Local News Online: Aggregators, Geo-Targeting and the Market for Local News[*]

Lisa M. George[†]

Hunter College and the Graduate Center, CUNY

Christiaan Hogendorn[‡]

Wesleyan University

April 24, 2019

Abstract

We examine how placement of geo-targeted local news links on Google News affected local news consumption. Using a sample of news visits by 36,876 US households before and after a Google News design change, we find that aggregation increased the level and share of local news consumed online. Magnitudes are modest: consumption of local news among heavy Google News users rose by 25% after the redesign, but from a low base. We find that geo-targeting increased weekly consumption variety and reduced concentration among local outlets. However, we find no evidence for an aggregate increase in variety, suggesting that aggregators play a limited role in product discovery.

JEL Codes: D22, L13, L82

Keywords: Communications, Media, Internet, News, Platforms, Advertising

---

[†]Lisa M. George, Department of Economics, Hunter College and the Graduate Center, CUNY, 695 Park Ave., New York, NY 10065: lisa.george@hunter.cuny.edu .

[‡]Christiaan Hogendorn, Economics Department, Wesleyan University, 238 Church Street, Middletown, CT 06459: chogendorn@wesleyan.edu .

1

1 Introduction

By any measure, aggregators have transformed the market for news. On the demand side, survey evidence indicates that more than half of consumers access news through search, social, and aggregator platforms each week, with over 10% through Google News alone (Newman et al., 2017). For suppliers, Pew reports that Google News and Google Search account for 30% of traffic to major US news outlets (Heimlich, 2011).

Despite the growing role of aggregators, their influence on news markets remains poorly understood, and highly controversial. News industry rhetoric blames digital platforms for loss of advertising revenue and financial distress of newspapers. Technology firms emphasize the value of customization and discovery in the face of seemingly unlimited digital content. From an industrial organization perspective, the broad welfare trade-off in this debate balances lower search costs for consumers, which become increasingly salient as content proliferates online, against weakened incentives for suppliers to invest in news. However it is also the case that in reducing search costs, aggregators increase competition across outlets for reader attention. Heightened competition can incentivize quality, potentially offsetting effects of lost advertising revenue.[1]

Economic theory has made some progress in structuring these welfare trade-offs. George and Hogendorn (2012) show how news aggregators increase switching behavior of news viewers, increasing the number of sites visited and generally benefitting news outlets with few loyal viewers while hurting those with many loyal website visitors. Rutt (2011) studies price and quality competition in the presence of news aggregators. Alaoui and Germano (2015) studies how time constraints impact news consumption in aggregation and search environments. In Dellarocas et al. (2013), aggregators intensify competition between content providers, who improve the quality of their news in order to obtain the traffic brought by the aggregators. Hong (2014) develops a model in which some users trade off the convenience of news aggregators for a decrease in quality, creating herding behavior in the process. Jeon and Nasr (2016) model content provider incentives to specialize in specific topics so that they can attract more traffic. They find that although the quality of articles improves, the effect on profits depends on the relative strengths of the "business-stealing" and the "readership-expansion" effects. In the content bundling model of de Corniere and Sarvary (2017), consumers access news directly through their social media feed as well as through a newspapers website. They

---

[1]The role of aggregators in this sense is related to the IO literature on experts and other intermediaries whose presence introduces agency problems that may or may not be offset by heightened competition. See (Robles-Garcia, 2019) for a recent example.

find that even though the platform shares revenues with newspapers, newspaper profits fall. The impact on the quality of news depends on the personalization level of the social media feed. These theoretical models all contrast positive and negative effects of news aggregators on news outlets, and most of them find that the magnitudes of these contrasting effects depend on the size and/or reader loyalty of the news outlet. Thus they raise questions about preference parameters that need to be explored empirically to find the net effect on different types of news outlets.

Empirical work thus far offers only limited insights. From the outlet perspective, Chiou and Tucker (2017) suggest complementarity between aggregator links and outlet visits: when Google briefly removed Associated Press (AP) content from Google News during a contract dispute, online visits to AP articles fell. But the AP and other wire services have limited history and reputation as stand-alone news outlets, increasing the importance of search and aggregation. The positive link between aggregation and outlet visits thus does not readily generalize to the larger news market. On the consumer side, Athey and Mobius (2012) show that users who adopt a localization feature of Google News in France increase visits to local news outlets, but over time most additional local news consumption derives from increased use of Google News. Again, it is difficult to generalize from this result because the opt-in nature of localization leaves the broader question of complementarity unresolved. More closely related to this study is Athey and Mobius (2017), which shows that the shutdown of Google News in Spain reduced news consumption by a set of treated users by about 20% and that this that was not fully replaced by outside sources.

