Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure

Ana S. Cardenal¹,², Carlos Aguilar-Paredes³, Carol Galais⁴, and Mario Pérez-Montoro³

Abstract
This paper analyzes the role of different origins to news media in selective exposure. We rely on a unique web-tracking online dataset from Spain to identify points of access to news outlets and study the influence of direct navigation and news-referred platforms (i.e., from Facebook and Google) on selective exposure. We also explore cross-level interactions between origins to news and political interest and ideology. We find that direct navigation increases selective exposure while Google reduces it. We also find that the relationship between origins to news and selective exposure is strongly moderated by ideology, suggesting that search engines and social media are not content neutral. Our findings suggest a rather complex picture regarding selective exposure online.

Keywords
digital technologies, online selective exposure, media exposure, platforms, filter bubbles

Exposure to news media is increasingly mediated by digital technologies. Today, online media is the preferred news source in most advanced democracies, above TV and well above printed media. Moreover, two-thirds (65 percent) of online media users prefer to

¹Universitat Oberta de Catalunya, Barcelona, Spain
²Universitat Pompeu Fabra, Barcelona, Spain
³University of Barcelona, Barcelona, Spain
⁴Universitat Autònoma de Barcelona, Barcelona, Spain

Corresponding Author:
Ana S. Cardenal, Universitat Oberta de Catalunya, Parc Mediterrani de la Tecnologia (edifici B3), Av. Carl Friedrich Gauss 5, Castelldefels, 08860 Barcelona, Spain.
Email: acardenal@uoc.edu
use a side door of some kind (e.g., social media, search engines, news aggregators) to access news outlets (Newman et al. 2017). These new ways of accessing news have shifted content curation from editorial boards to individuals, their social networks, and algorithmic information sorting (Bakshy et al. 2015; Flaxman et al. 2016). Although many have seen in these technologies the potential for increasing exposure to diverse news media and opinions (e.g., Benkler 2006), the success of movements that appear to be immune to any factual evidence has reinvigorated claims concerning the potential of these technologies for personalizing information (Bruns 2017:1), locking people in echo-chambers (Sunstein 2009) and filter bubbles (Pariser 2011).

Much recent research has contributed to this debate questioning the prevalence of selective exposure in audiences’ online news consumption behavior (Dubois and Blank 2018; Dvir-Gvirsman et al. 2016; Flaxman et al. 2016; Gentzkow and Shapiro 2010; Nelson and Webster 2017; Weeks et al. 2016). While these are valuable contributions, with very rare exceptions (e.g., Flaxman et al. 2016), they have not addressed the contribution of specific technologies (e.g., social media, search engines) on observed selective exposure nor the mechanisms underlying selectivity. Therefore, even if from current research we can conclude that there is very little evidence of selective exposure in online news consumption, we still do not know whether this is the outcome primarily of “choice” (or voluntary exposure) or of “algorithmic filtering” (or involuntary exposure).¹

In this paper, we inquire about these different mechanisms by looking at the role of different origins (direct visits, social media, and search engines) in selective exposure. We use a unique dataset that combines survey and online web-tracking data from Spain to study the impact on selective exposure of three different points of access to news: direct (1), referred from Facebook (2), and referred from Google (3). We focus on these platforms because they are the largest within their kind (i.e., social media and search engines) and they account for an important part of news consumption (see the discussion below). Using a multilevel approach, we also analyze the impact of some individual attributes (e.g., ideology, political interest) whose influence in selective exposure is still unclear and explore cross-level interactions between these individual-level attributes and origins to news media.

We find that direct navigation increases selective exposure compared to most referred-based navigation—though the effect is small; that Facebook has no direct effect on selective exposure; and that Google considerably reduces selective exposure. More interestingly, we find that the relationship between points of access to news (e.g., direct, Facebook, Google) and selective exposure is strongly moderated by ideology, suggesting that search engines and social media are not content neutral. Our findings suggest a rather complex picture regarding selective exposure online.

**Digital Technologies and Selective Exposure**

Selective exposure is a core concept in communication and media studies, which states that given the chance, individuals will choose to consume media that reinforces their previous beliefs (Katz et al. 1955; Klapper 1960). This behavior is caused by cognitive
dissonance (Festinger 1957) and could be responsible for the minimal effects of the media on changes in attitude and opinion (e.g., Lazarsfeld et al. 1944).

According to mainstream accounts, technology may exacerbate selective exposure in the current media environment mainly through two mechanisms: choice and algorithmic filtering (Dubois and Blank 2018: 731). Assuming that individuals have a preference for like-minded information, choice prompts selective exposure by giving individuals the opportunity to choose, among some range of diverse opinions, the one that best matches their previous beliefs (Iyengar and Hahn 2009; Mutz and Martin 2001). Hence, when individuals are given some choice, selective exposure is expected to be the outcome of voluntary action.

