Marketing Science Institute Working Paper Series 2024

Report No. 24-104

Attention Spillovers from News to Ads: Evidence from an Eye-Tracking Experiment

Andrey Simonov, Tommaso Valletti and Andre Veiga

"Attention Spillovers from News to Ads: Evidence from an Eye-Tracking Experiment" © 2024

Andrey Simonov, Tommaso Valletti and Andre Veiga

MSI Working Papers are Distributed for the benefit of MSI corporate and academic members and the general public. Reports are not to be reproduced or published in any form or by any means, electronic or mechanical, without written permission.

Attention Spillovers from News to Ads: Evidence from an Eye-Tracking Experiment*

Andrey Simonov† Tommaso Valletti‡ Andre Veiga§

December 18, 2023

Abstract

We investigate the impact of online news content on the effectiveness of display advertising. In a randomized online experiment, participants read news articles randomly paired with brand advertisements. Leveraging non-intrusive eye-tracking technology, we measure individual attention to both articles and ads. We then measure ad recall, and participants make choices between cash and brand-specific vouchers. We show that heightened attention to articles results in "spillover" attention to ads on the same page which, in turn, increases both brand recall and purchase probability. We also consider the effect of news content type, differentiating between "hard" and "soft" news. We find that advertising next to hard news is at least as effective as advertising next to soft news. This provides evidence against the blunt implementation of "block lists" for sensitive news topics by advertisers. We discuss the implications of attention spillovers for firms contemplating investments in engaging news content within the digital advertising landscape.

JEL Classification Codes: M37, C91, L86

Keywords: Online Advertising, Online News, Experiments, Attention, E-commerce

*We thank Daniel Ershov, Garrett Johnson, Stephan Seiler, Anna Tuchman, Gokhan Yildirim, seminar participants at the Bank of Italy, Imperial College London, CESifo Conference on the Economics of Digitization, Fox Temple Conference on AI, ML, and Business Analytics, University of Milan, University of Bologna, Cergy Paris, and Yale-TSE Regulation of Digital Economy Workshop for helpful comments. Rosa Sanchis-Guerrer and Prashant Garg provided excellent research assistance. We thank the whole team of Lumen Research for giving us access to their eye-tracking technology. We gratefully acknowledge financial help from Imperial College COVID-19 Response Fund and from the NET Institute.

†Columbia University and CEPR, as5443@gsb.columbia.edu

‡Imperial College London, NHH and CEPR, t.valletti@imperial.ac.uk

§Imperial College London, a.veiga@imperial.ac.uk

1 Introduction

Spring 2020, the beginning of the COVID pandemic in the U.S., was characterized by an unusual dynamic for digital advertising. While visits to online media sites and news consumption increased by almost 50% (ComScore, 2020), digital advertising – a rapidly growing area of spending for companies over the last two decades (Statista, 2023) – experienced a 25-35% decline (emarketer.com, 2020). Part of this dip is explained by the overall uncertainty of companies due to the pandemic, but it was also largely driven by "block lists": advertising companies actively avoiding placing ads on pages with pandemic-related news content (digiday.com, 2020). Such avoidance of "hard news," i.e., news that is thought to be potentially sensitive and upsetting to some readers, is driven by the perception on the part of advertisers that placing ads with hard content could lead to negative associations with their brand, hurting the brand's image and dissuading readers from purchasing the advertised brand. In turn, this practice discourages news publishers from investing into "hard news" stories, leading to a potential under-provision of content that could have high societal benefits (e.g. The Guardian, 2020; IAB UK, 2020).

We examine the effect of news content on the effectiveness of advertising via an online experiment that uses non-intrusive eye-tracking technology. By "ad effectiveness" we mean the incremental impact of an additional second of attention to ads on the probability of brand recall and purchase. In our study, subjects were exposed to a random sequence of online articles from well-known news outlets. Individuals were also shown realistic ads for well-known brands. Importantly, the pairing of ads and articles was also randomized. Articles were selected to cover different news topics (e.g., "hard" and "soft" news). Eye-tracking allows us to directly measure the attention paid by individuals to articles and to the ads placed next to them. After reading the articles, individuals were asked to recall the advertised brands and make purchase decisions (choose between a voucher for each advertised brand and cash). The purchase decision was incentivized: individuals received the outcome of one of their choices, selected at random. Since articles vary in how interesting they are for each user, random matching of articles to ads created experimental variation in the attention paid to each ad; we use this variation to determine the impact of ad attention on brand recall and purchases.

We consider an empirical model of attention where individual attention devoted to articles can "spillover" to ads, and vice versa. We allow these spillovers to be positive or negative. For instance, if a reader's eyes randomly move between articles and ads, more time spent on the article increases exposure to the ad (a positive spillover). Alternatively, if individuals are focused on an interesting article, more time might imply less attention devoted to the ad (a negative spillover). The attention consumers devote to ads ultimately can impact ad recall and purchase probabilities of the advertised brand.

Our estimates show that the attention readers devote to articles has a positive spillover effect on the attention to ads displayed on the page. Moreover, this incremental attention to ads increases ad recall and purchase probability (i.e., the probability of choosing a brand-specific voucher over a cash reward). Thus, more captivating news content – one that attracts more attention from readers – increases recall and purchase probabilities of brands whose ads are shown on the same page.

Based on our preferred specification (OLS, using the entire sample), one additional second of attention to a brand's ad results in a 3.4 percentage points higher probability of recall and 0.7 percentage points higher probability of choosing that brand's gift card over cash. The latter estimate is confirmed by an IV specification where we only use the incremental attention to ads generated by spillovers from the attention to news content.

We further show that at least some "hard news" content – articles about the COVID-19 pandemic or the Black Lives Matter (BLM) movement in the summer of 2020 that we use in the study – does not detectably impact ad effectiveness. We find that readers spend less time on articles covering "hard news" – and because of this, devote less attention to ads shown next to hard news articles. However, ad effectiveness (the effect of incremental attention to ads on recall and purchases) is not statistically different for articles with "hard" versus "soft" news. If anything, ad effectiveness is 18-43% higher (albeit not significantly different) when article content is "hard news," which is confirmed throughout all OLS and IV specifications. On balance, this higher ad effectiveness compensates for the lower amounts of attention that readers devote to ads next to hard news articles. In sum, we find no evidence that advertising next to hard news is less effective than advertising next to soft news.

Our results have important implications for both news producers and advertisers. Regarding news producers, we show that the key dimension to be optimized is how captivating news content is, whereas the exact content of articles is less important. Similarly, on the advertisers' side, we show that a key metric to keep in mind when allocating display advertising is the overall engagement of users with the webpage, not necessarily the specific content on the page. As a result, our results suggest one should revisit the practice of blunt "block lists" of hard articles, providing an opportunity for optimizing ad allocation decisions for advertisers and marketing managers.

Apart from the substantive results, we provide a novel empirical strategy to measure advertising effectiveness using non-intrusive eye-tracking tools that have recently become more widely available. These tools allow us to run eye-tracking studies through a standard laptop or smartphone web camera, greatly reducing the costs of eye-tracking studies that are typically done in lab settings. This approach allows for the study of how users engage with online content in a realistic way.

2 Related Literature

This paper contributes to the vast literature that studies the effectiveness of online advertising. Relative to that literature, we make three key contributions.

Our first contribution is to show how more captivating news content creates attention spillovers towards ads and increases ad effectiveness. Two sets of papers are closest to ours. First, this paper builds on the sub-stream of the literature that has examined how the time spent on a webpage with an ad affects the memory and ad recall of users (e.g. Danaher and Mullarkey, 2003; Goldstein et al., 2011, 2015; Uhl et al., 2020).[1] Compared to these, we use eye-tracking to explicitly show the spillover from attention to webpage content towards the ads presented. Separately measuring the respondents' eye-sight dwell on article text and on ads allows us to rule out reverse causality as an alternative explanation (Becker and Murphy, 1993; Tuchman et al., 2018). We are also able to link the incremental attention users devote to ads to user willingness to pay for brands, going beyond the more upstream metric of ad recall.

Our work is also related to the eye-tracking literature that examines advertising effectiveness. However, to the best of our knowledge, there is no evidence of the effect of news content on advertising effectiveness. A sub-stream of this literature leverages eye-tracking to study the psychological mechanisms behind advertising effectiveness (e.g. Wedel and Pieters, 2000; Wedel et al., 2008; Aribarg et al., 2010; Higgins et al., 2014). Another sub-stream studies how different features and designs of advertisements increase viewers' attention (e.g. Nixon, 1924; MacKenzie, 1986; Pieters and Wedel, 2004; Pieters et al., 2007, 2010; Lee and Ahn, 2012; Scott et al., 2016; Zhang and Yuan, 2018). A third sub-stream discusses how viewers' involvement and familiarity with the brand (effects typically grouped by the literature as "top-down") affect attention to advertising (e.g. Treistman and Gregg, 1979; Rayner et al., 2001; Pieters and Wedel, 2007).[2]

Our contribution relative to this literature is that we employ eye-tracking data to examine how readers' attention to news content spills over to the advertising presented on the same page, allowing us to measure the causal effects of news content on attention to

[1] Other related papers include the literature that relates online engagement and advertising effectiveness. For instance, see Kilger and Romer (2007); Calder et al. (2009).

[2] Apart from these areas of inquiry related to advertising effectiveness, eye-tracking has been used in the marketing literature to further our understanding of consideration sets formation (e.g. Chandon et al., 2009), how consumers search and choose products (e.g. Russo and Leclerc, 1994; Lohse, 1997; Janiszewski, 1998; Meißner et al., 2016; Shi and Trusov, 2021), and survey design (e.g. Redline and Lankford, 2001). More broadly, eye-tracking has been used in many fields, including marketing, psychology, and economics, to study individual choices (e.g. Camerer et al., 1993; Armel et al., 2008; Brasel and Gips, 2008; Knoepfle et al., 2009; Reutskaja et al., 2011; Brocas et al., 2014; Pärnamets et al., 2015; Ghaffari and Fiedler, 2018). See Wedel and Pieters (2007) and Wedel (2015) for reviews.

ads and thus assess the importance of investment in high-quality engaging content.[3] We write down an empirical model of attention allocation to interpret this spillover effect and to disentangle this effect from ad avoidance of consumers. We also link this incremental attention to ads to subsequent ad recall and willingness to pay for the advertised brands, thereby providing a needed link between the incremental visual attention and a downstream brand choice measure, called for by Wedel and Pieters (2007).[4] Our analysis is further related to Brasel and Gips (2008); Teixeira et al. (2010) who use eye-tracking data to examine the determinants of attention to TV commercials.[5]

Our second contribution is to examine the effect of the news content on ad effectiveness. We find that more engaging news content increases the amount of attention the reader devotes to display advertising, adding to the results on the effect of page content on ad effectiveness (e.g. Goldfarb and Tucker, 2011). Yet, beyond the effect of devoting more attention to the news page, news content does not have any detectable additional effect on ad effectiveness.[6] In other words, once one statistically controls for attention to the article, whether the article is "hard news" or not has no impact on purchase. This result cautions against the practice of blank blacklisting certain news content for the purposes of targeted advertising (e.g. The Guardian, 2020). Our results on the drivers of attention to online news contribute to the broader literature understanding what makes people engage with news (e.g. Holmqvist et al., 2003; Pitler and Nenkova, 2008; Kazai et al., 2016; Lagun and Lalmas, 2016; Berger et al., 2019).

The third contribution of this article is to validate ad visibility – the amount of time that each ad is visible on the consumer's screen – as a reliable proxy of attention. For

[3]One mechanism behind the spillover of attention can be a visual distraction (e.g. Navalpakkam et al. (2011)). Such distraction has a negative effect on news content consumption (Yan et al., 2020).

[4]See the discussion on page 144 of Wedel and Pieters (2007). Treistman and Gregg (1979) is the closest paper that compares the designs of two commercials and links higher attention to more sales. Zhang et al. (2009) shows that ad features (e.g., size, color, and location of the ad) influence product sales by affecting consumer attention (measured through gaze duration), and van der Lans et al. (2021) shows that online advertising can speed up product search by visually suppressing competing products.

[5]Other recent studies of attention to TV ads include McGranaghan et al. (2022) and Liu et al. (2021).

[6]This suggests a limited interplay of information diagnosticity and accessibility between news content and ads (e.g. Lynch Jr et al., 1988).

this, we first measure attention using scalable and non-intrusive eye-tracking technology, and validate its precision on both desktop and mobile devices. We then show that our main analysis is robust to using attention metrics based on ad visibility metrics. While eye-tracking is a more accurate measure of consumer attention, ad visibility is significantly more likely to be available to researchers and practitioners, expanding the potential application of our paper.

More broadly, our work is related to other papers that have shown links between user exposure to ads and later purchase choices. Several papers link exposure to users becoming aware of the ad (e.g. Danaher and Mullarkey, 2003; Wilson et al., 2015; Elsen et al., 2016). Other articles explore the link between exposure, awareness and purchase (e.g. Hoyer and Brown, 1990; Macdonald and Sharp, 2000; Khurram et al., 2018; Martins et al., 2019). Another literature examines the effectiveness of online advertising on product sales using natural experiments (e.g. Rutz et al., 2012; Narayanan and Kalyanam, 2015; Jeziorski and Moorthy, 2018; Simonov and Hill, 2021) and field experiments (e.g. Lewis and Reiley, 2014; Hoban and Bucklin, 2015; Sahni, 2015; Johnson et al., 2017b,a; Simonov et al., 2018; Gordon et al., 2023).

3 Experimental Setting

In late July and early August 2020, we recruited 1,013 individuals, stratified evenly across two countries (UK and US) and two device types (desktop and smartphone). Respondents matched the UK/US online population in terms of age, gender, income, and location. They were recruited via a specialist supplier of research panels, Panelbase.[7]

The experiment proceeded as follows. First, we confirmed the viewer's consent. At the start of the experiment, participants were told only they were a part of "an academic study about media consumption," but were not given additional details. At this stage, participants were asked to report their age, education, income, gender, and postcode.

Each participant was then invited to read articles from two online newspapers. In

[7] See https://www.panelbase.net/.

each country, we chose outlets with a wide online readership: *The Guardian* and *Daily Mail* for UK participants, *The New York Times* and *USA Today* in the US.

We presented each individual with 9 articles. All articles had been published in the short time window prior to the experiment taking place, to maximize the probability that the articles were relevant and interesting. Within each newspaper, articles were split between soft and hard news. To select the latter, we followed the advice of industry experts and focused on articles about the COVID-19 pandemic and the BLM protests of the summer 2020. These two topics were frequently "blacklisted" by advertisers.[8] The text of the articles shown on desktop and mobile was the same. However, in our analysis, we consider these to be different articles, since the format of the text is quite different across devices.

Eight out of nine articles were accompanied by ads from well-known and widely available brands; one of the articles was randomly shown with blank spaces in the location where ads would be otherwise shown. In each country, we chose 8 prominent brands (see Appendix A). All ads accompanying a given article were for the same brand, inserted at fixed points along the article's page. We included one horizontal "billboard" ad before the text of the article, and two smaller "side" ads, on the side of the article text (desktop) or in-between paragraphs of the text (mobile). Our goal was to approximate, as much as possible, the typical reading experience online.

In each country, each participant was exposed to all 9 articles and all 8 brands. Each article and brand was shown only once. We randomized the order in which articles were presented to individuals and the pairing between articles and brands. Individuals were allowed to read the articles at their preferred pace.

For each individual, we obtained two measures of the attention devoted to each article and ad. First, the amount of time the article and the ad were visible on screen, which

[8]We provide article titles and links to the articles we used in Appendix A. We validate our categorization of articles as "hard" or "soft" news using an independent survey on Amazon Mechanical Turk (AMT), described in Appendix B.

does not require eye-tracking. Second, we recorded, via eye-tracking, the time that each individual's sight *dwelled* on each article and ad, referred to as dwell time.

After reading all articles, individuals were asked if they could remember the brands whose ads they had seen. Individuals were presented with a list containing the eight brands shown and eight "decoy" brands, in a random order. The decoy brands were chosen to be well known in each country, and of the same industries as the shown brands. All brands (shown and decoy) were presented to the participant simultaneously, and participants selected which of the 16 brands they remembered seeing.[9]

After the recall task, participants were asked to make purchase decisions. For each of the brands whose ads were shown, individuals were offered to choose between (i) an e-voucher worth £10 (in the UK) or $10 (in the US) specific to one of the brands shown, or (ii) a randomized amount of cash (£3-7 in the UK and $3-7 in the US). Individuals were informed that one of their e-voucher versus cash choices – selected later on at random – would be sent to them. As a result, purchase decisions were incentivized.

In addition to the voucher/cash reward, participants were paid a fixed participation fee. Participants were anonymous to the research team, with all payments delivered via the recruiting firm. The study protocol received ethical approval prior to the start of the experiment (see Appendix A.4).

We do not use a standard between-subjects experimental design. This is because our main goal is not to measure the extensive margin, i.e., the effect of the presence of ads relative to their absence. Instead, we aim to study the intensive margin: how incremental attention to articles results in incremental attention spillovers to ads. In our experiment, exogenous variation in attention to ads was induced by the random pairing of articles and ads. Some articles are more interesting than others, leading participants

[9]Immediately before collecting this "aided" recall measure, we have also collected a measure of "unaided" recall, where we asked participants to write the names of the brands they recalled seeing. These two measures of recall are highly correlated (65%) and all results are robust to using either measure. For brevity, we only report results using the aided recall measure. The robustness of our results to an unaided recall measure suggests that additional attention to ads leads to short-term memory activation – unaided recall requires participants to remember the advertised brands.

to devote more attention to those articles, which then influences the attention devoted to the ads placed next to them. It is this exogenous variation in attention to ads that we use to discuss the causal effect of attention on recall and purchase. This method for identifying the causal effect of attention to ads closely tracks our research question – the possible complementarity of the news content and ads – and is, to our knowledge, a novel way to measure ad effectiveness.[10]

For the purposes of this study, an online experiment provided several advantages. It allows for a large data collection effort, across multiple countries and devices, at a relatively low cost. It also allows us to show recently published articles to a large number of individuals, which would have been challenging in a lab environment. Our setting is also closer to the conditions under which individuals normally engage with online content.

The eye-tracking technology used was supplied by Lumen Research, a specialist advertising research agency.[11] The technology employs software that uses the camera of a desktop or mobile phone to measure where on the screen the retina of the participant is focused. No additional hardware is needed. See Appendix A for more details on the eye-tracking technology, its calibration and validation.

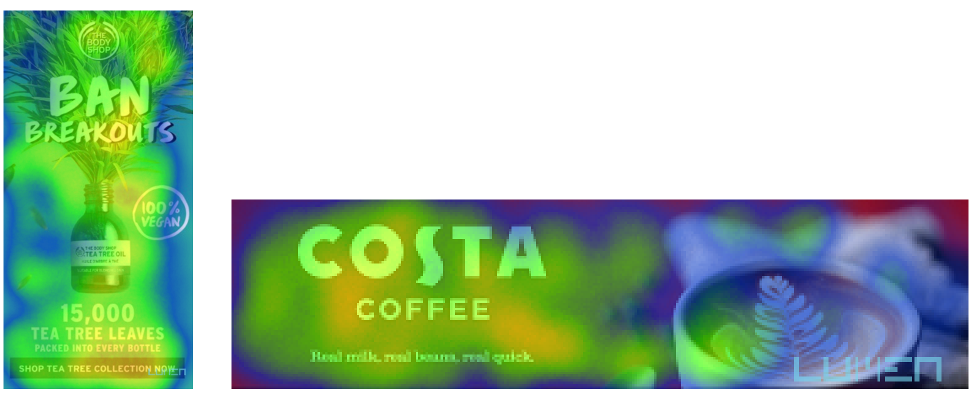
The heat map provided in Figure 1 Panel (a) is an example of how these metrics are constructed. The figure shows an article, as well as the ads (a "billboard" ad and two "side" ads) for one brand. The map highlights the regions on the screen that were actively dwelled upon by the participant. In Figure 1 Panel (b), we present examples of heatmaps for ads of two different brands.

[10] Goldstein et al. (2011) also randomize pairings of articles and ads in their first study, but they force ads to always be visible on the page and do not measure attention to ads via eye-tracking.

[11] See https://lumen-research.com/.

Figure 1: Example of Heat Maps

(a) Heat Map of a Page



(b) Heat Maps of Two Ads

In the heat map, red color means more attention and blue color means less attention.

4 Data

4.1 Variables

Table 1 presents summary statistics for our sample, which is at the individual×article level. About half of the observations occur on desktops (56%), correspond to female participants (55%), are from the US (48%), and have “hard” news articles (55%).

Table 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Desktop	6,431	0.563	0.496	0	1
Female	6,431	0.556	0.497	0	1
U.S.	6,431	0.483	0.500	0	1
Hard News	6,431	0.550	0.498	0	1
Article Visible (s)	6,431	143.301	169.341	20.130	1,894.635
Ad Visible (s)	5,707	19.027	17.371	0.000	291.905
Price (USD/GBP)	5,707	5.017	1.436	3.000	7.000
Recall	5,707	0.484	0.500	0.000	1.000
Buy	5,707	0.347	0.476	0.000	1.000
Article Dwell (s)	4,426	74.813	97.918	0.112	966.945
Ad Dwell (s)	3,925	2.755	3.161	0.000	40.214

Each observation is at the individual x article level.

Visibility Measures The variable *Article Visible* reports the number of seconds any part of the article was visible on screen (sample mean is about 2 minutes and 23 seconds), while *Ad Visible* reports the total number of seconds that any ad on the page was visible on the screen. To measure *Ad Visible* , we used Media Rating Council standards: an ad is considered visible if at least 50% of the pixels of the ad are displayed on the screen for 1 continuous second or more. The sample mean is approximately 19 seconds per article. These measures do not use eye-tracking.

Eye-tracking Measures The variable *Article Dwell* is the total time an article was actively being read, recorded via eye-tracking. The sample mean is about 1 minute and 15 seconds per article. Similarly, *Ad Dwell* reports the total time that all ads associated with an article were actively looked at (i.e., the sum of the dwell time of the 3 ads shown on each page). The sample mean is just short of 3 seconds. Intuitively, *Article Dwell* and *Ad Dwell* should be be lower than *Article Visible* and *Ad Visible* , respectively, since articles and ads can be present on the screen but not actively being looked at.

Purchase Participants in the US (UK) were offered choices between vouchers worth \$10 (£10) for each of the advertised brands and random amounts of cash. The amount of cash offered to individuals is captured by the variable *Price*, since this is the opportunity cost of choosing the voucher. For about 35% of observations, individuals chose the voucher (measured by the dummy variable *Buy*), while the rest opted for cash.

Recall About 48% of observations had individuals recall the associated brand (measured by the dummy variable *Recall*). In contrast to brand purchase choices, recall was not incentivized, but this measure is commonly used in marketing literature (e.g., Danaher and Mullarkey, 2003; Elsen et al., 2016).

Individual Demographics We recorded the following individual-level demographics: gender, age, education, income, country and device type. We also asked individuals about their self-report political leaning (liberal, conservative or moderate) but only at the end of the experiment, so as not to prime their responses to the articles.[12]The sample characteristics are similar when data is split by device type (mobile vs. desktop) and country (UK vs. US). In Appendix Tables 19 and 20, we replicate Table 1 for mobile and desktop devices separately and find demographic composition, news types, prices, purchasing, and brand recall summaries to be consistent. The only notable difference is that ads are more visible on desktop computers (average of 23.4 seconds) than on mobile phones (13.4 seconds). *Ad Dwell* is around 2.7 seconds on average on both device types. Consumers also spend more time reading articles on desktops (average 83 seconds) than on mobile phones (65 seconds).

Article Characteristics The main article characteristic we consider is whether the article constitutes "hard news." As described above, for this purpose we selected articles

[12]For brevity, we omit most of the demographic variables from Table 1 and present them in Appendix Figure 21.

focusing on the COVID-19 pandemic and the BLM protests during the summer of 2020. For some robustness checks, we also use the article's word count.

