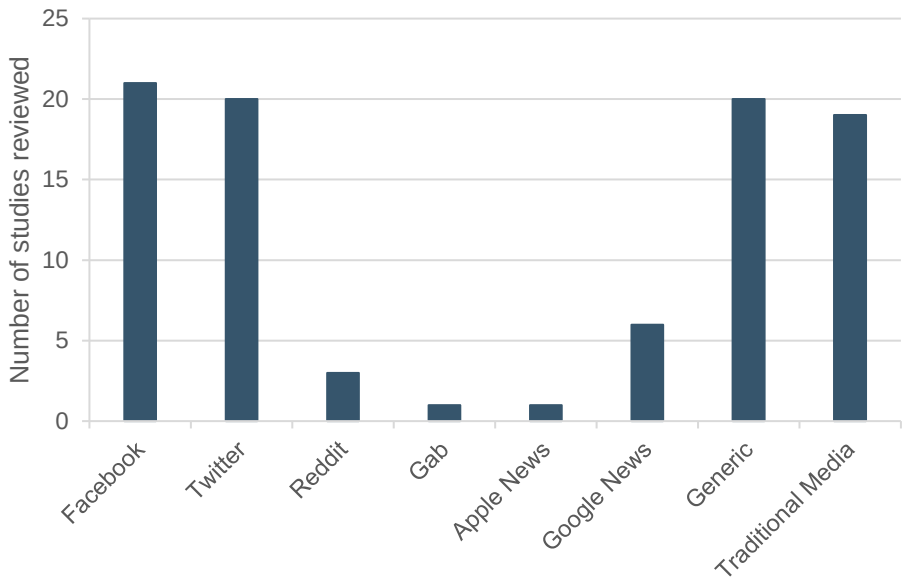
Source: CRA analysis.

Note: Multiple refers to studies that are carried out in multiple countries. Other refers to a study carried out in a specific country alone once (Argentina, Australia, Brazil, Germany, Poland, Spain and Switzerland).

Platforms analysed. Figure 12 details the platforms reviewed by the studies, showing that Facebook and Twitter were the main focus. This likely reflects their greater focus in the policy debate but may also reflect practical issues with implementing research designs (e.g., multiple studies attracted participants by advertising on the platforms themselves).

As we develop below, different platforms may exhibit different tendencies towards the theories of harm set out above. Even within a given platform there may be changes over time due to changes in policy (e.g., around algorithm design or moderation policies).

Figure 12: Platforms reviewed by the studies



Source: CRA analysis.

Note: Some studies analyse multiple platforms. Traditional Media refers to online and offline media and face-to-face interactions. Generic refers to purpose-built platforms, theoretical models or broader studies e.g., a study of social media platforms in general.

Having summarised these studies at a high level, in the four Sections that follow we present our assessment of the existing literature on the theories of harm of interest. Our goal is to provide a critical assessment of the current weight of evidence and to determine the methodological lessons and evidence gaps in this literature.

Or goal is to go beyond simply summarising the findings in the literature, but to take a view on which studies are most compelling and have the most to tell us about the theories of harm. Generally, we place most weight on those studies that can shed light on causal relationships as opposed to simply revealing correlations or other descriptive analysis.

4. POLARISING EFFECTS OF CONSUMING NEWS VIA ONLINE INTERMEDIARIES

4.1. Learnings from the literature on polarisation

We begin by considering the messages of the literature on the impact of consumption via online intermediaries on political polarisation. Polarisation can be defined in different ways, but accepted definitions include segregation between groups over: underlying facts; views on the motivation of political opponents; and willingness to compromise on political issues.

Overall, **the literature provides firm evidence that consuming news via social media can increase polarisation.** This is the message from the majority of studies we reviewed, and in particular from the studies which we consider most sophisticated and reliable. However, **it also identifies offsetting benefits in other areas such as increasing how well-informed users are about news events.** This finding is true based on a simple review of the weight of conclusions across studies and the conclusion is enhanced if one focusses attention on the studies we consider to be most well-designed and credible.

However, **the literature also indicates different impacts for different types of intermediated consumption with consumption of news via search engines appearing not to have analogous effects to social media.**

The key caveat is that the most compelling evidence of a link between social media and polarization comes from the US and is specific to Facebook. The academic research focusses on the US, but the UK has its own dimensions of polarisation (e.g., Brexit) and its own competitive and regulatory environment (e.g. as a result of regulation of broadcast news and the presence of prominent public sector broadcasters).

If one simply counts the number of studies which find online intermediaries increasing polarisation, these studies outnumber those that find the opposite or no impact by more than 4 to 1. This is based on 23 studies discussing polarisation and online intermediaries.

We consider the most reliable studies on this topic are Allcott et al. (2020)[50] and Levy (2021).[51] They stand out in terms of their focus on determining causal effects through an experimental design, their relative recency, and the fact they focus squarely on the theory of harm of interest. The downside of these studies, however, is that they are both based on US data.

Allcott, Braghieri, Eichmeyer and Gentzkow (2020). This study investigates the impact Facebook has on a range of media plurality issues, including political polarisation, alongside a broader set of welfare questions such as the impact of social media usage on subjective wellbeing. The overall message for media plurality is that the causal impact of Facebook is to make users better informed, but also more politically polarised. Unlike many of the studies of these issues which focus on correlations and on observed consumption patterns, the study uses an experimental design to investigate causal links between platform design and political polarisation and to track through the impact of Facebook usage on the outcome variables of relevance for political outcomes.

The study is based on recruiting a suitable sample of users and subjecting them to four rounds of surveys ("baseline", "midline", "endline", and "post-endline") and daily text messages.[52] The thrust of the methodology was as follows. First, a sample of Facebook users were recruited using Facebook advertising and given a screening survey and consent form. Those who passed the screening survey were asked to complete the baseline survey on a set of outcome variables (discussed below) and were asked if they would be willing to deactivate their Facebook account for 24 hours.

Next, the respondents who provided the required information and were willing to tolerate a 24-hour deactivation were then asked to deactivate their Facebook account for 24 hours and were subjected to a mid-line survey designed to elicit the cash value they would be willing to accept (WTA) to deactivate their Facebook accounts for a longer period of four weeks ending just after the November 2018 U.S. midterm elections. The authors then *randomly assigned* a share of those who had a willingness-to-accept of less than $102 to

50 Allcott, H. Braghieri, L. Eichmeyer. S. Gentzkow, M. 2020. "The Welfare Effects of Social Media", *American Economic Review.*

51 Levy, R. 2021. "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment", *American Economic Review.*

52 The daily text messages asked for measures of subjective wellbeing. This related to issues other than media plurality, so we do not discuss them further.

either a treatment group that was paid to deactivate, or to a control group that was not required to do so. The treatment and control group were then tracked over time. This took the form of periodic checks to see if their Facebook account was indeed deactivated.

- After the 4-week period both treatment and control group were sent an "endline" survey which asked them questions about the same outcome variables as in the baseline survey. Participants were also subjected to another assessment of their WTA to stop using Facebook. The outcome variables covered various metrics (e.g., subjective wellbeing and happiness) but of greatest interest for our purposes were: **News sources, news knowledge**, **political engagement**, and **political polarisation**.[53]
- Lastly, after the main survey two follow up emails were sent containing links to articles on limiting smartphone usage or on various political causes. This allowed the authors to measure click through rates as a metric for interest in these topics as between treatment and control group. A final survey was then sent asking users how much they were using Facebook according to the data recorded on their smartphones.

Having gathered this data, the authors estimated effects on the outcome variables of interest by looking at the relationship between the outcome variables and the proportion of days each subject's account was deactivated. To obtain causal estimates, the latter variable was instrumented for with the randomly assigned treatment variable.[54] The analysis controlled for the baseline level of the outcome variables so that the results could be interpreted as the impact relative to the baseline level reported at the start of the experiment.

As can be seen below, the authors found that:

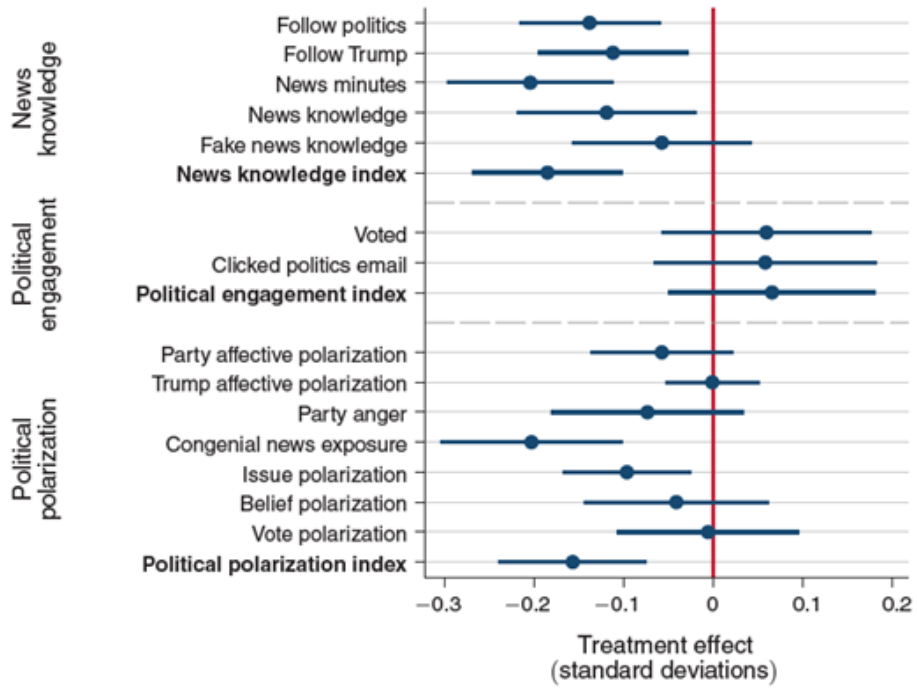
53 For news sources, at "baseline" they asked over the past four weeks how often respondents got news from a variety of sources, including Facebook, reading newspapers in print, other social media sites than Facebook, and from news websites or apps other than social media. The "endline" question asks in the last four weeks, relative to what is typical for the respondent whether they spent more or less time on the same sources as those in the baseline survey. For news knowledge, they asked over the previous four weeks how closely respondents followed US politics, news about Trump, and how many minutes they spent watching, reading, or listening to the news. They also asked at "baseline" and "endline", *"Of the following news events, which ones do you think are true, and which ones do you think are false? [True, False, Unsure]"*, including true statements, false statements, and statements to measure fake news knowledge. The news knowledge questions were updated at "endline". To measure political engagement, they recorded if respondents voted in the 2018 midterm election and whether they clicked on any link in the "post-endline" politics email. To measure political polarisation, they asked questions on party affective polarisation, Trump affective polarisation, party anger, congenial news exposure, issue polarisation, belief polarisation, and vote polarisation. For example, a question they asked to measure congenial news exposure was *"Thinking back over the last 4 weeks, how often did you see news that made you better understand the point of view of the [Republican/Democratic] Party? [Never, Once, Two or three times, Four times or more]".*

54 The idea was that the initial 24-hour period meant both treatment and control were exposed to the idea of deactivating Facebook and the treatment was the higher probability of people continuing their "detox". The final results can then be interpreted as "local average treatment effects" (i.e., the average effect of the treatment (turning off Facebook) amongst those users who were induced to turn off Facebook as a result of the intervention).

- **Deactivating Facebook caused a statistically significant *reduction* in news consumption and knowledge.** The study found that deactivating Facebook reduced consumption of news through social media and there was no offsetting change in non-social media online news consumption. The reduction in news consumption was around 8 minutes per day (15% of the average for the control group). This news reduction had a measurable negative impact on news knowledge: accuracy on the news knowledge quiz questions fell from 7.26 out of 10 to 7.12.
- **Deactivating Facebook caused generally statistically significant reductions in polarisation.** The most notable reduction was on the *"congenial news exposure"* metric with individuals reporting post deactivation that they were more likely to have seen *"news that made them better understand the point of view of their own political party relative to the other party".*
- **Deactivating Facebook caused a limited and generally not statistically significant increase in tweeting and consumption of other news sources** (albeit, as above, these increases were insufficient to offset the decline in news consumption on Facebook). There was also no detectable effect on various metrics of political engagement.

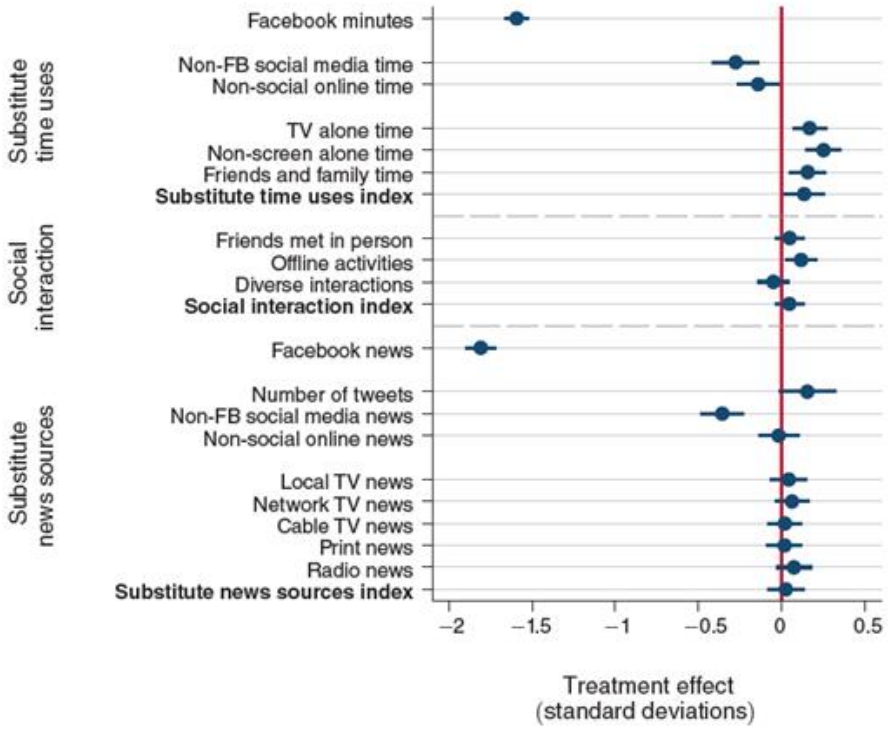
In other words, the study found that using Facebook makes users increase news consumption but increases their political polarisation. These impacts are also shown in the two figures below, alongside the study's other measured impacts. For each outcome of interest, the blue dots show the estimated size of the effect while the blue line shows the confidence interval around these estimates. If the confidence interval does not overlap with the red vertical line this indicates that the effect was statistically significant.

Figure 13: Impact of Facebook deactivation on news knowledge and political engagement/polarisation questions



Source: Allcott, Braghieri, Eichmeyer and Gentzkow (2020), Figure 3

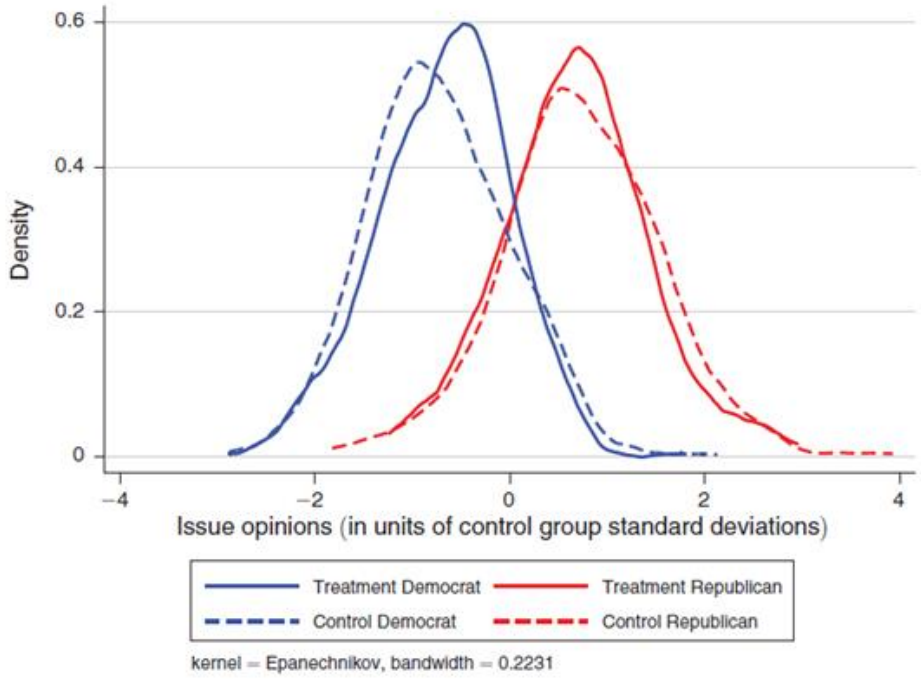
Figure 14: Impact of Facebook deactivation on time spent on other activities and substitute news sources



Source: Allcott, Braghieri, Eichmeyer and Gentzkow (2020), Figure 2

Looking at positions on left/right policy issues the study also found that deactivating Facebook led to distributions of political beliefs that were more centrist. This is shown below where the dotted distributions for users who were not encouraged to drop Facebook for 4 weeks have more “mass” in the left-hand extremes for Democrats and right-hand extremes for Republicans.