Algorithmic curation, in contrast, refers to systems of information selection that are automatized and beyond individual control (Bakshy et al. 2015; Dubois and Blank 2018; Zuiderveen Borgesius et al. 2016). Although algorithms are based on users’ past choices and tastes, they are seen to be conducive to selective exposure from involuntary action and without users’ consent (Zuiderveen Borgesius et al. 2016: 3). Selective exposure resulting from algorithmic filtering is largely known in the literature as the filter bubble argument (Dubois and Blank 2018; Pariser 2011; Zuiderveen Borgesius et al. 2016).

Some have depicted these two mechanisms in different ways. For example, Zuiderveen Borgesius et al. (2016) refer to choice as “self-selected personalization” and to algorithmic curation as “pre-selected personalization.” Here, we use these concepts interchangeably.

Direct Access to News

Under direct access to news media, selective exposure is likely to be the outcome of choice or self-selected personalization. Direct visits to news media online should not be very different from old ways of accessing news through traditional media, in which newsgathering has typically been seen as the outcome of conscious choice (Tewksbury et al. 2001: 533), and to rely on brand (or source) as the main heuristic criterion for information selection (Messing and Westwood 2012; Sundar et al. 2007).

There is considerable evidence of partisan selectivity when people are asked to choose among media labels that they associate with different (partisan) biases (e.g., Iyengar and Hahn 2009; Messing and Westwood 2012; Stroud 2008; Turner 2007). Based on these findings, many have claimed that the expansion of choice that is characteristic of the new media environment will only exacerbate partisan selective exposure (e.g., Bennett and Iyengar 2008; Iyengar and Hahn 2009). According to these accounts, a high-choice media environment coupled with the proliferation of partisan outlets will lead partisan motivations to play a more active role in online news selection (Iyengar and Hahn 2009; Mutz and Martin 2001; Stroud 2008).

This view, however, has been contested by recent work. Data from real-life news consumption behavior show that when consuming news outlets online, people tend to rely on their favorite outlets, which tend be mainstream and centrist (Flaxman et al. 2016; Gentzkow and Shapiro 2010; Guess 2016). More importantly for our purposes,
this behavior—that is, consuming mainstream, centrist outlets—has been shown to be the prevailing one when people access news directly (Flaxman et al. 2016).

In spite of these findings, there are several reasons to expect selective exposure under direct navigation. First, even if recent studies do not find evidence of selectivity in online news consumption, experimental and survey studies consistently find evidence of a confirmation bias in information selection (e.g., Garrett 2009a, 2009b; Hart et al. 2009; Iyengar and Hahn 2009; Knobloch-Westerwick and Meng 2009; Stroud 2008). Second, most of these studies do not discriminate between different technologies; hence, most of the diversity they find could be the result of accidental exposure to information, which might be most clearly promoted by search engines and social media (we discuss this below). Finally, most of the evidence is U.S.-based and might be influenced by this country’s media system. Some have argued that selective exposure is not only influenced by choice but also by the media system (Goldman and Mutz 2011; Mutz and Martin 2001). According to this view, a polarized pluralistic media system, such as the Spanish one (Hallin and Mancini 2004), would make it easier for people to exercise (partisan) selective exposure than a liberal system, such as the United States.

Social Media

Aside from algorithmic filtering, another mechanism at work in social media with the potential for personalizing news in an involuntary way is homophily. It refers to the tendency that people have to connect with similar others (McPherson et al. 2001). Although homophily involves choice (people choose their friends and peers in social media), it leads to unintended forms of exposure because users “transfer” news exposure decisions to friends who pre-select news stories for them (Bright 2016; Singer 2014).

Although most studies on echo-chambers and filters bubbles in social media have focused on Twitter (e.g., Barberá et al. 2015; Colleoni et al. 2014; Conover et al. 2011), here we focus on Facebook because it is by far the largest platform and the most frequently used for getting news. In 2017, 47 percent of Spanish web users declared using Facebook for getting news in the last week and a similar percentage of U.S. users (47 percent) reported the same (Newman et al. 2017).² Most studies focus on Twitter because data are easier to access. So an additional reason to study Facebook is that, in spite of being the largest platform for getting news, we know much less about the dynamics of news exposure on this platform.