Final Sample Our final dataset comprises 6,431 observations at the individual×article level. This is less than the originally targeted 9 observations per person, for two reasons. First, due to connectivity issues, no data was recorded for around 30% of individual-article pairs. These missing observations are slightly more prominent on mobile phones (43%) than on desktop computers (13.5%), and in later steps of the study. We confirm this does not introduce bias in our analysis by showing that there is no selection bias in terms of which brands' and articles' observations experienced connectivity issues (see Appendix D and Appendix Figures 10 and 11). Second, for a subset of participants, eye-tracking quality was poor. High-quality eye-tracking relies on minimal head movement for continuous tracking of the individual's retina. We only include in our analysis individuals with high-quality eye-tracking data. This explains why we have fewer observations (around 70%) with eye-tracking than visibility data. In our main sample, observations with low-quality eye-tracking data are identified using metrics typically used by the eye-tracking technology provider. Appendix A.3, D, and C provide additional information about the final sample, show that there is no selection bias in terms of brands' and articles' observations, and offer robustness checks using alternative metrics of eye-tracking data quality.

4.2 Descriptive Statistics

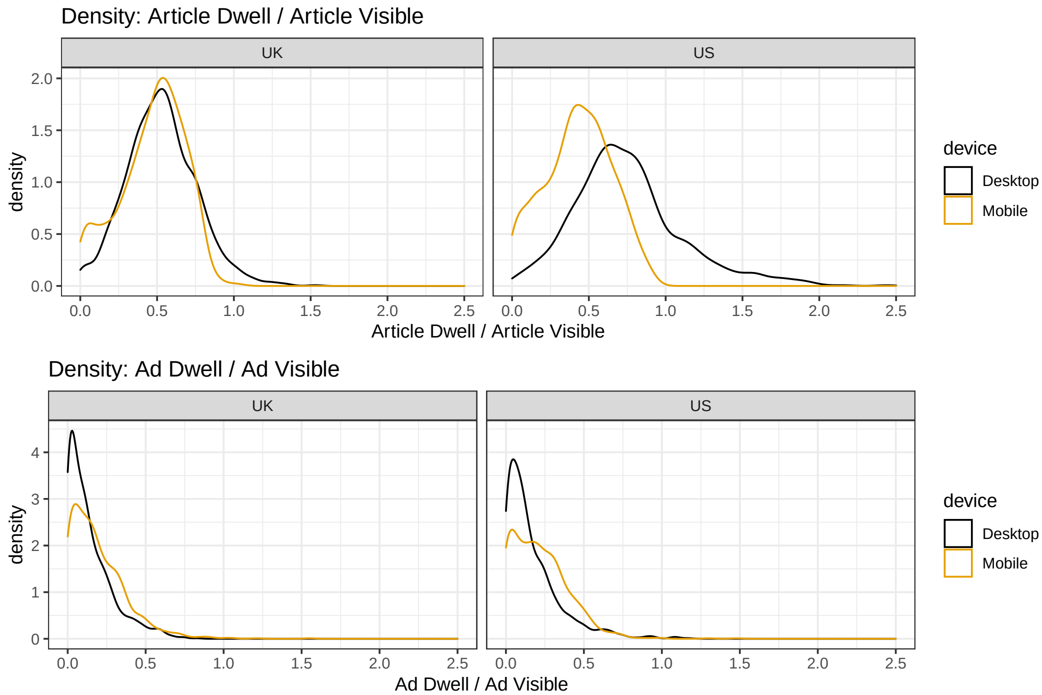
Distributions of Attention Measures

Figure 2 compares our measures of attention (*Ad Dwell*, *Article Dwell*) and visibility (*Ad Visible*, *Article Visible*).[13] In the top section of the figure, we present ratios of *Article Dwell* to *Article Visible* for each country and device type. On average, *Article Dwell* is around 50% of *Article Visible*, indicating that an average reader looks at the article 50%

[13] Appendix Figures 13 and 14 present marginal distributions of attention (*Article Dwell*, *Ad Dwell*) and visibility measures (*Article Visible*, *Ad Visible*), for all observations and averaged per consumer.

of their time when the page is loaded.[14]

Figure 2: Dwell to Visible Ratios, by Country and Device



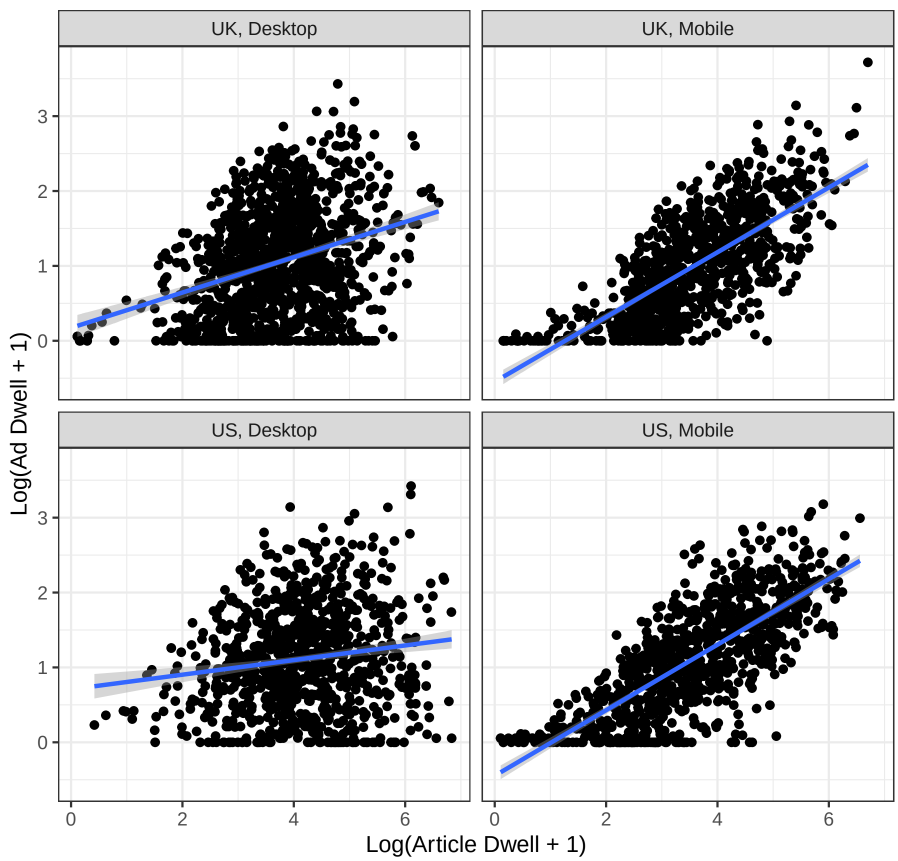
The plots show the ratio between time spent dwelling and time visible, for both articles and ads, computed across all observations in the data.

The lower part of Figure 2 presents ratios of *Ad Dwell* to *Ad Visible*. The average ratio is much lower compared to the analogous ratio for articles (around 18% instead of 50%). This agrees with the existing literature finding that TV ads can be visible for around 55% of viewers – meaning that viewers stay in the room for commercials – but only 7.7% of viewers actually devote visual attention to TV commercials (McGranaghan et al., 2022). The ratio of *Ad Dwell* to *Ad Visible* is slightly higher for mobile devices (21%) than desktops (15%). This reflects different prominence of display ads on desktop and mobile devices, and, in particular, the difference in prominence of "side" ads: on desktops, side ads are on the right side of the page, visible but easy not to devote attention to, whereas

[14]The average is slightly lower for mobile devices (45%) as compared to desktops (62%). This is largely explained by the desktop page design of *USA Today*, that shows only a small fraction of the article at first and therefore under-counts *Article Visible*. If we exclude *USA Today* articles, the average ratio of dwell-to-visible measures is 48% for mobile and 54% for desktops. The desktop page design of *USA Today* also explains almost all of the (rare) cases where the ratio of article dwell to article visible is greater than one.

on mobile phones they occupy blocks between the text in the center of the screen.

Figure 3: Positive Correlation in Article and Ad Dwell



Correlation between attention to article and ad, split by country (UK, US) and device type (desktop, mobile). Ad and article dwell times are transformed into the logarithmic scale to improve visualization. The blue line corresponds to the best linear prediction of the variable on the vertical axis by the variable on the horizontal axis. The grey area corresponds to 95% confidence intervals

Attention decreases throughout the experiment on both mobile and desktop devices. Appendix Figure 16 illustrates *Article Dwell* and *Ad Dwell* for the 9 experimental steps (e.g., the third article shown corresponds to step 3). On average, *Article Dwell* is 117 seconds on desktops in the first step, decreasing to 64 seconds in the last step (99 to 42 seconds on mobile devices, respectively). *Ad Dwell* is approximately 4 seconds in the first step on both device types, decreasing to 2.2 seconds in the last step.

There is a robust positive correlation between attention devoted to articles and their associated ads. Figure 3 displays a scatter plot of *Ad Dwell* and *Article Dwell* . Across countries and device types, a positive correlation of 0.36 is observed. The correlation is more pronounced for mobile devices (0.65) compared to desktops (0.16). Also, Appendix Figure 17 shows that this positive correlation persists within each article. Even after controlling for country, device, step-order, and demographic fixed effects (henceforth "FE"), the positive correlation between *Article Dwell* and *Ad Dwell* remains robust.[15] In Appendix C, we show that this positive correlation is robust to a battery of checks accounting for potential measurement error in attention.

Purchase and Recall Measures

Our key outcome measures are the recall of the shown brand and participants' choices in an incentivized "purchase" scenario (i.e., choosing a \$10/£10 voucher for the brand over a lower cash amount). We ensure the meaningfulness of these purchase choices by examining demand curves derived from randomly assigned cash amounts (Appendix Figure 19). Here, we describe the average share of US and UK consumers opting for the brand voucher over cash. On average, 52% of US consumers and 40% of UK consumers choose the brand voucher over \$3/£3. This preference diminishes as the cash amount increases, with only 34% of US consumers and 20% of UK consumers selecting the brand voucher over \$7/£7. In Appendix Figure 20, we separately estimate demand curves for each brand, confirming that this pattern is not influenced by outlier brands.

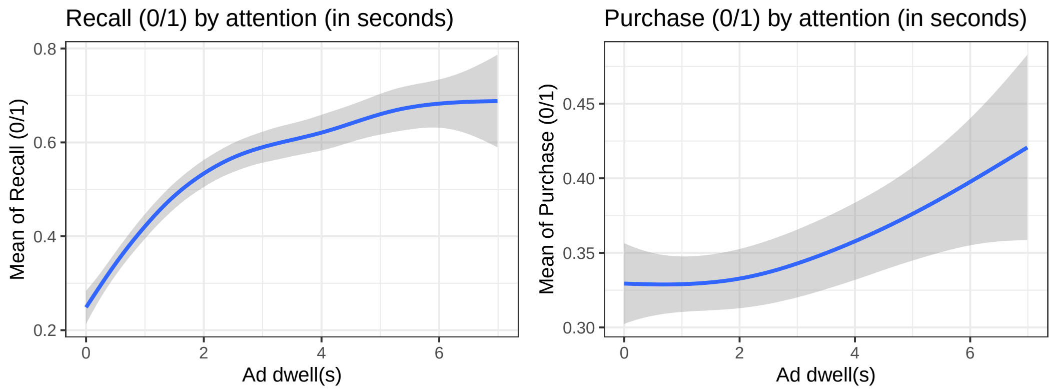
In Figure 4, we examine how recall and purchase correlate with attention devoted to ads, *Ad Dwell* . The percentage of individuals who chose the voucher and recalled seeing the brand increases with the amount of attention devoted to the ad.

Appendix E contains a number of additional data descriptives, including marginal densities of our attention metrics, attention by types of ads (side vs top), how attention changes with step order, demand curves for the different brands and distributions of

[15] Appendix Figure 18 presents a version of Figure 3 using only residualized variation in attention measures.

Marketing Science Institute Working Paper Series

Figure 4: Purchase and Recall Increase in Ad Dwell Time



The panels show non-parametric regressions of purchase/recall on *Ad Dwell* , together with 95% confidence intervals. The automatic optimal bandwidth is used. The range of the x-axis is capped at 7 seconds, which is approximately the 90th percentile of the distribution.

other individual characteristics (gender, education, age, income, and political leaning).

5 Determinants of Attention Allocation

5.1 Empirical Framework

Consider consumer i who decides to devote x^{art}_{ijks} seconds of attention to article j, which is paired with an ad for brand k, in experimental step $s = 1, .., 9$. Moreover, she also decides to devote x^{ad}_{ijks} seconds of attention to the corresponding ad. We assume that $x^{\text{art}}_{ijks}, x^{\text{ad}}_{ijks}$ are defined by the following set of simultaneous equations:

$$x^{\text{ad}}_{ijks} = \mathbb{1}_{ijks}(\delta_{iks} + \gamma \cdot x^{\text{art}}_{ijks} + \epsilon^{\text{ad}}_{ijks}), \tag{1}$$

$$x^{\text{art}}_{ijks} = \alpha_{ijs} + \mathbb{1}_{ijks}(-\beta + \gamma \cdot x^{\text{ad}}_{ijks}) + \epsilon^{\text{art}}_{ijks}. \tag{2}$$

The reader's attention to the ad, x^{ad}_{ijks}, is formed from three components. First, δ_{iks} is individual i's preference for devoting attention to the ad of brand k in step s. We estimate δ_{iks} as a flexible function of individual characteristics, ad FE, and experimental step FE. Controlling for individual characteristics accounts for the fact that some types of individuals might be more drawn to ads than others. Ad FE account for the fact that certain ads might be more appealing than others. We control for the experimental step to account for the potential fatigue of participants in the study.

Second, the attention devoted to an ad can be influenced by the attention devoted to the article, as captured by γ. This corresponds to possible "attention spillovers" between articles and ads. For instance, as an individual reads the article, small movements of the retina or peripheral attention might allow them to perceive the ad next to it. If so, attention to articles creates positive spillovers of attention to ads ($\gamma > 0$), corresponding to a model of "bottom-up" attention (Koch and Ullman, 1987; Itti et al., 1998; Pieters and Wedel, 2007; Cerf et al., 2007; Milosavljevic and Cerf, 2008). Alternatively, if articles absorb the individual's attention, high attention to articles would correspond to low attention to ads. Then, attention spillovers are negative ($\gamma < 0$), corresponding to a model of "top-down" attention (Drèze and Hussherr, 2003; Stenfors et al., 2003; Simola et al., 2011). We assume that this effect is homogeneous across individuals, articles and ads.

Finally, $\epsilon^{\text{ad}}_{ijks}$ is an independent idiosyncratic error term that can impact the attention the individual devotes to the ads. Since attention to an ad is zero when the ad is not present, (1) is multiplied by $\mathbb{1}_{ijks}$ – an indicator that equals one if the ad of brand k was shown next to article j for participant i in step s, and zero otherwise. Therefore, to estimate (1), we use only data for which articles were matched with ads.

Similarly, there are three components that determine the reader's attention to the article, x^{art}_{ijks}. First, α_{ijs} captures reader i's interest in article j during experimental step s. Below we define α_{ijs} as a flexible function of individual characteristics, article FE, and experimental step FE. The variable $\epsilon^{\text{art}}_{ijks}$ is an independent idiosyncratic error term that can affect the attention the individual devotes to the article.

Finally, if the ad is present ($\mathbb{1}_{ijks} = 1$), it affects the reader's attention to the article in two ways. First, the coefficient β is the reader's disutility of attention to the article when any ad is shown next to it (or utility if $-\beta > 0$). This captures the fact that ads can be distracting and therefore reduce the amount of attention devoted to articles.[16]

The coefficient γ captures the same "attention spillover" between articles and ads as in Equation 1. For simplicity, our main specifications assume that the attention spillover effect is symmetric – an extra second of attention to the article leads to γ seconds of attention to the ad, and vice versa. This is a natural assumption if attention spillovers are driven by peripheral attention or small movements of the retina due to distractions while looking at page elements. However, all our subsequent results are robust to the spillover from ads to news being of a different order of magnitude than the spillovers from news to ads. We show that the incremental spillover of attention from ads to articles is negligible because an average respondent spends 27 times more time paying attention to news (74.8 seconds) than ads (2.8 seconds), per Table 1.

In Appendix H we further ground our empirical specification in a simple attention allocation model, where individuals choose attention to maximize their utility. We show that (1)-(2) correspond to the solution of this utility maximization problem.

5.2 Identification

There are two main coefficients of interest. The sign of β reflects whether readers are "ad avoiders" (e.g. Wilbur, 2008, 2016; Huang et al., 2018) or "ad lovers" (e.g. Kaiser and Wright, 2006), while γ determines the sign and magnitude of attention spillovers.

The parameter γ can be consistently estimated because of the random matching between articles and ads. For a given ad, our experimental design randomly pairs it with a more or less interesting article. This creates an exogenous shock to the amount of attention a consumer devotes to the content paired with the ad, which we use to estimate the spillover effect of attention to news on the attention to ads.

We estimate γ in two ways. First, to isolate exogenous attention to articles, we estimate Equation (1) instrumenting x^{art}_{ijks} with the average amount of attention devoted to

[16]We assume that β is the same for all individuals and articles. This is a simplification since, in reality, some ads can be particularly distracting. In principle, we could allow β to vary by article and ad. In practice, we are underpowered to estimate those coefficients.

that article by all *other* individuals in the sample. We refer to this instrumental variable (IV) as the "Leave One Out" (L1O) mean of article attention.[17]

Second, we show that an OLS regression of x_{ijks}^{art} on x_{ijks}^{ad} in (1) leads to statistically similar estimates of γ as the L1O IV regression. In a general setting, OLS estimates of γ from (1) should be biased due to reverse causality since γ is also present in (2). However, in our context this simultaneity bias is negligible. This is because the true γ is precisely estimated at around 0.008, and the average consumer allocates around 27 times more attention to the article (74.8 seconds) than the ad (2.8 seconds). Thus, the "feedback" of attention spillovers from ads to articles is approximately $0.008 \cdot 2.8 = 0.022$ seconds, or $0.022/74.8 = 0.029\%$ of the average attention individuals devote to an article.

Given the estimates of γ, the parameter β is identified by comparing attention to articles shown with and without ads after controlling for the estimated $\hat{\gamma} \cdot x_{ijks}^{\text{ad}}$ in (2). Readers' overall tastes for devoting attention to articles and brands' ads $(\delta_{iks}, \alpha_{ijs})$ are identified from the average attention consumers spend on articles and ads at different experimental steps. Article, ad, and individual FE are not necessary for consistent estimates of γ, β, but we show specifications where they are included to check robustness of the estimates.

5.3 Estimation

We estimate the parameters described above in two steps. First, we use (1) to estimate γ and δ_{iks}. We estimate γ both by using the L1O IV (described above) and by an OLS regression, including varying sets of FE (step, brand, and individual) to increase precision and to check the robustness of the estimates. To estimate (1), we use only observations when an ad is present on the page, since otherwise $x_{ijks}^{\text{ad}} = 0$ mechanically.

[17] This "jackknife" instrument is similar to the use of article fixed effect as an instrument, but eliminates the bias associated with including the current individual when computing the fixed effect, as discussed by Angrist et al. (1999); Kolesar (2013). This IV approach is similar to the "random judges" instruments used in Dahl et al. (2014); Dobbie et al. (2018). In this case, each article is a "judge", and ads are randomly assigned to articles. Articles that vary in attractiveness play the role of judges who vary in leniency. Our results are also robust to using only attention to articles from individuals for whom the article was *not* paired with the same ad as the target person.

We then use (2) to estimate α_{ij}, β. We estimate an OLS regression of $x^{\text{art}}_{ijks} - \mathbb{1}_{ijks} \cdot \hat{\gamma} \cdot x^{\text{ad}}_{ijks}$ (i.e., attention paid to the article net of spillover effects from the attention to ads) on the indicator $\mathbb{1}_{ijks}$. We use the entire sample including the articles shown without ads. Here, $\hat{\gamma}$ is the estimate of γ from the first step. In our empirical context, we find that the effect of $\hat{\gamma} \cdot x^{\text{ad}}_{ijks}$ on x^{art}_{ijks} is negligible due to both a low estimate of $\hat{\gamma}$ (around 0.008) and relatively low attention to ads relative to articles, as discussed above. Our estimates of α_{ij} and β would be virtually identical if we assumed $\hat{\gamma} = 0$ or a value of $\hat{\gamma}$ that is 10 times larger than what we estimate. We include varying sets of FE (step, article, and individual) for robustness.

In both steps, we cluster standard errors at the individual level.

5.4 Results

Table 2 presents the estimates under alternative specifications of α_{ijs} and δ_{iks}. Columns (1-3) present the estimates using the L1O IV for estimating γ in Equation (1). The first stage relationship is highly significant across all specifications, with an incremental F-statistic of 65.9-128.2.

In Column (1), we assume that α_{ijs} and δ_{iks} are only a function of the experimental step. Formally, we assume that $\alpha_{ijs} = \alpha_s$ and $\delta_{iks} = \delta_s$. Parameters $\hat{\alpha}_1$ and $\hat{\delta}_1$ show the attention devoted, on average, to the article and ad shown to each individual in experimental step $s = 1$. In the first article-ad pair presented, individuals allocate on average approximately $\hat{\alpha}_1 \approx 106$ seconds of attention to the article and $\hat{\delta}_1 \approx 3$ seconds of attention to the ad. An extra second spent looking at the article increases the amount of time individuals look at the ad by $\hat{\gamma} = 0.008$ seconds. Thus, the 106 seconds spent (on average) looking at the first article creates a total of $106 \cdot 0.008 = 0.848$ seconds of positive spillover attention to the ad, or a $100 \cdot 0.848/3 = 28.2\%$ increase in ad attention. The magnitude of the reverse effect is very small: the 3 seconds of attention (on average) devoted to the ad create $3 \cdot 0.008 = 0.024$ seconds extra attention to the article, or a $100 \cdot 0.024/106 = 0.02\%$ increase. Having no ad next to the article increases the amount

Table 2: Estimates of attention spillovers and ad avoidence

	Ad Dwell					
	IV			OLS		
Panel I	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}_1$	3.083***			2.715***		
	(0.306)			(0.197)		
$\hat{\gamma}$	0.008***	0.007***	0.009***	0.011***	0.011***	0.011***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
1st Stage Incr. F-Stat	65.86	80.15	128.23			
Observations	3,925	3,925	3,925	3,925	3,925	3,925
R^2	0.135	0.135	0.202	0.145	0.152	0.205
	Article Dwell - $\hat{\gamma}$ Ad Dwell					
Panel II	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\alpha}_1$	105.907***			105.894***		
	(4.521)			(4.521)		
$\hat{\beta}$	7.015*	6.832*	7.906**	7.024*	6.845*	7.913**
	(3.919)	(3.741)	(3.536)	(3.919)	(3.741)	(3.536)
Observations	4,426	4,426	4,423	4,426	4,426	4,423
R^2	0.030	0.112	0.500	0.030	0.112	0.500
FE:						
Step Order	Y	Y	Y	Y	Y	Y
Article	N	Y	Y	N	Y	Y
Brand	N	Y	Y	N	Y	Y
Country x Device	N	N	Y	N	N	Y
Dem. Controls	N	N	Y	N	N	Y
Poly(Duration, 4)	N	N	Y	N	N	Y

*p<0.1; **p<0.05; ***p<0.01

All specifications include step order fixed effects, with step order = 1 normalized to zero. Estimates in Panel I represent coefficients from a regression of Ad Dwell on Article Dwell. In the IV specification, Article Dwell is instrumented for by the average amount of attention devoted to that article by all but this individual (Leave One Out IV). Estimates in Panel II represent coefficients from an OLS regression of Article Dwell on an indicator of whether the ad is present on the news article. We subtract $\hat{\gamma}$ Ad Dwell from Article Dwell in Panel II to control for the attention spillover from ad to article. Demographic controls include income, gender, education, age, and self-reported political leaning. Poly(Duration, 4) corresponds to a quartic polynomial in log of average time that an average article was visible for by each individual. Standard errors clustered at the individual level.

of time readers devote to the article by approximately $\hat{\beta} \approx 7$ seconds, showing that the average consumer is an ad-avoider.[18]

In Column (2) of Table 2, we allow α_{ijs} and δ_{iks} to vary across articles and ads by including article and ad FE: we assume $\alpha_{ijs} = \alpha_j + \alpha_s$ and $\delta_{iks} = \delta_k + \delta_s$. The baseline levels of $\hat{\alpha}$ and $\hat{\delta}$ are now subsumed by these FE, so we omit them from Column (2). However, the estimates $\hat{\beta}$ and $\hat{\gamma}$ are nearly identical to those in Column (1).