Figure 15: Impact of Facebook deactivation treatment on political opinions



Source: Allcott, Braghieri, Eichmeyer and Gentzkow (2020), Figure 4

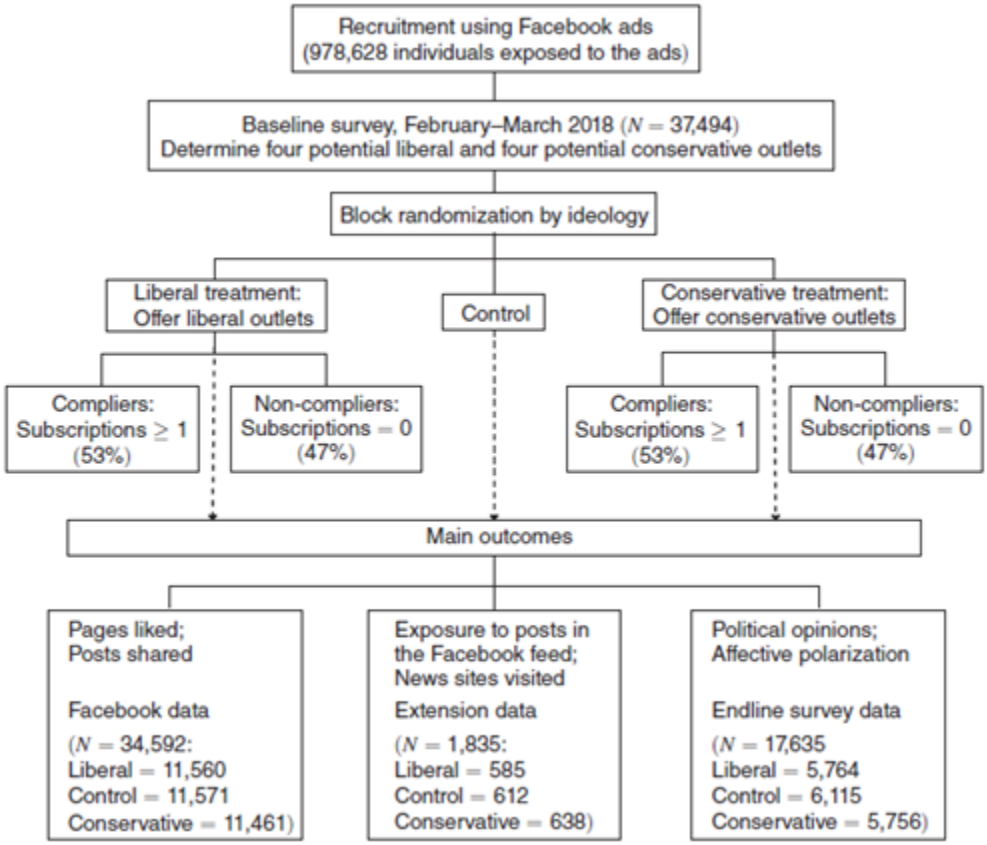
Levy (2021). This study also looks at the impact of social media on polarisation based on an experimental design. It also goes further to look at the composition of news content shown to users and how the design decisions made by Facebook in terms of what content to show users may contribute to effects on polarisation.

This study estimates the effects of social media news exposure by conducting a large field experiment randomly offering participants to subscribe to or "like" conservative or liberal news outlets. It employs over 37,000 participants between February to March 2018 using Facebook ads. This large sample relies on 978,628 who people saw the ads, 87,648 people who clicked the links in the ads, and approximately one-half of those who began the survey.

The study design is shown in the Figure below and essentially involved grouping individuals based on their reported ideology and then randomly assigning them to either a liberal or conservative "treatment" in which they were offered to subscribe to outlets of this persuasion.

We consider this study as reliable, due to the large sample and its experimental approach, which means it can speak to causality between the intermediary's ranking of news sources and the observed outcomes in polarisation.

Figure 16: Experimental design of Levy (2021)



Source: Levy (2021), Figure 1.

The study reports four main findings.

First, random variation in exposure to news on social media substantially affects the slant of news sites that individuals visit. This implies users' consumption of news on social media (and potentially other intermediaries) is not just due to individual choice. Rather, decisions by platforms, including around algorithmic design, can influence consumer news consumption. This therefore implies choices by online intermediaries can impact polarization.

Second, exposure to counter-attitudinal news decreases negative attitudes toward the opposing political party. This finding is consistent with the first in that the online intermediaries have the ability to lessen polarisation through the news content they surface.

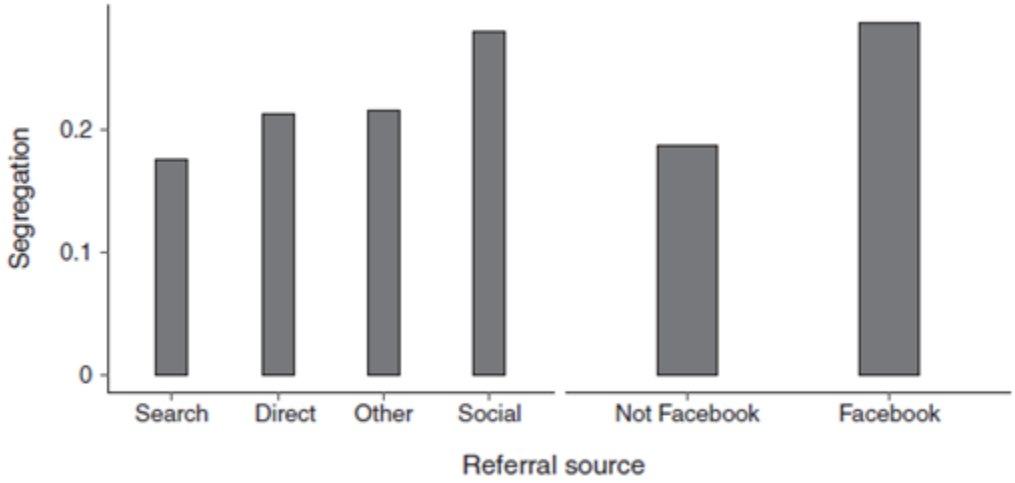
Third, in contrast to the effect on attitudes, the author finds no evidence that the political leanings of news outlets consumed affect political opinions. The participants in the conservative/liberal treatment were offered to subscribe to their four potential conservative/liberal outlets, while the control group was not offered any outlets (see figure above for the experimental design). The treatment was found to have no impact on the subsequent self-reported political opinion.

Fourth, Facebook's algorithm is found to be less likely to supply individuals with posts from counter-attitudinal outlets, conditional on individuals subscribing to them.

Lastly, not part of the experimental approach but the study also reports comScore data on 8,882 individuals all of whom consume news through both Facebook and other channels.

The study reports that the referral source though which news sites are visited is an important distinction on the level of segregation, [55] with social media referral leading to higher segregation than direct access to the news, and online search leading to lowest levels of polarisation. This is shown in the figure below reporting segregation by referral source. Segregation in this sense is similar to a broad definition of polarisation, as it measures the difference in the slant of news sites by the channel used.

Figure 17: Levy (2021) figure on polarisation in news sites visited by referral source



Source: Levy (2021), Figure 3 relying on comScore data.

Together, the results suggest that social media algorithms have the ability to limit exposure to counter-attitudinal news and thus increase polarisation. The main limitation of the study is its focus on the United States and on the specific case of Facebook. Like the Alcott et al. study, it also does not allow us to assess the extent to which the polarising effects of social media could be mitigated by consumption of news through other sources. This is an important evidence gap for the UK context where broadcast news is regulated and PSB providers are more prominent than in the US.

The study also has implications for methodology and future evidence gathering (a topic we return to in Section 8) by showing how survey and tracking-based methodologies can be used in parallel to gather a more detailed picture on consumer behaviour to answer the policy questions Ofcom is interested in.

4.2. Details on studies which are more sceptical as to the link between intermediated news consumption and polarisation

We find the results above compelling, but there are other studies questioning whether social media use increases polarisation. As we develop in this Section, the empirical findings of these other studies are not necessarily in conflict with the two studies above.

The first category of studies reaching somewhat different conclusions are those questioning the mechanism by which social media can increase polarisation or which question the premise of whether polarisation is in fact increasing. For example, a literature review by

55 Segregation is defined as the scaled standard deviation of partisan news exposure. This can be interpreted as the expected square distance between the slant of news sites visited by two random participants in the sample.

Arguedas et al. for the Royal Society argues that concerns around echo chambers and filter bubbles are overstated and that *"most people have relatively diverse media diets"* and that use of platforms "*generally lead to slightly more diverse news use"*, arguing that this implies more limited effects on polarisation.

The same study also notes that, on some metrics polarisation has declined in Europe in recent years.[56] Similarly, Boxell, Shapiro and Gentzkow (2017) find that the growth in polarisation in the US in recent years is largest for the demographic groups *least likely* to use the internet and social media. For example, the increase in polarisation is higher for those older than 75 than for those aged 18–39. The authors conclude that *"none of this is to say that the rise of digital technologies is not important"* but conclude that the data supports polarisation resulting from a broader range of factors than just the rise of social media.

We do not see these findings as necessarily being in conflict with the findings from the experimental studies described above for the following reasons:

- Whether consumption of news via social media increases polarization is a distinct question from whether polarization in society is increasing or whether there are other factors contributing to these trends. The experimental studies above indicate that such a causal effect exists, but this does not mean that it is the only factor at work. For example, higher polarization in older individuals in the US could reflect other factors such as the growth of cable news channels.[57]
- Similarly, the causal effect of social media usage on political polarization is a distinct question from whether online consumers see a breadth of news sources online. It might well be that most users of social media see news from a breadth of different sources, but that the operation of their recommendation algorithms still have an impact on the sources they see and the level of polarisation. Indeed, this is the finding of the Levy study above.
- A further issue is that some of the studies which are sceptical of polarization effects[58] have focussed on the extent of "echo chambers" in a relatively narrow sense (instances where consumers see *only* content from a particular part of the ideological spectrum). While it is comforting to know that such instances are rare, this does not preclude polarizing effects through other channels.

56 Arguedas, Robertson, Fletcher and Nielsen (2022).

57 See, for example, Delavigna, S. Kaplan. E. 2007. "The Fox News Effect: Media Bias and Voting", *Quarterly Journal of Economics*.

58 For example, Fletcher, Robertson and Nielsen (2021) measure echo chamber by "the percentage who say they *only* get news from left-wing or right-wing online news outlets in that country" (original emphasis).

- Similarly, analysis of echo chambers based on the diet of consumption within social media platforms may be subject to measurement effects. It may be that social media increases the diversity and breadth of *outlets* consumed, but still results in a mix of news *content* that has polarising effects on individuals. For example, Fletcher, Kalogeropoulos, and Nielsen (2021) find that search engines and social media are associated with a more diverse news diet but with more partisan outlets featuring more prominently (partly explained by the fact that BBC, considered less partisan, accounts for about half of direct online news consumption – so direct consumption is less partisan but also associated with fewer outlets). Further evidence in this fine is provided by the work of D'Amico and Tabellini (2022) discussed further below.

A second category of studies are those finding quite different results for other online intermediaries and, in particular, search engines like Google. While there are no parallel experimental studies looking at consumption of news via Search engines, the observational studies that exist provide a rationale for why consumption of news via search engines could be less polarising than through social media platforms:[59]

- Nechushtai and Lewis (2019) find that the Google News algorithm does not lead to echo chambers, there is little personalisation, and little evidence of ideological or geographic bias in the "Top 5" article ranking. This is based on observing 168 individuals' online news consumption behaviour, using Google News for stories related to the 2016 U.S. presidential campaigns. The observations looked at the top 5 stories for individuals with either political lean as well as other factors like geography. The study finds that "*Users with different political leanings from different states were recommended very similar news, challenging the assumption that algorithms necessarily encourage echo chambers"* and that "*the news agenda constructed on Google News replicates traditional industry structures more than disrupts them*". At the same time, the study finds that five organisations alone account for about a half of all recommendations, thereby leading to a concentration in outlet reach.
- Fletcher and Nielsen (2018)[60] find that *"those [users] who find news via search engines (i) on average use more sources of online news, (ii) are more likely to use both left-leaning and right-leaning online news sources, and (iii) have more balanced news repertoires in terms of using similar numbers of left-leaning and right-leaning sources."* This finding is consistent with Nechushtai and Lewis (2018) but relies on a survey of participants from the UK as well as the US, Germany and Spain, expands beyond Google to Yahoo and Bing. It uses the 2017 Reuters Institute Digital News Report data of about 2800 observations that measures the news diet of two groups: those that use search engines to search for news topics, and those that do not. The main limitation of this study is the lack of causal inference studied – i.e., the work does not explain how the algorithm plays a role in the polarisation. Similarly, Levy (2021) also reports that search engines are associated with lower levels of polarisation (see Figure 17).

These studies are smaller scale (e.g., the Nechushtai and Lewis study contains just 168 participants) and are based on observational data rather than causal designs. They are

59 Other less notable studies reviewed with a similar conclusion are Zuiderveen Borgesius et al. (2016); Dutton, Reisdorf, and Blank (2017); Haim, Graefe, and Brosius (2018).

60 The authors are associated with the Reuters Institute for the Study of Journalism, which is partly funded by Google, however, the authors indicate that the funders were not involved with the analysis or writing of the research.

also impacted by some of our broader commentaries on gaps in the literature.[61] However, it seems intuitive that search engines have different effects on polarisation than social media as a result of lessened personalisation and a greater tendency to link to more prominent news outlets or coverage of breaking news rather than opinion. Compared to social media platforms, search engines have fewer datapoints on the user (e.g., a search engine does not have access to a user's self-reported interests and may only deduce the characteristics from the search history or their interaction with other products operated by the Search provider), which likely lowers their ability to personalise the search results.

4.3. Gaps in the evidence of the effects of intermediated consumption on polarisation

Overall, there is a compelling body of evidence that news consumption intermediated by social media can lead to higher levels of polarisation.[62] However, several key questions are not yet answered in a robust manner:

Better evaluating the overall magnitude of these effects and how they relate to other factors. The outcomes in polarisation are shown to be caused by social media consumption but there may be other confound factors playing a role – for example, generational effects. Broader demographic studies find that polarisation has been on the rise since at least the 1960s and the UK is, for example, a less polarised society than the US (see Arguedas et al. (2022) for a discussion).

How much could these issues be mitigated by consumption of news through traditional channels? This is an important implication for Ofcom's current survey design as it can say whether sampling of non-intermediated news consumption is biased towards lower polarisation. Wojcieszak et al. (2021b) partly answer this by showing that using search engines besides social media leads to a more diverse diet (less polarising). Dubois and Blank (2018) also show that consuming traditional media significantly lowers the echo chambers but neither of these studies involve a detailed investigation of the potential mitigation of polarisation through consumption of news through traditional channels.

Understanding better the mechanisms by which social media use intensifies polarisation. Allcott et al. (2020) study provides a "reduced form" message that use of social media raises polarisation on various metrics. The Levy study provides information on a potential mechanism: algorithms not showing counter-attitudinal information as much as they could do. However, there are elements of these mechanisms which need to be understood better. Most notably, much of the analysis of consumption is relatively high level relying on outlet-level classifications of ideology. Going forward, one would want a more granular understanding of the ideological content of information promoted by platforms and consumed by users. The mechanisms at work could also have different implications for policy. For example, a potential mechanism is that social media increases consumers' news knowledge, which then leads to a higher polarisation. If this were an important driver it would introduce a trade-off for policy.

61 For example, they too code the ideology of news sources at an outlet level when it might be more relevant to do so at the level of individual articles or pieces of content.

62 For a breakdown of social media platforms and search engines covered by the studies reviewed, see Appendix C.

Understanding better the degree to which polarisation effects of intermediated consumption is due to individual choices vs algorithmic drivers. The literature also does not yet fully explain the relative importance of individual decision making vs. platform decision making in determining the diet of consumption viewed by users of online intermediaries. The Levy study provides compelling evidence that platform decisions play an important role (e.g., because the amount of counter-attitudinal news shown is limited without intervention by the user and that subscribing to a service has less impact on its visibility if it differs from the user's ideology), but a precise quantification of these different effects is unavailable. Some other studies, for instance, Bakshy et al. (2015) show that individuals' choices and their network play a larger role in determining the content they see.

Evolution in platform design and validity of past studies. A necessary restriction of the studies reviewed is that it takes as given the internal operations of the online intermediary (i.e., takes Facebook, Twitter, Google, etc as in the form at the time of the study). This can tell us the impact the online intermediaries have *in their current form* but cannot tell us the impact of potential targeted interventions. For example, it could be that changes in Facebook's algorithm design and curation could allow the benefits of greater news knowledge identified by Allcott et al. (2020) to be achieved without accompanying increases in polarisation.

Generalising findings across platforms. As above, polarisation effects can be expected to vary by platform, however, the majority of the literature studies Facebook or Twitter. Cinelli et al. (2021) analyse content from Facebook, Twitter, Reddit and Gab and find higher levels of segregation on Facebook compared to Reddit – however, the study does not explain in depth how the platform design affects this outcome in a way that would be meaningful for the issues at hand. Kitchens, Johnson and Gray (2020) also discuss the differing effects on polarisation across platforms and find that increased use of Facebook is associated with increased source diversity and more partisan news consumption, while Twitter does not change either. This study however also fails to explain the causal linkage between the platform design and the outcomes in polarisation.

Geographical differences. Many of the studies we consider are focussed on the United States which of course has its own features of partisan politics and issues around polarisation. The results are intuitive, but it is an empirical question the extent to which they read across to the UK. The US market is associated with a higher degree of polarisation than in the UK according to Arguedas, Robertson, Fletcher and Nielsen (2022). Sunstein (2017) also points to evidence that polarisation, or "partyism", has been increasing in the US since the 1960, though the rise of social media has increased it.

5. EVIDENCE ON ECHO CHAMBERS AND FILTER BUBBLES

5.1. Learnings from the literature on echo chambers and filter bubbles

We now turn to the question of echo chambers and filter bubbles, where the concern is that consumption of news via online intermediaries results in consumers only accessing news that conforms to their existing beliefs either because of user choice and association with likeminded people (echo chambers) or through the effect of algorithms surfacing content they agree with (filter bubbles).