Despite the potential for news personalization on Facebook, existing evidence does not support the filter bubble hypothesis (e.g., Bakshy et al. 2015; Fletcher and Nielsen 2018; Nechushtai and Lewis 2019; Zuiderveen Borgesius et al. 2016). In contrast, several factors may account for cross-cutting exposure in social media. First, even though social media reproduce homophily (e.g., Colleoni et al. 2014), they are also more likely to prompt weak ties (e.g., Barberá 2014), which increase opportunities to encounter novel information (Granovetter 1977; Prior 2008) and have been found to be fertile ground fostering cross-cutting exposure (Mutz and Mondak 2006). Second,
even if social media are increasingly popular for getting news, given their socially
dominant character, in these media people are more likely to encounter news while
doing other things (e.g., Kim et al. 2013; Gottfried and Shearer 2016; Valeriani and
Vaccari 2015), and unintended exposure has been found to increase cross-cutting
political exchanges and deliberation (Brundidge 2010; Wojcieszak and Mutz 2009).
Finally, social media make available alternative cues for information selection—for
example, social recommendations—that may potentially override partisan cues
(Knobloch-Westerwick et al. 2005; Messing and Westwood 2012; Nelson and Webster
2017). Moreover, using experimental design, Messing and Westwood (2012) have
shown that, given a choice, people prefer relying on social recommendations rather
than on partisan cues. They contend that information utility, which increases willing-
ness to engage with attitude-discrepant messages (Knobloch-Westerwick Klei
man 2012; Valentino et al. 2009), might explain this result. All these factors are more likely
to be at work on platforms with a strong social character, such as Facebook, than in
more specialized, news-oriented platforms, such as Twitter.

Search Engines

Google and other search engines rely on algorithms for selecting among thousands of
billions of information stored in the Internet, which they then present and recommend
to the viewer. Although there are a number of search engines in the market (e.g.,
Yahoo, Bing, AOL) to help users find information, Google is by far the search engine
with the greatest market share worldwide. In Spain, Google’s market share is 90.67
percent, and in the United States, its share is 78.81 percent (http://gs.statcounter.com).
In addition, search engines (i.e., Google) have become increasingly popular platforms
for accessing news. For 25 percent of the respondents in a study, search engines are the
preferred way of accessing news, and search engines and social media together are
preferred gateways to news for almost 50 percent of web users (Newman et al. 2017).
Recent research shows that Google search results may influence the evaluation of
sources in many fields of behavior, including voting (e.g., Epstein and Robertson
2015). Worries about personalization are thus warranted not only because many rely
on Google to be informed but also because Google search results have been shown to
directly affect political behavior.

Despite these worries, research has found no evidence of substantial personaliza-
tion in Google searches (e.g., Hannak et al. 2013; Haim et al. 2018; Hoang et al. 2015;
Fletcher and Nielsen 2018; Nechushtai and Lewis 2019; Puschmann 2017). In con-
trast, several factors might help to explain cross-cutting exposure in Google. First,
most people see search engines as a fair and unbiased source of information (Fallows
2005; Dutton et al. 2017), as neutral and ideologically blind (Sundar and Nass 2001),
and neutral sources are associated with higher source credibility (Sundar and Nass
2001), which in turn might increase the likelihood of exposure to cross-ideological
information by making users less defensive about media content (Druckman et al.
2012; Knobloch-Westerwick et al. 2015). Second, studies examining the use of heuris-
tics in Google searches have found that rank order is by far the most important cue in
information selection (Chitika-Insights 2013; Haas and Unkel 2017; Murphy et al. 2006) and heuristics that are not overtly partisan may increase opportunities for cross-ideological exposure (Messing and Westwood 2012). Third, Google might help users find sources that are not mainstream and contribute to diversifying people’s news media diets (Athey et al. 2017; Puschmann 2017).

From this discussion it follows that direct navigation should increase selective exposure compared to referred-based navigation and that Facebook and Google should decrease it compared to direct navigation. However, we would also expect some interaction effects of news origins with individuals’ political orientations, notably political interest and ideology. Politically motivated individuals tend to have stronger opinions and more coherent views of the political world (Converse 1964; Zaller 1992), which might lead them to rely more often on partisan cues in their news consumption decisions (Barberá and Sood 2014; Iyengar and Hahn 2009). If, as some studies have found (e.g., Dutton et al. 2017), politically motivated individuals are more likely to rely on popular news-referred platforms such as Google and Facebook to search political information, political interest might positively moderate the relationship between exposure to news media via these platforms and selective exposure.

Moreover, recent studies begin to debate whether news-referred platforms such as Facebook and Google are content neutral (e.g., Dutton et al. 2017; Hancock et al. 2018; Haim et al. 2018; Puschmann 2017). Some U.S.-based studies have found that liberals are more active in social media than conservatives (Anderson and Jiang 2018; Bakshy et al. 2015). Moreover, a recent YouGov study found that in the United Kingdom users perceive left-wing content to be more widespread in Facebook (Dahlgreen 2016). Following these works, we might expect ideology also to interact with origins in shaping selective exposure. In particular, based on existing works about Facebook (e.g., Dahlgreen 2016) we would expect a left-wing ideology to positively moderate the relationship between exposure to news media via Facebook and selective exposure.

To sum up, the following would be our testing hypotheses:

**Hypothesis 1 (H1):** Direct navigation will increase selective exposure compared to referred-based navigation.

**Hypothesis 2 (H2):** Exposure to news media through Facebook will decrease selective exposure compared to direct navigation.