In Column (3), we further allow α_{ijs} and δ_{iks} to vary across a variety of individual characteristics and controls, denoted X_i. We include country-by-device FE and socio-demographic variables (e.g. income, age, and gender) also included as FE.[19] We further include a quartic (4th order) polynomial of the time each individual spent on an average article (measured by how long each page was visible) to capture the fact that some individuals read more slowly or are intrinsically more engaged by articles than other individuals. In total, we include 31 additional covariates in the regression. Again, the estimates of $\hat{\beta}$ and $\hat{\gamma}$ are statistically indistinguishable from those in Column (1), while their precision is higher.

Further, our results are robust to adding individual FE that subsume X_i (see Appendix Table 21). Adding individual FE increases the number of controls by around 700 additional covariates, substantially reducing our statistical power in a sample of around 3,900 observations. Because of this, for our main analysis, we prefer a specification with individual covariates as controls.[20]

The magnitude of $\hat{\gamma}$ estimated in Columns (1-3) implies that reverse causality – the effect of attention to ads on attention to articles – is negligible. To confirm this further, in Columns (4-6) we present the results of estimating (1) by OLS. The parameter estimates

[18] We omit the step-order fixed effect estimates from Table 2 to improve readability. Articles and ads shown in later steps of the experiment obtained less attention from participants, as illustrated in Appendix Figure 16.

[19] For instance, individuals reported their age in bins of 10 years, so we include an indicator for each such bin.

[20] We further note that all our results are robust to a higher and lower order polynomial of the time individuals spent on an average article.

are statistically indistinguishable from those in Columns (1-3).[21]

6 Determinants of Recall and Purchase

So far we have estimated how individuals devote their attention. After reading the articles, we asked individuals if they recall the advertised brands. We also asked individuals to make a purchase choice for each brand they have seen.[22] We now estimate the effect of ad attention on consumers' brand recall and purchase decisions.

6.1 Empirical Framework

Consider whether consumer i recalls seeing an ad for brand k, shown next to article j in experimental step s after devoting attention x^{ad}_{ijks}. Let $r_{ijks} \in \{0,1\}$ be an indicator that takes value 1 when recall is correct, and 0 otherwise. We assume that this recall process follows the following linear probability model (Heckman and Snyder Jr, 1997):

$$r_{ijks} = \theta^r_s + \eta^r_k + \mu^r X_i + \rho x^{\text{ad}}_{ijks} + \epsilon^r_{ijks}. \tag{3}$$

The coefficients θ^r_s are FE for the experimental step s, capturing the effect on recall of seeing an ad later or earlier in the experiment.[23] The coefficients η^r_k are brand FE: some brands might be more memorable than others. Finally, X_i adds individual-level controls similar to the ones used in Columns (3) and (6) of Table 2 – country-by-device FE, socio-demographic characteristics, and a proxy for each individual's reading speed (a 4th degree polynomial in the average time taken to read an article). Finally, ϵ^r_{ijks} captures other idiosyncratic shocks that determine consumer i's recall of brand k. Our main parameter of interest is ρ, which captures recall ad effectiveness: the effect of additional

[21] For instance, Column (4) reports that an extra second of attention devoted to the article increases the amount of time devoted to the ad by $\hat{\gamma} = 0.011$ seconds (s.e. 0.002), statistically similar to the 0.008 seconds estimate in Column (1) (s.e. 0.003).

[22] We convert UK prices to US dollars at Purchasing Power Parity at the time of the experiment (July 2020), £1 = $ 1.66. Recall that the price p_{ik} is the random amount of cash offered to individual i as an alternative to choosing the voucher for brand k.

[23] The effect is a priori ambiguous. Ads shown later may receive less attention due to fatigue, but also might be more vivid in the participant's memory at the point when they are asked about their recall.

attention to the ad of brand k (x^{ad}_{ijks}) on the brand's recall.

Similarly, let $v_{ijks} \in \{0, 1\}$ be an indicator for the individual purchasing the voucher for brand k. We assume that the probability that the individual purchases the voucher at a price p_{ik} after devoting attention x^{ad}_{ijks} to the ad for brand k is:

$$v_{ijks} = \theta^v_s + \eta^v_{k,p_{ik}} + \mu^v X_i + \lambda x^{\text{ad}}_{ijks} + \epsilon^v_{ijks}. \tag{4}$$

Equation (4) is analogous to (3), with a few small differences. As above, θ^v_s are step-order FE. In (3), we assume that price does not affect consumer recall. However, in (4), we allow for price to potentially affect purchase decisions. Therefore, $\eta^v_{k,p_{ik}}$ are brand×price FE, which allows the price elasticity to vary flexibly along the demand curve for each brand and allows demand curves to differ across brands. Finally, ϵ^v_{ijk} captures other idiosyncratic shocks affecting the individual's purchase probability. The parameter λ is the purchase ad effectiveness: the effect of additional attention to the ad of brand k on the decision to purchase that brand.

6.2 Identification and Estimation

We are interested in the effects of attention to advertising on consumer recall and purchase decisions, captured by the parameters ρ, λ in (3)-(4). We rely on two empirical strategies to estimate these parameters.

First, we rely on accounting for potential sources of endogeneity biases by including several controls when estimating (3)-(4): step order FE, brand FE and individual characteristics. These controls account for the same information included in δ_{iks} in Equation (1), capturing the main potential sources of endogeneity – e.g. experimental step might be correlated with both ad recall and attention due to respondents' fatigue, and brand FE account for potentially higher quality ads by the more popular brands.[24] With these controls, the residual variation in the amount of time consumers allocate to ads, x^{ad}_{ijks},

[24]While our main specification includes individual demographics as controls, Appendix Table 23 shows that the estimates are robust to including individual FE.

is plausibly driven by individuals' idiosyncratic decisions of how much attention to devote to ads that appear randomly throughout our study. If this residual variation in x^{ad}_{ijks} is uncorrelated with the idiosyncratic shocks that influence consumers' recall and purchase outcomes $(\epsilon^r_{ijks}, \epsilon^v_{ijks})$, then OLS produces consistent estimates of $\hat{\rho}, \hat{\lambda}$.

Second, we leverage the random pairing of ads and articles in our experiment to further relax the assumption that x^{ad}_{ijks} is uncorrelated with $\epsilon^r_{ijks}, \epsilon^v_{ijks}$ conditional on controls. One possible remaining concern is reverse causality – perhaps individuals devote more attention to ads of brands that they are familiar with and particularly like, and therefore are more likely to recall and purchase. Formally, high x^{ad}_{ijks} is due to a high $\epsilon^{\text{ad}}_{ijks}$, which in turn might be correlated with $\epsilon^r_{ijks}, \epsilon^v_{ijks}$. This argument is in line with the model of Becker and Murphy (1993); Tuchman et al. (2018), where advertising has consumption value and enters viewers' utilities. To address these concerns, we instrument the amount of attention a reader devotes to an ad (x^{ad}_{ijks}) with the amount of attention she devotes to the article randomly paired with that ad (x^{art}_{ijks}). In Section 5.4, we have shown that there is a strong positive spillover in the consumer's attention from article to ads, making x^{art}_{ijks} a relevant instrument. Moreover, we have also shown that the "feedback" effect of ads on articles is minuscule and robust to using a L1O IV strategy, validating the exogeneity of the instrument. Therefore, this identification strategy uses only the incremental exposure to ads due to positive spillovers of attention from randomly paired articles (that can be more or less interesting to consumers) to measure the effect of ad exposure on recall and purchase decisions.

In all specifications, we cluster standard errors at the individual level.

6.3 Results: OLS

We begin by presenting the results of OLS regressions. Table 3 shows the OLS estimates of the effect of ad attention – measured both with *Ad Visible* and *Ad Dwell* – on recall and purchase. Column (1) reports the estimates of recall ad effectiveness $(\hat{\rho})$ based on all observations in the sample. Panel I considers attention measured by *Ad Visible* . If a brand's ad is visible for 1 extra second, this increases the probability of the individual remembering that brand by about 0.32 percentage points. An increase in *Ad Visible* of one standard deviation (17.37 seconds, from Table 1) is associated with an increase in recall of $100 \cdot 0.0032 \cdot 17.37 = 5.55$ percentage points. Panel II considers attention measured by *Ad Dwell* . If the individual devotes an additional second of attention to an ad, this increases the probability of recall by 3.43 percentage points. An increase in one standard deviation of *Ad Dwell* (3.16 seconds) increases recall probability by 10.83 percentage points (relative to the average recall probability of 48%). In line with intuition, the magnitudes of the estimates are larger when attention is measured using the time individuals actually spend engaging with the ad (which we measure using eye-tracking). *Ad Dwell* explains an additional 5.2% of the variation in recall than *Ad Visible* – R^2 is 0.091 in Column (1) of Panel I and 0.143 in Panel II. This highlights the value of eye-tracking as a more direct measure of attention.

Table 3: Estimates of advertising effects on recall and purchase: OLS

	Recall ($\hat{\rho}$)					Purchase ($\hat{\lambda}$)				
	All	Device		News Type		All	Device		News Type	
		Mobile	Desktop	Hard	Soft		Mobile	Desktop	Hard	Soft
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Visible	0.003^{***}	0.003^{***}	0.004^{***}	0.004^{***}	0.002^{***}	0.001^{**}	0.002^{**}	0.001^{*}	0.001^{**}	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	5,707	2,495	3,212	3,154	2,553	5,707	2,495	3,212	3,154	2,553
R^2	0.091	0.103	0.120	0.102	0.096	0.130	0.164	0.147	0.147	0.147
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Dwell	0.034^{***}	0.028^{***}	0.036^{***}	0.041^{***}	0.029^{***}	0.007^{**}	0.009^{**}	0.008^{**}	0.009^{**}	0.005
	(0.004)	(0.006)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	3,925	1,824	2,101	2,165	1,760	3,925	1,824	2,101	2,165	1,760
R^2	0.143	0.133	0.188	0.167	0.139	0.136	0.200	0.153	0.168	0.159

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand x price (for purchase) fixed effects. Standard errors clustered at the individual level.

Column (6) of Table 3 reports estimates of purchase ad effectiveness, the effect of ad attention on the incentivized purchase behavior ($\hat{\lambda}$). Panel I shows that, if an article is visible for an extra second, purchase probability increases by 0.13 percentage points. An increase of one standard deviation of *Ad Visible* increases purchase probability by 2.26 percentage points (relative to the average purchase probability of 35%). If an ad is actually looked at for an extra second, the probability of purchase increases by 0.73 percentage points. An increase of one standard deviation in *Ad Dwell* leads to a 2.31 percentage points higher purchase probability. As in the case of recall, *Ad Dwell* has a better predictive power of the outcome measure than *Ad Visible* : R^2 increases from 0.130 in Column (6) of Panel I to 0.136 in Panel II.

News Content Type Columns (4-5) and (9-10) of Table 3 report the estimates of ρ and λ separately for ads that were randomly matched to "hard" and "soft" news articles. As discussed at the outset, industry practitioners are wary of advertising next to "hard news" articles because of the perceived negative effect on their brand. This should imply that, for "hard" news, we should see smaller or even negative estimates of ρ, λ. In contrast, we find that estimates of the recall and purchase ad effectiveness are qualitatively similar across news types. If anything, we find that the magnitudes of the estimates are slightly *higher* for ads shown next hard news. For instance, a one-second increase in *Ad Dwell* for ads next to hard news articles increases purchase probability by 0.9 percentage points, whereas a similar estimate for ads on soft articles is 0.5 percentage points.

In order to better understand the interaction between article content, ad effectiveness, and the total amount of ad attention, we regress our four attention variables (*Ad Dwell* , *Ad Visible* , *Article Dwell*, *Article Visible*) on an indicator of whether the article is classified as hard news. To keep the estimates consistent, we include the same controls as in Table 3. Further, to keep articles comparable, we control for their length by including the number of words as a control.[25] Results are shown in Table 4.

[25]This addresses the concern that "hard news" articles might systematically be longer or shorter than other articles, which would mechanically affect attention.

Table 4: Attention and Hard News

	Measure of attention:			
	Ad Visible	Ad Dwell	Article Visible	Article Dwell
	(1)	(2)	(3)	(4)
Hard News	-0.9049^{***}	-0.4902^{***}	-10.1627^{***}	-6.7510^{***}
	(0.3168)	(0.0828)	(2.6201)	(2.0077)
Observations	5,707	3,925	5,707	3,925
R^2	0.4187	0.1461	0.6355	0.4705

Note: *p<0.1; **p<0.05; ***p<0.01. Fixed Effects: Individual covariates (income, gender, education, age, politics), Step Order, Brand, Country x Device. Includes a quartic polynomial in total time an average page is visible for each individual. Includes a linear control for number of words in article. Standard errors clustered at the individual level.

Hard news articles, and the ads randomly shown next to these articles, receive less attention than other ads and articles. Individuals spend less time looking at the ads (Columns 1 and 2), and also less time looking at the article itself (Columns 3 and 4). In terms of *Article Dwell* , there is a reduction of almost 7 seconds (about 10% of the median), and a reduction of 0.49 seconds for *Ad Dwell* (about 15% of the median).

These results should be interpreted with caution. There were many hard news stories on the topics of COVID and BLM in the press at the time of the experiment (July 24 - August 6, 2020), so individuals could already be broadly informed about the topic in the articles we chose (the experiment did not allow testing for pre-experiment knowledge). Alternatively, individuals might have wearied of such stories. We cannot say whether our finding is due to participants disliking hard news or because they were shown articles on topics they were already aware of, resulting in quick skimming.

Importantly, even if we interpret the lower attention that consumers devote to hard news as causal, our results suggest that advertising next to hard news is still at least as effective as advertising next to soft news. To determine this, we combine the negative effects of hard news on the amount of attention readers devote to ads (Table 4) and the positive effect of incremental attention on purchases (Columns (9-10) in Part II of Table 3). From Table 1, the average attention to ads (irrespective of news content type) is 2.76 seconds. From Table 3, devoting this amount of attention to ads next to "soft news" articles increases the purchase probability by $0.5 \cdot 2.76 = 1.38$ percentage points. For hard articles, the same effect is $0.9 \cdot (2.76 - 0.49) = 2.04$ percentage points. Hard news articles on average induce 0.49 seconds less attention to ads but have higher (0.9 instead of 0.5) ad effectiveness per second of devoted attention. On balance, this implies that the benefits of advertising next to hard news are similar, if not higher, compared to advertising on soft news articles.

Device Type Columns (2-3) and (7-8) of Table 3 report $\hat{\rho}, \hat{\lambda}$ separately for consumers participating in the experiment from mobile devices and desktop computers. Across all subsamples, estimates of the purchase and recall ad effectiveness are nearly identical, validating the importance of advertising both on mobile and desktop devices.

Robustness We present three additional robustness checks of the estimates. First, previous work has documented attention fatigue and decay (Goldstein et al., 2011; Ahn et al., 2018). In Appendix Table 22, we test for this by allowing outcomes to be a function of a quadratic polynomial in attention in Equations (3) and (4). We indeed find diminishing returns to attention; for all specifications, the quadratic term on the advertising attention is negative. However, within our sample, the non-linear effects on our main outcome variable (incentivized purchase) are economically and statistically small.

Second, our estimates are similar (although less precise) in a more demanding specification that includes individual FE instead of individual-level covariates (Appendix Table 23). Finally, our results are also robust to using a logit specification instead of the linear probability model (Appendix Table 24).

6.4 Results: IV

Table 5 presents the estimates of ρ, λ from IV regressions of Equations (3)-(4), where we instrument for *Ad Visible* and *Ad Dwell* with *Article Dwell*.[26] Panel I presents the results with *Ad Visible* as the measure of attention to ads. The first stage results are presented in the bottom part of Panel I. For all specifications, we have strong instruments – incremental F-statistics vary from 31.9 to 79. The first stage regressions confirm strong positive attention spillovers between the article and ads, described in Table 2. The second stage IV estimates are presented at the top part of Panel I.

For the outcome of recall (Columns 1-5), the estimates $\hat{\rho}$ are too imprecise to conclude that they are different from the OLS estimates or zero. For instance, when we include all observations (Column 1), $\hat{\rho} = 0.0003$ is smaller than the OLS estimate of 0.003 – but the standard error is 0.002, making the difference statistically insignificant.

For the purchase outcome (Columns 6-10), the estimates of $\hat{\lambda}$ are positive and statistically significant – the spillover attention to ads due to an interesting article leads to a higher purchase probability of the brand that was advertised. The estimated magnitudes of $\hat{\lambda}$ are larger for the IV case – although differences between the IV and OLS estimates are only marginally significant, due to larger standard errors of the IV estimates. The fact that the IV estimate of $\hat{\lambda}$ is larger than the OLS estimate suggests that reverse causality is not a big concern in this case, since consumer behavior *à la* Becker and Murphy (1993); Tuchman et al. (2018) would lead to an upward bias of the OLS estimates (and we find the opposite).

Panel II of Table 5 presents the results with *Ad Dwell* as the measure of attention to ads. All conclusions are the same as in the case of *Ad Visible*. The first stage results confirm a strong complementarity between attention devoted to articles and ads. Incremental F-statistics are between 10 and 183 across specifications, with the strongest

[26] We cannot use *Article Visible* as an instrument since it is mechanically influenced by the attention to ads when both article and ad are visible on the page.

Table 5: Estimates of advertising effects on recall and purchase: Article Dwell IV

	Recall ($\hat{\rho}$)					Purchase ($\hat{\lambda}$)				
	All	Device		News Type		All	Device		News Type	
		Mobile	Desktop	Hard	Soft		Mobile	Desktop	Hard	Soft
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Visible	0.0003	0.008	−0.001	0.001	−0.0004	0.006***	0.013**	0.004*	0.007**	0.005*
	(0.002)	(0.007)	(0.002)	(0.003)	(0.003)	(0.002)	(0.006)	(0.002)	(0.003)	(0.002)
Observations	3,925	1,824	2,101	2,165	1,760	3,925	1,824	2,101	2,165	1,760
R^2	0.105	0.097	0.138	0.122	0.103	0.123	0.128	0.146	0.148	0.149
	First Stage									
Article Dwell	0.058***	0.024***	0.076***	0.057***	0.059***	0.058***	0.026***	0.075***	0.057***	0.059***
	(0.007)	(0.007)	(0.009)	(0.010)	(0.007)	(0.007)	(0.007)	(0.008)	(0.010)	(0.007)
Observations	3,925	1,824	2,101	2,165	1,760	3,925	1,824	2,101	2,165	1,760
R^2	0.459	0.282	0.566	0.430	0.516	0.467	0.307	0.582	0.447	0.531
1st Stage Incr. F-Stat	75.85	12.83	76.71	31.92	71.4	77.74	14.53	79.01	33.5	70.79
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Dwell	0.001	0.008	−0.009	0.005	−0.002	0.028**	0.015**	0.052	0.036**	0.025*
	(0.010)	(0.007)	(0.032)	(0.014)	(0.013)	(0.011)	(0.007)	(0.037)	(0.016)	(0.013)
Observations	3,925	1,824	2,101	2,165	1,760	3,925	1,824	2,101	2,165	1,760
R^2	0.107	0.122	0.117	0.130	0.100	0.119	0.199	0.079	0.145	0.143
	First Stage									
Article Dwell	0.011***	0.022***	0.005***	0.011***	0.011***	0.011***	0.023***	0.005***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	3,925	1,824	2,101	2,165	1,760	3,925	1,824	2,101	2,165	1,760
R^2	0.205	0.476	0.129	0.208	0.232	0.220	0.500	0.156	0.235	0.262
1st Stage Incr. F-Stat	48.23	173.37	11.96	34.63	34.09	48.52	183.4	10.41	36.78	34.52

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand x price (for purchase) fixed effects. Standard errors clustered at the individual level.

relationship for mobile and the weakest for desktop devices. The estimates $\hat{\rho}$ are imprecise across the specifications (Columns 1-5), while the estimates of $\hat{\lambda}$ are positive and statistically significant (Columns 6-10). The IV estimates of $\hat{\lambda}$ are larger than the OLS estimates, but the difference is not statistically significant.

In Columns (7-10) of Table 5, λ estimates are presented separately for different devices and news types, confirming findings from the OLS analysis. For both mobile and desktop devices, λ estimates show no statistical difference when using *Ad Visible* or *Ad Dwell* as attention metrics. Likewise, the impact of advertising on recall and purchase for "hard" and "soft" news is qualitatively similar. Furthermore, the estimated magnitudes are slightly *higher* for ads displayed alongside articles featuring hard news.

Robustness As an additional robustness check, we use as an instrumental variable the L1O article attention (the average amount of attention devoted to the article by all *other* individuals in the sample, as defined in Section 5.2) to instrument for attention to ads. From Table 2, the L1O attention significantly shifts the amount of time a consumer devotes to the article, which in turn has a positive spillover effect on the attention to ads on the page. We present these results in Appendix Table 25. Once again, the first stage estimates confirm the positive spillover of attention from articles to ads, although the strength of the instrument is weaker (e.g. for *Ad Dwell* as a measure of ad attention, incremental F-statistics vary from 1.8 to 13.3). Because of the lower statistical power of this instrument, the second stage (IV) coefficients are also estimated imprecisely although, reassuringly, they have the same magnitude as previous OLS and IV results.

Finally, we consider an alternative shifter of consumers' attention to articles: the (mis)alignment between consumers' and newspapers' political views. We construct a measure of political alignment of consumers and news outlets by asking participants about their political views. Independently, we classify news outlets as left, center, or right-wing. A misalignment strongly predicts consumers' attention to articles – going from fully aligned views to completely misaligned views decreases the time people read the article by around 15 seconds. This, in turn, decreases the attention people devote to ads on the page, with ads becoming visible for 1.26 seconds less (s.e. of 0.64 seconds) and attracting 0.22 seconds less active attention dwell time (s.e. of 0.17 seconds). The magnitudes match the previous results on attention spillovers well – e.g., the first stage results in Table 5 (Column 1) implies that an extra 15 seconds of *Article Dwell* increase *Ad Visible* and *Ad Dwell* by 0.87 and 0.17 seconds, respectively. However, the effect of political mismatch on the attention to ads is too imprecise to produce conclusive estimates of ρ, λ. We present details of the analysis and discuss the results in Appendix G.3.

7 Managerial Implications

Our paper has four sets of findings that lend themselves to managerial implications. First, attention to ads can be measured and leads to higher ad recall and brand purchases, providing a way to measure ad effectiveness and price display advertising. Second, attention to articles has a positive spillover to ads placed next to them, highlighting the value of high-quality content. Third, "hard news" article content does not make ads less effective, cautioning against the practice of blunt " block lists" of advertisers. Fourth, ad visibility is a more imprecise metric of consumer attention, but still a valuable one for researchers. We consider each of these in turn.

We can use our results to calculate rough estimates of the costs and benefits of online display ads. First, we discuss the benefits. In our experiment, the ads on each page had an average dwell time of about 2.76 seconds per individual (i.e., the time individuals are attentive to the ad). At the mean, this attention increases the probability of purchase by $2.76 \times 0.007 \approx 0.02$, or about 2%.[27] In the US, for instance, the opportunity cost to individuals of acquiring the voucher (the amount of cash individuals had to forgo, or the

[27] Here we use the OLS estimates from Table 3. The IV estimates would imply an even higher value of advertising.

"price" of the voucher) was on average \$5. Therefore, we take the revenue for the brand from purchase to be \$5. This implies that an ad is worth $5 \times 0.02 = 10$ cents of revenue per person exposed to the ad, or \$100 for 1,000 people.[28] We note that these estimates might overestimate advertising effectiveness of display advertising since individuals make an *immediate* purchase decision following exposure to ads, when information about the brands is more easily retrievable from the memory (Keller, 1987).