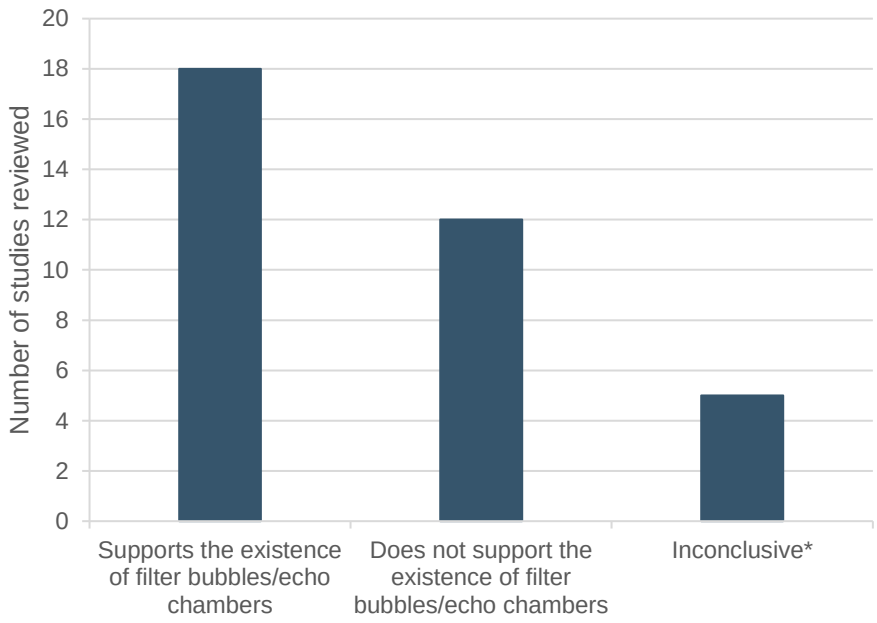
Echo chambers and filter bubbles are one of the mechanisms via which the polarisation effects discussed in the previous Section might occur. However, the literature gives a more

mixed picture. The main takeaways from the literature on echo chambers and filter bubbles are:

- **The theoretical economics literature identifies plausible mechanisms to induce echo chambers/filter bubbles.**
- **The empirical literature supports consumption via social media increasing the tendency for consumers to see like-minded content. However, the extreme case of consumers *only* seeing content that supports their views is rare.**
- **The evidence in respect of search engines is less detailed but does not support them having filter bubble effects** (which also indicates less of a concern for polarisation).
- **The literature differs as to whether the primary issue is demand-side effects (user-induced echo chambers) or supply-side effects (algorithm-induced filter bubbles), but supports both effects being present. There is causal evidence that algorithm design impacts the pattern of consumption on Facebook.**
- **UK-specific evidence from studies relying on survey and passive web-tracking data implies that online intermediaries do not lead to echo chambers but may still result in greater polarisation.** A UK-specific mechanism is that, in the direct channel, the BBC accounts for about half of news consumption, while its share of consumption via intermediaries is much lower and the more partisan outlets feature more prominently.
- **Echo chamber/filter bubble effects may be moderated by consumption of news through non-intermediated channels, but the size of these moderating effects is unclear.**
- **There remain important gaps in the evidence with the most critical being that the most reliable studies are US based and that most analysis has focussed on measuring ideology at the *outlet* level and has not looked at whether effects may manifest at the level of the individual *content* shown to users.**

The empirical studies find evidence of echo chambers and filter bubbles being associated with news intermediation. On a pure count of the number of studies, the chart below shows that the majority of studies show social media leading to effects on these lines, in the sense that social media leads to consumption of more like-minded content even it does not result in them *only* consuming like-minded content. The studies do not all use a consistent definition of echo chamber and filter bubbles, so the counts used are to be taken as a broad indication rather than an exact assessment. However, there is a material number of studies finding evidence against the existence of such effects.

Figure 18: Studies examining the existence of echo chambers and/or filter bubbles**



Source: CRA analysis. *Inconclusive refers to studies analysing polarisation, echo chambers and/or filter bubbles but do not provide a firm conclusion on the online intermediary effect on echo chambers or filter bubbles. ** Since the studies do not all define echo chambers and filter bubbles in a consistent and strict way, we group these together for the purposes of this chart.

What is the theoretical basis for echo chambers and filter bubbles to arise when news consumption is via online intermediaries?

The underlying premise for these concerns is that consumers prefer news content that supports their prior beliefs and this either results in them self-selecting into news sources that pander to these beliefs (echo chambers) or to algorithms endogenously serving up more such content (filter bubbles).

The economic literature studying traditional news consumption provides two mechanisms for how there might be an incentive to serve up news that pander to pre-existing beliefs.

First, such incentives can arise if consumers have intrinsic preferences for like-minded content. Mullainathan and Shleifer (2005) show how, if consumers have such tastes, competition between profit-maximizing media outlets results in them choosing ideologically slanted positions so as to appeal to different segments of the market and competitively differentiate themselves from one-another. This is consistent with the work of Gentzkow and Shapiro (2010) who estimate the slant that would be chosen if newspapers independently maximized their own profits and compare these profit-maximizing points with firms' actual choices. They find that readers have an economically significant preference for like-minded news, which accounts for about 20% of the variation in the news slant observed across newspapers. Mullainathan and Shleifer ultimately conclude that

competition between news outlets may counterintuitively worsen polarisation even while it reduces prices, a sobering proposition that has some empirical support.[63]

Second, such incentives can also arise even if consumers are fully rational. Gentzkow and Shapiro (2006) present a theory model in which consumers prefer to consume "high quality" news and infer that a news outlet which mirrors their prior beliefs is more likely to be high quality, something which low-quality news outlets can then exploit by telling users what they want to hear.

In the context of intermediated consumption, the proposition would be that platforms can use personalised algorithms to tailor consumption on an individual basis and maximizing the differentiation in content consumed across individuals either because this is what individuals enjoy reading, or because consumers will interpret this diet of news as more reliable or useful. While we are not aware of studies formalizing these effects in the case of online intermediaries, the basic hypothesis is taken seriously in the literature.

Consistent with this intuition, Bryanov, et al. (2020) use a field experiment in which they invited participants to use a custom-built news portal and randomised whether they were shown stories based on Google searches or adjusted to show more "congruent" content that matched the participants' prior beliefs. They found that participants shown the more ideological content used the portal more and were less likely to click on mainstream news sites. While there are inevitably limitations in reading across from such an analysis to real-life platforms it provides an indication of the potential profit maximizing incentives for platforms to display content that reflects users' existing views.

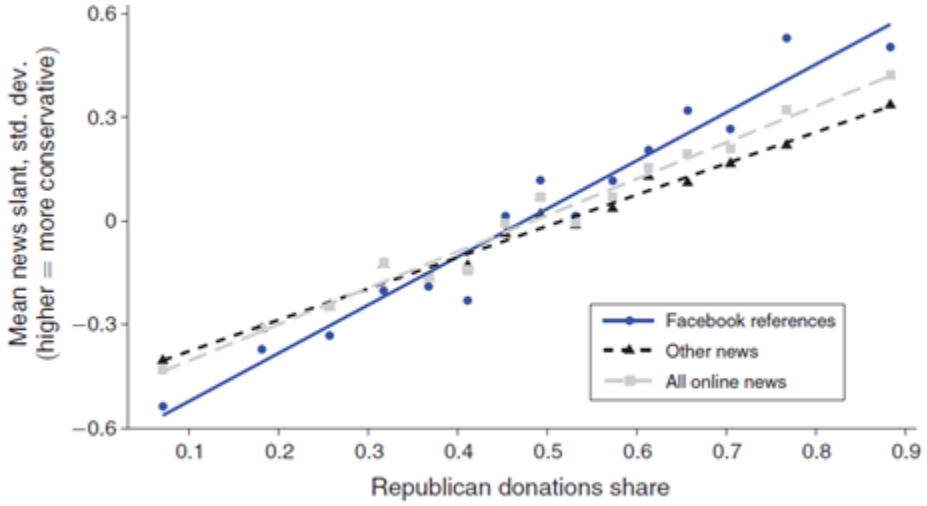
What is the extent of echo chambers/filter bubbles in absolute and relative terms?

Studies have provided different estimates as to the extent of echo chambers/filter bubbles and the extent to which they are more prevalent on online intermediaries than elsewhere.

We have seen above that Levy (2021) found using ComScore data for a panel of users that consumption via social media was more ideologically segregated than direct consumption of news or consumption of news via search. He found also that there was a correlation between the ideology of individuals (as measured by the level of party-political donations in their home ZIP code) and the slant of the news they consumed and that this correlation was stronger for consumption via Facebook than for their consumption of online news in general (see Figure below). This speaks in favour of the filter bubble hypothesis as does the finding that, even when a consumer likes a non-affiliated source, they are less likely to see this source than someone who shares its affiliation.

63 For example, there is evidence that entry of competing newspapers led to more slanted coverage and polarised views (Gentzkow, Shapiro and Sinkinson 2011) and the rise of TV network news (which could not be targeted in the same way) led to more political apathy and lower turnout (Gentzkow 2006).

Figure 19: Finding of Levy (2021) that news consumption is correlated with consumer ideology and that this correlation is stronger for Facebook news consumption than consumption of news through other channels



Source: Levy (2021), Figure 5.

We note that older studies have found different results with Gentzkow and Shapiro (2011) finding that, while "segregation" in online news consumption was higher than offline news consumption, it was low in absolute terms and as compared to offline social interactions. A possible explanation for this discrepancy is the rise of intermediated consumption (although the 2011 study did not find evidence of upward trends in polarisation).

There are also more current studies which argue that echo chamber/filter bubble effects are limited and not warranting of the current level of policy attention. Fletcher, Robertson and Nielsen (2021) report that, in the UK, only around 2% of individuals are in a left-leaning echo chamber while only around 5% are in a right-leaning echo chamber. They conclude that the US is the only country with material shares of echo chamber activity at respectively 10 and 3%.

These findings are of interest for ruling out the most extreme form of the filter bubbles/echo chamber hypothesis, but we do not view them as ruling out the scope for negative effects on media plurality through this channel.

First, the study deliberately defines echo chambers in narrow sense (consumers only seeing like-minded content).[64] It may be that these "pure" echo chambers are indeed rare, but that intermediated consumption through online intermediaries still increases the consumption of conforming output at the margin. As such the findings of the study are not inconsistent with other studies (e.g. the Levy 2021 study discussed above) which find social media algorithms indeed have such effects.

64 The study is based on asking consumers in each country if they consumed news from each of around 30 outlets in each country. These outlets are then ranked by ideology and placed into ideological sets using a classification algorithm with the authors then working out what proportion of users report only consuming output from outlets in a particular ideological set.

Second, and as with much of the rest of the literature, the analysis may also be affected by measuring ideology at the outlet level rather than with reference to individual pieces of content which may overstate the breadth of content that people are consuming.

Similarly, a mix of survey and passive tracking studies finds lack of evidence for echo chambers in intermediated news consumption (by increasing the diversity of information consumers are exposed to) despite the online intermediaries' polarising effects. Arguedas et al. (2022) note that the median number of different sources of news that people in the UK use on a weekly basis offline is two, and just one online. So, while search engines and social media do not vastly expand this number, they do lead consumers to a slightly more diverse news diet than they would on their own. For the UK specifically, the BBC represents a large share of the direct channel, but social media also gives prominence to the more partisan sources, thereby increasing diversity of the channels. Online intermediaries can also expose readers to a higher number of sources (including the so-called "automated serendipity"), but it is not clear whether a higher diversity of outlets is always associated with a higher diversity of content. The empirical studies discussing this in more detail are:

- Fletcher and Nielsen (2018) use survey data of 2000 participants in 36 countries (2017 Reuters Institute Digital News Report survey) tracking the use of search engines and the number and political slant of news sources consumed. The authors find that search engine use is associated with more diverse and more balanced news consumption due to "automated serendipity" where automated news selection results in exposure of news the users would not normally see. They conclude that search engines do not lead to echo chambers and filter bubbles.
- Wojcieszak et al. (2021b) use an observational data on web browsing history records from 636 Facebook users in the US paired with survey self-reports. While the participants were recruited via Facebook ads, the data allows the authors to study usage of both search engines and social media platforms. They find that search engines and social media are drivers of more diverse exposure, rather than generating like-minded echo chambers or algorithmically curated filter bubbles.
- DuBois and Blank (2018) use 2017 survey data on the UK with 2000 participants. They show that high diversity in news consumption is associated with avoidance in echo chambers. The authors do not distinguish between the online and offline users (and hence do not make conclusions on the algorithmic selection) but find that the broader media environment, rather than any single social media platform, there is little apparent echo chamber.
- Flaxman et al. (2016) use observational web-browsing records of 50,000 US users in 2013. The authors find that social media and search engine usage is associated with higher levels of political polarisation. But perhaps counterintuitively, the authors also find that these channels are associated with a greater exposure to opposing perspectives. The authors do not observe whether segregation is the effect of algorithmic filtering of the news stories appearing in one's social feed, or the result of ideological similarity among one's social contacts.
- Cardenal et al. (2019) utilise survey and web-tracking data from Spain of 408 respondents in 2015. They find little support for echo chambers in online environments and report that selective exposure intensifies as news consumption increases. This raises additional behavioural considerations around key events in the news.

- Fletcher, Kalogeropoulos, and Nielsen (2021) use passive web tracking data from 13,709 participants in the UK in 2017. The authors also find that search engines and social media are associated with a more diverse news diet but with more partisan outlets featuring more prominently. This is partly explained by the fact that in the direct online news consumption, BBC accounts for about a half of the visits, while BBC's share is much lower when news is accessed via the intermediaries (and other more partisan outlets take a higher share). Hence, mediated news consumption is associated with lack of echo chambers and higher polarisation.

As discussed above, the drawback of these studies is lack of their ability to identify a causal link and to robustly distinguish between self-selection vs algorithmic selection resulting in echo chambers.

The role of algorithmic design vs. social network effects?

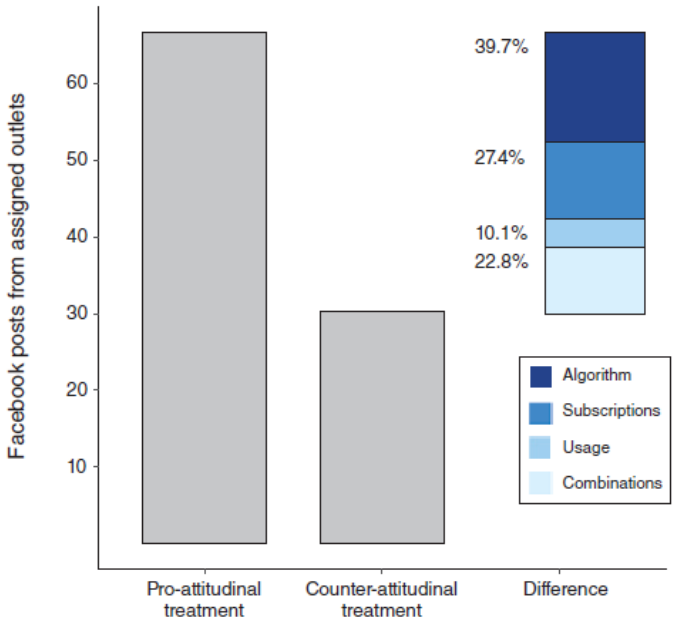
The next question is the breakdown of any observed effects as between echo chamber effects (consumers *self-selecting* into like-minded content or being induced to do so because their friends or the people they follow share similar views) or filter bubble effects (platform algorithms disproportionately serving such content to consumers. i.e., *algorithmic selection*). Put differently, do the effects result primarily from the *demand* or *supply*-side?

Understanding this mechanism is important as it has implications for the policy response: a supply-side issue might be mitigated by simply reducing the focus of algorithms on maximizing user engagement, whereas a demand-side issue might not be fully addressed by such an intervention. Our read of the literature is that it shows that both effects are likely to be present but does not yet permit a precise partitioning as between them.

Levy (2021) explores this issue and finds that supply-side effects are predominant. Levy finds that *"Facebook's algorithm is more likely to expose individuals to news matching their ideology, conditional on subscription"*, with the algorithm having a larger effect than the sharing behaviour from the user's friends.[65] This is summarised in the figure below, showing that the strongest force associated with participants' increased exposure to pro-attitudinal news is the algorithm. Even when individuals are willing to subscribe to outlets with a different point of view, Levy finds that Facebook's algorithm is less likely to show them content from those outlets.

65 Figure 6 of the study breaks down the treatment effects on echo chambers within subscriptions, Facebook news exposure, browsing behaviour and sharing behaviour. The treatment was found not to have a statistically significant impact on the browsing and sharing behaviour.

Figure 20: The "filter bubble" of Facebook's algorithm



Source: Levy (2021), figure 10.

Bakshy, Messing and Adamic (2015)[66] find that the network of friends is the primary driver of what content a user sees on social media, in contrast to Levy who finds that the pro- or counter-attitudinal treatment does not have a significant impact on the browsing and sharing behaviour.[67] Compared with algorithmic ranking, individuals' choices played a stronger role in limiting exposure to content that conflicted with users' existing beliefs. The authors arrive at this finding by analysing Facebook data of more than 10 million active US users 7 million of which self-reported a political affiliation. The data allowed them to (i) compare the ideological diversity of the broad set of news and opinion shared on Facebook with that shared by individuals' friend networks, (ii) compare this with the subset of stories that appear in individuals' algorithmically ranked News Feeds, and (iii) observe what information individuals choose to consume, given exposure on News Feed. This study finds evidence of individual actions playing a strong role in echo chambers, albeit it does not allow us to judge how important individual actions are as compared to algorithmic recommendations.

Homophily in social networks plays a role in the extent to which the social networks promote self-selection into echo chambers and filter bubbles. Cinelli et al. (2021) analyse content from Facebook, Twitter, Reddit and Gab and find higher levels of segregation on Facebook compared to Reddit. The authors find that the aggregation of users in homophilic clusters dominate online interactions on Facebook and Twitter. Colleoni, Rozza, and Arvidsson (2014) study homophily on Twitter users in the US. The authors find that if one looks at Twitter as a social medium, higher levels of homophily and a more echo chamber-like structure of communication is apparent. But if one instead looks at Twitter as a news medium, looking at information diffusion regardless of social ties, then

66 It should also be noted that the authors of the study were at the time of the article's publications employees of Facebook.