**Hypothesis 2.1 (H2.1):** Interest in politics will positively moderate the relationship between exposure to news media via Facebook and selective exposure.

**Hypothesis 2.2 (H2.2):** Left ideology will positively moderate the relationship between exposure to news media via Facebook and selective exposure.

**Hypothesis 3 (H3):** Exposure to news media through Google will decrease selective exposure compared to direct navigation.

**Hypothesis 3.1 (H3.1):** Interest in politics will positively moderate the relationship between exposure to news media via Google and selective exposure.
Data and Measures

Design

This analysis is based on a combination of survey and online web-tracking data for 408 individuals coming from a pool of the Spanish online population. Survey data were used to tap participants’ left–right self-placement, political interest, and other relevant sociodemographic traits. Web-tracking online data were used to trace participants’ online navigation behavior during a period of 3 months, extending from January 26 to April 27, 2015, which has been estimated to give an accurate description of peoples’ online media exposure habits (Athey et al. 2017). Navigation data enable us to rely on direct measures of exposure and also, given its granularity, to analyze the role of different origins to news media on selective exposure. For our study, we targeted individuals who consumed a minimum of information online and who had their residence in all-Spanish regions except for Catalonia.

Panel

Participants in our study are part of an opt-in panel of a Spanish market-research firm, which worked with us on all aspects of the sample and the implementation of the survey. Recruitment was done using online contacts and offering incentives for completing structured questionnaires on their personal electronic devices (home computers, tablets, or cell phones). We targeted a sample of 1,000 people and retained 40.8 percent of the subjects for our tracked sample ($N = 408$). Hence, a total of 408 explicitly agreed to share their anonymized browsing history for our study, a figure that accords with previous studies analyzing audience online news consumption (Guess 2016).

Participants in our panel match the characteristics of the Spanish online population in most sociodemographics and relevant attitudes (Robles et al. 2012). They tend to be younger, more educated, and more politically interested than the average Spanish citizen, although ideologically they are slightly to the right of the average Spanish citizen (see the Supplementary Information file for more details). To further check the representativeness of our sample, we compared the list of the top 20 most-visited news sites by our tracked sample with that provided by Alexa, one of the most widely used meters of online news audiences (Alexa Internet 2014), for the Spanish online population. We obtained a strong Pearson rank-order correlation score (.81). Hence, participants in our study tend to visit most of the same outlets (at least 80 percent), and with a similar frequency, as the general Spanish online population, testifying further to the representativeness of our sample with regard to a central element of our study such as online news consumption habits. In the Supplementary Information file, we include additional analyses testing to the sample’s representativeness.

Despite these similarities, we cannot make overgeneralizations from our final sample to the Spanish online population because people who voluntarily accept being tracked are generally less concerned about privacy. Yet we see this attitude as an advantage and assume that they will not modify their news consumption routines as a
result of our study. Notably, our subjects agreed to being tracked long before we started the study, which may have also helped to mitigate any initial change in their regular behavior.

**Tracking Data and Coding Procedure**

We used a dataset on navigation data on desktop devices for a period of 92 days, which contains a total of 1,024,026 URLs of visited sites corresponding to 79,071 domains for 408 unique users. These observed data for online navigation was filtered in order to identify news exposure for the top 42 most-visited news outlets in Spain according to Alexa rankings. Our final dataset is composed of the visits to the 42 top most-visited news outlets (our units of observation), which amount to 40,683 visits, and 3.97 percent of total site visits during our period of study. More information concerning the collection of the tracking data and its characteristics is provided in the Supplementary Information file.

**Measurements**

**Selective exposure.** The dependent variable in this study is selective exposure, which is defined at the visit level. To operationalize this variable, we first had to classify outlets and individuals according to their ideologies (see below). Visits to news media that are consistent with the visitor’s ideology are coded as 1, and as 0 otherwise. Table 1 includes the description of this and all other key variables.

**Media slant.** To measure media slant, we relied on reported media positions from the survey. Respondents provided the ideological position of each media outlet they usually visit (left, right, or neutral). We classified an outlet as partisan if at least 50 percent of respondents perceived it as right- or left-leaning. Our classification is shown in Table A5 of the Supplementary Information file. In our classification, almost half of the news media included in our list of 42 are classified as partisan ($M = 0.51$). More importantly, the list of partisan media includes the two major Spanish news outlets, *El Mundo* and *El País*, a result that matches previous classifications and research (Hallin and Mancini 2004; Newman et al. 2017). This is consistent with the characteristics of the Spanish media system, which has been classified as polarized pluralism (Hallin and Mancini 2004), in contrast, for example, to liberal media systems, such as the U.S., where, according to previous research (Flaxman et al. 2016; Gentzkow and Shapiro 2011; Guess 2016), mainstream outlets and news portals tend to be classified as neutral, and partisan media are generally small in size.