On the cost side, the advertising industry typically uses the metric of a "cost per mille" (CPM, or cost per 1,000 impressions). For a digital inventory, this is difficult to assess because it is the result of an auction every time an ad is available rather than the setting of a price in general. Things are further complicated because advertisers tend to pay for targeting information (e.g., to ensure that a particular ad is shown to individuals who, based on their known characteristics, are likely to be interested in the brand), which further influences the cost. Still, Lumen Research shared with us their estimate of the cost per *attentive* 1,000 views (aCPM), which is £21.88 ($\approx \$30$) on desktops and £13.54 ($\approx \$19$) on mobile devices. On top of this, we would have to include technology and agency fees – that is, the cost of creating the ads and employing marketers. However, on the whole, these figures suggest that advertising is likely worth its cost. The magnitudes of the implied return-on-investments are larger than those typically reported in the literature (e.g. Lewis and Reiley, 2014; Kireyev et al., 2016), likely due to the immediacy of the purchase decision of consumers in our context.

Our second set of results shows that there is a positive attention spillover from articles to ads. These results emphasize the value of good, captivating news content – not only does such content drive more visitors to news outlets and increase their reputation, it also increases the effectiveness of advertising on news outlets' web pages. Thus, by investing in the quality of news content, publishers can charge higher CPM rates to advertisers. These findings provide business justifications against the practice of "click-

[28] We are considering only revenue, not profit, since we have no estimate of the cost to the brand of producing and supplying the goods.

bait" (using catchy titles or images to entice users to visit low-quality articles that are then immediately skipped). Instead, the result suggests that publishers should be incentivized to invest in more captivating and high-quality news content, even when only considering ad revenue. Our results on the "political mismatch" between outlet and readers further corroborate this idea: newspapers that cater to their audiences attract valuable attention to the article that spills over to the ad.

The third managerial implication that arises from our results is a word of caution when it comes to block lists that often do not allow ads to be placed next to "hard news." In our experiment, these were articles associated with the COVID-19 pandemic and the BLM protests. Our results reject the hypothesis of a negative effect of hard news *per se* on either ad recall or brand purchase. There may still be other reasons, such as brand safety (marketingweek.com, 2017), or preferences and career concerns of brand managers (Gordon et al., 2021), to limit exposure of ads to certain types of content. However, our results suggest that the current system might be too blunt or exhibit excessive risk aversion. Limiting the practice of block lists is particularly important at times of major societal events – e.g. pandemics, wars, and the fight against climate change – since block lists penalize news outlets for providing detailed coverage of these important issues and informing citizens.

Our last set of results concerns alternative ways of measuring attention. We used two metrics of attention to articles and ads, visibility and actual dwell time. Dwell time is a more precise and accurate measure, as it measures the amount of time a person actually looks at webpage objects. However, to produce that measure, one needs access to (costly) eye-tracking software. We found that the simpler measure, visibility, still produces reliable results when measuring the impact of attention on brand recall and purchase. This is reassuring and has important repercussions. Depending on the question at hand, research teams without access to eye-tracking software can still obtain robust answers by using non-eye-tracking-based measures of attention to ads, such as the time ads are visible on the page.

Our results validate the importance of the attention of website visitors for display advertising effectiveness, which can be “priced in” by the publishers and platforms. This view is aligned with the current thinking in the media industry. For instance, Mail Metro Media, which represents several UK’s media brands (such as The Daily Mail and The Telegraph) created a “high attention” package of advertising, for which they charge a price premium to brands (dmgmedia.co.uk, 2021). A similar program is run by Ozone, the aggregated selling house used by The Guardian. Again, they charge a premium on their advertising inventory which is justified in part by the higher attention their ads receive because of the intensity of the engagement with the content (ozoneproject.com, 2021). This does not seem to be only a sell-side or online news phenomenon. Havas, one of the biggest media buying networks in the world, has adopted an explicit position that it will pay more for the quality of attention an ad receives (The Media Leader, 2022). Like our method, these pricing strategies and measures of ad effectiveness benefit from a novel approach of leveraging the intensive margin of attention to ads, rather than an extensive margin of showing or not showing an ad on the page. McGranaghan et al. (2022) discuss similar strategies for incorporating attention metrics into the measures of ad effectiveness and pricing for TV ads.

8 Conclusions

This paper has used measures of attention obtained with eye-tracking to estimate advertising effectiveness in online markets. We run an experiment that focuses on display advertising online, in which ads are shown next to articles. We showed that more engaging articles generate positive spillovers of attention from the news to the ads. This incremental ad attention increases the probability that the advertised brands are correctly recalled and subsequently purchased.

There are several important caveats to keep in mind regarding the external validity of our results, typical for similar experimental settings. First, we asked individuals to make an *immediate* purchase decision, so we are likely overestimating the effect a real ad would have on purchases. We note, however, that the brand-specific vouchers that individuals could obtain were valid for 1 year or more, so *consumption* does not need to be immediate, and hence possibly mitigates this bias.

Second, we may be underestimating the impact of ads, since our ads are not targeted to specific individuals. We relied on the representativeness of the panel selected by a specialist supplier of research and marketing panels, and we chose brands that are of sufficient appeal to large audiences. We cannot estimate the effectiveness of targeted ads (and this was not the goal of our experiment) – this would require access to an algorithm that assigns ads to readers online, which we do not possess.

Notwithstanding these limitations, we hope that this work will prompt more research on the drivers and effects of digital attention, including an extended model of the links between attention and recall and a more detailed investigation into the underlying mechanisms. These models can be motivated and informed by the recent literature in neuromarketing and neuroscience; e.g. Plassmann et al. (2012) discuss recent advances in the literature and suggest avenues of theory generation that build on consumer neuroscience. Tools such as eye-tracking software are now increasingly precise and available at scale in realistic settings to measure this.

References

Jae-Hyeon Ahn, Yoon-Soo Bae, Jaehyeon Ju, and Wonseok Oh. Attention adjustment, renewal, and equilibrium seeking in online search: An eye-tracking approach. Journal of Management Information Systems, 35(4):1218–1250, 2018.

Joshua David Angrist, Guido W Imbens, and Alan B Krueger. Jackknife instrumental variables estimation. Journal of Applied Econometrics, 14(1):57–67, 1999.

Anocha Aribarg, Rik Pieters, and Michel Wedel. Raising the bar: Bias adjustment of recognition tests in advertising. Journal of Marketing Research, 47(3):387–400, 2010.

K Carrie Armel, Aurelie Beaumel, and Antonio Rangel. Biasing simple choices by manipulating relative visual attention. Judgment and Decision Making, 3(5):396–403, 2008.

Gary S Becker and Kevin M Murphy. A simple theory of advertising as a good or bad. The Quarterly Journal of Economics, 108(4):941–964, 1993.

Jonah Berger, Wendy W Moe, and David Schweidel. What leads to longer reads? psychological drivers of reading online content. ACR North American Advances, 2019.

S Adam Brasel and James Gips. Breaking through fast-forwarding: Brand information and visual attention. Journal of Marketing, 72(6):31–48, 2008.

Isabelle Brocas, Juan D Carrillo, Stephanie W Wang, and Colin F Camerer. Imperfect choice or imperfect attention? understanding strategic thinking in private information games. Review of Economic Studies, 81(3):944–970, 2014.

Bobby J Calder, Edward C Malthouse, and Ute Schaedel. An experimental study of the relationship between online engagement and advertising effectiveness. Journal of Interactive Marketing, 23(4):321–331, 2009.

Colin F Camerer, Eric Johnson, Talia Rymon, and Sankar Sen. Cognition and framing in sequential bargaining for gains and losses. Frontiers of Game Theory, 104:27–47, 1993.

Moran Cerf, Jonathan Harel, Wolfgang Einhäuser, and Christof Koch. Predicting human gaze using low-level saliency combined with face detection. Advances in Neural Information Processing Systems, 20, 2007.

Pierre Chandon, J Wesley Hutchinson, Eric T Bradlow, and Scott H Young. Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. Journal of Marketing, 73(6):1–17, 2009.

ComScore. Revisited: Media consumption during the coronavirus pandemic. https://www.comscore.com/Insights/Blog/Revisited-Media-Consumption-during-the-Coronavirus-Pandemic [Acccessed: January 17, 2023], 2020.

Gordon B Dahl, Andreas Ravndal Kostøl, and Magne Mogstad. Family welfare cultures. The Quarterly Journal of Economics, 129(4):1711–1752, 2014.

Peter J Danaher and Guy W Mullarkey. Factors affecting online advertising recall: A study of students. Journal of Advertising Research, 43(3):252–267, 2003.

digiday.com. Coronavirus climbs up keyword block lists, squeezing news publishers' programmatic revenues. https://digiday.com/media/coronavirus-climbs-keyword-block-lists-squeezing-news-publishers-programmatic-revenues/ [Acccessed: January 17, 2023], 2020.

dmgmedia.co.uk. Mail metro media unveils centre for attention. https://www.dmgmedia.co.uk/news/mail-metro-media-unveils-centre-for-attention/ [Acccessed: January 31, 2023], 2021.

Will Dobbie, Jacob Goldin, and Crystal S Yang. The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. American Economic Review, 108(2):201–40, 2018.

Xavier Drèze and François-Xavier Hussherr. Internet advertising: Is anybody watching? Journal of Interactive Marketing, 17(4):8–23, 2003.

Millie Elsen, Rik Pieters, and Michel Wedel. Thin slice impressions: how advertising evaluation depends on exposure duration. Journal of Marketing Research, 53(4):563–579, 2016.

emarketer.com. Estimated short-term change in average digital ad spending due to the coronavirus pandemic. https://www.emarketer.com/chart/235710/estimated-short-term-change-average-digital-ad-spending-due-coronavirus-pandemic-according-us-agency-brand-ad-buyers-by-format-march-2020-change [Accessed: January 17, 2023], 2020.

Minou Ghaffari and Susann Fiedler. The power of attention: Using eye gaze to predict other-regarding and moral choices. Psychological Science, 29(11):1878–1889, 2018.

Avi Goldfarb and Catherine Tucker. Online display advertising: Targeting and obtrusiveness. Marketing Science, 30(3):389–404, 2011.

Daniel G Goldstein, R Preston McAfee, and Siddharth Suri. The effects of exposure time on memory of display advertisements. Proceedings of the 12th ACM Conference on Electronic Commerce, pages 49–58, 2011.

Daniel G Goldstein, R Preston McAfee, and Siddharth Suri. Improving the effectiveness of time-based display advertising. ACM Transactions on Economics and Computation (TEAC), 3(2):1–20, 2015.

Brett R Gordon, Kinshuk Jerath, Zsolt Katona, Sridhar Narayanan, Jiwoong Shin, and Kenneth C Wilbur. Inefficiencies in digital advertising markets. Journal of Marketing, 85(1):7–25, 2021.

Brett R Gordon, Robert Moakler, and Florian Zettelmeyer. Close enough? a large-scale exploration of non-experimental approaches to advertising measurement. Marketing Science, 42(4):768–793, 2023.

James J Heckman and James M Snyder Jr. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. The Rand Journal of Economics, 28:S142, 1997.

Emily Higgins, Mallorie Leinenger, and Keith Rayner. Eye movements when viewing advertisements. Frontiers in Psychology, 5:210, 2014.

Paul R Hoban and Randolph E Bucklin. Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment. Journal of Marketing Research, 52(3):375–393, 2015.

Kenneth Holmqvist, Jana Holsanova, Mari Barthelson, and Daniel Lundqvist. Reading or scanning? a study of newspaper and net paper reading. In The Mind's Eye, pages 657–670. Elsevier, 2003.

Wayne D Hoyer and Steven P Brown. Effects of brand awareness on choice for a common, repeat-purchase product. Journal of Consumer Research, 17(2):141–148, 1990.

Jason Huang, David Reiley, and Nick Riabov. Measuring consumer sensitivity to audio advertising: A field experiment on pandora internet radio. Available at SSRN 3166676, 2018.

IAB UK. Covid-19 x keyword blocking: 9 tips for advertisers. https://www.iabuk.com/news-article/covid-19-x-keyword-blocking-9-tips-advertisers [Acccessed: January 17, 2023], 2020.

Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254–1259, 1998.

Chris Janiszewski. The influence of display characteristics on visual exploratory search behavior. Journal of Consumer Research, 25(3):290–301, 1998.

Przemysław Jeziorski and Sridhar Moorthy. Advertiser prominence effects in search advertising. Management Science, 64(3):1365–1383, 2018.

Garrett Johnson, Randall A Lewis, and Elmar Nubbemeyer. The online display ad effectiveness funnel & carryover: Lessons from 432 field experiments. Available at SSRN 2701578, 2017a.

Garrett A Johnson, Randall A Lewis, and Elmar I Nubbemeyer. Ghost ads: Improving the economics of measuring online ad effectiveness. Journal of Marketing Research, 54 (6):867–884, 2017b.

Ulrich Kaiser and Julian Wright. Price structure in two-sided markets: Evidence from the magazine industry. International Journal of Industrial Organization, 24(1):1–28, 2006.

Gabriella Kazai, Iskander Yusof, and Daoud Clarke. Personalised news and blog recommendations based on user location, facebook and twitter user profiling. In Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, pages 1129–1132, 2016.

Kevin Lane Keller. Memory factors in advertising: The effect of advertising retrieval cues on brand evaluations. Journal of Consumer Research, 14(3):316–333, 1987.

Mehreen Khurram, Faisal Qadeer, and Muhammad Sheeraz. The role of brand recall, brand recognition and price consciousness in understanding actual purchase. Journal of Research in Social Sciences, 6(2):219–241, 2018.

Max Kilger and Ellen Romer. Do measures of media engagement correlate with product purchase likelihood? Journal of Advertising Research, 47(3):313–325, 2007.

Pavel Kireyev, Koen Pauwels, and Sunil Gupta. Do display ads influence search? attribution and dynamics in online advertising. International Journal of Research in Marketing, 33(3):475–490, 2016.

Daniel T Knoepfle, Joseph Tao-yi Wang, and Colin F Camerer. Studying learning in games using eye-tracking. Journal of the European Economic Association, 7(2-3):388–398, 2009.

Christof Koch and Shimon Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. In Matters of Intelligence, pages 115–141. Springer, 1987.

Michal Kolesar. Estimation in an instrumental variables model with treatment effect heterogeneity. Unpublished Working Paper, 2013.

Dmitry Lagun and Mounia Lalmas. Understanding user attention and engagement in online news reading. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, pages 113–122, 2016.

JooWon Lee and Jae-Hyeon Ahn. Attention to banner ads and their effectiveness: An eye-tracking approach. International Journal of Electronic Commerce, 17(1):119–137, 2012.

Randall A Lewis and David H Reiley. Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on yahoo! Quantitative Marketing and Economics, 12(3):235–266, 2014.

Xiau Liu, Matthew Shum, and Kosuke Ueteke. Passive and active attention to baseball telecasts: Implications for content (re-)design. Available on SSRN (https://ssrn.com/abstract=3717894), 2021.

Gerald L Lohse. Consumer eye movement patterns on yellow pages advertising. Journal of Advertising, 26(1):61–73, 1997.

John G Lynch Jr, Howard Marmorstein, and Michael F Weigold. Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. Journal of Consumer Research, 15(2):169–184, 1988.

Emma K Macdonald and Byron M Sharp. Brand awareness effects on consumer decision making for a common, repeat purchase product: A replication. Journal of Business Research, 48(1):5–15, 2000.

Scott B MacKenzie. The role of attention in mediating the effect of advertising on attribute importance. Journal of Consumer Research, 13(2):174–195, 1986.

marketingweek.com. Should brands be pulling digital ad spend over "funding terror" claims? https://www.marketingweek.com/brands-pulling-digital-ad-spend-terror-funding-claims/ [Acccessed: January 17, 2023], 2017.

José Martins, Catarina Costa, Tiago Oliveira, Ramiro Gonçalves, and Frederico Branco. How smartphone advertising influences consumers' purchase intention. Journal of Business Research, 94:378–387, 2019.

Matthew McGranaghan, Jura Liaukonyte, and Kenneth C Wilbur. How viewer tuning, presence, and attention respond to ad content and predict brand search lift. Marketing Science, 2022.

Martin Meißner, Andres Musalem, and Joel Huber. Eye tracking reveals processes that enable conjoint choices to become increasingly efficient with practice. Journal of Marketing Research, 53(1):1–17, 2016.

Milica Milosavljevic and Moran Cerf. First attention then intention: Insights from computational neuroscience of vision. International Journal of Advertising, 27(3):381–398, 2008.

Sridhar Narayanan and Kirthi Kalyanam. Position effects in search advertising and their moderators: A regression discontinuity approach. Marketing Science, 34(3):388–407, 2015.

Vidhya Navalpakkam, Justin Rao, and Malcolm Slaney. Using gaze patterns to study and predict reading struggles due to distraction. In CHI'11 Extended Abstracts on Human Factors in Computing Systems, pages 1705–1710. 2011.

Howard Kenneth Nixon. Attention and interest in advertising. Number 72. 1924.

ozoneproject.com. Ozone study shows clear link between attention paid to quality content and accompanying ads. https://www.ozoneproject.com/blog/ozone-study-shows-clear-link-between-attention-paid-to-quality-content-and-accompanying-ads [Acccessed: January 31, 2023], 2021.

Philip Pärnamets, Petter Johansson, Lars Hall, Christian Balkenius, Michael J Spivey, and Daniel C Richardson. Biasing moral decisions by exploiting the dynamics of eye gaze. Proceedings of the National Academy of Sciences, 112(13):4170–4175, 2015.

Rik Pieters and Michel Wedel. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. Journal of Marketing, 68(2):36–50, 2004.

Rik Pieters and Michel Wedel. Goal control of attention to advertising: The yarbus implication. Journal of Consumer Research, 34(2):224–233, 2007.

Rik Pieters, Michel Wedel, and Jie Zhang. Optimal feature advertising design under competitive clutter. Management Science, 53(11):1815–1828, 2007.

Rik Pieters, Michel Wedel, and Rajeev Batra. The stopping power of advertising: Measures and effects of visual complexity. Journal of Marketing, 74(5):48–60, 2010.

Emily Pitler and Ani Nenkova. Revisiting readability: A unified framework for predicting text quality. In Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, pages 186–195, 2008.

Hilke Plassmann, Thomas Zoëga Ramsøy, and Milica Milosavljevic. Branding the brain: A critical review and outlook. Journal of Consumer Psychology, 22(1):18–36, 2012.

Keith Rayner, Caren M Rotello, Andrew J Stewart, Jessica Keir, and Susan A Duffy. Integrating text and pictorial information: eye movements when looking at print advertisements. Journal of Experimental Psychology: Applied, 7(3):219, 2001.

Cleo D Redline and Christopher P Lankford. Eye-movement analysis: a new tool for evaluating the design of visually administered instruments (paper and web). 2001.

Elena Reutskaja, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel. Search dynamics in consumer choice under time pressure: An eye-tracking study. American Economic Review, 101(2):900–926, 2011.

J Edward Russo and France Leclerc. An eye-fixation analysis of choice processes for consumer nondurables. Journal of Consumer Research, 21(2):274–290, 1994.

Oliver J Rutz, Randolph E Bucklin, and Garrett P Sonnier. A latent instrumental variables approach to modeling keyword conversion in paid search advertising. Journal of Marketing Research, 49(3):306–319, 2012.

Navdeep S Sahni. Effect of temporal spacing between advertising exposures: Evidence from online field experiments. Quantitative Marketing and Economics, 13(3):203–247, 2015.

Noel Scott, Christine Green, and Sheranne Fairley. Investigation of the use of eye tracking to examine tourism advertising effectiveness. Current Issues in Tourism, 19(7):634–642, 2016.

Savannah Wei Shi and Michael Trusov. The path to click: Are you on it? Marketing Science, 40(2):344–365, 2021.

Jaana Simola, Jarmo Kuisma, Anssi Öörni, Liisa Uusitalo, and Jukka Hyönä. The impact of salient advertisements on reading and attention on web pages. Journal of Experimental Psychology: Applied, 17(2):174, 2011.

Andrey Simonov and Shawndra Hill. Competitive advertising on brand search: Traffic stealing and click quality. Marketing Science, 40(5):923–945, 2021.

Andrey Simonov, Chris Nosko, and Justin M Rao. Competition and crowd-out for brand keywords in sponsored search. Marketing Science, 37(2):200–215, 2018.

Statista. Online advertising revenue in the united states from 2000 to 2021. https://www.statista.com/statistics/183816/us-online-advertising-revenue-since-2000/ [Acccessed: January 17, 2023], 2023.

Iréne Stenfors, Jan Morén, and Christian Balkenius. Behavioural strategies in web interaction: A view from eye-movement research. In The Mind's Eye, pages 633–644. Elsevier, 2003.

Thales S Teixeira, Michel Wedel, and Rik Pieters. Moment-to-moment optimal branding in tv commercials: Preventing avoidance by pulsing. Marketing Science, 29(5):783–804, 2010.

The Guardian. Uk publishers losing digital ad revenue due to content 'blacklists'. https://www.theguardian.com/media/2020/jan/20/uk-publishers-losing-digital-ad-revenue-due-to-content-blacklists [Acccessed: January 17, 2023], 2020.

The Media Leader. Havas: news, science and sports sites get most user engagement. https://the-media-leader.com/havas-news-science-and-sports-sites-get-most-user-engagement/ [Acccessed: January 31, 2023], 2022.

Joan Treistman and John P Gregg. Visual, verbal, and sales responses to print ads. Journal of Advertising Research, 19(4):41–47, 1979.

Anna E Tuchman, Harikesh S Nair, and Pedro M Gardete. Television ad-skipping, consumption complementarities and the consumer demand for advertising. Quantitative Marketing and Economics, 16(2):111–174, 2018.

Christina Uhl, Nadia Abou Nabout, and Klaus Miller. How much ad viewability is enough? the effect of display ad viewability on advertising effectiveness. arXiv preprint arXiv:2008.12132, 2020.

Ralf van der Lans, Rik Pieters, and Michel Wedel. Online advertising suppresses visual competition during planned purchases. Journal of Consumer Research, 48(3):374–393, 2021.

M Wedel and R Pieters. A review of eye-tracking research in marketing. Review of Marketing Research, Volume 4, pages 123–146, 2007.

Michel Wedel. Attention research in marketing: A review of eye-tracking studies. 2015.

Michel Wedel and Rik Pieters. Eye fixations on advertisements and memory for brands: A model and findings. Marketing Science, 19(4):297–312, 2000.

Michel Wedel, Rik Pieters, and John Liechty. Attention switching during scene perception: how goals influence the time course of eye movements across advertisements. Journal of Experimental Psychology: Applied, 14(2):129, 2008.

Kenneth C Wilbur. A two-sided, empirical model of television advertising and viewing markets. Marketing Science, 27(3):356–378, 2008.

Kenneth C Wilbur. Advertising content and television advertising avoidance. Journal of Media Economics, 29(2):51–72, 2016.

Rick T Wilson, Daniel W Baack, and Brian D Till. Creativity, attention and the memory for brands: an outdoor advertising field study. International Journal of Advertising, 34(2):232–261, 2015.

Shunyao Yan, Klaus M Miller, and Bernd Skiera. How does the adoption of ad-blockers affect news consumption? Available at SSRN, 2020.

Jie Zhang, Michel Wedel, and Rik Pieters. Sales effects of attention to feature advertisements: A bayesian mediation analysis. Journal of Marketing Research, 46(5):669–681, 2009.

Xuebai Zhang and Shyan-Ming Yuan. An eye tracking analysis for video advertising: Relationship between advertisement elements and effectiveness. IEEE access, 6:10699–10707, 2018.

[FOR ONLINE PUBLICATION]

Web Appendix

A Experimental Details

A.1 Branded Ads

We displayed ads for general-interest brands, rather than specific products, chosen for broad appeal and easy redemption with e-vouchers. These brands were selected based on availability of brand-specific vouchers, purchased from GiftPay (https://www.giftpay.co.uk/) and Tango Card (https://www.tangocard.com/). We also ensured similarity in brand categories between the two countries. The table below lists the selected brands.

Type of brand/Country	US	UK
Coffee shop	Starbucks	Starbucks
Coffee shop	Dunkin' Donuts	Costa
Clothing	Banana Republic	Primark
Clothing	GAP	H&M
Food	Domino's Pizza	Pizza Express
Food	Burger King	Wagamama
Bath products	Bath & Body Works	The Body Shop
DIY/Home improvement	Home Depot	B&Q

A.2 Articles

We report below the headlines of the articles that were chosen, split by country and by newspaper. We indicate with an asterisk (*) those articles that we classified as hard news. We provide the URL to retrieve the full article (click on the headline).