67 See Levy (2021), Figure 6.

lower levels of homophily and a more public sphere-like scenario is apparent. The authors propose that social behaviour needs to be considered (e.g., the Democratic and Republican political culture that the study investigates). Halberstam and Knight (2016) study information on the links between voters within the observed network of voters ahead of the 2012 election. The authors find strong evidence of homophily, with conservatives more likely to link to conservatives and liberals more likely to link to liberals. They conclude that social media may be a force for further exacerbating the majority-minority gap and echo chambers for all groups.

Can echo chamber effects through online intermediaries be mediated by consumption of other sources?

A key policy question for Ofcom is the extent to which echo chamber effects might be mitigated by consumption of news through other sources.

The theoretical economic literature provides support for the notion that access to alternative sources and an ability to "cross check" information can act to discipline biased information transmission. In Mullainathan and Shleifer (2005) theoretical model, which assumes bias arises from the demand side as a direct result of consumer preferences, introducing the ability for consumers to cross check news sources increases the informativeness of the information they receive. Similarly, in Gentzkow and Shapiro (2006) coverage, where bias results from consumers rationally inferring information, they agree with is higher quality, the ability to choose between different sources tempers the ability and incentive of publishers to introduce bias. While the literature pre-dates the rise of intermediated consumption it does provide a source for hope that the most malign effects can be avoided via consumption of other sources alongside intermediated consumption.

One study in this vein is Dubois and Blank (2018), which uses a nationally representative survey of adult internet users in the United Kingdom (N = 2000). The authors argue that considering users' broader media consumption implies that *"only a small segment of the population are likely to find themselves in an echo chamber"*.

Other studies have found that consumption of news via non-social media platforms such as search engines may have moderating effects. Levy (2021), for example, arrives at this finding while noting that search engines may still have other negative impacts on news publishers, including evidence that it leads to the concentration in the number of publishers being ranked in search results.

However, we consider that all of these conclusions are somewhat tentative as they are still largely based on descriptive analysis of consumption patterns rather than analysis of causal effects. The ideal analysis would be to conduct experimental assessments of the impact of intermediated consumption on different categories of consumer and to determine whether the effects differ substantially as between those who compete a broad range of non-intermediated news vs. those who consume primarily via online intermediaries. To our knowledge no such study has been conducted.

5.2. Gaps in the evidence on echo chambers and filter bubbles

There is a large number of studies assessing the existence of filter bubbles and echo chambers. However, there remain important gaps:

- **Lack of experimental evidence for the UK.** While there are UK studies looking at this issue, the most compelling experimental studies remain US-based.

- **Lack of targeted analysis to assess the cross-cutting effect of consumption via non-intermediated channels.** Ideally, we would want to know more about how consuming news outside of social media or other intermediaries moderates potential echo chamber effects.
- **Potential lack of representativeness of studies based on user tracking.** Many of the studies that use tracking data track desktop or laptop use rather than tracking mobile usage. This means that their findings do not necessarily carry over for news consumption on mobile devices. This is highlighted in Fletcher, Kalogeropoulos, and Nielsen (2021), noting that their *"findings cannot necessarily be extended to the world of mobile news"*. Dunaway, Searles, Sui, Paul (2018) conduct two lab experiments and one observational study to show that mobile news access is high but attention per session is relatively lower as most sessions involve less time and lower recall by the users. At the same time, a large part of the news access is coming from mobile devices, the study reports comScore data[68] that 27% use desktop to access news, while 36% use a tablet and another 36% use a mobile device to access news.
- **Lack of access to platform algorithms.** Platforms are understandably protective of the details of the workings of their algorithms. In a world without data constraints, one would ideally do more direct experimental studies (e.g., by making targeted changes to platform algorithms to a random sample of users and tracking the impact on user consumption patterns and behaviour). Such studies are unlikely to be available to academics but could be available to regulators (or the online intermediaries themselves), an issue we return to in Section 8.
- **Reliance on outlet-level ideology measures.** Another key gap in the literature addressing echo chambers and filter bubbles is that while intermediation may increase the number of news outlets one is exposed to, these might still present one-sided views and not necessarily increase diversity in terms of ideological content. Indeed, the large majority of studies reviewed measure ideology of content at the outlet level. As such, the ideology of *content* of the individual articles an outlet produces, and a user consumes is not measured directly.

  A notable, and recent, exception is D'Amico and Tabellini (2022). This study looks at the content of news articles commented on by users on Reddit. The study finds that Reddit users on opposite partisan sides comment on the same news sources, but on different news, suggesting that an analysis based on outlet-level ideology measures risks giving an incomplete picture. The authors find in particular that users are more likely to engage and comment on news stories which are negative towards political candidates they disagree with. As such, the ideology of content of the individual articles an outlet produces, and a user consumes is not measured.

68 Dunaway, Searles, Sui, Paul (2018), table A2.

- **Measuring ultimate impacts on political outcomes.** Finally, there is also a gap in the literature measuring the impact of echo chambers on the broader democratic process. While Alcott et al. (2020) does measure Facebook's impact on political engagement, it does not do so in a way that is more "reduced form" and cannot speak to directly to the extent to which the impacts on political outcomes are due to echo chamber effects vs. some other feature of consumption via social media. Levy (2021) finds little support of exposing Facebook users to counter-attitudinal news outlets having a significant impact on political opinions, although the study does not report on any longer-term effects (e.g., if one was exposed to counter-attitudinal outlets over a longer time period than a few weeks).

6. IMPACT ON MISINFORMATION AND RESULTING INFLUENCE ON POLITICAL OUTCOMES

There is a large number of studies looking at misinformation and fake news, either in the context of echo chambers or polarisation. We summarise the main findings here.

The theoretical literature shows how platform incentives in respect of echo chambers and filter bubbles can also facilitate the spread of misinformation. Acemoglu, Ozdaglar, and Siderius (2020) set out a theoretical model showing how the same incentives that contribute to echo chambers and filter bubbles can also encourage the spread of misinformation. The study considers a stylized model of a social network and assumes that participants in the network enjoy seeing the content they share "go viral" but dislike being caught spreading false information. Its main finding is that users are more likely to share misinformation within their social network if it exhibits a high degree of homophily (i.e., users are connected to others who are ideologically similar). This is because, in a situation of homophily, users expect positive social media feedback and little negative reactions and so are more likely to share content even if it is of low quality or contains misinformation. Users would in turn be much less likely to share misinformation within a network with a low degree of homophily because other users would be more likely to dissent with content they disagree with in the first place. This implies that the same forces which might encourage social media platforms to encourage homophily could also propagate misinformation.

By contrast, the authors establish that high-reliability content (i.e., content on the opposite end of misinformation on the reliability spectrum) tends to spread anyway because most users recognise it as such and share it. Therefore, if platforms' incentives cause them to promote homophily then this will tend to increase the spread of unreliable information relative to reliable information.

One counterargument to these findings would be that the paper fails to model the incentives for content moderation, for instance along the lines of Jimenez-Duran (2021). The paper shows that content moderation can actually increase advertising revenue, highlighting that platform incentives are more complex in a two-sided pricing setting.

There are studies which cast doubt on the ultimate impact of misinformation on the outcome of elections, albeit that these studies are in the context of particular events. Misinformation is also studied in the context of elections, and the 2016 U.S. presidential election specifically. There are few studies with a more robust methodology on this topic, but those that exist indicate that misinformation and fake news had a limited effect on the

outcome of the 2016 U.S. elections.[69] One such study is Allcott and Gentzkow (2017), which estimates that if one fake news article were about as persuasive as a single TV campaign ad, the fake news in the study's database would have changed vote shares by an amount on the order of hundredths of a percentage point. This is consistent with the impact of misinformation being limited. The authors note that an average US adult read and remembered several fake news articles during the election period and that few adults believed these articles to be true.

Other studies have found that misinformation and fake news is concentrated among a small share of users.[70] Changes to Facebook's platform design may also have contributed to lessening the magnitude of sharing misinformation.[71]

It is not clear that social media usage influences users' ability to identify false information. Alcott et al. (2020) estimates Facebook's impact on "fake news knowledge" (the awareness of fake news stories and ability to identify them)[72] as well as on users' knowledge of the news overall (as measured by a news knowledge quiz discussed above). The study finds that deactivating Facebook reduces news knowledge overall, but has no statistically significant impact on fake news knowledge.

We note that while this study provides good metrics to measure the platforms' short-run impact on news knowledge, the experiment is not designed to measure the broader outcome, for instance a longer-term effect of consuming the news via the intermediaries vs by direct channels.

There is evidence that decisions by platforms can moderate the spread of misinformation. There is evidence also that platforms can reduce the spread of misinformation. Allcott, Gentzkow and Yu (2019) study the proliferation of 9,540 URLs containing misinformation on Facebook and Twitter between 2015 and 2018 and find that there is an absolute reduction in engagement with these on Facebook starting in 2016 as well as a reduction relative to the level on Twitter. Acemoglu, Ozdagar, and Siderius (2020) also place agency on the algorithmic design to either enable or contain the spread of misinformation.

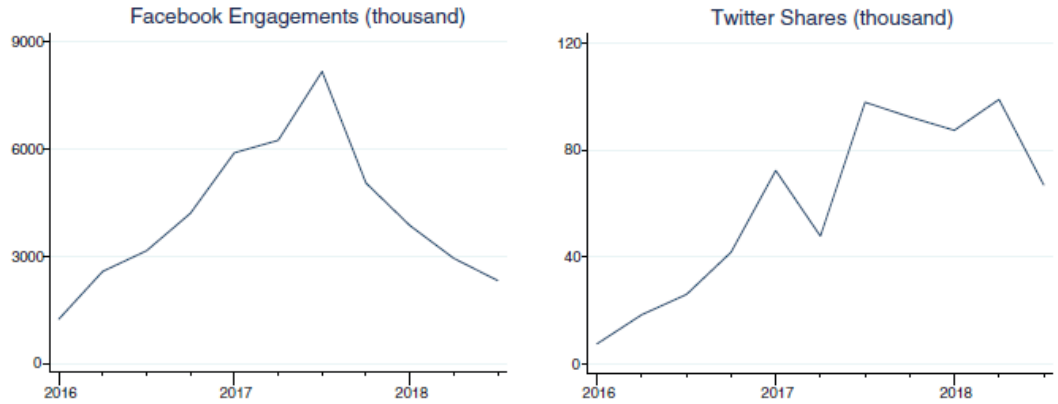
69 Allcott and Gentzkow (2017); Allen, Howland, Mobius, Rothschild, Watts (2020); Guess, Nyhan, and Reifler. (2020).

70 This is also consistent with Grinberg, Friedland, et al. (2019) and Azzimonti and Fernandes (2021)

71 Allcott, Gentzkow and Yu. (2019)

72 This is done by asking the study participants to identify the correct statements among some statements on current affairs, including "statements from fake news stories, rated false by third-party fact-checkers Snopes.com and Factcheck.org, that circulated heavily within a four-week period before the survey." The fake news knowledge variable is measured by the count of fake statements correctly rated as "false" plus one-half for every statement about which the respondent was unsure. This is then assessed against the control group which has not had its Facebook deactivated.

Figure 21: Difference in the prevalence of misinformation across Facebook and Twitter over time



Source: Figure 3 of Allcott, Gentzkow, and Yu. 2019. "Trends in the diffusion of misinformation on social media", *Research and Politics.*

Figure 21 shows that misinformation can spread differently across different platforms, a finding which is important both because it shows policy choices by platforms can address the issue of misinformation and because it means that the other empirical results in the literature are likely to depend on the policies that were in place at the time.

7. ALGORITHMIC BIAS INCLUDING THROUGH FINANCIAL LINKS BETWEEN PUBLISHERS AND INTERMEDIARIES

We now consider additional theories of harm in relation to algorithmic bias. Such biases could arise deliberately because of financial incentives or other biases on the part of intermediaries online intermediary (e.g., overt political biases). Alternatively, the biases may arise inadvertently without any particular motivation, as seems to have been the case according to Twitter's internal study by Huszár et al. (2021) which found that its recommender system inadvertently promoted right wing news more than left wing news.

Bias due to financial incentives. Here the concern is that financial links between platforms and news outlets could distort the positioning and display given to different news outlets with potential knock-on effects for media plurality (e.g., if it results in the marginalisation or exit of news outlets with a resultant reduction in the diversity of viewpoints consumed by users). There are multiple ways such links could arise.

First, such links could arise if platforms owned, or had financial stakes in, individual news outlets. In this case, there might be a financial incentive to divert traffic to these brands over those of third parties. For the time being at least, this concern is hypothetical: we are not aware of any investments by major online intermediaries in news outlets either in the UK or elsewhere.[73] However, one can envisage a role for media plurality considerations in the

73 There are examples of similar scenarios, e.g., Jeff Bezos is executive chairman of Amazon and owner of the Washington Post. However, this falls outside of the hypothetical concern at issue here because Amazon is likely to be a more limited source of traffic for news outlets than is a major search engine or social media platform and because this example involves an individual being invested in two related firms rather than a direct investment by a platform in a news outlet.

case of major platforms vertically integrating into news output either via organic growth or acquisition. Indeed, at the time of finalizing this report Twitter was in the process of potentially being sold to Elon Musk showing that changes in ownership of online intermediaries area not just a hypothetical.

There are close parallels with the regulatory debate around "self-preferencing" by tech platforms and their potential incentive to promote their own products and services over those of third parties. For example, Google has been found to have abused its dominance in search by giving prominent placement to its shopping comparison service while simultaneously demoting rivals in organic search results.[74] In some cases the algorithmic bias has been overt with Google using the appearance of specific competitors in its organic search results as a trigger to show its own service at the top of the results page.[75] The academic literature, however, has pointed out that platforms self-preferencing their own content is not necessarily welfare-decreasing.[76]

Second, such links could arise indirectly if news outlets make use of platforms' other products and services in a way that causes an incremental unit of traffic being sent to one news outlet generating more revenue and profit than a unit of traffic sent to another. For example, Google operates both a search engine and an "ad tech" business which publishers use to display ads on their pages. Google might generate more revenue from traffic sent to a news site which displays more ads on its pages or makes greater use of its ad tech services. Indeed, there have been accusations by subscription news publishers that they have been disadvantaged in search results.[77] Similar issues could also arise for other platforms (e.g., around Facebook's advertising business on third-party ads) and could also consider non-pecuniary links (e.g., an intermediary might give more prominence to publishers where it has greater ability to gather data and/or track users behaviour).

The concern for media plurality is that such incentives can encourage news outlets towards an ad-funded business model which might have knock on effects for media plurality by, for example, reducing incentives for investment in content generation. There is potential for further knock-on effects if such conducts reduced competition in online advertising and thereby further reduce publishers' ability to monetize.

Consistent with these concerns, the CMA has also reported that publishers think that *"in some cases algorithm changes may be commercially motivated to favour the platforms or affiliated parties at the expense of other publishers."*[78] This is echoed by the responses to

74 For instance, as reported in November 2021: https://techcrunch.com/2021/11/10/google-fails-to-overturn-eus-e2-42bn-shopping-antitrust-decision/. We declare an interest having been involved with antitrust cases around this issue.

75 This is reported in a leaked FTC memo on Google's conduct in specialised search. https://www.politico.com/news/2021/03/16/google-files-ftc-antitrust-investigation-475573

76 See, for example, Hagiu, Teh, and Wright (2020), Anderson and Bedre-Defolie (2021), and Yang and Muir (2022).

77 We declare an interest having been involved with antitrust cases around this issue. For coverage of complaints by the Wall Street Journal about its visibility in search and decision to prevent its content from being indexed by Google see: https://searchengineland.com/wsj-first-click-free-269545.

78 https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers

the Ofcom's 2021 consultation on media plurality[79] and there is ongoing litigation by the owners of the Daily Mail against Google making this accusation alongside accusations of ideological bias discussed further below.[80]

Third, there may be incentives for the platforms to favour other products or technologies they operate. For instance, there have been accusations that Google and Facebook have made prominence on their platforms contingent on news outlets displaying their articles in proprietary formats (e.g., Google's "Accelerated Mobile Pages" or Facebook's "Instant Articles").[81] Again, there could be knock on effects for media plurality if, for example, use of such formats reduces news outlets' ability to monetize or invest, help increase the market power of online intermediaries in these adjacent services, or result in a commodification of news consumption.

Other forms of algorithmic bias. As well as biases in relation to financial incentives, there have been accusations of other sources of bias within platforms' algorithms.[82] First, there have been high-profile accusations of political bias in the outlets shown by tech platforms. For example, the Daily Mail has brought a case against Google alleging that it is biased against right of centre outlets[83] and similar allegations have been made by conservative outlets in the US.[84] The key debate here is whether this lower visibility reflects ideological bias per se or that conservative outlets differ on other dimensions that warrant lower visibility. A study by The Economist magazine argued that the apparent bias against conservative outlets disappeared once one controlled for measures of outlet's "trustworthiness" and the total volume of articles published on particular topics.[85] We discuss below experimental studies on this topic.