**Political leanings.** To measure people’s political leanings, we used ideological self-placement in a left–right scale from 0 (extreme left) to 10 (extreme right). Participants placing themselves in positions 6 to 10 of the scale are coded as right, and those placing themselves in positions 0 to 4 as left. Following other studies on selective exposure
(e.g., Gentzkow and Shapiro 2011), moderates—those self-placed at 5 on the scale in our study—were eliminated from the analysis.

**Origins.** We consider a visit to be *direct* if there is no overlapping between the current and the previous session or if when visiting a new site the previous site session has been closed. More technically, for a visit to be coded as direct, the difference between the starting time of a visit to outlet \( j \) \((SN_{t+1})\) and the starting time of the previous visit \((SN_t)\) must be larger than the duration of navigation in the site of origin (we add an extra second to account for the opening and closing session of the site) or

\[
(SN_{t+1}) - (SN_t) > (DN_t) + 1
\]

Conversely, we consider a visit to be *referred* if the opposite condition holds, that is, if the difference between the starting time of a visit to outlet \( j \) \((SN_{t+1})\) and the starting time of the previous visit \((SN_t)\) is smaller than the duration of navigation in the site of origin. Hence, if the following condition holds, we code a visit as referred (see Note 6):

\[
(SN_{t+1}) - (SN_t) < (DN_t) + 1
\]

We consider a visit to a news outlet \( j \) to have its origin in Facebook if (1) the site that immediately precedes the visit to \( j \) is Facebook and (2) it fulfills the condition of being referred.8 The same coding procedure is applied for Google. All other indirect referrals (e.g., Twitter) have been coded as “others.”

**Political interest.** Political interest is assessed by asking: “How much you would say you are interested in politics: Very much (3), quite interested (2), hardly interested (1), or not at all (0)?” The mean political interest in our panel is 1.97 \((SD = 0.76)\).

**Education.** We have included a third variable measured at the individual level that accounts for educational level. Given the skewed distribution of this variable (as the sample was highly educated), we have coded it as having attended college against all the other possibilities.
The nature of our dependent variable requires a logistic estimation to test the relationship between origins and selective exposure. Yet, given the nested structure of our data (where visits are nested in individuals), we cannot use a simple logistic regression because we would be violating some fundamental assumptions of regression; for example, the observations must be independent, and errors must not be correlated.

Using a multilevel approach allows us not only to explore the effect of individual (second) level variables (i.e., political interest, ideology) in the outcome variable (i.e., selective exposure), but also to explore cross-level interactions between first- and second-level variables and thus, to model varying-intercept and varying-slope effects.

**Results**

To test our hypotheses, we ran a series of hierarchical models. The results are displayed in Table 2. The first column in Table 2 reproduces a null model, followed by a model that only includes first-level variables and a model that includes both first- and second-level variables. The first model includes information for the variance of individual-level (Level 2, in our analysis) intercepts. Most importantly, the intra-class correlation (ICC) figure suggests that the amount of variation at the visits level due to factors related to the upper level of analysis (individuals, in our case) is as high as 34 percent. This is more than enough to justify a multilevel analysis.

The coefficients in Model 2 read as regular logistic coefficients, estimated with proper standard errors. We ran the same models twice because our hypotheses are established against different categories of reference: referred-based (H1) and direct navigation (all the others). Here, we only show the results for the models using direct navigation as the reference category because it is the relevant reference category for most of the hypotheses (H2, H2.1, H2.2, H3, and H3.1). However, the results with other referrals as the reference category can be found in the Supplementary Information file.

Models 2 and 3 show results for first-level and first- and second-level predictors, respectively, assuming heterogeneous intercepts (different average levels of selective exposure across individuals). As we can see from these models, of all the origins, only Google has a significant effect on selective exposure and in the expected direction—that is, it reduces the probability of selective exposure (H3). Neither Facebook nor other referrals have a significant effect on selective exposure compared to direct navigation.