The following articles were sourced from *The New York Times* (US):

Trump Aides Undercut Fauci as He Speaks Up on Virus Concerns*

Qualified Immunity Protection for Police Emerges as Flash Point Amid Protests*

Technology Bridges the Gap to Better Sight

What if the U.S. Bans TikTok?

The following articles were sourced from *USA Today* (US):

CDC adds runny nose, nausea to the growing list of COVID-19 symptoms*

'I thought this was a hoax': Patient, 30, dies after attending 'COVID party,' doctor says*

California officer under investigation for allegedly sharing 'vulgar image' of George Floyd; NAACP San Diego calls for his firing*

Johnny Depp accuses Amber Heard of hitting him with 'roundhouse punch' near end of their marriage

Pour by phone: Coca-Cola introduces contactless technology to pour your beverage

The following articles were sourced from *The Guardian* (UK):

NHS data reveals 'huge variation' in Covid-19 death rates across England*

Boris Johnson says face masks should be worn in shops in England*

Police apologise to woman told to cover up anti-Boris Johnson T-shirt*

Johnny Depp tells high court libel case how he lost $650m in earnings

How we met: 'It's 1,300 miles to Romania – the same as the number of pounds my phone bill was'

The following articles were sourced from the *Daily Mail* (UK):

People living in England's poorest areas are TWICE as likely to die of coronavirus than those in the wealthiest neighbourhoods, statistics show*

Two-thirds of Britons back Boris Johnson's refusal to 'take the knee' because people should not be 'bullied' into making 'gestures'*

Scooby Who? Great Dane's popularity falls to its lowest level in 50 years after peaking in the 1980s thanks to the Scooby Doo TV series

Are you a victim of 'batterygate?' Users with older iPhones may be eligible for a $25 settlement if their device was covertly slowed by the tech giant

A.3 Eye-tracking Technology

Details of the eye-tracking technology are summarized in the top panel of Figure 5. No hardware is needed. The software is Javascript code which is entirely removed from the participant's device after completion of the experiment. Before an eye-tracking session is started, the user is taken through a calibration procedure. During this procedure, the eye-tracker measures characteristics of the user's eyes and uses them together with an anatomical 3D eye model to calculate the gaze data. During the calibration, the user is asked to look at specific points on the screen (calibration dots). The first calibration dot appears in the middle of the screen and then sequentially moves between four corners of the screen, covering the entire screen perimeter in the end. Several images of the eyes are collected and analyzed. The resulting information is then integrated into the eye model and the gaze point for each image sample is calculated. When the procedure is finished, the calibration process is illustrated by green lines of varying length (see the lower panel of Figure 5 for an example).

During the calibration procedure at the beginning of the experiment (twice), as well as at two other points in the study – after the third and sixth articles – users do another validation procedure where the software measures the reliability of the data collected by eye-tracking. During these validation steps, the user is asked to look at five specific dots on the screen sequentially. Three main metrics determine the reliability of the eye-tracking data: "accuracy", "precision", and "gaze duration" (which are standard industry metrics of eye-tracking quality; see section C for definitions). If on average per respondent one of these metrics climbs above the pre-specified thresholds (for instance, 300 CSS pixels for accuracy and precision, and 100 milliseconds for gaze dura-

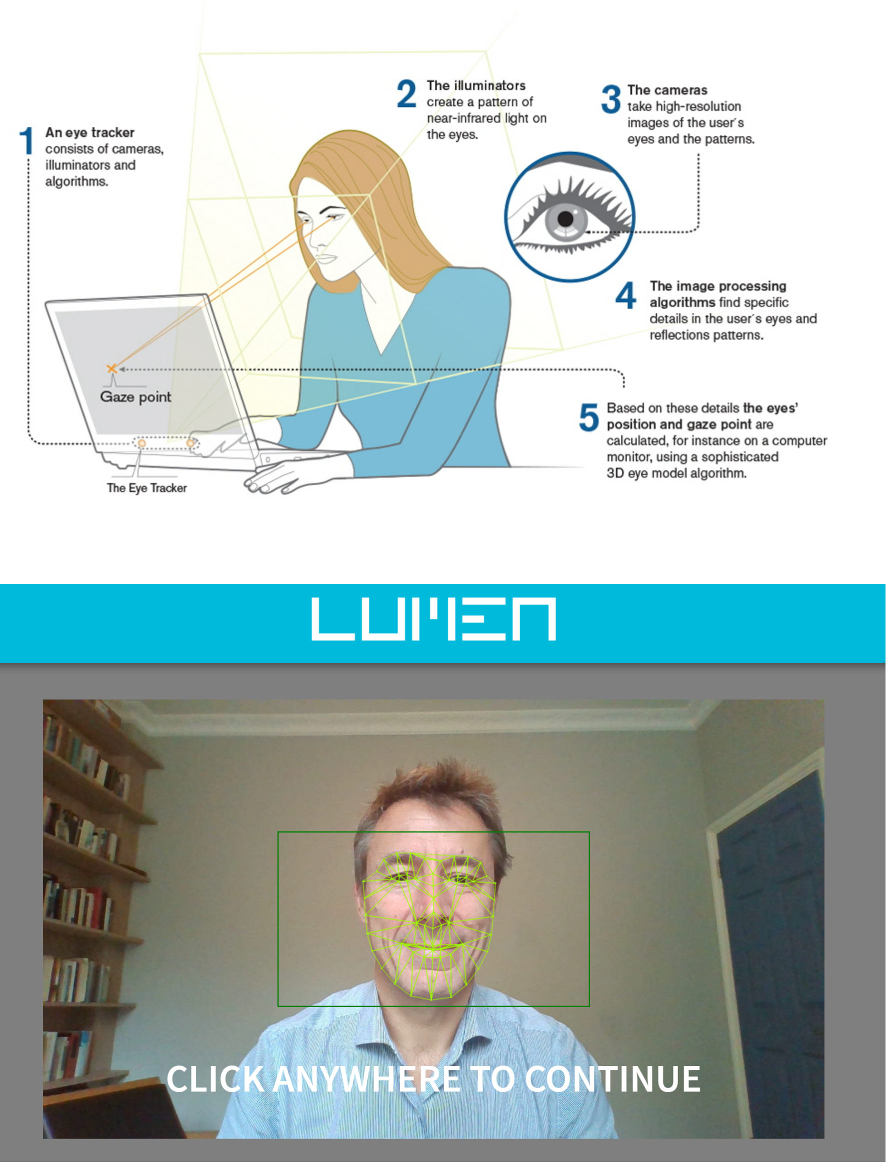
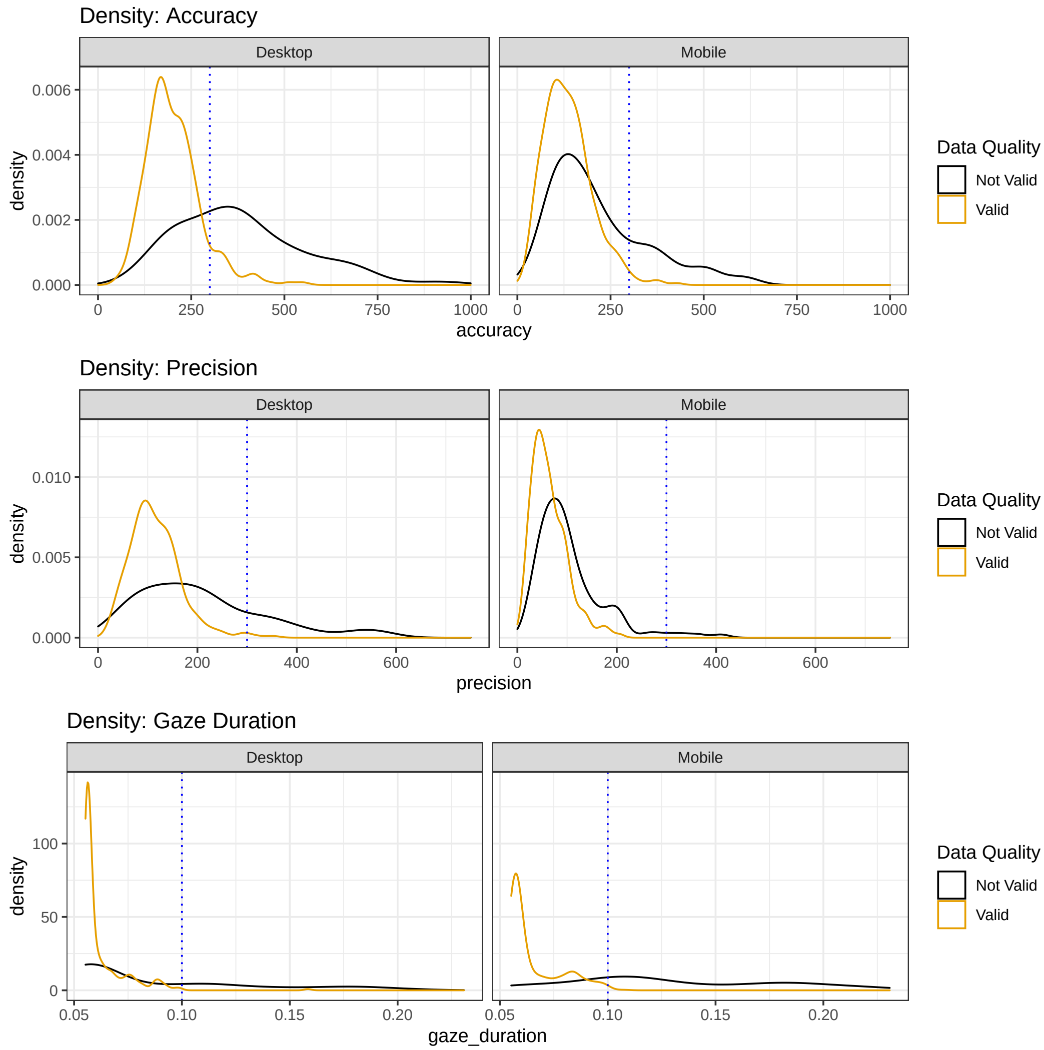


Figure 5: Eye-tracking technology

tion) the eye-tracking data in the articles before and after the validation step is considered "invalid" and is not used for the analysis. For our main sample, we rely on rules commonly used by the provider of eye-tracking technology (Lumen Research) for flagging participants for whom the data is deemed invalid. We then confirm the robustness of our results to alternative eye-tracking quality metrics.

Figure 6 presents the densities of accuracy, precision, and gaze duration metrics for data deemed valid and invalid by Lumen Research. Figures on the left correspond to

Figure 6: Densities of precision, accuracy, and gaze duration



Values of accuracy, precision, and gaze duration are computed as an average across the available validation steps per respondent. Values of accuracy and precision are in CSS pixels. Values of gaze duration are in seconds. The blue dotted line corresponds to thresholds above which Lumen typically flags data as invalid for analysis.

desktop devices, and figures on the right to mobile devices. For desktop devices, respondents flagged by Lumen Research as invalid have distributions systematically to the right of the respondents labeled as valid, confirming low-quality data across all three metrics. A substantial share of users deemed to have low-quality eye-tracking data tend to have values of accuracy and gaze duration higher than the thresholds, and some have precision higher than the threshold. In contrast, the vast majority of users flagged as valid have average values of accuracy, precision, and gaze duration below the thresholds.

Similarly, on mobile devices, users who were deemed invalid tend to have worse metrics of accuracy, precision, and gaze duration compared to the respondents who are flagged as valid. For mobile users, gaze duration is the primary metric that crosses the quality threshold and disqualifies users from the analysis.

Comparing precision metrics across device types, we can see that eye-tracking can capture gazes more accurately and precisely on mobile devices – on average, accuracy for respondents flagged as valid on mobile devices is 137, compared to 202 on desktops, and precision is 64 on mobile devices, compared to 115 on desktops. However, if we normalize the metrics by devices' screen sizes, relative accuracy and precision on mobile devices are approximately the same as on desktops.

Table 6: Average Eye-Tracking Metrics across Devices for Valid Respondents

Device	Accuracy	Precision	Gaze_Duration
Desktop	201.75	115.13	0.06
Mobile	136.66	64.47	0.07

Values are computed as an average across the available validation steps per valid respondent. Values of accuracy and precision are in CSS pixels, and in seconds for gaze duration.

Table 7 presents the quality metrics of eye-tracking data across validation steps. As the study progresses, the accuracy and precision of eye-tracking data decrease. Average accuracy drops from 128 to 240, and the average precision drops from 51 to 144. This is explained by the fact that even though respondents were instructed to sit still, eventually, as respondents read articles, they changed their position. We check the robustness of our main results by accounting for the measurement error induced by this noise in Section C.

Table 7: Average Eye-Tracking Metrics across Validation Steps for Valid Respondents

Validation Steps	Accuracy	Precision	Gaze_Duration
1	127.56	51.47	0.06
2	182.96	102.67	0.07
3	207.67	132.12	0.07
4	240.21	144.11	0.07

Values are computed as an average with the available validation steps per valid respondent. Values of accuracy and precision are in CSS pixels, and in seconds for gaze duration.

A.4 Ethics Approval

The protocol received ethical approval from the Imperial College Research Ethics Committee (ICREC) and the Science Engineering Technology Research Ethics Committee (SETREC); SETREC reference: 20IC6104. The study was approved by SETREC on 12/06/20 and by the Joint Research Compliance Office on 19/06/20. The study was registered in the AEA RCT Registry as RCT ID AEARCTR-0006010.

B Validation: Hard News and Political Slant

In this section, we detail the validation process for our assessment of hard news and political slant. A survey was administered on Amazon Mechanical Turk (AMT), involving 250 participants from the UK and another 250 from the US. All participants utilized their desktops for survey participation.

Within each country, participants were instructed to read the same articles employed in the original experiment, presented in desktop format. To mitigate potential fatigue effects, each individual was assigned to read a random subset of 4 articles, resulting in approximately 2,000 observations.

Articles presented to participants were devoid of ads, and individuals were prompted to express their opinions on each article along three dimensions. First, participants rated the "upsetting" nature of the article on a Likert scale from 1 to 5. Second, they assessed whether the article was "interesting," again on a Likert scale from 1 to 5. Third, participants indicated their perception of the political slant of each article as Left, Neutral, or Right. The political slant of each article was then computed by assigning a score of +1, 0, or -1, corresponding to a participant's perception of a right-wing, neutral, or left-wing slant, respectively.

Subsequently, we calculated the average responses from the AMT survey to determine the extent to which each article was perceived as upsetting (hard news), interesting, and right-wing slanted. The outcomes of the AMT survey are presented in Table 8.

Table 8 shows, for each article, the corresponding newspaper and its classification as hard news. Additionally, the table presents the mean and standard deviation of responses from the AMT survey, capturing the perceived levels of upsetting, right-wing, and interesting aspects of each article. The variables for "upsetting" and "interesting" articles have been rescaled to a range between 0 and 1.

Table 9 displays a regression analysis of the "upsetting" scores given by participants in the AMT survey, specifically examining their correlation with our subjective definition

Table 8: Validation Summary Statistics

Newspaper	Hard	E[Upset]	SD[Upset]	E[Rightwing]	SD[Rightwing]	E[Interest]	SD[Interest]
Guardian	FALSE	0.304	0.268	-0.121	0.600	0.539	0.268
Guardian	FALSE	0.394	0.291	0.111	0.708	0.521	0.249
Guardian	TRUE	0.218	0.278	0.039	0.572	0.480	0.220
Guardian	TRUE	0.243	0.293	-0.418	0.731	0.583	0.237
Guardian	TRUE	0.481	0.304	-0.194	0.623	0.588	0.238
Mail	FALSE	0.183	0.258	-0.114	0.689	0.526	0.274
Mail	FALSE	0.180	0.287	0.130	0.626	0.426	0.296
Mail	TRUE	0.478	0.290	-0.237	0.788	0.573	0.240
Mail	TRUE	0.276	0.287	0.331	0.838	0.450	0.276
NYT	FALSE	0.223	0.266	-0.214	0.717	0.500	0.307
NYT	FALSE	0.156	0.263	-0.077	0.478	0.680	0.283
NYT	TRUE	0.502	0.313	-0.542	0.608	0.609	0.296
NYT	TRUE	0.401	0.313	-0.225	0.825	0.522	0.281
USAT	FALSE	0.135	0.265	-0.061	0.551	0.653	0.264
USAT	FALSE	0.384	0.315	-0.148	0.656	0.503	0.305
USAT	TRUE	0.408	0.303	-0.074	0.581	0.595	0.266
USAT	TRUE	0.571	0.328	-0.291	0.734	0.579	0.297
USAT	TRUE	0.605	0.312	-0.250	0.638	0.652	0.295

of hard news. The observations are at the article-respondent pair level, and the positive coefficient validates our subjective definition.

Table 9: Validation of Hard News Measure from AMT survey

	Dependent variable:
	Article Upsetting (0-1)
Hard News	0.1804***
	(0.0139)
Constant	0.2360***
	(0.0103)
Observations	2,016
R^2	0.0774

Note: *p<0.1; **p<0.05; ***p<0.01. One observation per article-respondent pair.

A similar procedure was employed to assess the political leaning of each newspaper (where right-wingness takes values from -1 to 1). Table 10 reveals that, in the US, *USA Today* is perceived as more right-wing compared to *The New York Times* (omitted). Similarly, in the UK, the *Daily Mail* is considered more right-wing than *The Guardian* (omitted). Notably, the "political distance" between *The Guardian* and the *Daily Mail* is greater than the distance between *The New York Times* and *USA Today*. These findings align with our data coding (for coding, see Appendix G.3). Importantly, these results remain consistent when respondent FE are included.

Table 10: Validation of Political Leaning

	Article is Right-wing (from -1 to 1)	
	US	UK
	(1)	(2)
USAT	0.1121**	
	(0.0489)	
Mail		0.2130***
		(0.0621)
Constant	−0.2881***	−0.1634***
	(0.0331)	(0.0430)
Observations	778	588
R^2	0.0067	0.0197

Note: *p<0.1; **p<0.05; ***p<0.01. One observation per article-respondent pair.

C Robustness of Eye-Tracking Measurements

C.1 Definition of eye-tracking quality metrics

In this section, we address the potential measurement error in eye-tracking. Throughout the study, consumers undergo multiple validation procedures (initially, and after the third and sixth articles) to assess the quality of the eye-tracking data and to re-calibrate the eye-tracking software. During each validation, individuals are asked to focus on a moving point on the screen, and their eye gaze is measured (see Appendix A.3).

We primarily consider three measures of eye-tracking data quality, extensively used by the provider of the eye-tracking software (Lumen Research) and in the literature (e.g. Semmelmann and Weigelt, 2018; Schneegans et al., 2021; Yang and Krajbich, 2021). First, ”accuracy”: the average Euclidean distance between the instructed gaze point on the screen and the recorded gaze points. Second, ”precision”: the re-scaled standard deviation of gaze points around the instructed gaze point. Third, ”gaze duration”: a measure of how frequently the camera records eye movement. Note that these are *inverse* measures of eye-tracking quality (higher values imply lower quality).

C.2 Validation of eye-tracking quality metrics

We assess eye-tracking data quality using the three measures described above (accuracy, precision and gaze duration). If, for a given respondent×article, the eye-tracking data is deemed “invalid,” we drop the eye-tracking data for that observation, as discussed in Section 4 and further illustrated below. For these observations, we maintain the visibility data, since that does not require eye-tracking. For our main analysis, we rely on Lumen’s algorithm to identify invalid observations. We validate that participants identified through this procedure indeed exhibit lower-quality eye-tracking measures in Appendix A.3 and establish robustness with respect to all these measures below.

For individuals with valid eye-tracking data, we verify the high quality of the data. These individuals exhibit an average accuracy of 201 and 137 CSS pixels (on desktop and mobile, respectively) and an average precision of 115 and 65 CSS pixels (on desktop and mobile devices, respectively).[29] These accuracy and precision values, while low compared to the ad sizes in our study, are sufficient to capture respondents' gazes within the interior of ads for all types except mobile billboard ads, which have minimal contribution to total ad dwell, as shown in Figure 15 in Section E.3.

C.3 Eye-tracking quality and sample balance

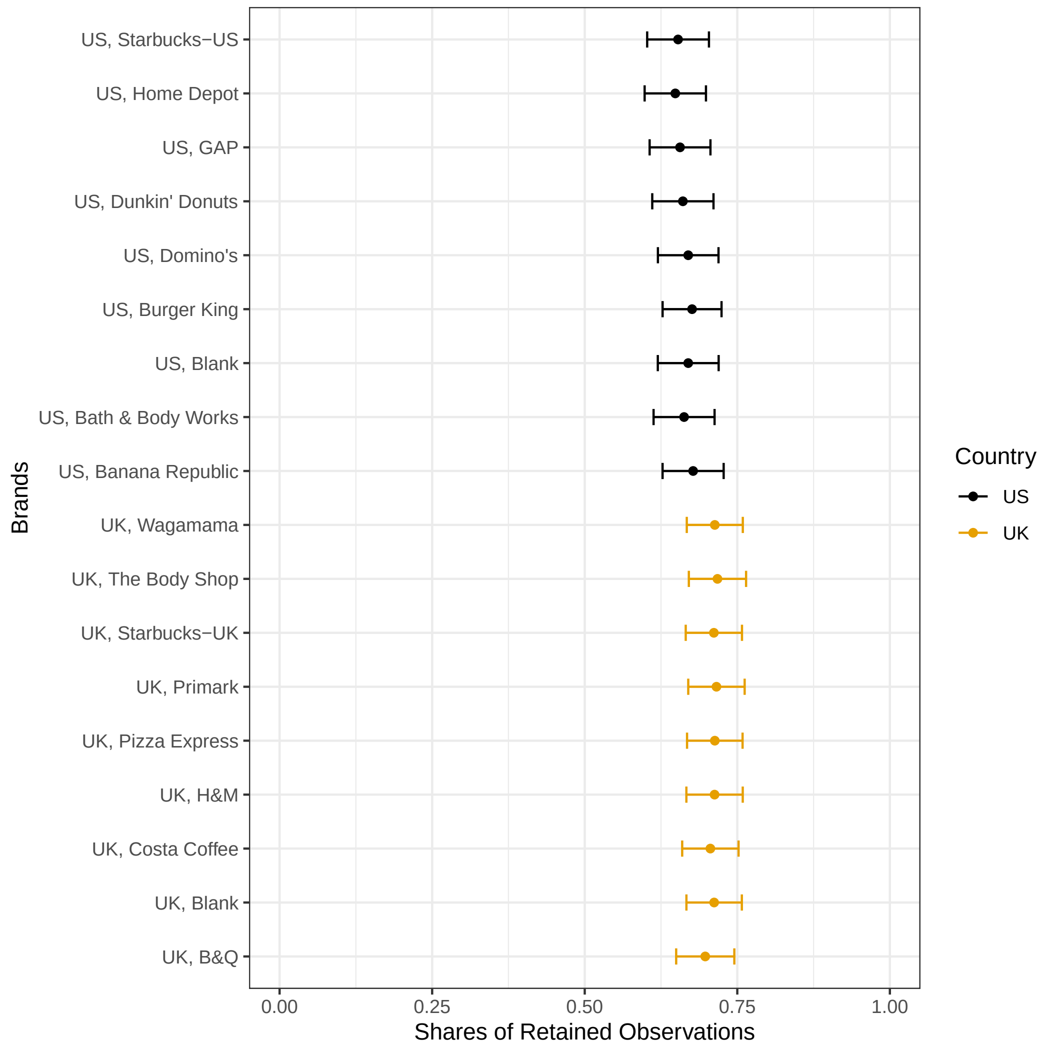
In Table 11, we conduct a comparison between two sub-samples: individuals with valid eye-tracking data and those with low-quality eye-tracking data, for whom we use only visibility data. This situation arises primarily when accuracy, precision, and gaze duration frequency measures are excessively high, often resulting from excessive head movement, as detailed in the main text and Appendix A.3.

Our findings indicate that individuals with valid eye-tracking data are less likely to be on desktop and exhibit slightly lower *Ad Visible* (but not lower *Article Visible*). Importantly, apart from these differences, the two subsamples demonstrate balance on other observable characteristics.