79 https://www.ofcom.org.uk/consultations-and-statements/category-2/future-media-plurality-uk?showall=1

80 https://www.thetimes.co.uk/article/daily-mail-cries-foul-over-google-bias-zzk0qq3dj

81 The issue of Google Accelerated Mobile Pages (AMP, framework created by Google as a competitor to Facebook Instant Articles and Apple News) (where we also declare an interest) was discussed in depth in the US House Judiciary Committee's report on antitrust issues around big tech. See: https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf?utm_campaign=4493-519

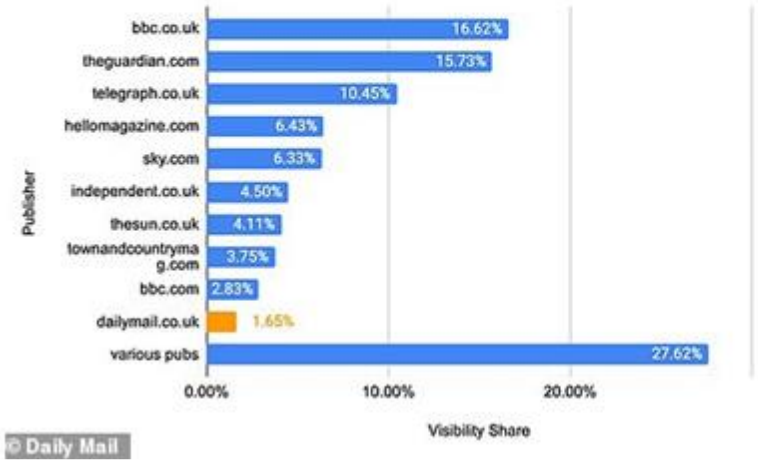
82 Related to the intermediaries' market power, there are concerns that intermediaries may alter prominence of certain publishers or news stories for political reasons and personal preferences of the platform owner. Nevertheless, we have only been able to identify such concerns being raised by the publishers or other regulatory agencies, but no clear manifestation of harm has been evidenced. For example, the literature reviewed discusses the concerns around platforms' political bias in the 2016 and 2020 elections, however, the algorithms have not been shown to actively promote political biased news.

83 https://www.dailymail.co.uk/news/article-9491607/Mail-files-antitrust-lawsuit-against-Google.html

84 https://www.foxnews.com/opinion/dan-gainor-yes-google-censors-conservatives-even-liberal-journalists-now-admit-it

85 https://www.economist.com/united-states/2018/08/30/googling-the-news

Figure 22: Daily Mail antitrust lawsuit against Google alleges that it is demoted vs other publishers



Source: Daily Mail

They detail that changing the algorithms has direct financial consequences, such as lost monetisation through advertising due to falls in traffic[86] but also noted how these sudden changes can lead to significant and potentially wasteful spending to understand the algorithms and to optimise content to appear high up in the rankings.

Second, concerns have been raised that algorithms might give insufficient credit for outlets generating original content and too much credit for "freshness". The concern is that this disincentivises investment in investigative journalism because it permits free riding by other outlets.[87] There are also concerns that it encourages investment by publishers in so-called "churnalism" whereby there is an emphasis on producing content that will rank well in search or social media by, for example, publishing new versions of the same article with minor changes or cross-linking between multiple similar articles to achieve higher ranking. The Cairncross Review (2019) also notes that Google and Facebook increasingly control the distribution of publishers' content online meaning that as a result of their position they can impose terms on publishers without needing to consult or negotiate with them; something which has the potential to threaten the viability of news publishers' online businesses.

86 The Daily Mail's SEO director said that the dailymail.co.uk *"lost 50% of daily traffic*" the day after Google updated its core search algorithm and that it also saw a 90% drop in Google Discover fed traffic. See: https://searchengineland.com/daily-mail-seo-says-site-lost-big-after-june-google-update-asks-community-for-help-317926 Similar reports claim that a change to Facebook's curation algorithm in 2016 to show users more stories from friends and family resulted in a significant decline in traffic to publishers' sites, up to 25 percent for some. https://www.nytimes.com/2016/06/30/technology/facebook-to-change-news-feed-to-focus-on-friends-and-family.html and https://www.theguardian.com/technology/2016/jun/29/facebook-news-feed-algorithm-change-traffic-publishing

87 For example, Twitter has previously taken an initiative to increase relevance over recency https://www.thedrum.com/news/2016/02/11/how-will-twitters-new-relevancy-over-recency-algorithm-affect-brands. Westerman, Spence and Van Der Heide (2014) find that users associate recency with credibility, explaining the increasing role of recency in news on social media.

7.1. Learnings from the literature

Bias due to financial links

While there is a significant policy debate around these issues, the amount of formal theoretical and empirical work is more limited, especially if one focusses only on studies focussed specifically on news consumption. However, there are broader parallels from the literature studying platform incentives more generally and, in particular, the incentives of platforms to artificially promote first-party content. The key lessons from this literature are as follows:

The literature identifies multiple mechanisms by which intermediaries may have financial incentives to "steer" consumers to less relevant but more profitable content. In the context of search engines, de Corniere and Taylor (2013) develop a theoretical framework modelling consumers, publishers and advertisers. Their stylised search engine (i) allocates users across publishers, and (ii) competes with publishers to attract advertisers. They show that a search engine has incentives to send less traffic to sites who are selling ads which compete with its own, even if it is not directly integrated into any of these sites. This is because search engines have an incentive to factor in the price at which it can sell its sponsored links, so the search engine will penalise publishers that display many ads more than consumers would wish.[88]

Burguet, Caminal, and Ellman (2015) also look at the search engine financial links. They study how ad market incentives affect the quality of organic and sponsored links. The authors find that integration of the engine with a small fraction of content reinforces the engine's incentives to divert business from independent publishers to affiliated or owned publishers. This effect may outweigh positive effects of integration and result in lower consumer and total welfare.

Bourreau and Gaudin (2021) present a model of content platforms (motivated primarily by music and video streaming) in which platforms have an incentive to direct users to content which has lower marginal cost (e.g., because it is produced in-house and hence has zero marginal cost or because the content producer has agreed a lower royalty for its content).

Finally, platforms may have incentives to send less content to services which constitute a "dynamic threat" to their core business.[89] This literature was developed in the context of antitrust cases against Microsoft (where the intuition was that a dominant operating system might promote an inferior in-house web browser for fear that third-party browsers could evolve into substitutes for the functions served by the operating system). In the context of news provision, platforms might have incentive to send less traffic to publishers who are better able to keep users on their own sites (e.g., by encouraging them to install an app or navigate to the site direct) and thereby reduce the attention given to their intermediation services in the future.

88 The authors also find that, integration can, but need not, lead to own-content bias integration may lead to own-content bias. This is because own content bias (i) may merely be offsetting the advertising-related bias, (ii) increasing the expected per-user revenue of the search engine, integration increases the incentive to provide high-quality search results, which further disciplines the search engine, and (iii) the integrated publisher has an incentive to lower its supply of advertising space, which benefits users and justifies making this website more prominent.

89 See for example, Carlton and Waldman (2002).

Incentives for "steering" depend on the business models and monetisation strategies of platforms. While the literature identifies circumstances where incentives can exist, an important message is that such incentives are context-specific and depend on the business models and monetization strategies of the platforms in question. For example, Etro (2021a, 2021b) shows that ad-funded (zero price) business models can generate greater incentives for self-preferencing. The intuition is that because "free" (ad-funded) services do not charge users directly for each "click", they have less of an incentive to take full account[90] of the value consumers place on different pieces of content and more incentive to show lower-quality content which generates revenue through other channels (e.g., advertising).[91]

Algorithmic bias due to financial links will tend to be moderated by competition. A recurring theme of the literature above is that incentives for bias can be moderated by competition. The intuition is that showing less relevant content for financial reasons involves a "quality sacrifice". If users have alternatives to the online intermediary in question, then this quality sacrifice is more likely to result in reduced usage of the online intermediary, increasing the costs of such strategies. This indicates that concerns around algorithmic bias due to financial links should be focussed on platforms that plausibly have market power.

The literature identifies a prospect of using algorithm design to exert leverage over content suppliers, which could become relevant with future regulatory initiatives. As above, Bourreau and Gaudin (2021) show in the context of music streaming that platforms have incentives to lead users to content associated with lower costs (e.g., in terms of per stream royalties). They then show how platforms can in turn use the *threat* of algorithmic demotion to obtain lower royalties from content suppliers. While the welfare effects of this in their model can be positive if content producers have market power, it illustrates a potential mechanism of concern and especially so if one believes that the extent of bargaining power between platforms and news producers is already asymmetric and there are social spillovers from content production.

Aggregate ideological bias in algorithms

We have seen in the discussion of filter bubbles above that algorithms can significantly influence the news content users see. Most of the literature has looked at *personalised* biases in algorithms (i.e., that the bias would be towards users' pre-conceived notions whatever they might be), but some studies have looked also for evidence of *aggregate* ideological bias (i.e., that an intermediary might on average send users towards disproportionate parts of the political spectrum). One such study is Nechushtai and Lewis (2019) which finds little support for personalised bias in Google News but a high degree of homogeneity and concentration in the news recommendations with five outlets accounting for 49% of recommendations.

The most credible such study in our view is Huszár et al. (2021). This study, authored in part by individuals working at Twitter, employed a randomised control trial in which some users saw the "traditional" Twitter feed of chronologically ordered tweets by people they

90 Consumers' willingness to pay attention to ads may also decrease as content quality falls, so platforms may internalize these effects.

91 A further discussion on how business models matter for the regulation of digital platforms is available here: https://voxeu.org/article/designing-regulation-digital-platforms

follow, and others were randomly assigned to see tweets selected by Twitter's recommendation algorithm. The study found that, in 6 out of 7 countries examined, *"the mainstream political right enjoys higher algorithmic amplification than the mainstream political left" but* did not find any evidence that the algorithms promoted extreme voices at each end of the spectrum.[92]

We note that such bias need not reflect deliberate ideological choices by platforms (although this claim has been made and it in principle could occur). Rather it could simply be that the vagaries of recommendation algorithms cause them to have this effect and particularly so when it comes to specific topics or issues. Even such "inadvertent bias" could raise media plurality concerns if it relates to platforms operating on such a large scale that even inadvertent slant in their recommendations might be expected to have an impact on views and behaviours in society at large.

Churnalism

To our knowledge the concern that algorithms promote derivative content and "churnalism" is more anecdotal in nature.[93] [94] It is certainly true that algorithms promote "freshness" (such that they may show most prominently a copy of an original article that has been published more recently),[95] but weight given to other factors such as brand reputation may offset this issue to a degree. More work is needed to determine whether current algorithm design appropriately rewards and attributes journalistic contributions by news outlets.

7.2. Gaps in the evidence on algorithmic bias

There is no specific literature looking at online intermediaries' incentives to bias traffic between news sources based on financial considerations and the literature on aggregate ideological bias in algorithms is also small and nascent. While inferences can be made from broader literature and experience on self-preferencing, this is an area where UK and platform-specific evidence in the specific area of news consumption would shed significant light on policy priorities. This is an area where new data gathering initiatives, discussed in the next Section, could play an important role.

8. HOW COULD THE GAPS IN THE LITERATURE BE FILLED?

The objective of this Section is to identify analyses and information gathering that could be done to fill the gaps in the literature discussed above, and to explain what insights these initiatives would bring. Because many of these initiatives can shed light on multiple theories of harm, we set out first the data/information we would propose to collect and then discuss

92 This study was reported more broadly, for instance by the BBC: https://www.bbc.com/news/technology-59011271

93 https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth

94 For example, Twitter has previously taken an initiative to increase relevancy over recency https://www.thedrum.com/news/2016/02/11/how-will-twitters-new-relevancy-over-recency-algorithm-affect-brands. Westerman, Spence and Van Der Heide (2014) find that users associate recency with credibility, explaining the increasing role of recency in news on social media

95 https://support.google.com/news/publisher-center/answer/9606702?hl=en-GB and https://www.forbes.com/sites/denispinsky/2018/01/11/google-news/?sh=5f3ba2ed3d50

how it could be used to test each of the topics at issue. We also discuss how this new data would help fill the evidence gaps identified in the previous Section.

The initiatives we propose can be grouped into the following main categories: user traffic and engagement data to be gathered from platforms and publishers; information and A/B tests to assess the operation and impact of platforms' algorithms; extensions to the existing NCS; replicating the key experimental studies on UK data; and new initiatives to measure the potential extent of filter bubbles and echo chambers based on the *content of the actual content being consumed* rather than the characteristics of the news outlet it comes from or the individual sharing it. Overall, our view is that there is a substantial amount of "low hanging fruit": there is a paucity of information on basic measurement questions (e.g., on the relative visibility of different news outlets on platforms and the pattern of consumption on platforms) and scope for relatively targeted data requests to significantly increase Ofcom's visibility around the theories of harm we have considered.

That said, we recognise that some of these proposals would require the involvement and co-operation of the online intermediaries or information gathering powers. We have sought to set out what we see as effective ways to measure the concerns at issue and leave to one side the question of the legislative or other changes required to make this information accessible to Ofcom. An overview of the proposed initiatives is shown in the table below.

Table 1: Overview of proposed additional information gathering initiatives to address the gaps in the literature

Initiative	Implications for the involvement of online intermediaries
Online intermediaries' data on impressions and "link out" (Dataset 1)	Requires online intermediary involvement
Online intermediaries' data tracking a randomly selected panel of users (Dataset 2)	
Data from publishers on traffic and revenue (Dataset 3)	Does not require any additional online intermediary involvement
Information on online intermediaries' algorithm objective functions and signals (Dataset 4)	Requires online intermediary involvement
Online intermediaries' telemetry and worked examples data (Dataset 5)	
Amendments to the existing NCS	Does not require any additional online intermediary involvement
Experiments on UK data	

8.1. Proposed data gathering from the online intermediaries

8.1.1. Measurement of consumer behaviour on platforms

We understand that Ofcom is already able to obtain referrals data from publishers or by using third-party services that track the web browsing behaviour of users arriving at publishers' websites.[96] However, these data are limited.

- First, referrals data from publishers is restricted to tracking traffic referrals from the platforms, and so cannot measure consumption that is occurring on the platforms themselves.
- Second, both sources of data are limited in terms of understanding the information displayed to users when they were on the platform. As such, one cannot distinguish between traffic resulting from different use cases (e.g., a click on a news outlet the consumer has chosen to follow, vs. an article shared by a friend, an article shared by an influencer or someone the consumer has chosen to follow, or a news outlet that is endogenously surfaced by the platform's algorithm).
- Third, and relatedly, there is therefore very limited ability to understand the role of decisions by the platform (e.g., in terms of algorithm design) in determining the pattern of consumption by consumers.
- Fourth, there is limited ability to understand the composition across users with different demographics (e.g., the extent to which traffic from a given platform to a given news outlet comes from users with particular characteristics). While there are services (e.g., ComScore, Ipsos IRIS) that can provide user tracking linked to a panel of underlying users, we understand such data is expensive to generate and may also suffer from representativeness issues as the panels tracked by these services are by definition self-selected.

For all these reasons, a high priority evidence gathering initiative would be to gather data from the major platforms which provides fuller and more granular data on the traffic they send to news outlets; and provides an insight on how consumers engage with their platform. We propose gathering two datasets to achieve this.

Dataset 1: impressions and "link out" data by news source ideally split by type of link, consumer characteristics, and characteristics of the content. Under this proposal, platforms (both major social media platforms and search engines) would provide aggregate data on the engagement with, and traffic to, different news outlets on their platform. In particular they would:

- Provide a time series of daily data on the page views (impressions) and click volumes generated by news source (ideally covering both traditional news outlets as well as other news sources such as political parties, prominent politicians, independent publishers, or posts that talk about news topics) on mobile and desktop devices.[97]

96 For instance, a user arriving at the publisher's website can contain a "referrer URL" which tells the address of the webpage where a person clicked a link that sent them to the publisher's page.

97 We focus here on impressions and clicks, but, depending on the platform one could track other measures of engagement/sentiment (e.g., number of comments or likes).

- A further split of this time series according to the way the news source was displayed when it generated the click/impression (e.g., whether the link was organic or an ad, whether it was a news source the user had chosen to follow, whether the link was shared by a friend, whether the link was shared by someone the user followed etc.).[98]
- A further split of the resulting clicks as between those which led to consumption still within the platform (e.g., a Google Accelerated Mobile Page or a Facebook Instant Article) vs. a link out to a publisher web page.
- A further split with any further categorization of the content used by the platform: e.g., the topic of the article (politics, sports), whether it related to a trending topic.
- A further split by demographics, to the extent available. For example, for services like Facebook the data could be split according to the self-reported ideology of the user clicking/liking the content.[99]
- Data on any revenues and profits generated in respect of visits/views of third-party content including revenue from complementary services (e.g., ads sold on third-party websites).
- Data on the number of news sources over time and their reach (posts, likes, comments received).

Dataset 1 would allow exploring of the following issues:

- Even in its simplest form of providing traffic data without the further sub-splits, the data would allow one to understand the traffic volumes sent to different publishers and potential variation over time and assess whether the pattern of traffic indicates any lean or preferential treatment of specific publishers (e.g., those with financial links to the online intermediary). Because the proposed data is daily one could also look at days with major breaking news to see which news outlets benefit most in case of breaking news and the extent to which originators of new stories receive a larger share of any traffic uplift. Gathering consistent data across platforms would also permit comparisons in the composition of traffic (e.g., do different platforms send traffic that is more concentrated to traditional news outlets or to outlets with different ideological positions or levels of reliability/trust).
- The first and second of our proposed additional splits would allow an understanding of the pattern of consumption on the platform. This would allow one to understand the proportion of intermediated consumption which reflects active consumption decisions (e.g., to follow particular news outlets) vs. more algorithmically generated ones. It would also allow one to understand the importance of the social dimension of different platforms (e.g., how much news consumption is driven by articles shared by friends).

98 This description focusses on a social media context. In a search setting, one would rather want to know about whether the link was organic or paid, whether it appeared in a news carousel or a normal “blue link” and so on.

99 We would anticipate this information to be available on a more complete basis for some platforms than others (e.g., it would be more complete for social media platforms requiring a more detailed profile).

- The split of clicks between those which remain "on platform" and those which go out to the wider web could be used to build a reliable picture of how much news is being consumed on each platform (e.g., reading snippets/articles within Facebook), as well as their relative importance as a source of referrals to publishers' own services.
- The split by user demographics would allow one to assess the question of echo chambers and filter bubbles because it would allow one to look at the extent to which (say) right-of-centre users are more likely to navigate to right of centre websites than left-of-centre users. Because the proposed dataset covers both impressions and clicks it would also allow explain of *why* any such differences in the pattern of consumption were occurring (i.e., was the different share of traffic because consumers were more likely to be *shown* content from outlets they agreed with or to *engage with* this content when they saw it?).
- The accompanying financial data would permit an assessment of whether there are financial links that would incentivise sending traffic to particular publishers over others.
- The number of news sources allows a better understanding of whether social media increases competition between news publishers.