Models 2 and 3, however, assume that the effects of origins on selective exposure are homogenous across individuals (i.e., they assume invariable slope effects), which may yield biased estimates if within-slope effects are in fact not homogenous. When freeing the slope of the variable origin (Model 4), we see some changes, implying that the homogeneity (slope) assumption is probably wrong. First, the coefficient for Google increases, suggesting a larger Google effect. Second, we see for the first time a significant negative coefficient for other referrals. This suggests that other referrals have a negative effect on selective exposure, or, symmetrically, that accessing news media directly increases the probability of selective exposure relative to arriving from
## Table 2. Multilevel Logistic Estimation of Selective Exposure.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective exposure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Origin: Facebook</strong></td>
<td>0.087</td>
<td>0.088</td>
<td>0.102</td>
<td>−0.449**</td>
<td>−0.139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.064)</td>
<td>(0.100)</td>
<td>(0.185)</td>
<td>(0.281)</td>
<td></td>
</tr>
<tr>
<td><strong>Origin: Google</strong></td>
<td>−0.284***</td>
<td>−0.286***</td>
<td>−0.351***</td>
<td>0.216</td>
<td>−0.352</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.056)</td>
<td>(0.097)</td>
<td>(0.160)</td>
<td>(0.263)</td>
<td></td>
</tr>
<tr>
<td><strong>Origin: Other (referrals)</strong></td>
<td>−0.035</td>
<td>−0.034</td>
<td>−0.108*</td>
<td>−0.182</td>
<td>−0.173</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.063)</td>
<td>(0.119)</td>
<td>(0.171)</td>
<td></td>
</tr>
<tr>
<td><strong>Ideology (left)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.724***</td>
<td>−0.719***</td>
<td>−0.690***</td>
<td>−0.719***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.170)</td>
<td>(0.182)</td>
<td>(0.180)</td>
<td>(0.182)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Political interest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.561***</td>
<td>0.562***</td>
<td>0.563***</td>
<td>0.538***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.103)</td>
<td>(0.102)</td>
<td>(0.109)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>College</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.074</td>
<td>−0.059</td>
<td>−0.062</td>
<td>−0.073</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td>(0.156)</td>
<td>(0.155)</td>
<td>(0.156)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Facebook × Left</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.764***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.215)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Google × Left</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−0.830***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.195)</td>
<td></td>
</tr>
<tr>
<td><strong>Other (referrals) × Left</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.139)</td>
<td></td>
</tr>
<tr>
<td><strong>Facebook × Political Interest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.127</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.134)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 2. (continued)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable:</td>
<td>Selective exposure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google × Political Interest</td>
<td>-0.003*** (0.127)</td>
<td>0.033 (0.082)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (referrals) × Political Interest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.882*** (0.084)</td>
<td>-0.851*** (0.085)</td>
<td>-1.327*** (0.252)</td>
<td>-1.346*** (0.257)</td>
<td>-1.367*** (0.256)</td>
<td>-1.291*** (0.266)</td>
</tr>
<tr>
<td>N (groups)</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Intra-class correlation</td>
<td>.34</td>
<td>.34</td>
<td>.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept SD</td>
<td>1.29</td>
<td>1.29</td>
<td>1.23</td>
<td>1.24</td>
<td>1.24</td>
<td>1.24</td>
</tr>
<tr>
<td>Random slope</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Akaike information criterion</td>
<td>32,349.460</td>
<td>32,336.220</td>
<td>32,289.470</td>
<td>31,808.320</td>
<td>31,784.110</td>
<td>31,813.360</td>
</tr>
<tr>
<td>Bayesian information criterion</td>
<td>32,366.210</td>
<td>32,368.080</td>
<td>32,356.460</td>
<td>31,950.680</td>
<td>31,959.960</td>
<td>31,980.830</td>
</tr>
</tbody>
</table>

*p < .1, **p < .05, ***p < .01.
referrals other than Google and Facebook. This can be seen more clearly from Table A6 in the Supplementary Information file, which presents the same models with other referrals as the category of reference. From Model 3 in Table A6, we have estimated that direct navigation increases the probability of selective exposure by almost 3 percent at the maximum, compared to referred-based navigation. Hence, when we assume heterogeneous slope effects, we find a positive effect of direct navigation on selective exposure—albeit not a strong one—and some support for H1.

The fifth and sixth models show the cross-level interactions. As we can see from the coefficients and standard errors in Model 6, political interest is not a relevant moderator in the relationship between origins and selective exposure, lending no support to H2.1 and H3.1. However, as Models 2 to 6 show, it has a strong direct effect in selective exposure, increasing the probability of selectivity by 14 percent, at most. In contrast, left–right self-placement (and, more particularly, being a leftist) not only is a strong direct predictor of selective exposure but also comes out as a powerful moderator between most origins (e.g., Facebook and Google) and selective exposure. In particular, as expected (H2.2), among left-wing individuals, Facebook referred-news navigation increases the probability of selective exposure compared to direct navigation by 7 percent.

Figure 1 represents graphically the interaction included in Model 5 (Table 2). It plots predicted probabilities of selective exposure for different origins by ideological group.

As we can see, for left-leaners, the predicted probability of selective exposure increases from 27 to 34 percent when the origin of news changes from direct access to Facebook, while it drops to 17 percent when it changes to Google. We can also see that the opposite is true for right-leaning individuals: The probability of selective exposure drops from 43 to 32 percent when the origin changes from direct access to Facebook, while it increases to 48 percent when it changes to Google. In other words, origins have opposite effects on left- and right-wing ideologues: Facebook increases selective exposure among left-leaners and decreases it among right-leaners, while Google decreases it among left-leaners and increases it among right-leaners.