We further compare the balance of the observations with reliable and unreliable eye-tracking data across articles and brands. Since the order of articles and the ads was random, observations should be balanced along both dimensions after we account for the country and device type. Figures 7 and 8 confirm this balance. Figure 9 further shows that eye-tracking issues were slightly more pronounced later on in the study, although differences are not statistically significant.

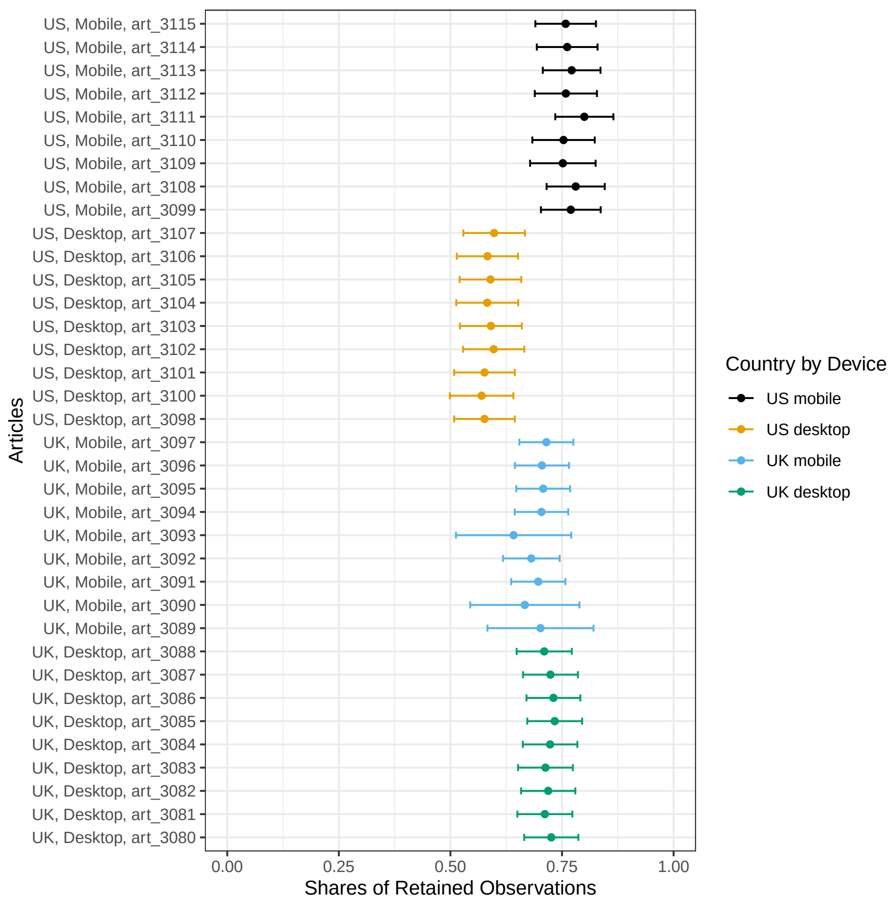
[29] A CSS pixel is a metric used in web browsers to ensure that web objects consistently occupy the same proportion of the screen, regardless of the device's physical pixel density. See https://www.w3.org/Style/Examples/007/units.en.html for details.

Figure 7: Balance of Observations per Brand Retained in the Study due to the Eye-Tracking Issues



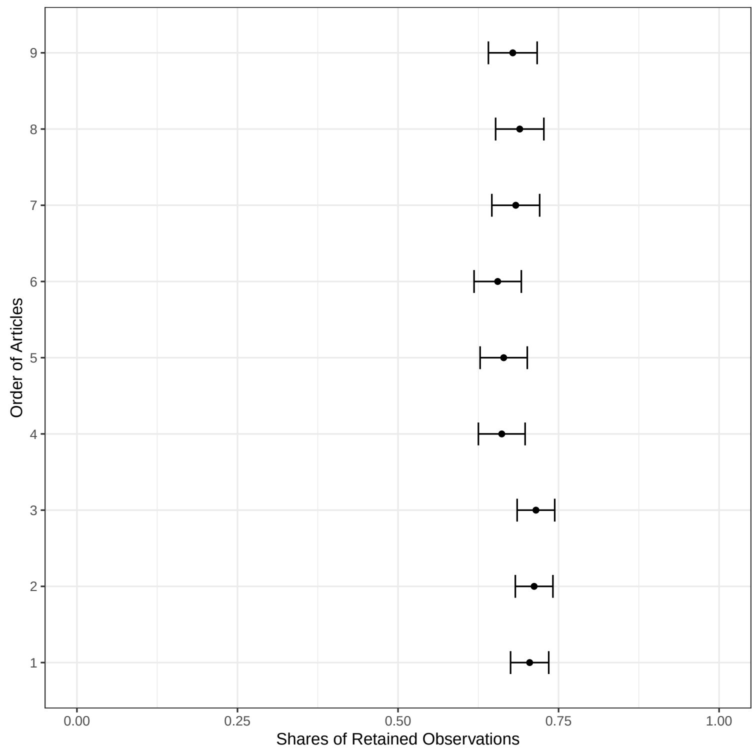
Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

Figure 8: Balance of Observations per Article Retained in the Study due to the Eye-Tracking Issues



Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

Figure 9: Balance of Observations per Step-order Retained in the Study due to Eye-Tracking Issues



Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

Table 11: Individuals with invalid and valid eye-tracking data.

	Valid		Invalid	
Variables	N	Mean	N	Mean
Desktop	708	0.41	296	0.57
		(0.02)		(0.03)
Female	708	0.58	296	0.52
		(0.02)		(0.03)
U.S.	708	0.44	296	0.5
		(0.02)		(0.03)
Hard News	708	0.55	296	0.56
		(0.01)		(0.01)
Article Visible (s)	708	142.11	296	155.59
		(4.67)		(8.46)
Ad Visible (s)	708	17.51	296	20.51
		(0.48)		(0.89)
Price (USD/GBP)	708	5.01	296	5.06
		(0.03)		(0.04)
Recall	708	0.46	296	0.47
		(0.01)		(0.02)
Buy	708	0.34	296	0.35
		(0.01)		(0.02)
Article Dwell (s)	708	76.32	0	NaN
		(2.87)		(NA)
Ad Dwell (s)	708	2.79	0	NaN
		(0.09)		(NA)
All observations	708	0.46	296	0.47
		(0.02)		(0.03)

Standard errors are in brackets. One observation per individual.

C.4 Robustness of OLS and IV results

To ensure that our results are not influenced by imprecision in our *Ad Dwell* measure, Tables 12 and 13 reiterate our main analysis, re-weighting observations by the (inverse of the) three eye-tracking data quality measures described above, along with two additional measures: the "hit rate," representing the share of gaze points within 200 CSS pixels from the calibration point during validation, and the inverse of the time since the study's beginning, accounting for the deterioration in eye-tracking metrics as the study progresses (see Appendix Table 7). Each observation is assigned the quality metrics from the last validation step before the article is shown.

Moreover, we use raw eye-tracking data to create alternative *Ad Dwell* measures, minimizing potential measurement error. In the first alternative, we consider only gazes within the "interior" of the ad, excluding borders of 25% of the ad's height and width. For the second adjusted *Ad Dwell* measure, we exclude gazes on the 50% of the ad's surface closest to the main text. Both measures are rescaled to align their means with the original *Ad Dwell*, facilitating the comparability of coefficient magnitudes. These adjusted attention variables alleviate concerns regarding the potential influence of peripheral attention. While eye-tracking is established as a measure of central attention (Holmqvist et al., 2003), there is some evidence suggesting it may underestimate the effect of peripheral attention, where individuals become aware of information even without directing their eye gaze towards it.

Table 12: Estimates of advertising effects on recall and purchase: OLS, Robustness in Ad Dwell Measurements

	Recall ($\hat{\rho}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.034***	0.017*	0.036***	0.033***	0.034***	0.029***	0.026***	0.026***
	(0.004)	(0.010)	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)
Observations	3,925	3,875	3,875	3,925	3,925	3,925	3,925	3,925
R^2	0.143	0.176	0.153	0.147	0.146	0.161	0.142	0.141
	Purchase ($\hat{\lambda}$)							
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.007**	0.007*	0.007**	0.006*	0.008**	0.005	0.005**	0.005**
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)
Observations	3,925	3,875	3,875	3,925	3,925	3,925	3,925	3,925
R^2	0.136	0.281	0.174	0.137	0.168	0.174	0.136	0.136

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand x price (for purchase) fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

Table 13: Estimates of advertising effects on recall and purchase: Article Dwell IV, Robustness in Ad Dwell Measurements

	Purchase ($\hat{\lambda}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.028**	0.017*	0.023**	0.028**	0.019*	0.022	0.030**	0.038**
	(0.011)	(0.010)	(0.011)	(0.012)	(0.011)	(0.014)	(0.012)	(0.015)
Observations	3,925	3,875	3,875	3,925	3,925	3,925	3,925	3,925
R^2	0.119	0.277	0.165	0.118	0.164	0.164	0.096	0.073
	First Stage							
Article Dwell	0.011***	0.018***	0.011***	0.011***	0.012***	0.011***	0.011***	0.009***
	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	3,925	3,875	3,875	3,925	3,925	3,925	3,925	3,925
R^2	0.220	0.691	0.240	0.217	0.234	0.272	0.144	0.128
1st Stage Incr. F-Stat	48.52	34.82	72.81	44.07	100.4	61.32	33.06	36.5

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand x price fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

Tables 12 and 13 replicate our main results with re-weighted samples and alternative measures of *Ad Dwell*. In Table 12 we replicate the main OLS estimates for *Ad Dwell* from Table 3, Columns (1) and (6) in Panel II. We report these estimates in Column (1) of Table 12 to make comparisons of estimates easier. We find that the magnitude of coefficients is the same across all seven alternative specifications. For the recall outcome (Panel I), the estimates vary from 0.017 to 0.036, but none are significantly different from the main estimate of 0.034 (considering the standard errors). For the purchase outcome (Panel II), the estimates vary from 0.005 to 0.008, but again none are significantly different from the main estimate of 0.007.

In Table 13 we replicate the main IV estimates for *Ad Dwell* for the purchase outcome from Table 5 (since we obtained statistically significant estimates only on the purchase outcome). Once again, we report the main estimates in Column (1) of Table 13 to make comparisons easier. Both the IV and the first stage estimates are not statistically different from our main results. IV estimates vary from 0.017 to 0.038 across specifications, whereas in Section 6.4 we obtained an estimate of 0.028. Similarly, the first-stage estimates vary from 0.009 to 0.018, whereas in Section 6.4 the estimate was 0.011. The stability of the first stage coefficients is especially re-assuring, since these coefficients correspond to the $\hat{\gamma}$ estimate, the attention spillover coefficient from Table 2. Thus, we show that the complementarity in consumers' attention to articles and ads is not an artifact of the measurement error, and instead captures true spillover in attention.

C.5 Robustness by Device Type

Tables 14–17 replicate the robustness checks separately for desktop and mobile devices.

Table 14: Estimates of advertising effects on recall and purchase: OLS, Robustness in Ad Dwell Measurements, Mobile

	Recall ($\hat{\rho}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.028***	0.007	0.032***	0.026***	0.030***	0.027***	0.026***	0.027***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)
Observations	1,824	1,785	1,785	1,824	1,824	1,824	1,824	1,824
R^2	0.133	0.239	0.144	0.140	0.144	0.163	0.141	0.143
	Purchase ($\hat{\lambda}$)							
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.009**	0.005	0.011***	0.009**	0.012**	0.006	0.010***	0.008**
	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)	(0.006)	(0.003)	(0.004)
Observations	1,824	1,785	1,785	1,824	1,824	1,824	1,824	1,824
R^2	0.200	0.383	0.241	0.201	0.222	0.252	0.202	0.200

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand x price (for purchase) fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

Table 15: Estimates of advertising effects on recall and purchase: OLS, Robustness in Ad Dwell Measurements, Desktop

	Recall ($\hat{\rho}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.036***	0.037***	0.036***	0.036***	0.033***	0.026***	0.024***	0.023***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.003)	(0.003)
Observations	2,101	2,090	2,090	2,101	2,101	2,101	2,101	2,101
R^2	0.188	0.186	0.205	0.192	0.188	0.209	0.179	0.177
	Purchase ($\hat{\lambda}$)							
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.008**	0.012***	0.008*	0.007*	0.009*	0.009	0.003	0.005
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)
Observations	2,101	2,090	2,090	2,101	2,101	2,101	2,101	2,101
R^2	0.153	0.211	0.190	0.155	0.197	0.218	0.151	0.152

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand by price (for purchase) fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

Table 16: Estimates of advertising effects on recall and purchase: Article Dwell IV, Robustness in Ad Dwell Measurements, Mobile

	Purchase ($\hat{\lambda}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.015**	0.009	0.015*	0.014**	0.013	0.006	0.018**	0.019**
	(0.007)	(0.010)	(0.009)	(0.007)	(0.008)	(0.012)	(0.008)	(0.009)
Observations	1,824	1,785	1,785	1,824	1,824	1,824	1,824	1,824
R^2	0.199	0.382	0.241	0.201	0.221	0.252	0.198	0.195
	First Stage							
Article Dwell	0.023***	0.026***	0.020***	0.023***	0.020***	0.019***	0.018***	0.018***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)
Observations	1,824	1,785	1,785	1,824	1,824	1,824	1,824	1,824
R^2	0.500	0.843	0.474	0.503	0.463	0.497	0.269	0.291
1st Stage Incr. F-Stat	183.4	90.43	155.16	169.06	233.08	158.62	114.08	48.22

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand by price fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

Table 17: Estimates of advertising effects on recall and purchase: Article Dwell IV, Robustness in Ad Dwell Measurements, Desktop

	Purchase ($\hat{\lambda}$)							
	Main	Re-weighted observations by					Adjusted Ad Dwell	
		1/precision	1/accuracy	1/gaze duration	Hit Rate	1/time since calibr.	Interior	Away from Text
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ad Dwell	0.052	0.052	0.036	0.054	0.020	0.044	0.038	0.071
	(0.037)	(0.036)	(0.033)	(0.040)	(0.031)	(0.030)	(0.027)	(0.059)
Observations	2,101	2,090	2,090	2,101	2,101	2,101	2,101	2,101
R^2	0.079	0.154	0.160	0.072	0.193	0.168	0.065	−0.157
	First Stage							
Article Dwell	0.005***	0.005***	0.005***	0.005***	0.006***	0.007***	0.007***	0.004**
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Observations	2,101	2,090	2,090	2,101	2,101	2,101	2,101	2,101
R^2	0.156	0.178	0.186	0.158	0.194	0.246	0.128	0.111
1st Stage Incr. F-Stat	10.41	16.05	16	9.87	19.08	15.29	4.55	12

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand by price fixed effects. A few observations have missing values of precision and accuracy variables; we drop those observations from the analysis in specifications with 1/precision and 1/accuracy weights. Standard errors clustered at the individual level.

D Sample Balance: Connectivity

In this section we show that the data is balanced on the basis of which observations were retained due to connectivity issues.

In Table 18, we conduct a comparison between two sub-samples of individuals. The first sub-sample consists of individuals for whom we possess data on all the articles presented to them. In contrast, the second sub-sample comprises individuals for whom we only have a subset of articles due to connectivity issues. As outlined in the main text, an internet connectivity issue during the experiment may lead to some lost data for that particular individual. However, it is crucial to note that even when connectivity issues occurred, each individual was still exposed to all articles and made all choices; only some of that data was not recorded.

Table 18 highlights that individuals who did not encounter any connectivity issues are more likely to be on desktops, more likely to be in the US, exhibit slightly higher measures of *Ad Visible* (but not *Ad Dwell*), and are slightly more likely to recall the ad. Nevertheless, it is important to emphasize that the two sub-samples remain balanced on observables.

We further assess the balance between retained and missed observations for each article and advertised brand. Given the random order of articles and ads, a balanced distribution is expected along both dimensions, accounting for the country of origin and device type. The results are presented in Figures 10 and 11. Upon comparison within a country and device, observations exhibit well-balanced characteristics across brands and articles.

It is noteworthy that one exception arises concerning three articles on mobile phones in the UK (articles 3093, 3090, and 3089), all consistently experiencing a higher frequency of connectivity issues (approximately 20% of data retained) compared to other UK mobile articles (retaining around 70% of data). We can confirm that our results remain robust when excluding observations related to these specific articles from the

Table 18: Individuals with and without connectivity issues.

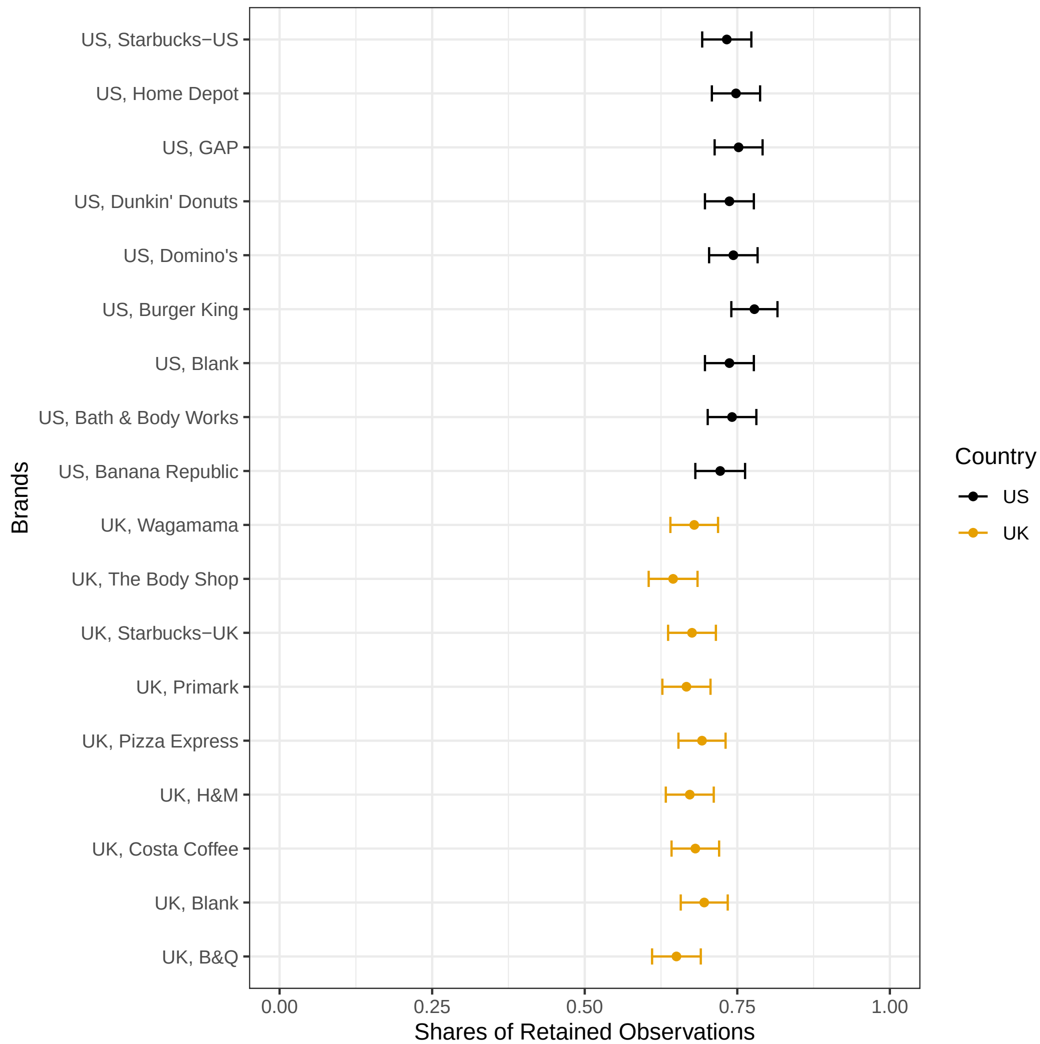
	With issues		Without issues	
Variables	N	Mean	N	Mean
Desktop	538	0.22	466	0.73
		(0.02)		(0.02)
Female	538	0.55	466	0.58
		(0.02)		(0.02)
U.S.	538	0.39	466	0.54
		(0.02)		(0.02)
Hard News	538	0.55	466	0.55
		(0.01)		(0)
Article Visible (s)	538	149.42	466	142.24
		(5.58)		(6.16)
Ad Visible (s)	538	16.78	466	20.26
		(0.58)		(0.62)
Price (USD/GBP)	538	5.03	466	5.02
		(0.04)		(0.02)
Recall	538	0.42	466	0.51
		(0.02)		(0.02)
Buy	538	0.33	466	0.35
		(0.01)		(0.01)
Article Dwell (s)	381	79.17	327	73
		(4.17)		(3.89)
Ad Dwell (s)	381	2.87	327	2.7
		(0.13)		(0.11)
All valid eyetracking	538	0.71	466	0.7
		(0.02)		(0.02)

Standard errors are in brackets. One observation per individual.

main analysis.

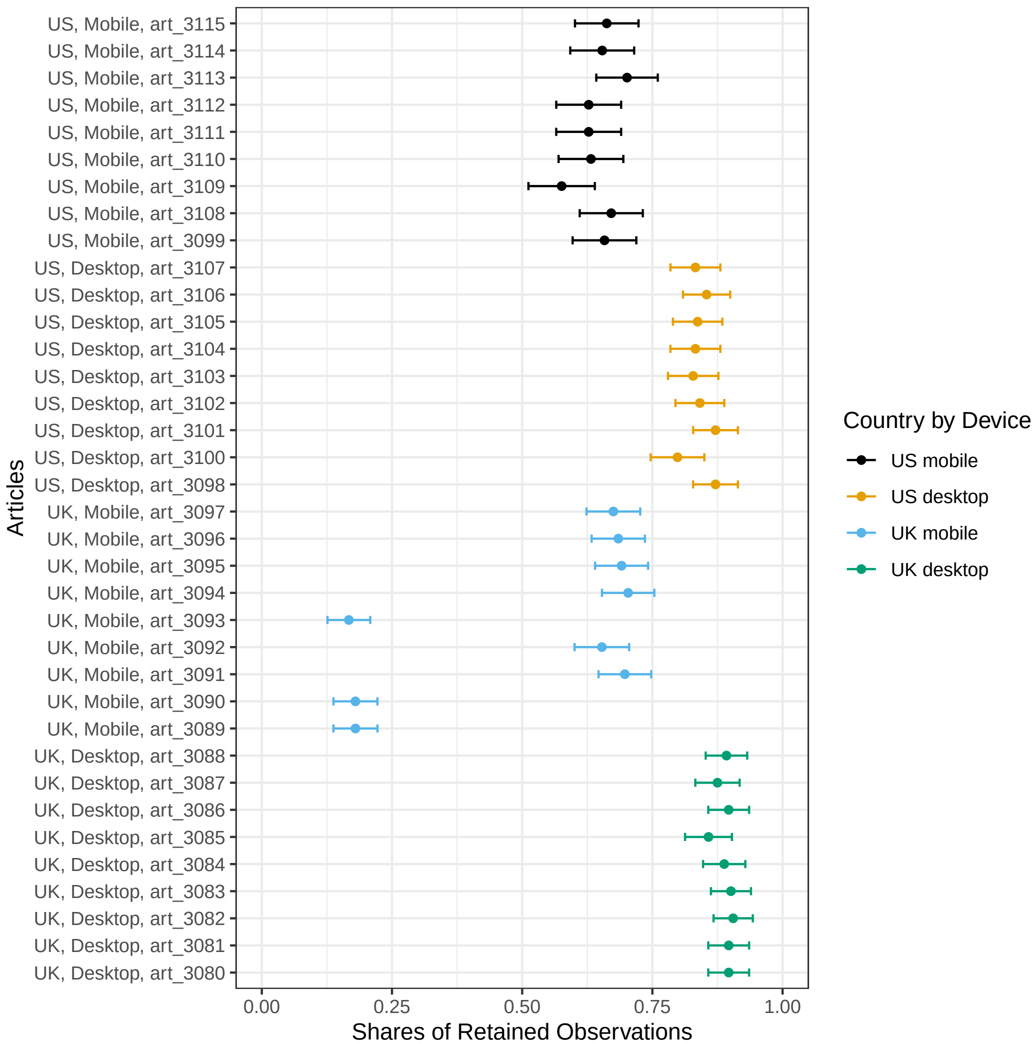
Figure 12 illustrates that connectivity issues became more pronounced as the study progressed. Specifically, about 90% of observations were retained during the readership of the first three articles (before the second validation step), approximately 62% of observations were retained before the third validation step (after the sixth article), and around 58% of observations were retained from the last three articles.