We know even less about the impact of news aggregators on the supply and demand for variety. A substantial literature in digital entertainment markets explores how the "long-tail" of niche products interacts with popularity of superstars. Optimism about increased variety at the start of the Internet era was followed by more recent evidence of superstar effects (Napoli, 2016). But there is little research on the long tail in news, and fundamental differences in product characteristics limit generalization from other areas to news, especially because news quality is not fully observable. While news aggregators almost by definition increase variety available to consumers at any particular moment, the tendency of aggregators to promote niche versus popular outlets over time has not been studied. The issue is particularly salient with local outlets, as they typically report news poorly suited to a mass audience but with potentially high demand among a narrower set of readers.

The role of aggregators in news markets is highly relevant to current policy debate,

3

with proposals under consideration in both the US and EU to curtail the ability of platforms to excerpt, index, and link content.[2] Policy toward aggregation has in the past been framed in the language of intellectual property, especially fair use. However, more recent regulatory attention has emphasized privacy and allowable uses of consumer data, especially in the EU. Since news aggregators rely on individual consumer data for targeting, these proposals would substantially impact the ability of platforms to deliver tailored content. Facts related to the benefits and costs of privacy restrictions in news markets are scant.[3]

This paper exploits a major redesign of the Google News page to study the impact of aggregation on news consumption directly. On July 1, 2010, Google placed a permanent strip of geo-targeted local news links onto the Google News front page. Local links were chosen using consumer IP addresses and could not be removed through customization. This redesign marks the first wide-scale use of consumer data to target content on a news platform. We use the redesign to study the impact of aggregation on local news consumption, the share of attention devoted to local outlets, and consumption diversity. Our goal is to establish new empirical facts that can guide theory and policy.

Using a a sample of news visits by 36,876 US households in the six months before and after the design change, we find that adding geo-targeted links increased the level and share of visits to outlets local to households. The increase is positively correlated with intensity of Google News use before the redesign. Magnitudes are modest: local visits increase about 25% for heavy Google News users, but from a very low baseline. But contrary to many expectations, we find no evidence that consumers substituted away from direct outlet visits. We also study the impact of geo-targeting on consumption diversity, finding that households increased the number of local outlets visited each week. Concentration ratios among local outlets also fell after the redesign. However we cannot reject the hypothesis of no overall increase in consumption variety at the outlet

---

[2]EU copyright restrictions moving through the regulatory process are particularly sweeping, see https://www.cjr.org/the_new_gatekeepers/eu-copyright-law.php for a current overview. In the US, the FTC compiled a set of policy recommendations for supporting the news industry, see https://www.ftc.gov/sites/default/files/documents/public_events/how-will-journalism-survive-internet-age/new-staff-discussion.pdf.

[3]The EU's General Data Protection Regulation affects not only news aggregators but all advertiser funded news outlets, leading to speculation that privacy laws might ultimately strengthen the role of technology firms with deep pockets in competition with traditional outlets, for example https://www.nytimes.com/2018/04/23/technology/privacy-regulation-facebook-google.html. For preliminary coverage of the impact of the GDPR on the news industry see http://www.niemanlab.org/2018/08/more-than-1000-u-s-news-sites-are-still-unavailable-in-europe-two-months-after-gdpr-took-effect/.

4

level in the months after the redesign, suggesting that aggregators play a limited role in product discovery, at least in local markets.

Our results are consistent with theories that model intermediaries as reducing search costs, as consumers devote more attention to aggregated outlets after geo-targeting. In this sense our results relate to a literature on digitization that highlights how technology alters competition by re-allocating attention.[4] Our result that aggregators increase consumption diversity on a weekly basis suggests that platforms heighten competition at the article level, with readers choosing the most appealing content over a larger set of outlets than before the re-design. However the fact that aggregation increases high-frequency variety rather than a net increase in sources over time indicates that aggregators show limited potential to promote product discovery, at least in local markets, but instead increase competition across established brands.

From a welfare and policy perspective, our results point to benefits from news aggregation generally and targeted local news in particular. It is well established in the media literature that in the presence of fixed costs or other non-convexities, firms cater to the preferences of large groups. In the absence of geo-targeting, this implies news platforms would emphasize major markets over small ones.[5] Our results suggest that geo-targeting ameliorates large-market bias, consistent with long-standing communications policy goals to support localism in news media. From a welfare standpoint, shifts to newly-aggregated outlets should be efficient, made possible by reductions in transaction costs. Thus our results suggest a clear welfare gain from consumer data against which privacy concerns must be balanced. However at the same time, we find no evidence that aggregation leads to product discovery, at least with respect to local outlets. The result suggests that aggregation heightens competition among established brands rather than promoting policy goals of diversity.