**Discussion**

Numerous works have studied the prevalence of selective exposure in online news consumption. Yet, they have rarely addressed the contribution of news origins to observed selective exposure. This has hindered our ability to learn about the consequences of different pathways to news for selective exposure, and about the mechanisms underlying online selectivity. Relying on a unique dataset that combines survey and web-tracking online data from Spain, this paper has analyzed the impact of different origins to news media and of several moderators in an attempt to uncover a complex picture about selective exposure online and to bring light to the mechanisms underlying this phenomenon.

With regard to the (direct) impact of origins, we have expected direct navigation to increase selective exposure compared to referred-based navigation (H1) and Facebook
and Google to reduce selective exposure compared to direct navigation (H2 and H3). Yet, we have also expected political orientations, such as interest in politics and ideology, to interact between origins and selective exposure. In particular, we have expected political interest to increase selective exposure in popular news-referred platforms such as Facebook and Google (H2.1 and H3.1) compared to direct navigation, and that having a leftist ideology will increase selective exposure on Facebook compared to other forms of accessing news (H2.2).

Our study partially supports our expectations. Although we did not find support for H1 when we assumed homogenous slope effects, when we allowed the effects of origins (the slope) to vary across individuals, we found that direct navigation increases selective exposure, even though the effect is small (at most, it increases the probability of selective exposure by approximately 3 percent). As for the role of specific platforms, we found that Facebook has no direct effect in selective exposure, but that Google is a significant actor shaping online selective exposure. We estimated that Google reduces selective exposure at most 9 percent, which is a considerable effect. In line with other studies (e.g., Fletcher and Nielsen 2018), we show that Google not only does not lock people into filter bubbles but increases media consumption across ideological lines.

However, our most interesting results probably concern the cross-level interactions. Although we did not find political interest to moderate origins and selective exposure, we found a strong interaction effect between origins and ideology. In particular, we found ideology and news-referred platforms (i.e., Facebook and Google) to shape selective exposure in opposite ways—Facebook increased the probability of selective exposure.
exposure among left-leaners and decreased it among right-leaners, and Google decreased the probability of selective exposure among left-wing ideologues and increased it among right-wing ideologues. These results suggest at least two explanations.

The first explanation relates to group behavior: Left- and right-leaners might behave differently as active information seekers in these platforms. For example, conservatives might use Google to search directly the newspaper they normally read, while liberals might rely on search terms more often. In social media, left-leaners might be more active news consumers than right-leaners, prompting greater selectivity as consumption intensifies (Guess 2016; Prior 2013).

The second explanation relates to information availability in these platforms and suggests that neither Facebook nor Google might be neutral in information retrieving (e.g., Haim et al. 2018; Hancock et al. 2018; Puschmann 2017). There is some evidence that Facebook might be biased to the left in content retrieving because left-wing individuals are overrepresented on this platform (Mellon and Prosser 2017), and they are also more active on it (Anderson and Jiang 2018). Also, some works are beginning to discuss and study the neutrality of Google in content retrieving, suggesting that it might have an ideological bias (e.g., Haim et al. 2018; Hancock et al. 2018; Puschmann 2017). Although evidence is scarce and it is (still) too soon to know, the strong interaction effect that we find between ideology and news-referred platforms suggests that these platforms might not be content neutral, something that, given the increasing popularity of these platforms as news levers and their potential effects on voting behavior (Epstein and Robertson 2015), deserves much more attention from research in the future.

Finally, we found that political interest and ideology are strong direct predictors of selective exposure. In line with other studies (e.g., Barberá et al. 2015), we found that ideology matters when accounting for patterns of news media exposure and that left-leaning news consumers are significantly more likely to consume media from the other side of the political spectrum than right-leaning individuals. In particular, we estimated a probability change of at most 20 percent due to ideological leaning. In line with previous studies (e.g., Stroud 2011), we also found that political interest increases—not decreases—the probability of selective exposure. This suggests that politically interested people not only are more capable of identifying the content of political messages ex ante, but are also more likely to rely on this information (i.e., partisan cues) to select political information.

This study suffers from several limitations. First, exposure is measured at the media level, not at the content level. Second, exposure is observed only through desktop computers, leaving out news access through mobile devices. Although access to news through mobile devices has experienced the most important increase in recent years, desktop computers continue to be the leading device in online news consumption (Newman et al. 2018). Finally, measurement of news media exposure is limited to the top 42 most-visited outlets, according to Alexa, which might leave out consumption of niche media. In our favor, it is worth noting that in line with other studies (e.g., Gentzkow and Shapiro 2011), we found that, among our panelists, those who consumed niche media were also consumers of mainstream media.
In spite of these limitations, this study addresses the contingencies of online selective exposure through the use of unobtrusive data and a multilevel approach. It shows that more nuanced approaches are needed to tackle the complexity of selective exposure and opens a promising line of analysis for uncovering some of the contingencies influencing selectivity online.