Figure 10: Balance of Observations per Brand Retained in the Study due to Connectivity Issues



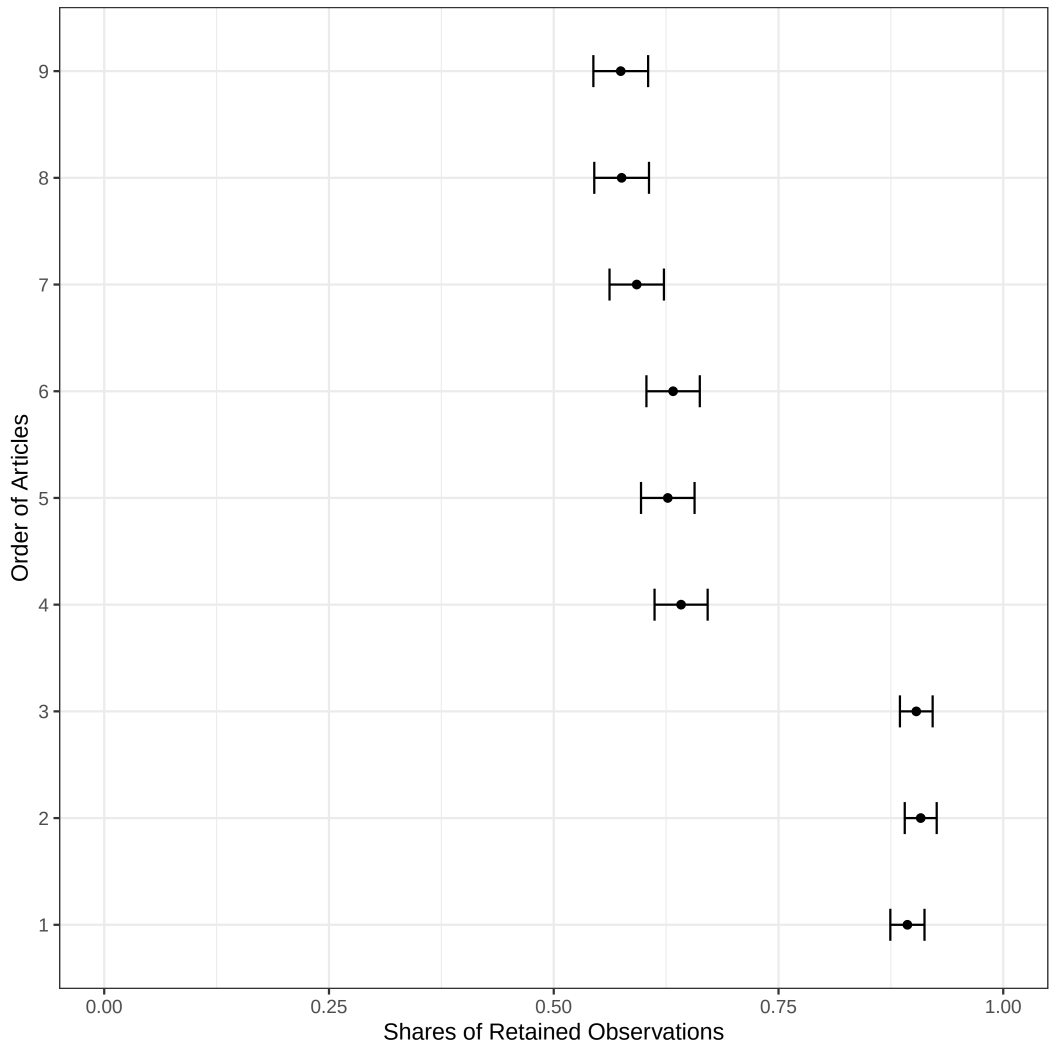
Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

Figure 11: Balance of Observations per Article Retained in the Study due to Connectivity Issues



Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

Figure 12: Balance of Observations per Step-order Retained in the Study due to Connectivity issues



Shares are computed directly from the data. Bars correspond to 95% confidence intervals.

E Additional Data Descriptives

In this section we provide additional descriptives of our raw data.

E.1 Summary Statistics By Device Type

Tables 19 and 20 show summary statistics for our sample by device type (in a way analogous to Table 1).

Table 19: Summary Statistics: Mobile

Statistic	N	Mean	St. Dev.	Min	Max
Desktop	2,810	0.000	0.000	0	0
Female	2,810	0.554	0.497	0	1
U.S.	2,810	0.478	0.500	0	1
Hard News	2,810	0.542	0.498	0	1
Article Visible (s)	2,810	144.957	171.568	20.130	1,894.635
Ad Visible (s)	2,495	13.436	12.608	0.000	291.905
Price (USD/GBP)	2,495	4.999	1.437	3.000	7.000
Recall	2,495	0.457	0.498	0.000	1.000
Buy	2,495	0.347	0.476	0.000	1.000
Article Dwell (s)	2,055	65.452	88.727	0.112	966.945
Ad Dwell (s)	1,824	2.690	3.134	0.000	40.214

Each observation is at the individual x article level.

Table 20: Summary Statistics: Desktop

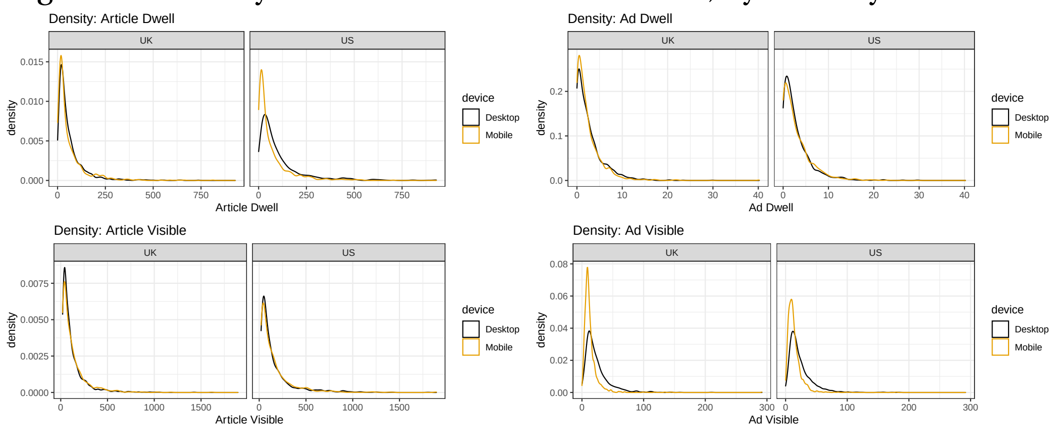
Statistic	N	Mean	St. Dev.	Min	Max
Desktop	3,621	1.000	0.000	1	1
Female	3,621	0.558	0.497	0	1
U.S.	3,621	0.487	0.500	0	1
Hard News	3,621	0.556	0.497	0	1
Article Visible (s)	3,621	142.015	167.606	22.599	1,805.904
Ad Visible (s)	3,212	23.370	19.226	0.000	210.103
Price (USD/GBP)	3,212	5.031	1.436	3.000	7.000
Recall	3,212	0.506	0.500	0.000	1.000
Buy	3,212	0.347	0.476	0.000	1.000
Article Dwell (s)	2,371	82.927	104.578	0.120	928.984
Ad Dwell (s)	2,101	2.810	3.184	0.000	29.895

Each observation is at the individual x article level.

E.2 Distribution of Attention

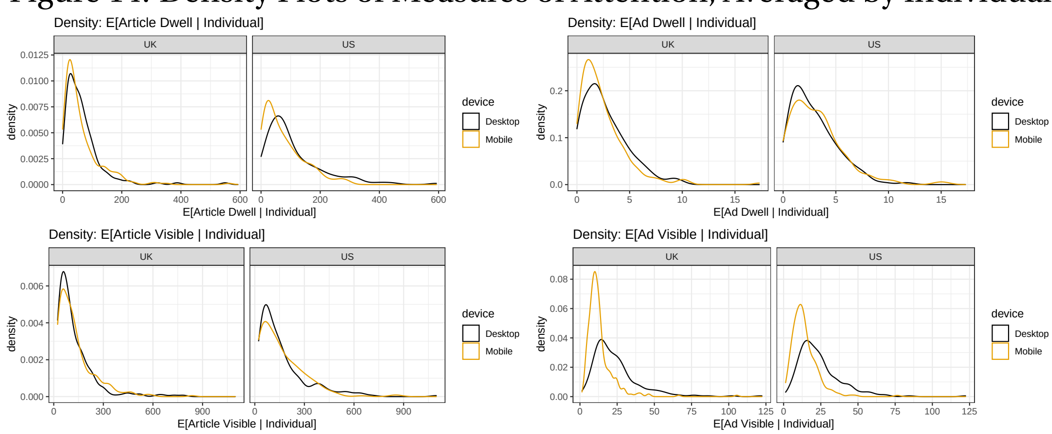
Figure 13 shows density plots of our attention measures, computed across all observations in the data. Figure 14 shows the same density plots, but averaged by individual.

Figure 13: Density Plots of Measures of Attention, by Country and Device



The plots show Article and Ad Dwell and Visible, computed across all observations in the data.

Figure 14: Density Plots of Measures of Attention, Averaged by Individual

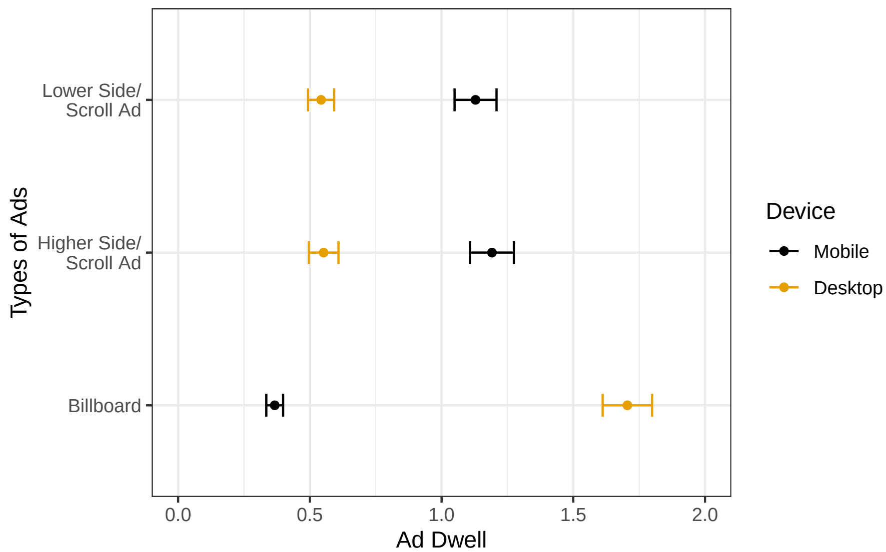


Density plots of measures of attention, averaged by individual and broken down by country and device.

E.3 Attention by Type of Ad

Figure 15 splits the attention paid to ads by their types – the "billboard" ad on top of the screen, the higher "side" located closer to the top of the page, and the lower "side" ad located further down the page. For mobile phones, "side" ads are shown in the center of the screen between paragraphs of text and therefore capture more of the consumer's attention than side ads on desktops. Consumers devote slightly less attention to the lower side ad on both types of devices. In contrast, billboard ads on desktops capture much more attention, because the wide format of desktop computers allows for a longer presence and share of the screen occupied by billboard ads.

Figure 15: Attention by Types of Ads



Attention to billboard, higher side, and lower side ads, by device type. Billboard ads receive more attention on desktop devices, but less attention on mobile phones. Bars correspond to 95% confidence intervals.

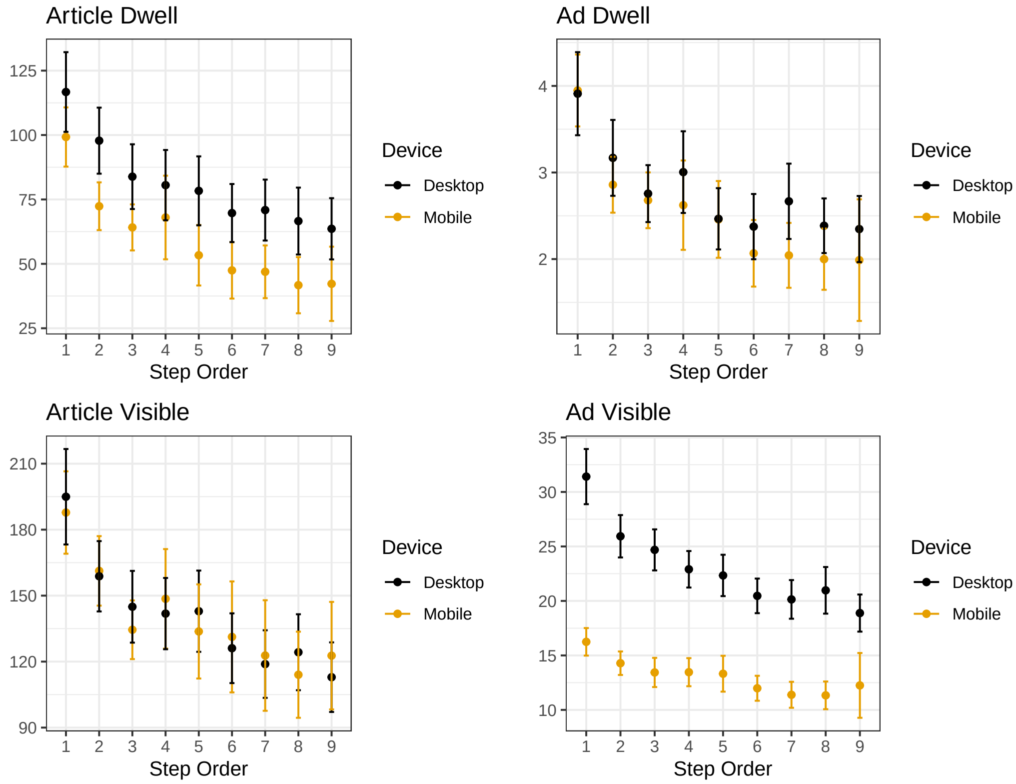
E.4 Attention by Experimental Step

Figure 16 illustrates how all metrics of attention are decreasing over the duration of the experiment (i.e., as a function of the experimental "step"), split by device type.

E.5 Correlation between Attention Metrics

Figure 17 shows the correlation between attention to article (*Article Dwell*) and attention to ad (*Ad Dwell*), for each article in our sample. Recall that we consider the same content shown on different devices as different articles, since the information is displayed in a significantly different way. Figure 18 is analogous to Figure 17, but shows attention residualized using the same FE as in our main empirical specification (e.g. Table 3).

Figure 16: Attention to Articles and Ads is Decreasing Throughout the Study



Attention devoted to articles and ads as a function of the "step" at which they are shown in the experiment. "Step-order" refers to the order in the experiment in which an article and ad were shown. Bars correspond to 95% confidence intervals.

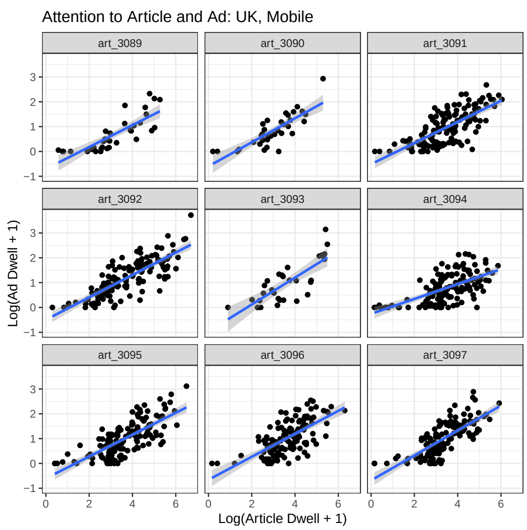
E.6 Demand Curves

Figures 19 and 20 show the relationship between price and probability of purchase (i.e., demand curves), first averaged for all brands in a country, and then for each brand individually.

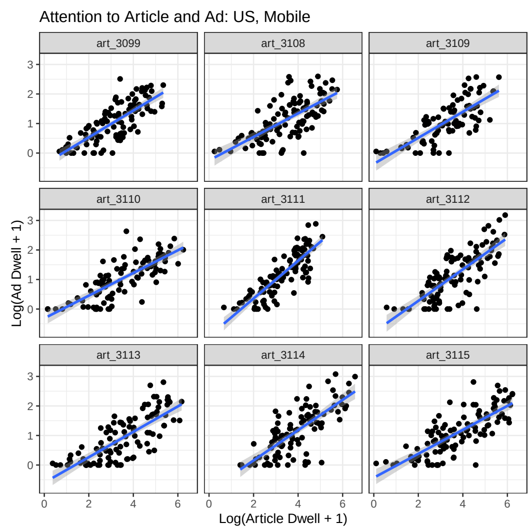
E.7 Distribution of Individual Demographics

Figure 21 shows additional detail on the distribution individual characteristics, namely gender, age, education, income and political leaning. These were collected as categorical variables.

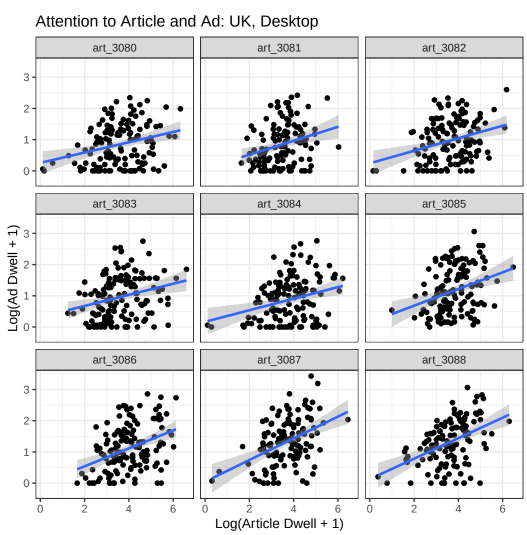
Figure 17: Correlation in Article and Ad Dwell, By Articles



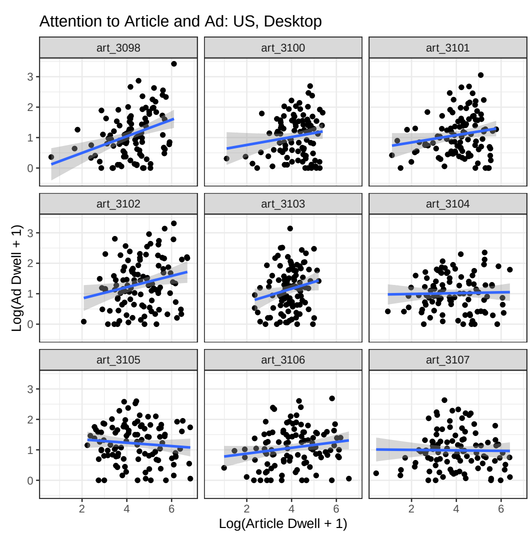
(a) UK, Mobile



(b) US, Mobile



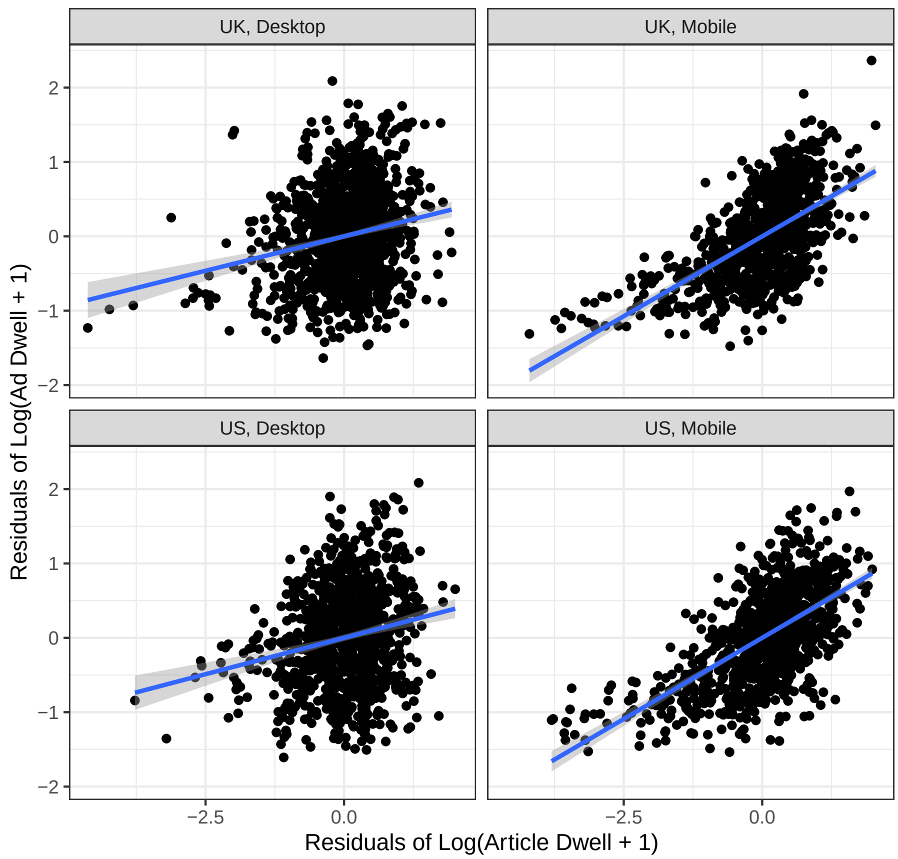
(c) UK, Desktop



(d) US, Desktop

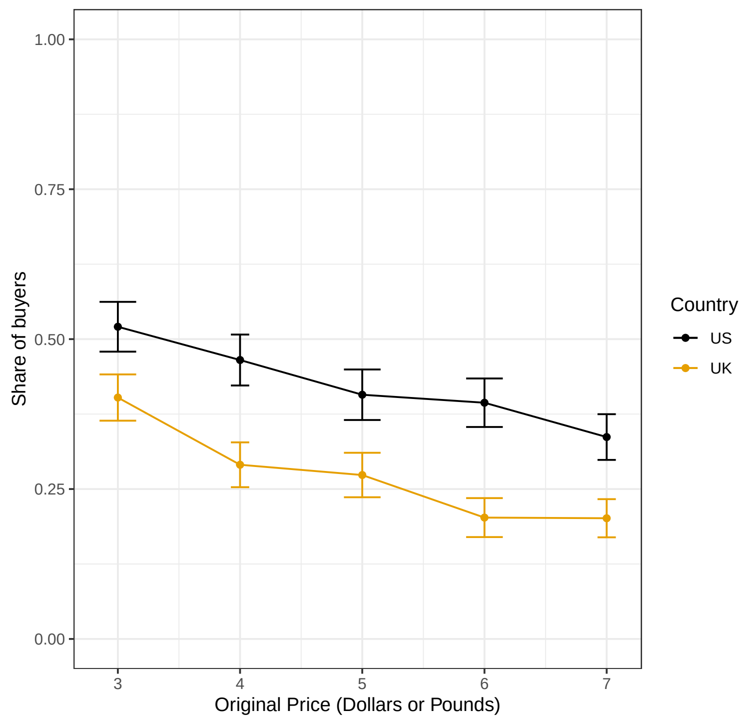
Correlation between attention to article and attention to ad, split by country (UK, US) and device type (desktop, mobile). Each panel corresponds to a single article. Ad and article dwell times are transformed into the logarithmic scale to make the visualization easier to read. Blue lines correspond to the best linear prediction of the variable on the vertical axis by the variable on the horizontal axis. Shaded regions correspond to 95% confidence intervals.