Our data also allow us to provide some evidence on the impact of aggregators on outlets. We find a modest association between aggregation and overall traffic, with a headline on Google News increasing visits by 3.4% overall compared to days when the outlet is not listed on the aggregation page. We find that the redesign alters the composition of the audience for local media, increasing the share of traffic from the

---

[4]For example, George (2008) showed how the spread of the internet altered the composition of the audience for traditional local newspapers, pulling younger and more educated readers out of local newspaper markets. George and Waldfogel (2006) showed how national expansion of the *New York Times* made possible by satellite printing attracted highly educated readers away from local newspapers, with consequences for local media markets and also for local voting George and Waldfogel (2008).

[5]We indeed see this in our raw data, see section 3 for details.

5

home market. As with household results, magnitudes are modest, suggesting that geo-targeting provides limited benefit to local outlets.

The paper proceeds as follows. Section 2 describes the research design. Section 3 outlines construction of the data. Section 4 describe the empirical strategy. Sections 5 and 6 present results for households and outlets, respectively. Section 7 concludes.

2   Research Design

Google News is a free news aggregator founded in 2002 and launched officially in 2006. Since its inception, the site has offered links to international and national news headlines as well as soft news items such as sports and entertainment. Major news stories from regions of the US were historically included in a "national" section of the site. Until 2008, stories from US regions only appeared on the site if they met criteria for inclusion in the national section, and then they would be visible to all users. In February 2008, news tailored to user preferences became available for individuals willing to sign in, register, and opt-in for customization features.

On June 30, 2010, Google introduced the first comprehensive redesign of the Google News page since the start of the service. (A second re-design launched in May 2018.) The redesign altered the presentation of content and allowed additional customization for registered users. Most important for this study, the redesign added a strip of content along the right side of the page with a set of headline links *local to the user.* In contrast to the primary news page, news on this sidebar was customized to the individual user, with location identified automatically through the IP address. Unlike earlier personalization features, after June 30, 2010 local content was shown automatically and could not be removed through customization.

Figure 1 shows a screen shot of a Google News page on June 28, 2010 before the design change and Figure 2 shows a screen shot on July 2, 2010 just after the change. The June 28 page reveals an area where viewers can enter content preferences. This screenshot, scraped from the Internet Archive's "Wayback Machine" shows a non-customized page. The screenshot in Figure 2, taken four days later, reflects the new format. The page still includes personalization options for users who log into the site, but local headlines and links are now fixed on the right side of the page, shown with a large arrow added for clarity. The location of the scraping server is identified as San Francisco, and the local content includes two local news headlines from the *San*

6

Francisco Chronicle and one from the *San Jose Mercury News.*

The basic empirical strategy in this paper is to measure the effect of adding localized news links to Google News on the level and share of local news consumption by sample households. We also study the effect of the treatment on consumption diversity. Using a modified difference-in-difference approach, we examine whether households with heavier Google News use before the implementation of geo-targeting show greater changes in consumption after re-design. We consider alternatives to our baseline identification strategy that compare response to the treatment among heavy Google relative to heavy Yahoo users. We supplement our primary results with anlaysis of the impact of geo-targeting on local news outlets.

A limitation of our analysis is that our data do not directly distinguish news visits referred from Google News from those referred from Google search. We use several strategies to compensate for this limitation. To construct our primary independent variable, we estimate Google News use intensity as the share of news visits before the redesign that are both referred from Google and are directed to outlets present on Google News the day visited. We calculate the household-level measure over January - June, 2010, the months leading up to the re-design. We construct alternate measures that do not rely on this imputation, but instead categorize users according to referrals from Google versus Yahoo (news and search), which we measure directly.

Another potential limitation of our analysis is that the June 2010 redesign of Google News expanded customization features at the same time geo-targeted local news was added to the site. These opt-in features allowed users to prioritize content categories on the main page. To the extent that these changes increase the appeal of non-local content on the main pages after geo-targeting, they work against finding increases in the share of local news consumed after the redesign.[6]

The empirical strategy has features in common with Athey and Mobius (2012), who study the effect of a Google News design change on local news consumption in France. In addition to covering the larger and more complex US news market, a key difference in our study is that the Google News redesign placed geo-targeted news links in front of all users simultaneously, rather than as an "opt-in" feature. Studying this comprehensive change allows for a simpler estimation strategy than with the introduction of opt-in

---

[6]Google's customization checklist focused on the major (and non-local) categories on the page at the time: World, U.S., Elections, Business, Science/Technology, Entertainment, Sports and Health. Although it was possible for users to suggest other topics, which might have included local keywords such as sports teams, the customization overall emphasized non-local topics.