The precise parameters of this dataset could be adjusted to ensure proportionality and reflect the internal metrics tracked by platforms, but it would be a useful tool to understand all of the theories of harm we have considered. The main downside of this data is that it is still somewhat aggregated and does not provide a full insight on how users see and interact with content. As such, it is more suited for providing (albeit still useful) high-level facts and correlations rather than a full picture of the effects of interest.

As such, we propose that gathering also a second, more disaggregated dataset.

Dataset 2: tracking for a randomly selected panel of (anonymised) users. Under this proposal platforms would provide granular user data for a representative panel of users. The goal would be to have more direct understanding of the content shown to users and how they engage with it. Ideally, platforms would provide:

- A panel dataset tracking the behaviour of a random sample of anonymised users.
- For each user over time the data set would provide telemetry data on each visit to the platform in question setting out:
  - The date of the session and whether they were accessing the dataset from desktop or mobile
  - What content was shown and in what order
  - Background information on the content shown: whether it was news or news content; whether there was a link to a content page or website and, if so, the organization it related to; the topic of the content; whether the content was shared by a friend, influencer or news outlet and so on;[100] whether the content was advertising
  - The intensity of consumption of each piece of content displayed (e.g., measurements of how much time the users spent looking at the content)

100 We know that Twitter, for example, has automated tagging of content by platform.

  - How the user interacted with the content (click throughs or subsequent websites visited, likes, comments sentiment, sharing, following the content producer etc.)
  - Tracking of comments/likes of the content/shared by a friend
- The platform would also provide information on each panel member's demographics (e.g., age, ethnicity, gender) and any self-reported political information (e.g., party affiliation)

Dataset 2 would allow one to develop its understanding of all the relevant theories of harm:

- Understand what news outlets are *shown* to users as opposed to just those which are clicked on/engaged with. This is relevant for all the theories of harm which relate to the role of algorithm design in determining behaviour including potential concerns around financial links between news outlets and platforms.
- Understand why consumers obtain news in different ways (e.g., by following links or reading paraphrased/commentary posts from friends vs. directly following news outlets).
- Explore whether individual users exhibit behaviour consistent with filter bubble concerns (e.g., by engaging only with the content that mirrors their ideology) and whether algorithms are exacerbating this issue (i.e., by showing users likeminded content).
- Extent to which other users' recommendations and sharing is being shown, explaining whether this content is similar or cross-attitudinal.
- Understand the extent to which people are more likely to engage with an outlet with comparable ideological position.
- Analyse the diversity of outlets shown to particular users and whether there are categories of users who see a more or less diverse bundle of news sources.
- Understand what sort of news content shows up on platform (e.g., is it more editorial or factual)
- How much traffic due to friend recommendations vs. other use cases.
- Understand whether advertising changes by types of users (e.g., more ads displayed to users who consume more news)
- Understand the volume of news content shown to users vs. other sources of content and whether this mix differs between users.
- Understand if platforms are merely intermediators of news or if they are adding additional value through the social sharing and discussing of news.

This data would provide for a much richer analysis of the evaluation of the impact that online intermediaries have on news consumption. As a practical step one it may be that this data can be provided in an automated fashion (e.g., via APIs) so that information can be tracked in real time and with more limited ongoing costs for the observer and the platforms concerned.

8.1.2. Corresponding data from publishers

Dataset 3: data from publishers. To accompany the data from the online intermediaries, we propose a further data gathering exercise from the publishers. The rationale is to gather side by side data on the mediated and non-mediated consumption of news once the user arrives at the publisher's website. This would entail

- Data on traffic by referral source
- Information on how long the reader interacts with the content on the page
- Subsequent pages visited, including whether the consumer reads more than one page or just a single article
- Revenue data, both the revenue from advertising, as well as revenue per impression for direct visits vs. intermediated ones

We would also propose gathering more qualitative information from publishers on how they are responding to the role of the online intermediaries: this would entail questions on the impact on the publisher's revenue, incentives in content creation (e.g., recency vs original content, global vs local, etc.), impact on the number of journalists being employed over time, etc.

Data from publishers would be a useful cross-check on the information received from platforms. However, we would anticipate the information gathered here to be less granular both because publishers are likely to have more limited data gathering capabilities and because the most important data for understanding the theories of harm at issue relate to data generated within the platforms themselves.

8.2. Developing understanding of the online intermediaries' algorithms

The data above would allow for an analysis of the outcomes of the intermediation on the news consumption. A complementary source of information would explain the workings of the intermediaries' algorithms, although practical considerations will be important as the algorithms often change and are proprietary (e.g., Google has been reported to be adjusting its algorithm 600 times a year).[101] This would involve three types of information:

Dataset 4: information on algorithm objective functions and "signals". We would suggest conducting an audit of the key algorithms which might be expected to significantly impact the news content shown to users. The challenge with doing so is that these algorithms will rely on machine learning and artificial intelligence techniques which are largely "black box" in nature.[102] To get around this and, as a starting point, we would suggest to focus on the most central algorithms for each platform (e.g., the news feed algorithm for Facebook) and request information on the high-level operations of these algorithms, the objective function they are seeking to optimize for (e.g., engagement or some other metric), the variables ("signals") taken account of by the algorithm, the aspects of the algorithm which are personalised for a given user, and detail of any manual or other

101 https://www.politico.com/magazine/story/2015/08/how-google-could-rig-the-2016-election-121548/

102 By this we mean that they will endogenously weight variables to achieve their objectives in a way that may be opaque to the designer and, as such, it may not be possible to fully understand or predict how they combine variables to achieve their outcome. This is distinct from more parametric approaches where it is possible to explicitly link the outputs of an operation to the inputs it relies upon.

ex-post adjustments made to the algorithm to determine the final results.[103] Another starting point would be to obtain archives of moderated data, specially including users or posts that are deleted or labelled for violating misinformation policies.[104] It would be useful to have metadata on whether the content is removed through algorithms or user reports, and to understand the "signals" that moderation algorithms use.

Conducting this sort of audit would speak to the following questions of relevance for the main theories of harm:

- Understanding the extent to which the algorithms have features which lend themselves to the filter bubble/echo chamber considerations above: e.g., are they heavily personalised, do they optimize for engagement, do they rely on "signals" (e.g., expressions of negative reactions) which could be expected to amplify polarising content.
- Understanding whether platforms are likely to differ in the strength of these effects (e.g., the literature review above suggests that Google News ranking algorithm does not personalise to the same degree that Facebook's news feed algorithms does, and some Twitter users may still be seeing news feeds showing results in chronological order without any algorithmic selection).
- Understanding whether the basic conditions are in place for the "financial links" theory to be an issue: e.g., whether the algorithms include profitability as a signal or if there are other signals which might endogenously cause the algorithm to prioritise publishers with whom the platform has a financial link or based on other non-pecuniary compensation (e.g., data access) received from publishers.
- Understanding the extent to which the algorithms rely on hardcoded adjustments which might create scope for platforms to exert influence over the content people see.
- Understanding the extent to which the algorithms focus on recency and whether they take account of originality (i.e., of who is the original source of the story). This is relevant for the question of whether algorithms are suitably rewarding and incentivising content generation and investment.
- Evaluating platforms' moderation policies (e.g., testing whether platforms induce political bias through their moderation practices).

For a fuller understanding of the algorithm's impact on the real-world news consumption, one would need to be able to test its hypotheses on the algorithm. This could be done by gathering more detailed telemetry to understand why algorithms output the results they do; or by conducting A/B tests to understand the impact of making changes to these algorithms.

Dataset 5: telemetry and worked examples data, would also shed light on how varying the signals/inputs for the algorithm affect the output of the algorithm.

103 For example, as discussed in Section 3 we understand that the quality scores delivered by Facebook's news feed algorithm are adjusted to deliver a mix of different types of content and to downweigh certain types of content that are deemed undesirable.

104 Some forms of moderation imply the disappearance of content, so it is not possible to recover it ex-post.

- Providing telemetry/worked examples for a random sample of visits to the homepage. We suggest requesting the platform to show the candidate content that was to be displayed to the user, the scores for the content that was shown and provide one with visibility as to why the selected content was selected.
- Providing telemetry/worked examples for a random sample of posts and of posts with high levels of engagement explaining how the scores for these posts were computed.

The telemetry and worked examples data would allow one to understand what factors drive prominence and whether algorithms are driving negative dynamics. This would also ensure for a better understanding of the financial incentives the algorithms may cater to (e.g., increasing engagement at the cost of polarisation).

This would also help explain how the algorithm's treatment of news content differ across users (e.g., do some consumers receive higher volumes of news information and if so, who are these consumer groups). It would also explain how the different consumer groups are targeted with advertising (e.g., those who consume more news may be more valuable than those who do not, implying different ad content focussing on news (and if so, is the advertised news biased).

A/B tests. We suggest running periodic A/B tests to assess the impact of different aspects of platforms' design decisions on outcome variables of interest. Ideally, one could specify the outcome variables of interest and the "treatments" it wanted the platforms to test. The platform would then be required to implement the treatment for a random sample of users and track the outcome variables of interest while also doing the same for a comparable control group.

Treatments of interest would potentially include:

- Treating some consumers with a non-algorithmic version of their news feed which simply showed information produced by their friends and people they follow on a chronological basis.[105]
- Measuring consumers' awareness or misperception of platform tools. For example: do Twitter users know that they can switch their feed to a chronological version?
- Tracking the impact of algorithm changes (e.g., by treating one group of individuals with historic implementations of an algorithm and others with a new version).
- Varying the level of personalisation of the algorithms.
- Varying the strength of "social signals" by, for example, varying the weight placed on friends' recommendations and activities. This would help determine the extent to which various aspects of consumer behaviour are driven by their social interactions would help determine strength of the network effect (e.g., as per Bakshy, Messing, Adamic (2015)).
- Varying the inclusion or weight given to specific algorithm inputs. For example, platforms could be required to run tests of the impact of only tracking some forms of engagement (e.g., "likes") and not others (e.g., the "negative face emoji") or to introduce new measures accounting for whether a news outlet is the original source for a particular news story.

105 This is equivalent to the internal studies performed by Twitter.

- Varying the algorithm to increase the amount of "non-aligned" content seen by individuals (along the lines of the Levy (2021) study).

Outcomes of interest would then include:

- Patterns of news consumption (e.g., the pattern of consumption across different outlets).
- Levels of user engagement.
- Reported levels of polarisation (perhaps gathered through pre- and post-treatment surveys similar to those in the Allcott et al. (2020) study discussed above).
- Willingness to pay for using platforms (similar to the one in the Allcott et al. (2020) study)

The exact list of treatments and control variables could be refined to reflect policy priorities but introducing a framework for such A/B tests to be performed would provide a powerful tool to better understand the effects at issue.

8.3. Amendments to the existing NCS

The NCS remains a useful tool for getting a representative picture on how individuals consume news and to get a sense of the influence of individual news *outlets*. It can also provide useful background on the scale of intermediated consumption of news and the extent to which consumers continue to consume intermediated news alongside traditional sources. However, as we discussed in Section 2, the NCS was not designed with intermediated news consumption in mind and there is a limit to how much it can be used to answer these questions.

In light of this, a more in-depth survey could be used to measure a range of outcomes (information, affective polarisation, whether the individual regularly reads counter attitudinal material, feels overwhelmed with information etc). This might be useful as a baseline, to understand which platforms lead to harm and to better identify the need for further research into any specific platform (e.g. the causal methodologies used by Allcott et al. (2020) or Levy (2021), as well as algorithm reviews). We propose the following additions.

Asking questions to understand how consumers access individual news outlets. As set out in Figure 3 above, the current NCS asks respondents about the aggregate split of their online news consumption as between direct consumption from news sources vs. social media. However, we consider that the NCS would benefit from asking consumers their pattern of news consumption *for each news outlet individually* (e.g., if a consumer says that they have seen articles from the Times, Huffington Post, Telegraph etc. they would be asked whether these predominantly came through direct or intermediated channels and which intermediaries were used (Facebook, Apple News etc.)).

Adding these additional questions would allow researchers to:

- Compute further correlations as to level of polarisation and how it relates to the patterns of consumption.

- Analyse the extent to which the share of reference (SoR) for news outlets may be disproportionately due to mediated vs. unmediated consumption and hence whether there may be differences in the level of agenda setting power across different news outlets.[106]
- Understand whether consumption of news via traditional channels diminishes the polarisation from the intermediated news (e.g., by comparing similar consumers, one group which uses intermediaries only and another who uses both intermediaries and direct traditional channels)
- Understand the extent to which social media undermines users' ability to attribute sources (which may lead to undermining news producers' branding, including their ability to raise ad revenue) by looking at the proportion of "don't know" responses to questions asking what users read and how they got there.

Asking questions on additional socioeconomic factors. The NCS already collects data on the individuals' gender, age, geography, race, political affiliation and religion. In addition, education level and income would be helpful additions. This would allow one to understand how news consumption, and in particular news consumption through intermediaries, differs across socioeconomic lines, as in Angelucci and Prat (2021).[107] It would also allow one to study the breadth of coverage of different intermediaries (e.g., the extent to which Twitter's reach may be more concentrated on higher-income individuals to Facebook).

Asking questions on the outcome measures of news knowledge and polarisation used in the literature. A more ambitious adjustment would be to ask respondents some of the survey questions used in the work of Allcott et al. (2020). The NCS already includes a question on political leaning, but could be expanded to cover the following topics:[108]

- **Political engagement:** this may be measured by the voter turnout in the last general elections. Allcott et al. (2020) also measure the likelihood of their experiment participants clicking on email links to support political causes, suggesting that higher click rate is indicative of higher political engagement. This metric would help answer the impact that the intermediaries have on the ultimate outcome of democratic decision-making and tying these outcomes to the harms around polarisation.
- **News knowledge**: questions on whether the individual is aware of current affairs and can distinguish between true and false statements.
- **Polarisation factors**: questions of measurement of the participants' polarisation replying on "diverse interactions": whether the individuals interact with someone who voted the opposite way as you in the last general election; interact with someone from another country.

106 As noted in Section 2, there are more general issues around Share of Reference and it needs to be accompanied with a more detailed analysis of individual consumption patterns and multi-homing (i.e., the extent to which different news sources are consumed alone or as part of a broader bundle of consumption).

107 Angelucci and Prat (2021) find that socioeconomic factors matter much more than partisanship in determining information levels – in other words, when individuals are exposed to misinformation, socioeconomic factors are the primary determinant of the extent to which the individual is prone to be polarised by it.

108 See Online Appendix B of Allcott et al. (2020) https://assets.aeaweb.org/asset-server/files/11625.pdf

- **Political polarisation:** party affective polarisation (how warm or cold the individual feels towards the parties and the government); congenial news exposure (how often see the news to understand the point of view of the different parties); issue polarisation (views on specific political topics), belief polarisation (agreement with partisan beliefs), vote polarisation.

While illuminating, we recognise that these questions would add considerable length of the NCS questionnaire and may not be feasible to include on an ongoing basis. If they were to be included, we would suggest doing so on a one-off basis (perhaps asking them to only a subset of respondents) and a preferable alternative in our view would be instead to track these issues using a more bespoke experimental design performed in parallel to the NCS.

8.4. Additional considerations

8.4.1. Building a UK-specific evidence base

As we have identified in our literature review, a key drawback is that the most influential (and most credible) studies are often focussed on US data. The concern would be that, while there are likely to be parallels in the economic mechanisms, there are UK-specific differences that may prevent the results (e.g., that social media use increases polarisation) from perfectly reading across. For instance, while the COVID-19 polarisation bears similarities across the different countries, divisive issues like Brexit are UK-specific. This means that the nature of polarisation differs, and it is not clear whether the UK consumers of news exhibit the same level of polarisation as those in the US. Similar geography-specific issues in the evidence base exist in the context of echo chambers and filter bubbles, financial links between intermediaries and publishers, as well as the intermediaries' impact on news production.

To deal with the gap the literature lacking specific studies on the UK context, we propose to build UK-specific evidence by replicating the two most seminal studies on UK data. We propose replicating one or both of the Allcott et al. (2020) or Levy (2021). These two studies offer a robust and established framework to test TOH1. Levy offers a framework to understand how the intermediaries' algorithms affect the slant and source of the news being consumed by their users, and how this compares to consumers accessing news through the traditional channels or both intermediated and traditional. At the same time, Allcott et al. (2021) framework can answer how absence of intermediated news affects consumer behaviour. Both of these studies are conducted on Facebook. However, the frameworks can be adjusted to other social media platforms, and Levy (2021) can be applied to the search platforms as well.

One possibility would be for Ofcom to (partially) fund a PhD student or other researcher to conduct this analysis, perhaps augmenting the existing studies to take account of additional evidence gaps in the literature.

8.4.2. Tracking ideology/plurality at the content level

Within media or an outlet, content is of course not uniformly slanted to a given ideology. This is especially so with opinion articles and independent journalist publications. Capturing the slant at content level rather than outlet level would better account for this issue and help explain whether intermediaries filter content at article level and not just outlet level. Intermediaries may rely on a diverse set of outlets but potentially serve a non-diverse content. For example, if the intermediaries present only climate sceptical articles from a

range of publishers, the approach above would not be able to identify the lack of diversity in content (and thereby a filter bubble) while concluding diversity in outlets.

A more ambitious approach to measuring the outcomes of media plurality, and the intermediaries' impact, would therefore look at the content of individual articles. This would allow for a closer study of echo chambers/filter bubbles and polarisation but would require a more complex method of measuring the content.

The issues relating to internal validity and measurement, and the possible approaches, are discussed in Section 3 above. In sum, this task is complicated by the fact that there is no single best practice approach indicated by the literature. Any content-level measurement requires manual input or relies on machine-learning approach which is prone to biases and mismeasurement as well.