Figure 18: Positive Correlation in the Residualized Article and Ad Dwell



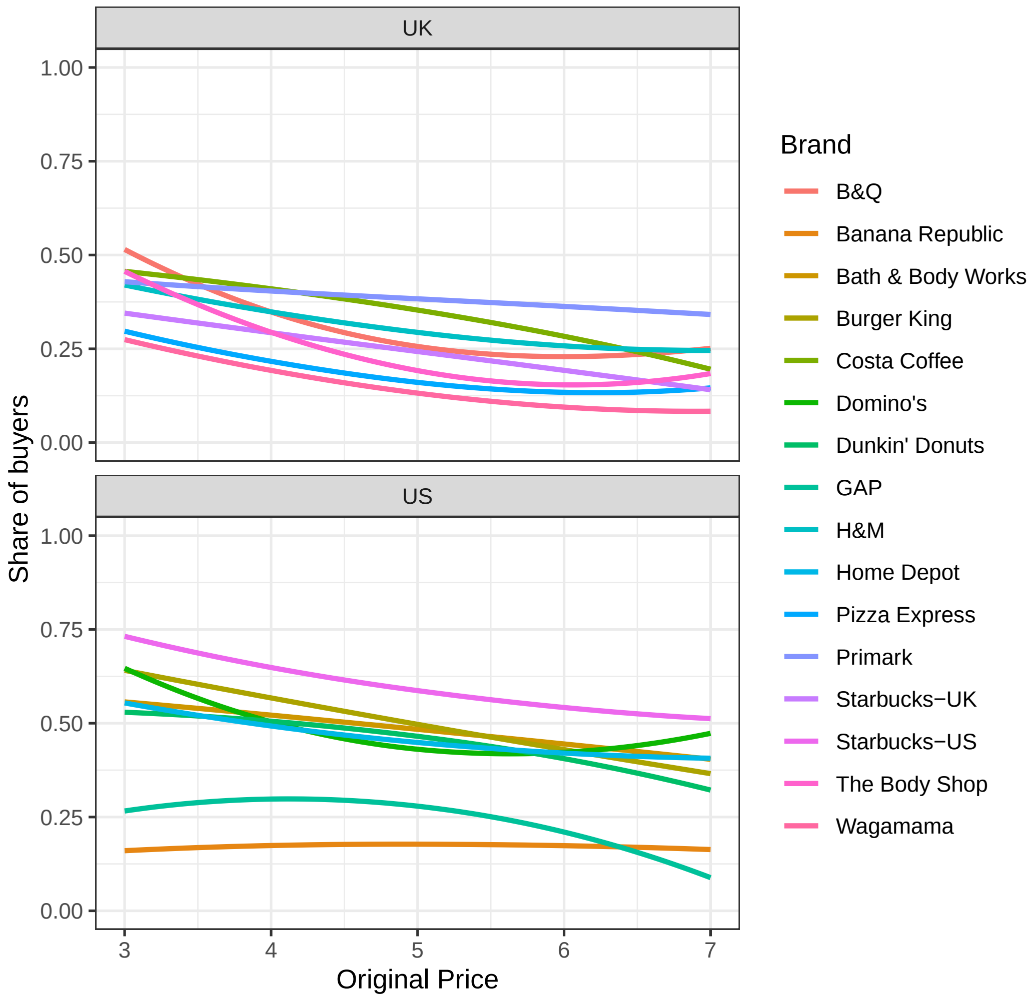
Correlation between attention to article and attention to ad, split by country (UK, US) and device type (desktop, mobile). Each panel corresponds to the set of articles in a country and device type. Ad and article dwell times are transformed into the logarithmic scale to make the visualization easier to read, and then residualized using the same FE as in our main empirical specification (e.g. Table 3). Blue line corresponds to the best linear prediction of the variable on the vertical axis by the variable on the horizontal axis.

Figure 19: Demand Curves by Country



Demand curves are computed in each country as an average purchase probability across brands. The currency is dollars in the US and pounds in the UK.

Figure 20: Demand curves by brands



Demand curves are computed for each brand in each country. The plots show a non-parametric regression of purchase decisions on the price shown to the individual. Demand curves are broadly downward sloping for each brand.

Figure 21: Additional Details on Individual-level Covariates

Details on the distribution of gender, age, education, income and political leaning, at the individual level and across countries.

F Additional Results for Attention Allocation Model

In this section, we present a robustness check of the results in Table 2 with individual fixed effects. Columns (3) and (6) of Table 21 show this specification.

Table 21: Estimates of attention spillovers and ad avoidence: Individual FEs

	Ad Dwell					
	IV			OLS		
Panel I	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}_1$	3.083***			2.715***		
	(0.306)			(0.197)		
$\hat{\gamma}$	0.008***	0.007***	0.009***	0.011***	0.011***	0.012***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
1st Stage Incr. F-Stat	65.86	80.15	128.23			
Observations	3,925	3,925	3,925	3,925	3,925	3,925
R^2	0.135	0.135	0.545	0.145	0.152	0.547

	Article Dwell - $\hat{\gamma}$ Ad Dwell					
Panel II	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\alpha}_1$	105.907***			105.894***		
	(4.521)			(4.521)		
$\hat{\beta}$	7.015*	6.832*	8.791**	7.024*	6.845*	8.798**
	(3.919)	(3.741)	(3.509)	(3.919)	(3.741)	(3.509)
Observations	4,426	4,426	4,426	4,426	4,426	4,426
R^2	0.030	0.112	0.640	0.030	0.112	0.640
FE:						
Step Order	Y	Y	Y	Y	Y	Y
Article	N	Y	Y	N	Y	Y
Brand	N	Y	Y	N	Y	Y
Individual	N	N	Y	N	N	Y

*p<0.1; **p<0.05; ***p<0.01

All specifications include step order fixed effects, with step order = 1 normalized to zero. Estimates in Panel I represent coefficients from a regression of Ad Dwell on Article Dwell. In the IV specification, Article Dwell is instrumented for by the average amount of attention devoted to that article by all but all other individuals who did not see the same article-ad pairing (Leave Many Out IV). Estimates in Panel II represent coefficients from an OLS regression of Article Dwell on an indicator of whether the ad is present on the news article. We subtract $\hat{\gamma}$ Ad Dwell from Article Dwell in Panel II to control for the attention spillover from ad to article. Standard errors clustered at the individual level.

G Effects of Advertising on Recall and Purchase

G.1 Functional Form Robustness

In this section, we present several functional form robustness checks of our results regarding the effect of attention on recall and purchase (analogous to Table 3). Table 22 presents a specification that allows for non-linear (in particular, quadratic) effects of attention. Table 23 presents the results of a specification with individual FE. Table 24 presents the results of a logit specification.

Table 22: Effect of Attention on Recall/Purchase

	Dependent variable:			
	Recall		Purchase	
	(1)	(2)	(3)	(4)
Ad Visible	0.0063***		0.0016*	
	(0.0009)		(0.0009)	
Ad Visible sqr.	−0.00003***		−0.000002	
	(0.00001)		(0.00001)	
Ad Dwell		0.0747***		0.0104**
		(0.0069)		(0.0053)
Ad Dwell sqr.		−0.0028***		−0.0002
		(0.0005)		(0.0003)
Brand FE	Y	Y		
Price x Brand FE			Y	Y
Observations	5,707	3,925	5,707	3,925
R^2	0.0946	0.1678	0.1300	0.1361

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include a quartic polynomial in log of average time that each article was visible for by each individual. All specifications include Step Order and Device x Country FE. All specifications include FE for individual covariates (income, gender, education, age, and self-reported political leaning). Standard errors clustered at the individual level.

G.2 Leave-1-Out (L1O) IV

In this section, we present our estimates of the effects of attention on purchase and recall, when we instrument attention (*Ad Visible* and *Ad Dwell*) using the L1O measure of attention (the average attention devoted to each article by all other individuals in our sample).

Table 23: Effect of Attention on Recall/Purchase

	Dependent variable:			
	Recall		Purchase	
	(1)	(2)	(3)	(4)
Ad Visible	0.0007		0.0010**	
	(0.0006)		(0.0005)	
Ad Dwell		0.0106***		0.0036
		(0.0033)		(0.0027)
Brand FE	Y	Y		
Price x Brand FE			Y	Y
Observations	5,707	3,925	5,707	3,925
R^2	0.5069	0.5092	0.4863	0.4789

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include a quartic polynomial in log of average time that each article was visible for by each individual. All specifications include Step Order and Device x Country FE. All specifications include individual FE. Standard errors clustered at the individual level.

Table 24: Effect of Attention on Recall/Purchase (Logit)

	Dependent variable:			
	Recall		Purchase	
	(1)	(2)	(3)	(4)
Ad Visible	0.0148***		0.0067***	
	(0.0023)		(0.0023)	
Ad Dwell		0.1780***		0.0342***
		(0.0142)		(0.0123)
Brand FE	Y	Y		
Price x Brand FE			Y	Y
Observations	5,707	3,925	5,707	3,925

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include a quartic polynomial in log of average time that each article was visible for by each individual. All specifications include Step Order and Device x Country FE. All specifications include FE for individual covariates (income, gender, education, age, and self-reported political leaning). Standard errors clustered at the individual level.

Table 25: Estimates of advertising effects on recall and purchase: L1O IV

	Recall ($\hat{\rho}$)					Purchase ($\hat{\lambda}$)				
	All	Device		News Type		All	Device		News Type	
		Mobile	Desktop	Hard	Soft		Mobile	Desktop	Hard	Soft
Panel I	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Visible	0.001	0.003	0.003	−0.022	0.004	0.002	0.009	−0.003	0.016	−0.0002
	(0.004)	(0.007)	(0.006)	(0.031)	(0.004)	(0.005)	(0.007)	(0.006)	(0.037)	(0.004)
Observations	3,925	2,165	1,760	1,824	2,101	3,925	2,165	1,760	1,824	2,101
R^2	0.136	0.167	0.156	−0.131	0.146	0.109	0.102	0.087	0.009	0.141
	First Stage									
L1O Article Dwell	0.068***	0.072***	0.074***	0.019	0.099***	0.069***	0.074***	0.070***	0.015	0.104***
	(0.010)	(0.016)	(0.011)	(0.013)	(0.013)	(0.010)	(0.016)	(0.011)	(0.013)	(0.014)
Observations	3,925	2,165	1,760	1,824	2,101	3,925	2,165	1,760	1,824	2,101
R^2	0.414	0.407	0.466	0.292	0.509	0.405	0.389	0.450	0.268	0.488
1st Stage Incr. F-Stat	49.1	19.47	44.25	2.21	57.24	48.94	20.67	42.04	1.24	57.28
Panel II	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ad Dwell	0.009	0.023	0.037	−0.050	0.046	0.016	0.068	−0.052	0.030	−0.002
	(0.037)	(0.058)	(0.078)	(0.061)	(0.057)	(0.040)	(0.050)	(0.109)	(0.065)	(0.049)
Observations	3,925	2,165	1,760	1,824	2,101	3,925	2,165	1,760	1,824	2,101
R^2	0.136	0.162	0.116	0.091	0.098	0.132	0.147	−0.125	0.132	0.136
	First Stage									
L1O Article Dwell	0.008***	0.009***	0.006*	0.009**	0.008***	0.008***	0.010***	0.004	0.008**	0.008***
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Observations	3,925	2,165	1,760	1,824	2,101	3,925	2,165	1,760	1,824	2,101
R^2	0.158	0.179	0.203	0.280	0.146	0.143	0.150	0.171	0.252	0.117
1st Stage Incr. F-Stat	12.95	10.34	2.78	5.52	6.77	12.73	13.27	1.78	5.31	7.83

*p<0.1; **p<0.05; ***p<0.01

All specifications include a quartic polynomial in log of average time that an average article was visible for by each individual, step order and device x country fixed effects, fixed effects for individual covariates (income, gender, education, age, and self-reported political leaning), and brand (for recall) or brand x price (for purchase) fixed effects. Standard errors clustered at the individual level.

G.3 Ideological Mismatch as IV

In this section, we investigate the impact of the mismatch between individuals and newspapers, particularly concerning their political ideologies, on the effectiveness of advertising. Essentially, we consider this mismatch as an alternative instrumental variable (IV) that can exogenously influence the attention individuals devote to articles. As established earlier, heightened attention to news content increases readers' exposure to display ads, subsequently enhancing ad effectiveness. Prior research suggests that alignment of readers' political beliefs could facilitate such increased attention, as news readers tend to prefer articles with an ideological slant aligned with their views (Schmuck et al., 2019).

To gauge the political alignment between news readers and articles, we rely on the political orientation of news outlets. Participants can see the newspaper from which each story originates, as a billboard displaying the news source is presented at the top of each article. Additionally, at the end of the experiment, respondents were asked about their political views.[30]

The selected outlets and their political slants are well-known in each country. In the UK, *The Guardian* is left-leaning, while the *Daily Mail* is right-leaning. In the US, *The New York Times* leans left, while *USA Today* is centrist. An independent survey on AMT validates these choices for each newspaper's political slant (see Appendix B).

Does an individual with self-reported conservative views react differently to news published by a newspaper that leans politically to the left? To address this, we construct an index of "right-wing-ness" for each newspaper and individual. For newspapers, *Daily Mail* is assigned +1, *USA Today* is assigned 0, and *The New York Times* and *The Guardian* are assigned -1.[31]

[30] This question was posed at the end of the experiment to prevent potential bias in participants' behavior.

[31] This classification is broadly confirmed by sites that regularly conduct media bias ratings, e.g., https://www.allsides.com/media-bias/media-bias-ratings.

Similarly, individuals self-identifying as Conservative, Moderate, and Liberal are assigned +1, 0, and -1, respectively. We then compute, for each observation, the "political mismatch" between each individual and newspaper article, defined as the absolute value of the difference between these two variables. A mismatch of 0 indicates no mismatch, occurring when a person places themselves to the right of the political spectrum when reading the *Daily Mail* (or a left-wing person reading *The Guardian*). Conversely, a large mismatch (mismatch = 2) occurs when that person is presented with an article from an outlet at the opposite end of the political spectrum, with intermediate cases (mismatch = 1) arising from other combinations.

With the relevant data defined, we first explore whether a politically mismatched individual-article pair is associated with the individual devoting less attention to the article and display advertising on this article's page. Since individual-article match is the primary source of identifying variation, we include article, brand, and step order fixed effects.[32] Table 26 presents the results. A higher political mismatch correlates with lower attention to the article: transitioning from completely misaligned views (mismatch = 2) to fully aligned views (mismatch = 0) leads to an increase in the time people spend reading the article by around 15 seconds (Columns 1 and 2). This, in turn, results in increased attention to ads on the page, with ads becoming visible for $0.63 \cdot 2 = 1.26$ seconds more (Column 3) and attracting $0.11 \cdot 2 = 0.22$ seconds more active attention time (Column 4).[33]

We proceed to validate that the incremental attention to advertising, resulting from the alignment between an article and a reader's ideological stance, translates into heightened advertising effectiveness. To achieve this, we recalibrate our instrumental variable (IV) specifications (Equations 3 and 4) by employing the political mismatch as an instrument for ad attention. We adopt the same specification as in Table 26.

[32] Results are robust if we include individual fixed effects.

[33] Table 26 has fewer observations than some of the previous tables. This is because we allowed individuals to opt out of reporting their political orientation. In these cases (about 6% of the data), the political mismatch could not be computed. The same holds for Tables 27 and 28.

Table 26: Attention and Political Mismatch

	Dependent variable:			
	Article Visible	Article Dwell	Ad Visible	Ad Dwell
	(1)	(2)	(3)	(4)
Political Mismatch (0/1/2)	-6.6857^{**}	-7.8402^{***}	-0.6265^{**}	-0.1072
	(3.0376)	(2.3433)	(0.3178)	(0.0885)
Observations	5,360	3,652	5,360	3,652
R^2	0.6473	0.5008	0.4413	0.1882

Note: *p<0.1; **p<0.05; ***p<0.01. Fixed Effects: Article, Brand, Step Order. Includes individual covariates FE (income, gender, education, age, politics). Includes quartic (degree 4) polynomial in total time page is visible for each individual. Includes Price x Brand FE. Excludes observations for which no brand was shown. Standard errors clustered at the individual level.

Table 27 presents the estimates delineating the effects of ad attention on purchase probabilities. Analogous to our previous approach, Columns 1 and 2 showcase the first stage and IV estimates, utilizing ad visibility as a metric for ad attention. Although the first stage exhibits a lower statistical power compared to our alternative IV (F-statistic of 3.87), the IV estimate still yields a statistically significant positive effect, aligning with our overarching findings regarding the pivotal role of incremental attention to display ads.

Columns 3 and 4 of Table 27 present the first stage and IV estimates using ad dwell as a measure of ad attention. In this specification, the first stage estimates are too imprecise to generate conclusive IV estimates.

We present the results for ad recall in Table 28. As before, these results are inconclusive on the effect of incremental ad attention on recall.

Table 27: 2SLS Regression of Purchase on Attention, IV: Pol Mismatch

	Dependent variable:			
	Ad Visible (1st stage)	Purchase (2SLS)	Ad Dwell (1st stage)	Purchase (2SLS)
	(1)	(2)	(3)	(4)
Pol Mismatch	-0.6265^{**}		-0.1072	
	(0.3183)		(0.0887)	
Ad Visible		0.0655^{*}		
		(0.0367)		
Ad Dwell				0.3582
				(0.3127)
1st Stage Incr. F-Stat	3.87		1.46	
Observations	5,360	5,360	3,652	3,652
R^2	0.4413	-3.0455	0.1882	-4.3809

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include a quartic polynomial in log of average time that each article was visible for by each individual. All specifications include Article and Step Order FE. All specifications include FE for individual covariates (income, gender, education, age, and self-reported political leaning). Standard errors clustered at the individual level. All specifications include Price x Brand FE.

Table 28: 2SLS Regression of Recall on Attention, IV: Pol Mismatch

	Dependent variable:			
	Ad Visible (1st stage)	Recall (2SLS)	Ad Dwell (1st stage)	Recall (2SLS)
	(1)	(2)	(3)	(4)
Pol Mismatch	-0.6510^{**}		-0.1386	
	(0.3220)		(0.0874)	
Ad Visible		0.0021		
		(0.0166)		
Ad Dwell				0.0599
				(0.0897)
1st Stage Incr. F-Stat	4.09		2.51	
Observations	5,360	5,360	3,652	3,652
R^2	0.4363	0.0982	0.1717	0.1296

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include a quartic polynomial in log of average time that each article was visible for by each individual. All specifications include Article and Step Order FE. All specifications include FE for individual covariates (income, gender, education, age, and self-reported political leaning). Standard errors clustered at the individual level. All specifications include Brand FE.

H A Stylized Model of Attention Allocation

We present here a simple model to microfound how individuals allocate their attention to articles and ads.

Consider reader i deciding how much attention to devote to article j (x_{art}) and display ads of brand k shown next to this article (x_{ad}). For notational simplicity, we drop references to i, j, and k in what follows – they are reintroduced for the empirical specification. The reader chooses x_{art} and x_{ad} to maximize the entertainment utility from examining the web page,

$$U(x_{\text{art}}, x_{\text{ad}}) = \alpha x_{\text{art}} - \frac{x_{\text{art}}^2}{2} + \mathbb{1}(-\beta x_{\text{art}} + \delta x_{\text{ad}} + \gamma x_{\text{art}} x_{\text{ad}} - \frac{x_{\text{ad}}^2}{2}). \tag{5}$$

Here, α captures reader i's interest in article j. The indicator $\mathbb{1}_{ijks}$ describes whether the ad of brand k was shown next to article j for participant i. The coefficient β is the reader's disutility from devoting attention to the article when any ad is shown next to it (or utility if $-\beta > 0$). The coefficient δ is the reader's preference for devoting attention to the ad of brand k. Finally, γ is the parameter that determines whether the reader prefers to spend more attention on the ad if they spend more attention on the article, and vice versa (i.e., it measures the complementarity or substitutability between the article and ad). We capture the reader's costs of devoting increasing attention to the article and ad by including negative quadratic terms $x_{\text{art}}^2/2$ and $x_{\text{ad}}^2/2$, which ensure an interior solution while keeping the setting simple.[34]

Maximizing utility with respect to $x_{\text{art}}, x_{\text{ad}}$, and denoting the solutions as x^{art} and x^{ad}, yields the following First Order Conditions:

$$x^{\text{ad}} = \mathbb{1}(\delta + \gamma x^{\text{art}}) \tag{6}$$

[34]The model can be easily extended to allow for an overall time constraint such that $x_{\text{art}} + x_{\text{ad}} \leq \overline{x}$, and for differential costs of attention for ads and articles. However, such extensions do not give additional insights.

,

$$x^{\text{art}} = \alpha + \mathbb{1}(-\beta + \gamma x^{\text{ad}}). \tag{7}$$

Notice that, after reintroducing the notation referring to an individual, article, brand, and step order (plus an error term), the First Order Conditions correspond directly to our empirical framework (1)-(2).

There are two coefficients of interest: β and γ. The sign of β reflects whether the reader is an "ad avoider" or "ad lover". It is possible that individuals can be "ad lovers" (e.g., this might be particularly likely in the context of car or beauty magazines, (Kaiser and Wright, 2006)). However, past literature has found that consumers are more likely to be ad avoiders (e.g. Wilbur, 2008, 2016; Huang et al., 2018), so we expect $\beta > 0$.

The coefficient γ determines whether articles and ads are substitutes or complements. A priori, both could happen: a more interesting article could grab the reader's attention more effectively due to voluntary ("top-down") attention and increase ad avoidance (Drèze and Hussherr, 2003; Stenfors et al., 2003; Simola et al., 2011), but more time spent on the page reading the article also provides more opportunities for the ad to distract the reader with its visual stimuli, working through a model of "bottom-up" attention (Koch and Ullman, 1987; Itti et al., 1998; Pieters and Wedel, 2007; Cerf et al., 2007; Milosavljevic and Cerf, 2008).

References

Moran Cerf, Jonathan Harel, Wolfgang Einhäuser, and Christof Koch. Predicting human gaze using low-level saliency combined with face detection. Advances in Neural Information Processing Systems, 20, 2007.

Xavier Drèze and François-Xavier Hussherr. Internet advertising: Is anybody watching? Journal of Interactive Marketing, 17(4):8–23, 2003.

Kenneth Holmqvist, Jana Holsanova, Mari Barthelson, and Daniel Lundqvist. Reading or scanning? a study of newspaper and net paper reading. In The Mind's Eye, pages 657–670. Elsevier, 2003.

Jason Huang, David Reiley, and Nick Riabov. Measuring consumer sensitivity to audio advertising: A field experiment on pandora internet radio. Available at SSRN 3166676, 2018.

Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254–1259, 1998.

Ulrich Kaiser and Julian Wright. Price structure in two-sided markets: Evidence from the magazine industry. International Journal of Industrial Organization, 24(1):1–28, 2006.

Christof Koch and Shimon Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. In Matters of Intelligence, pages 115–141. Springer, 1987.

Milica Milosavljevic and Moran Cerf. First attention then intention: Insights from computational neuroscience of vision. International Journal of Advertising, 27(3): 381–398, 2008.

Rik Pieters and Michel Wedel. Goal control of attention to advertising: The yarbus implication. Journal of Consumer Research, 34(2):224–233, 2007.

Desirée Schmuck, Miriam Tribastone, Jörg Matthes, Franziska Marquart, and Eva Maria Bergel. Avoiding the other side? an eye-tracking study of selective exposure and selective avoidance effects in response to political advertising. Journal of Media Psychology: Theories, Methods, and Applications, 2019.

Tim Schneegans, Matthew D Bachman, Scott A Huettel, and Hauke Heekeren. Exploring the potential of online webcam-based eye tracking in decision-making research and influence factors on data quality. 2021.

Kilian Semmelmann and Sarah Weigelt. Online webcam-based eye tracking in cognitive science: A first look. Behavior Research Methods, 50(2):451–465, 2018.

Jaana Simola, Jarmo Kuisma, Anssi Öörni, Liisa Uusitalo, and Jukka Hyönä. The impact of salient advertisements on reading and attention on web pages. Journal of Experimental Psychology: Applied, 17(2):174, 2011.

Iréne Stenfors, Jan Morén, and Christian Balkenius. Behavioural strategies in web interaction: A view from eye-movement research. In The Mind's Eye, pages 633–644. Elsevier, 2003.

Kenneth C Wilbur. A two-sided, empirical model of television advertising and viewing markets. Marketing Science, 27(3):356–378, 2008.

Kenneth C Wilbur. Advertising content and television advertising avoidance. Journal of Media Economics, 29(2):51–72, 2016.

Xiaozhi Yang and Ian Krajbich. Webcam-based online eye-tracking for behavioral research. Judgment and Decision Making, 16(6):1486, 2021.