7

local content. The comprehensive redesign also allows us to consider how aggregators impact local news outlets, which is not possible with the opt-in rollout. More generally, experimental conditions for testing whether aggregated links substitute for direct visits are uncommon, and this broad-based change to Google News allows estimates of how aggregation affects the composition of the audience for local news.

We offer a few final points before turning to the data. Our measure of news consumption is based on direct visits to outlets. In this paper we do not consider headline browsing as informative activity on its own. While it might be argued that browsing headlines delivers utility in imparting information, the same behavior can be interpreted as the cost of searching for content that matches preferences. Separating transaction costs from what might be called "transaction benefits" is an interesting area for future research, but our data is not well suited to the task. We focus on article consumption as most consistent with theoretical treatment of news aggregators to date and most closely related to policy debate about the impact of news aggregators on outlet traffic.

3   Data

Our basic working data are counts of online visits to news domains by 36,876 households over 47 weeks in 2010. For each household each week, we record the total number of news visits as well as the number of news visits to outlets local to the household. We also record the number of unique outlets visited each week, both overall and local to the household.

We classify visits as local if (1) the outlet serves a city, county or metropolitan area; and (2) the household resides in the Metropolitan Statistical Areas (MSA) served by the outlet. The number of local news visits, the local visit share, and the number of unique local outlets are the primary dependent variables in the regressions.

Our primary independent (treatment) variable is the imputed share of visits referred through Google News by each household over the first six months of 2010, before the Google News redesign. We supplement our primary treatment variable with the measured share of referrals from Google and Yahoo (news and search) over this time period.

The time period of analysis is calendar year 2010, excluding the launch week for the redesign and the partial week at the start of the year. We also exclude weeks after December 1 due to a widely-reported change in Google algorithms made in response to

a *New York Times* story on manipulation of search rankings.[7] In section 5 we examine the robustness of our results to a more restricted time window.

The working dataset is constructed from several sources, discussed in turn.

3.1 Visit Data

We construct visit counts from the 2010 Comscore "Web Behavior Database." The database logs the complete browsing history for a nationally-representative sample of approximately 50,000 households (machines) that have opted to allow tracking of internet use. The database also records the referring domain, if any, for each browsing instance. We count each browsing instance in the database as a visit, then aggregate visits and Google referrals to the household-domain-day. For our primary specifications we further aggregate visits to the machine-week. For our outlet analysis we aggregate to the domain-day or domain-week. As a robustness check in working with count data we replace visit counts with a weekly indicator variable equal to one for each domain visited by each household and zero otherwise.

The raw data records the zipcode for each household. We merge zipcodes to 2010 census records to identify household MSAs. As detailed below, we classify visits as local if the MSA targeted by the media outlet matches the MSA of the household. Households living outside of MSAs are not included in the analysis.[8]

3.2 News Outlets

The raw Comscore data for 2010 includes 3.3 million top-level domains. We exclude domains visited fewer than 5 times in 2010, which reduces the pool of domains to 717,724.[9] From this set, we identify news domains, broadly defined, from several pro-

---

[7]Google provides information on the change at https://googleblog.blogspot.com/2010/12/being-bad-to-your-customers-is-bad-for.html. The original *New York Times* story can be found at https://www.nytimes.com/2010/11/28/business/28borker.html?_r=1&pagewanted=all.

[8]We find the correlation between the number of households in our sample in each MSA and the MSA population reported by the census to be high, over 0.99, indicating that our Comscore sample is geographically representative of the US population. We also find no evidence of attrition in our sample, though users are considerably more active early and late in the calendar year. Comscore data are collected at the machine level, so they may reflect behavior of more than one user in each household. The data do not capture computer use on mobile devices or computers accessed outside the home. These limitations are common to most sources of computer use microdata and largely affect the generalization of results rather than estimation. Section 5 includes robustness checks on the sample. De Los Santos et al. (2012) provide additional background and detail on the composition of the Comscore sample.

[9]A substantial portion of excluded domains are advertisements, misspellings, and corrupted links.