Rather than an ongoing tracking of content-level news consumption, this could be undertaking as a one-off exercise. By selecting an existing methodology, or designing a new on, a particular topic of interest (e.g., vaccination, climate change) could be studied in order to understand whether filter bubbles exist at a content level even if intermediaries rely on a diverse set of publishers. An alternative and potentially more cost-effective approach would be to use any existing methods by the platforms to tag the content of posts using categories that are relevant to the study of echo chambers/filter bubbles and polarisation.

8.5. Summary of recommendations on evidence gathering

In light of the discussion above we would propose the following high-priority evidence gathering initiatives:

- **New datasets to be gathered from platforms.** This would inform the user behaviour on the platforms. It would include the "link out" data split by news source, consumer characteristics, and type of link, to tell us how users arrive at the news. We would also want a more detailed data showing what users are shown on the platform to understand the algorithm's ranking, friend suggestions, ads, and other stories that the user interacts on the platform but does not necessarily click through to a publisher's page.
- **New data to be gathered from publishers.** The publishers can be a rich source of information on the broader group of consumers, to understand the behaviour of those who access via intermediaries vs. those who access directly. Moreover, we would want to test other hypotheses around the impact on news production, such as ad revenue and recency vs originality of content being consumed by the different groups.
- **Algorithm review and A/B tests.** The rationale for this type of information is to better understand the workings of the algorithm. We would want to understand the input signals and how they affect the various measures of interest, such as polarisation of content. The data would include a description of these algorithms accompanied by telemetry and worked examples data. We would also want the platforms to run different A/B tests to analyse the real-world impact of varying the algorithm's design.

As medium priority items we would propose the following:

- **Additions to the NCS.** While the NCS continues to be valuable tool for measuring media plurality, extensions to better capture the impact of the intermediaries would allow for an ongoing measurement. The survey would also provide a more representative sample of the UK consumers, which the platform and publisher data cannot provide. The additions would include questions on the route to the news publication (e.g., direct v intermediated), political affiliation, news knowledge and their socioeconomic factors.
- **Experiments on UK data.** An important but more ambitious data gathering exercise would involve running experiments mimicking those by the key studies in the literature. This would help understand the drivers of the harm in a UK context, as well as how prone consumers are to variations in the algorithm. Some of these questions may already be answered by the suggested A/B tests above.

As a lower priority we would consider also studying content-level consumption. This is an important issue, telling us the extent to which the harms around polarisation arise at the outlet-level vs the article-level. However, the literature review suggests no best practice and an exercise that would be potentially very ambitious. We would suggest this as a follow up after the “lower-hanging fruit” initiatives above.

9. CONCLUSIONS

This report considered four categories of harm resulting from online intermediation of news consumption: increases in polarisation amongst users; echo chambers and filter bubbles; the spread of misinformation; and algorithmic biases (e.g. as a result of commercial links between online intermediaries and publishers).

We have identified and reviewed more than 70 relevant academic studies. Our key conclusions are:

- Key studies in the literature support a causal link between intermediated consumption via social media and polarisation. The main evidence comes from experimental studies which show how the use of Facebook results in more polarised consumption and attitudes than for a comparable group which uses other channels to access news. While this evidence is compelling, the main limitation is that the key studies are conducted in a US setting and specifically for Facebook.
- The evidence on intermediaries leading to the other harms is more nuanced. There is evidence of incentives for social media platforms to design algorithms that lead to echo chambers and filter bubbles, but the extreme scenario of consumers *only* seeing content that supports their views is rare. Similarly, the few studies that have looked for quantitative evidence on the impact of fake news and misinformation on electoral outcomes have found that the effects may be more muted than the public debate suggests and show also that platforms can take steps to mitigate these issues if incentivised to do so.

- There is less clarity on the extent to which algorithmic bias through commercial links are present in the news intermediation. There is evidence of bias, including some limited evidence showing that platforms promote one form of partisan news over another. However, the studies do not explain whether the bias is financially driven, and it may even be the case that there is a particular motive behind it. Although several publishers have raised this issue, we have found few compelling research studies to support the claims.
- Another key message is that the impact of these concerns is platform specific. Social media (the evidence is broadest on Twitter and Facebook which are also the two most commonly used platforms to access news in the UK) is shown to lead to higher levels of polarisation and the associated issues around echo chambers. In contrast, there is some evidence that online search (i.e., studies analysing Google) leads to a more cross-attitudinal consumption of news and lower polarisation, perhaps in part because search results are less personalised.

The key gaps in the literature relate to the lack of UK specific evidence. The focus of the literature, including the most reliable studies, is the US market, with few causal studies conducted on the UK. There are also other important areas for additional research:

- Lack of access to the workings of platform algorithms means that there is a gap in clarity of how algorithms function. Platforms are understandably protective of the details of the workings of their algorithms, but this makes it difficult to pinpoint the specific algorithmic biases and features leading to the negative outcomes.
- There is a lack of targeted analysis to assess the cross-cutting effect of consumption via non-intermediated channels. Ideally, we would want to know more about how consuming news outside of social media or other intermediaries moderates potential echo chamber effects, especially in the UK context where public broadcasting is shown to be more centric and widely used.
- The majority of the studies rely on measuring of plurality at the outlet level, meaning that the literature does not tell us how online intermediation may affect consumption *within* an outlet (e.g., investigative reporting pieces that may run contrary to the broader editorial slant). There is an open question whether the algorithm promotes article-level or post-level slant even if the content is from a plurality of outlets.[109]
- There is lack of literature looking at intermediaries' incentives to bias traffic between news sources based on financial considerations. Relatedly the literature on aggregate ideological bias in algorithms is also small and nascent.
- Finally, there is also a gap in the more direct link between online intermediation and the broader democratic process. Some studies show that social media can increase news knowledge and voting turnout, but this is still not well understood, especially in the UK context.

To fill these gaps it would be desirable to see four types of initiatives:

109 As noted above, an extremely recent study looking at these issues is D'Amico and Tabellini (2022).

- First, gathering additional data from the major platforms would provide an insight into how thew mediate the traffic they send to news outlets, and would explain how consumers engage with their platform. A corresponding data from the publishers would strengthen this understanding and tell us what effects online intermediation has on the UK news consumption.
- Second, to develop a good understanding of the intermediaries' algorithms and to help more directly assess its role in the potential consumer harms, one could consider gathering descriptive information from the intermediaries on the workings of their algorithms as well as intermediaries providing access to policymakers to telemetry data and specific A/B tests on the platform.
- Thirdly, the NCS could be amended to better track the role of the intermediaries or accompanied by complementary surveys targeted at these issues.
- Fourth, carrying out experimental studies in the UK context, by relying on or replicating the most compelling academic work, would illuminate the possible differences in the US-focussed literature vs the UK environment. While there are parallels in the economic mechanisms studied in the literature, there are UK-specific factors that would better inform the optimal policy design.

APPENDIX A: BACKGROUND ON THE AUTHORS

Oliver Latham. Oliver is a Vice President with Charles River Associates' European Competition Practice based in London. He specialises in antitrust issues relating to two-sided markets and innovative industries including online advertising, social media, and search. He also has broad experience on antitrust and merger cases in the broadcasting and media sector as well as in online advertising.

Within the field of media plurality specifically, he holds a PhD in Economics from the University of Cambridge which focussed on measuring bias in information transmission, including in media coverage by UK newspapers. One chapter of this thesis, "Lame Ducks and the Media" was published in the Economic Journal. He has also written policy papers analysing the BBC's citations of think tanks of different political persuasions for the Centre for Policy Studies and advised Fox on competition and media plurality issues arising from its proposed acquisition of Sky.

He has also advised multiple news publishers worldwide, including News Corp and Seven West, on issues around digital advertising and funding of news production. He has also worked with complainants in respect of antitrust investigations into Google's conduct in specialised search, online advertising, and the Android operating system and has also worked with other tech platforms and online advertising businesses including Amazon, Uber, ByteDance, Kelkoo, Outbrain, Taboola, Microsoft, and Yelp where his work has included analysing the impact of algorithm design on consumer outcomes.

Gaber Burnik. Gaber is an Associate Principal with Charles River Associates. He advises clients on a range of regulatory issues, including in the context of mergers requiring economic analysis of two-sided markets and digital platforms.

He also has extensive policy experience from his prior role at the Financial Conduct Authority, where he spent four years as an Economist and a Lead Specialist in regulatory strategy. This experience includes analysis of the platforms' role in the distribution of retail financial products, including algorithmic bias, and the application of financial intermediary standards to these platforms. He has previously worked with Ofcom during his secondment to the CMA on development of a fair pricing framework to tackle the "loyalty penalty".

Gaber holds an MPhil in Economics from University of Oxford.

Rafael Jiménez Durán. Rafael is an academic economist specialising in applied microeconomics in the context of digital platforms and social media. His Job Market Paper "The Economics of Content Moderation" includes a theoretical discussion and an experimental design of two field trials on Twitter to examine the effect of moderating hate speech on user behaviour and welfare. His other work includes the study of competition and content moderation and on the measurement of hate speech on social media. Rafael's publications are forthcoming in the Journal of Monetary Economics and the Journal of Financial Market Infrastructures. His work on social media and hate speech has been presented in front of several regulators.

Raphael is completing his PhD at the University of Chicago. He will join the Social Science Research Council as a Post-Doctoral Fellow in 2022 and Bocconi University as an Assistant Professor in 2023.

Hasnain Khaki. Hasnain is an Associate in Charles River Associates' European Competition Practice. Prior to joining CRA, he completed a BSc in Economics at The London School of Economics and Political Science.

APPENDIX B: BIBLIOGRAPHY

Acemoglu, Ozdaglar and Siderius. 2022. "A model of online misinformation", working paper, *National Bureau of Economic Research*.

Allcott, Braghieri, Eichmeyer and Gentzkow. 2020. "The Welfare Effects of Social Media", *American Economic Review.*

Allcott, Gentzkow and Yu. 2019. "Trends in the diffusion of misinformation on social media", *Research and Politics.*

Allcott, Gentzkow. 2017. "Social Media and Fake News in the 2016 Election", *Journal of Economic Perspectives.*

Allen, Howland, Mobius, Rothschild, Watts. 2020. "Evaluating the fake news problem at the scale of the information ecosystem", *Science Advances.*

Anderson, S. P. and O. Bedre-Defolie ¨ (2021): "Hybrid Platform Model," Working paper.

Angelucci and Prat. 2021. "Is Journalistic Truth Dead? Measuring How Informed Voters Are About Political News".

Arguedas, Robertson, Fletcher and Nielsen. 2022. "Echo Chambers, Filter Bubbles, and Polarisation: a Literature Review", *Reuters Institute for the Study of Journalism.*

Australian Competition & Consumer Commission. 2019. "Digital Platforms Inquiry Final Report" Available at: https://apo.org.au/sites/default/files/resource-files/2019-07/apo-nid250026_1.pdf.

Azzimonti and Fernandes. 2021. "Social Media Networks, Fake News, and Polarization", working paper. *National Bureau of Economic Research*.

Bail et al. 2018. "Exposure to opposing views on social media can increase political polarization", *Proceedings of the National Academy of Sciences.*

Bakshy, Messing, Adamic. 2015. "Political science. Exposure to ideologically diverse news and opinion on Facebook", *Science.*

Bandy and Diakopoulos. 2020. "Auditing News Curation Systems: A Case Study Examining Algorithmic and Editorial Logic in Apple News", *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media (ICWSM 2020).*

Besley, Prat. 2006. "Handcuffs for the Grabbing Hand", *American Economic Review.*

Boxell, Gentzkow, Shapiro. 2017. "Is the Internet Causing Political Polarization? Evidence from Demographics", working paper, *National Bureau of Economic Research*.

Bryanov, Watson, Pingree, Santia. 2020."Effects of Partisan Personalization in a News Portal Experiment", *Public Opinion Quarterly.*

Bourreau, Gaudin. 2020. "Streaming platform and strategic recommendation bias", *Journal of Economics and Management Strategy*.

Budak, Goel, Rao. 2016. "Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis", *Public Opinion Quarterly.*

Carlton, Waldman. 2002. “The strategic use of tying to preserve and create market power”, *RAND Journal of Economics*

The Cairncross review: A sustainable future for journalism. 2019.

Cardenal, Aguilar-Paredes, Cristancho, Majó-Vázquez. 2019. “Echo-chambers in online news consumption: Evidence from survey and navigation data in Spain”, *European Journal of Communication*.

Chen, Lerman, Ferrara. 2020. “Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set”, *JMIR Public Health and Surveillance.*

Chiang and Knight. 2011. “Media bias and influence: evidence from newspaper endorsements”, *Review of Economic Studies.*

Cinelli et al. 2021. “The echo chamber effect on social media”, *Proceedings of the National Academy of Sciences*.

Clayton et al. 2020. “Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media”, *Political Behavior.*

de Cornière, Greg Taylor. 2014. “Integration and search engine bias”, *The RAND Journal of Economics.*

Cota, Ferreira, Pator-Satorras, Starnini. 2019. “Quantifying echo chamber effects in information spreading over political communication networks”, *EPJ Data Science*.

Council of the European Union. 2020. Available at: https://data.consilium.europa.eu/doc/document/ST-12735-2020-INIT/en/pdf.

D’Amico, Tabellini. 2022. “Online Political Debates”, *CESifo working paper*.

Dalal, Adlim, Lesk. 2019. “How to measure relative bias in media coverage?”, *Significance.*

Dubois, Blank. 2018. “The echo chamber is overstated: the moderating effect of political interest and diverse media”, *Information, Communication & Society.*

Dunaway, Searles, Sui, Paul. 2018. “News Attention in a Mobile Era”, *Journal of Computer-Mediated Communication*.

Dutton, Reisdorf, Blank. 2017. “Social Shaping of the Politics of Internet Search and Networking: Moving Beyond Filter Bubbles, Echo Chambers, and Fake News”, *SSRN Electronic Journal.*

Etro. 2021a. “Device-funded vs ad-funded platforms”, *International Journal of Industrial Organization*.

Etro. 2021b. “Product selection in online marketplaces”, *Journal of Economics and Management Strategy*.

Ferrara, Jiang, Ren. 2021. “Social Media Polarization and Echo Chambers in the Context of COVID-19: Case Study”, *JMIRx Med.*

Flaxman, Goel, Rao. 2016. “Filter Bubbles, Echo Chambers, and Online News Consumption”, *Public Opinion Quarterly.*

Fletcher, Kalogeropoulos, Nielsen. 2021. “More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the United Kingdom”, *New Media & Society.*

Fletcher, Nielsen. 2018. “Automated Serendipity”, *Digital Journalism.*

Fletcher, Robertson, Nielsen. 2021. “How Many People Live in Politically Partisan Online News Echo Chambers in Different Countries?”, *Journal of Quantitative Description: Digital Media.*

Frenkel and Kang. 2021. “An Ugly Truth”, *Bridge Street Press.*

Gentzkow. 2006. “Television and Voter turnout”, *Quarterly Journal of Economics.*

Gentzkow, Shapiro. 2006. “Media Bias and Reputation”, *Journal of Political Economy.*

Gentzkow, Shapiro. 2010. “What Drives Media Slant? Evidence From U.S. Daily Newspapers”, *Econometrica.*

Gentzkow, Shapiro. 2011. “Ideological Segregation Online and Offline”, *Quarterly Journal of Economics.*

Gentzkow. Shapiro and Sinkinson. 2011. “The Effect of Newspaper Entry and Exit on Electoral Politics”*, American Economic Review.*

Gerber, Karlan, Bergan. 2009. “Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions”, *American Economic Journal.*

Gilardi, Gessler, Kubli and Müller. 2021. “Social Media and Political Agenda Setting”, *Political Communication.*

Grinberg, et al. 2019. “Fake news on Twitter during the 2016 U.S. presidential election”, *Science*.

Guess, Nyhan, and Reifler. 2020. “Exposure to untrustworthy websites in the 2016 U.S. election”, *Nature Human Behaviour.*

Hagiu, A., T.-H. Teh, and J. Wright (Forthcoming): “Should Platforms be Allowed to Sell on Their Own Marketplaces?” *RAND Journal of Economics.*

Haim, Graefe, and Brosius. 2018. “Burst of the Filter Bubble?: Effects of personalization on the diversity of Google News”, *Digital Journalism.*

Halberstam, Knight. 2016. “Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter”, *Journal of Public Economics.*

Helberger, Kleinen-von Königslöw and Van Der Noll. 2015. “Regulating the new information intermediaries as gatekeepers of information diversity”, *Info.*

Huszár et al. 2021. “Algorithmic Amplification of Politics on Twitter”, *Proceedings of the National Academy of Sciences.*

Jimenez-Duran. 2021. “The Economics of Content Moderation: Theory and experimental evidence from hate speech on Twitter”, working paper.

Jürgens, Stark. 2017. “The Power of Default on Reddit: A General Model to Measure the Influence of Information Intermediaries”, *Policy and Internet.*

Kitchens, Johnson and Gray. 2020. “Understanding echo chambers and filter bubbles: the impact of social media on diversification and partisan shifts in news consumption”, *MIS Quarterly*.

Kotkov, Wang, Veijalainen. 2016. “A Survey of Serendipity in Recommender Systems”, *Knowledge-based Systems.*

Langer, Gruber. 2021. “Political Agenda Setting in the Hybrid Media System: Why Legacy Media Still Matter a Great Deal”, *The International Journal of Press/Politics.*

Le, Shafiq, Srinivasan. 2017. “Scalable News Slant Measurement Using Twitter”, *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017).*

Leigh, Gans. 2011. “How partisan is the press? Multiple measures of media slant”, working *paper, IZA Discussion Papers, Institute for the Study of Labor (IZA).*

Leveson Inquiry. 2012. “An inquiry into the culture, practices and ethics of the press”.

Levy. 2021. “Social Media, News Consumption, and Polarization: Evidence from a Field Experiment”, *American Economic Review*.