**Declaration of Conflicting Interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this paper has been funded by the Spanish Ministry of Economy and Competitiveness under grant CSO2013-47082-P.

**ORCID iD**

Ana S. Cardenal https://orcid.org/0000-0002-1540-8004

**Supplemental Material**

Supplemental material for this article is available online.

**Notes**

1. In what follows, we use the term choice to refer to voluntary exposure and the term algorithmic filtering (or algorithmic curation) to refer to involuntary exposure (e.g., Dubois and Blank 2018; Zuiderveen Borgesius et al. 2016).

2. Users reporting to get news from Facebook dropped in the United States from 47 to 39 percent in the last year (Newman et al. 2018).

3. To guarantee a maximum of useful information, we aimed at selecting people for our panel who consumed a minimum of online information. We used four filter questions, ranging from less restrictive to more restrictive. The less restrictive question asked people if they followed current events or read news online. A positive answer in any of the four questions was enough to include the individual in our panel. Only four people were filtered out, which proves the inclusiveness of our questions and rules out potential problems of selection bias. The questionnaire can be accessed here: https://www.dropbox.com/sh/61j9own413ho1oq/AACzPaFJ5lq8yLJrPLBA0gsa?dl=0.

4. We excluded Catalonia from our study to keep the analysis as simple and tractable as possible. In contrast to Spain, where all issues collapse to a single left–right dimension and politics can be characterized as one-dimensional, in Catalonia (as well as in the Basque Country) the issue of politics is multidimensional because a second dimension related to national identity issues and the territorial organization of state also structures political conflict. Adding dimensions would have unnecessarily complicated a study about segregation. Both, size and politics informed our decision to exclude Catalonia. Catalonia represents a significant portion of the Spanish population (20 percent) and, during the period of study, its politics were dominated by national identity issues, as testified by the rise of the Catalan
pro-independence movement, which increased its support from 14 percent in 2006 to 41 percent in 2015 (Muñoz and Tormos 2015).

5. The Supplementary Information file is available on the journal’s website.

6. We selected the top 42 most-visited news outlets reported by Alexa because they represent an important percentage of the total audience for news according to both Alexa and ComScore. The strong correlation (.906) between media’s position in both these lists—Alexa and ComScore—proves the accuracy of our media sample. Media outlets that have not been included in this list are niche news providers that yield relatively low audience-reach figures according to ComScore. Our list of 42 media includes 99.85 percent of all reported visits to online news outlets in our sample, which adds to its exhaustiveness.

7. Table 1 shows that the number of people in our panel accessing news outlets from the two largest referral-platforms, Google and Facebook, amount to 7 and 5 percent, respectively. These numbers clearly stand in contrast to other research and market studies (e.g., Digital News Report), which using survey data report higher percentages of people getting news from Facebook (48 percent) and search engines (25 percent) in Spain (Newman et al. 2017). Yet they are very close to what other studies based on similar designs (i.e., web-tracking data) find for the United States (e.g., Flaxman et al. 2016). This disparity may be explained because the number of people who see headlines in their News Feed in Facebook is certainly much higher than the number of people actually clicking on the headline to access the outlet and read the whole story. Studies based on survey methodology probably tap the first type of behavior, while studies based on web-tracking data and observed exposure tap the second. Our figures thus provide a more accurate description of “getting news from Facebook or Google” since our methodology is probably a more stringent test of actual news consumption from distributed platforms.

8. Note that session overlapping is only a necessary condition for referral, not a sufficient one—for example, the user might have the previous session open without this necessarily implying that he or she has arrived to the next site through this (previous) one. Thus, even though session overlapping is the best measure at hand for referral, it is an imperfect one because we cannot be certain that a site—even if there is session overlapping—has actually worked as a referral for the next site.

References


Valeriani, Augusto, and Cristian Vaccari. 2015. “Accidental Exposure to Politics on Social Media as Online Participation Equalizer in Germany, Italy, and the United Kingdom.” New Media and Society 18 (9): 1857–74.


Author Biographies

Ana S. Cardenal is an associate professor in the Department of Law and Political Science at the Universitat Oberta de Catalunya (UOC) and a visiting professor at the Universitat Pompeu Fabra (UPF).

Carlos Aguilar-Paredes is an assistant professor in the Department of Information Science and Media Studies at the University of Barcelona (Spain) and vice dean of the Faculty of Information and Communication at the same university.
Carol Galais is a postdoctoral Fellow (Ramón y Cajal) in the Department of Public Law and Political Science at the Universitat Autònoma de Barcelona (Spain). She has also been a post-doctoral fellow at the UOC and Université de Montreal.

Mario Pérez-Montoro is a professor in the Department of Information Science and Media Studies at the University of Barcelona (Spain) and director of the PhD Program in Information and Communication at this University.