Marks et al. 2018. “Epistemic spillovers: Learning others’ political views reduces the ability to assess and use their expertise in nonpolitical domains”, *Cognition.*

Martin and Yurukoglu, A. 2017. “Bias in Cable News: Persuasion and Polarization”, *American Economic Review*

Mosquera et al. 2019. “The Economic Effects of Facebook”, *Experimental Economics.*

Mullainathan, Shleifer. 2005. “The Market for News”, *The American Economic Review.*

Nechushtai and Lewis. 2019. “What kind of news gatekeepers do we want machines to be?”, *Computers in Human Behaviour.*

Noam. 2016. “*Who Owns the World's Media?”*

Ofcom. 2021. “The future of media plurality in the UK”. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0019/228124/statement-future-of-media-plurality.pdf.

Pariser. 2011. “The Filter Bubble: What The Internet Is Hiding From You”.

Pennycook, Bear, Collins, Rand. 2020. “The Implied Truth Effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings”, *Management Science.*

Polo. 2005. “Regulation for pluralism in media markets”, published in 2009 in the book *The Economic Regulation of Broadcasting Markets.*

Prat. 2014. “How can we measure media power”. Available at: https://voxeu.org/article/how-can-we-measure-media-power.

Prat. 2018. “Media Power”, *Journal of Political Economy.*

Prat. 2020. “Measuring and Protecting Media Plurality in the digital Age: A Political Economy Approach”. Available at: https://knightcolumbia.org/content/measuring-and-protecting-media-plurality-in-the-digital-age.

Prior. 2009. “The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure”, *Public Opinion Quarterly.*

Schmidt et al. 2017. “Anatomy of news consumption on Facebook”, *Proceedings of the National Academy of Sciences*.

Singer et al. 2011. “Participatory Journalism. Guarding Open Gates at Online Newspapers”.

Snyder and Stromberg. 2010. “Press coverage and political accountability”, *Journal of Political Economy.*

Spohr. 2017. “Fake news and ideological polarization: Filter bubbles and selective exposure on social media”, *Business Information Review.*

Stier, Mangold, Scharkow, Breuer. 2021. “Post Post-Broadcast Democracy? News Exposure in the Age of Online Intermediaries”, *American Political Science Review.*

Stromberg. “Media Coverage and Political Accountability: Theory and Evidence” in Anderson, Stromberg, and Waldfogel (eds).2015. *“Handbook of Media and Economics”*.

Sunstein. 2017. “#Republic: Divided Democracy in the Age of Social Media”.

Tella, Galvez, Schargrodsky. 2021. “Does Social Media cause Polarization? Evidence from access to Twitter Echo Chambers during the 2019 Argentine Presidential Debate”, *National Bureau of Economic Research.*

Terren, Borge. 2021. “Echo Chambers on Social Media: A Systematic Review of the Literature”, *Review of Communication Research.*

The Cairncross review: A sustainable future for journalism. 2019.

Törnberg, 2018, “Echo chambers and viral misinformation: Modeling fake news as complex contagion”, *PLOS One.*

Waszak, Kasprzycka-Waszak, Kubanek. 2018. “The spread of medical fake news in social media – The pilot quantitative study”, *Health Policy and Technology.*

Westerman, Spence, Van Der Heide. 2014. “Social Media as Information Source: Recency of Updates and Credibility of Information.” *Journal of Computer-Mediated Communication.*

Wilding, Fray, Molitorisz, McKewon. 2018. “The Impact of Digital Platforms on News and Journalistic Content”.

Wojcieszak, Leeuw, Menchen-Trevino, Lee, Huang-Isherwood, Weeks. 2021a. “No Polarization from Partisan News: Over-Time Evidence from Trace Data”, *The International Journal of Press/Politics.*

Wojcieszak, Menchen-Trevino, Goncalves, Weeks. 2021b. “Avenues to News and Diverse News Exposure Online: Comparing Direct Navigation, Social Media, News Aggregators, Search Queries, and Article Hyperlinks”, *The International Journal of Press/Politics.*

Yang Kang, Muir (2022). “Contracting and vertical control by a dominant platform”. Working paper.

Zuiderveen Borgesius, et al. 2016. “Should we worry about filter bubbles?”, *Internet Policy Review.*

APPENDIX C: CLASSIFICATION OF STUDIES IN THE LITERATURE

Table 2: Classification of papers considering the main theories of harm

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
ACCC. 2019.	Australia	Facebook, Twitter, Google News	Survey data, literature review	Observational, literature review	Outlet	Negative impact on news production	n/a	Supports the existence of filter bubbles
Acemoglu, Ozdaglar and Siderius. 2022.	n/a	Generic	n/a	Theoretical	n/a	Polarisation/Echo chambers & filter bubbles	n/a	Supports the existence of filter bubbles
Allcott, Braghieri, Eichmeyer and Gentzkow. 2020.	U.S.	Facebook	Survey data	Experiment	n/a	Polarisation	Increases	Inconclusive
Allcott, Gentzkow and Yu. 2019.	US	Facebook, Twitter	Web browsing data	Observational	n/a	Polarisation	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Allcott, Gentzkow. 2017.	US	Facebook, Twitter	Survey data, Web browsing data	Observational	n/a	Polarisation	n/a	n/a
Allen et al. 2020.	U.S.	Traditional media, generic	Web browsing data, others	Observational	n/a	Polarisation	n/a	n/a
Angelucci and Prat. 2021.	US	Traditional media	Survey data	Observational	n/a	Polarisation	n/a	n/a
Arguedas, Robertson, Fletcher, Nielsen. 2022.	UK	Facebook, Twitter, Google News	Survey data, Literature Review	Literature review	Outlet	Polarisation/Echo chambers & filter bubbles	No impact	Doesn't support the existence of filter bubbles
Azzimonti and Fernandes. 2021.	n/a	Twitter	n/a	Theoretical	n/a	Polarisation	n/a	n/a
Bail et al. 2018.	U.S.	Twitter	Survey data	Experiment	n/a	Polarisation	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Bakshy et al. 2015.	US	Facebook	Post-level data	Observational	Individual	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Bandy and Diakopoulos. 2020.	US	Apple News	Web browsing data	Observational	Outlet	Platforms' financial links to publishers	Increases	Supports the existence of filter bubbles
Boxell, Gentzkow, Shapiro. 2017.	U.S.	Generic	Survey data	Observational	n/a	Polarisation	Increases	Inconclusive
Bryanov, Watson, Pingree, Santia. 2020.	US	Generic	Survey data, Web browsing data	Experiment	Outlet	Polarisation	Increases	Inconclusive
Budak, Goel, Rao. 2016.	U.S.	Traditional media	Web browsing data, Post-Level data	Observational	Post-level	Polarisation	n/a	n/a
Cardenal, Aguilar-Paredes, Cristancho, Majó-Vázquez. 2019.	Spain	Traditional media	Survey data, web browsing data	Observational, experiment	Outlet	Polarisation/Echo chambers & filter bubbles	Inconclusive	Doesn't support the existence of filter bubbles

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Cinelli et al. 2021.	Global	Facebook, Twitter, Reddit, Gab	Post-level data	Observational	Outlet	Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Clayton et al. 2020.	US	Facebook	Survey data	Observational	Outlet	Polarisation	n/a	n/a
Cota, Ferreira, Pator-Satorras, Starnini. 2019.	Brazil	Twitter	Post-level data	Observational	Post-level	Echo chambers & filter bubbles	n/a	Supports the existence of filter bubbles
Council of the European Union. 2020.	n/a	Generic	n/a	Literature review	n/a	Polarisation	n/a	n/a
Dalal, Adlim, Lesk. 2019.	U.S.	Traditional media	Post-level data	Observational	Post-level	n/a	n/a	n/a
Dubois, Blank. 2018.	UK	Generic	Survey data	Observational	n/a	Echo chambers & filter bubbles	n/a	Doesn't support the existence of filter bubbles

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Dutton, Reisdorf, Dubois, Blank. 2017.	UK, France, Germany, Italy, Poland, Spain, US	Generic, traditional media	Survey data, web browsing data	Observational, experiment	n/a	Polarisation/Echo chambers & filter bubbles	Lessens	Doesn't support the existence of filter bubbles
Ferrara, Jiang, Ren. 2021.	US	Twitter	Post-level data	Observational	Post-level/outlet	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Flaxman, Goel, Rao. 2016.	U.S.	Traditional media, Facebook, Twitter, Google News, Generic	Web browsing data	Observational	Outlet	Polarisation/Echo chambers & filter bubbles	Increases	Doesn't support the existence of filter bubbles
Fletcher, Kalogeropoulos, Nielsen. 2021.	UK	Traditional media, Facebook, Twitter, Google News, Generic	Web browsing data	Observational	Outlet	Echo chambers & filter bubbles	Inconclusive	Doesn't support the existence of filter bubbles
Fletcher, Nielsen. 2018.	UK, USA, Germany, Spain	Generic	Survey data	Observational	Individual	Polarisation/Echo chambers & filter bubbles	Lessens	Doesn't support the existence of filter bubbles

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Fletcher, Robertson, Nielsen. 2021.	Austria, Denmark, Germany, Norway, Spain, the UK, and the US	Generic	Survey data	Observational	Outlet	Echo chambers & filter bubbles	No impact	Supports the existence of filter bubbles - Finds echo chambers are real but only a minority inhabit them
Gentzkow and Shapiro. 2011.	US	Traditional media	Web browsing data, Survey data	Observational	Individual/Outlet	Polarisation	Increases	Supports the existence of filter bubbles
Gentzkow, Shapiro. 2010.	U.S.	Traditional media	Post-level data	Observational	Individual/Outlet	Polarisation/Echo chambers & filter bubbles	n/a	n/a
Gerber, Karlan, Bergan. 2009.	US, Virginia	Traditional media	Survey data	Experiment	Outlet	Polarisation	n/a	n/a
Gilardi, Gessler, Kubli and Müller. 2021.	Switzerland	Traditional media, Twitter	Post-level data	Observational	News article-level	Platforms' financial links to publishers/Negative impact on news production	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Grinberg et al. 2019.	US	Twitter	Post-level data	Observational	n/a	Polarisation	n/a	n/a
Guess, Nyhan, and Reifler. 2020.	US	Facebook	Survey data, Web browsing data	Observational	n/a	Polarisation	n/a	n/a
Haim, Graefe, and Brosius. 2018.	Germany	Google News	Post-level data	Observational	Outlet	Echo chambers & filter bubbles	Lessens	Doesn't support the existence of filter bubbles
Halberstam and Knight. 2016.	US	Twitter	Web browsing data, Post-Level data	Observational	Individual	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Helberger, Kleinen-von Königslöw and Van Der Noll. 2015.	n/a	Generic	Literature review	Literature review	n/a	Platforms' financial links to publishers, Negative impact on news production	n/a	n/a
Huszár et al. 2021.	First analysis: US, Japan, UK, France, Spain, Canada, Germany.	Twitter	Post-level data	Observational, experiment	Individual/Outlet	Polarisation	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
	Second analysis: US							
Jimenez-Duran. 2021.	Experiment 1: not specified Experiment 2: survey of U.S. twitter users	Twitter	Post-level data, survey data	Experiment	n/a	Polarisation	n/a	n/a
Jürgens, Stark. 2017.	Likely global	Reddit	Post-level data	Observational, experiment	n/a	Polarisation	n/a	n/a
Kitchens, Johnson and Gray. 2020.	US	Facebook, Twitter, Reddit	Web browsing data	Observational	Outlet	Polarisation/Echo chambers & filter bubbles	Increases	Doesn't support the existence of filter bubbles
Kotkov et al. 2016.	n/a	Generic	n/a	Theoretical	n/a	n/a	n/a	Supports the existence of filter bubbles
Langer, Gruber. 2021.	UK	Traditional media, Twitter	Survey data, post-level data, web browsing data	Observational	n/a	Negative impact on news production	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Le, Shafiq, Swinivasan. 2017.	Not specified	Twitter	Post-level data	Observational	Individual	Polarisation	n/a	n/a
Leigh, Gans. 2011.	Australia	Traditional media	Post-level data	Observational	Individual/Outlet/News article-level	Polarisation	n/a	n/a
Levy. 2021.	US	Facebook	Survey data, post-level data, web browsing data	Experiment	Outlet	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Marks et al. 2018.	US	Generic	Survey data	Experiment	n/a	Polarisation	n/a	n/a
Mosquera et al. 2019.	US	Facebook	Survey data	Experiment	n/a	Polarisation	n/a	n/a
Nechushtai and Lewis. 2019.	US	Google News	Post-level data	Experiment	Outlet	Platforms' financial links to publishers/Negative impact on news production	Lessens	Doesn't support the existence of filter bubbles
Pennycook, Bear, Collins, Rand. 2020.	US	Facebook	Survey data	Experiment	n/a	Polarisation	n/a	n/a

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Prat. 2014.	US	Traditional media	Survey data	Observational	n/a	n/a	n/a	n/a
Prat. 2018.	US	Traditional media	Survey data	Observational	n/a	n/a	n/a	n/a
Prat. 2020.	36 countries	Traditional media, Generic, Facebook	Survey data	Observational	n/a	n/a	n/a	n/a
Prior. 2009.	US	Traditional media	Survey data, other (Nielsen ratings)	Observational	n/a	n/a	n/a	n/a
Schmidt et al. 2017.	Global	Facebook	Post-level data	Observational	n/a	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Spohr. 2017.	n/a	Facebook	n/a	Theoretical	n/a	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Stier et al. 2021.	France, Germany, Italy, Spain, UK, US	Generic	Survey data, Web browsing data	Observational	Outlet	Polarisation	Increases	Supports the existence of filter bubbles

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Sunstein. 2017.	US context, draws on multiple geographies	Facebook	n/a	Literature review	Outlet	Polarisation/Echo chambers & filter bubbles	Increases	Supports the existence of filter bubbles
Tella, Galvez, Schargrodsky. 2021.	Argentina	Twitter	Survey data, Post-level data	Experiment, observational	Outlet	Polarisation/Echo chambers & filter bubbles	Inconclusive	Supports the existence of filter bubbles
Terren. Borge. 2021.	n/a	Generic	Literature review	Literature review	n/a	Polarisation/Echo chambers & filter bubbles	Inconclusive	Inconclusive
Törnberg. 2018.	n/a	Generic	n/a	Theoretical	n/a	Polarisation/Echo chambers & filter bubbles	n/a	Supports the existence of filter bubbles
Waszak, Kasprzycka-Waszak, Kubanek. 2018.	Poland	Facebook	Web browsing data	Observational	n/a	Polarisation	n/a	n/a
Wilding, Fray, Molitorisz, McKewon. 2018.	Australian context, draws on evidence from other geographies	Generic	Literature review	Literature review	n/a	Polarisation/Echo chambers & filter bubbles/Negative impact on news production	Inconclusive	Inconclusive

Study	Geography	Platform studied	Data type	Empirical design	Measurement of ideology	Relevant theory of harm	Implications of social media on polarisation	Evidence on the existence of filter bubbles
Wojcieszak et al. 2021a.	US	Traditional media	Survey data, Web browsing data	Observational	Outlet	Polarisation	n/a	n/a
Wojcieszak, et al. 2021b.	US	Facebook	Web browsing data	Observational	Outlet	Polarisation	Increases	Doesn't support the existence of filter bubbles
Zuiderveen Borgesius et al. 2016.	n/a	Generic	n/a	Literature review	n/a	Echo chambers & filter bubbles	Inconclusive	Doesn't support the existence of filter bubbles

Source: CRA analysis.

APPENDIX D: TECHNIQUES FOR CLASSIFYING THE IDEOLOGICAL CONTENT OF NEWS ARTICLES

Gentzkow, Shapiro (2010) develops an automated method to index media slant. They use text of debates in the US Congress to identify words and phrases that are correlated with a speaker being a Democrat and Republican and then rank the ideology of newspapers according to whether they use Democrat or Republican-leaning phrases.

Le, Shafiq, Swinivasan (2017) propose a method to measure slant of individual news articles by observing their sharing patterns on Twitter. The method monitors tweets about a news article and a set of landmark Democrat and Republican accounts on Twitter to estimate its slant. The results show that the slant of news articles estimated by the method matches the slant estimated by crowd sourced workers. At a high level, the idea is that if a news article is tweeted/retweeted by more Democrats than Republicans, it is likely to have a Democratic slant.

Leigh, Gans (2011) use 2 outlet measures and 1 individual measure. First, the authors use parliamentary mentions to code over 100 public intellectuals on a left-right scale and estimate slant by using the number of mentions that each public intellectual receives in each media outlet. Second, they use independent "raters" who separately code front-page election stories and headlines. Third, they tabulate the number of electoral endorsements that newspapers give to each side of politics in federal elections. The second measure of media slant relies on content analysis. After removing all identifying information (e.g., headline, newspaper name), they asked a team of "coders," to rate – on a left/right scale – all front-page newspaper articles on political topics that appeared during the month before the 2004 Australian election. Combining these ratings provides a proxy for the media slant of major journalists at these newspapers. They then estimate a measure of the media slant of editors. For this purpose, they asked the same team of coders to give a left/right rating to all front-page political stories in the 2004 election campaign. They also coded all newspaper editorials over this period and counted the number of endorsements that each newspaper gave to each political party.

Dalal, Adlim, Lesk (2019) propose a measure that pairs a news outlet article with the origination of the story on Reuters (which is more factual than the news outlets typically are). In their study, they use New York Times articles paired with Reuters by finding articles that are "closely matched to each other in terms of content, we removed stop words (such as "the", "a", "an", "in") and, after stemming (a process of reducing words down to their roots, so that, for example, "stemming" and "stemmed" become "stem"), we used a contiguous sequence of n words, for n = 1, 2 and 3, and then calculated and applied a statistic known as "term frequency–inverse document frequency" (TF/IDF)." They match 1,058 pairs of unique New York Times and Reuters articles and report the sentiment in New York Times by "adding up the words with a positive sentiment (e.g., "happy", "satisfactory") or a negative sentiment (e.g., "sad", "failure") and taking their net difference" and averaging all the sentences in the article. This enables the authors to analyse the stories across different dimension (e.g., politics content, business, international, Trump/non-Trump content, etc.).