9

prietary databases and public sources. We include in our sample print and broadcast media outlets listed in *Burrelle's Media Directory* (2000 & 2005 edition), Bulldog Reporter's *MediaPro Directory* (2008) and the Newspaper Association of America directory (2010). We supplement these directories with information from the Radio Locator search engine and TV Station Index.

We also include in our sample a set of outlets that do not appear in these databases but that are linked on Google News at least once in the years 2002-2010. To produce the Google News list, we scrape archived pages of Google News hosted on the Internet Archives "Wayback Machine". Our final working data include 8,240 media outlets, 2,626 that have appeared on Google News at some point (1,319 in 2010 prior to the redesign) and 5,614 that have not appeared in our Google News scrape. About 800 outlets appear on Google News but not in directory sources. These are largely online entertainment, sports and technology sites.[10]

We classify each outlet as local or non-local using the same sources listed above. For local outlets, we define the home market as the MSA of publication. As a check on our data and to fill in gaps, we calculate for each outlet the MSA from which the outlet receives the most traffic and the share of total traffic originating from the home MSA. Local media in our data are primarily daily and weekly newspapers plus radio and television stations, but also include local magazines, local web sites and local professional publications.

Our primary sample includes an expansive set of media outlets, including traditional newspapers, broadcast news stations, and specialty sites covering entertainment, sports, technology and other topics. Since our aim is to characterize how aggregators allocate consumer attention, we prefer a broad definition of news as this best aligns with the spectrum of outlets linked on the Google News pages. However, as a robustness check we exclude predominantly "soft news" outlets from our sample. We classify outlets as soft if more than 75% of Google News links to the outlet appear in the entertainment, technology, sports or health sub-sections of the aggregator site.

We note that the raw site visit data available for this study include only top-level domains. This limitation means that multiple media outlets co-hosted on the same domain cannot be distinguished. For example, local television and radio stations sometimes share a web site. Since co-hosted sites are local to the same market and household

---

[10]In supplementing our set of outlets from Google News, we exclude about 100 highly-visited domains that appear infrequently on the aggregator site. Excluded domains are largely platforms (*about.com*), media conglomerates (*tvguide.com, gamespot.com*) and sports organizations (*nfl.com, mlb.com*).

visits are aggregated over domains, the limitation has minimal impact on measurement of our dependent variables. A more challenging problem is distinguishing Google News referrals from Google search referrals for our independent variables, discussed in the subsection on intermediation, below.

3.3 Intermediation

The primary independent variable in our study is the estimated intensity of Google News use for each household in the first half of 2010, the months leading up to the June 30 redesign. We expect households with higher Google News usage before the redesign to show the largest effects after implementation.

Because we observe only top level domains, we cannot directly distinguish between referrals from Google News and Google search. We instead measure intensity of Google News use indirectly by linking Google referrals in our visit data each day to outlets appearing on Google News in our scrape data on corresponding days. Specifically, we count a visit as a Google News visit if: (1) the visited domain appears on Google News the day of the visit; and (2) the visit is referred by Google. For example, a visit to the *Atlanta Journal Constitution* referred by Google on a day the newspaper appeared on Google News would be coded as a Google News visit and on other days would not. A visit to the *Atlanta Journal Constitution* that was not referred by Google would not be classified as a Google News referral whether or not the paper appeared on Google News that day. We calculate intensity of use for each household as the ratio of estimated Google News visits to all news visits over the entire pre-implementation period, January through June 2010.

Our data on the composition of the Google News site is compiled from the scrape of headlines, domains and categories described above. We identify for each domain each day whether or not the outlet appeared on the front page of Google News. We merge this to visit data for each household each day and aggregate over the total six month period January - June 2010, calculating the share of all news visits to Google News domains over this period for each household.[11]

---

[11]Our scrapes are typically time-stamped early in the morning, with an average of 7:00 am. Some outlets are replaced on the page later in the day, introducing measurement error into our calculation. We observe multiple scrapes on some days, and from this estimate that 45% of outlets appearing first thing in the morning are replaced over a six hour period. Averaging over the six month period helps limit the impact of measurement error in our treatment variable.

3.4 Sample Statistics

We summarize our working data in several tables and figures. We begin with information on the supply side, summarizing links on Google News in the months before the redesign. We then turn to demand information in our household sample, first summarizing news visits at the domain level then characterizing household behavior more generally.

Figure 3 reports the supply of links on the Google News site for the 25 most-linked domains on the aggregator platform from January through June 2010. The left panel reports the total number of links for each outlet and the right panel reports the number of days each domain appears on the Google News page. We restrict attention to the months before the re-design, so the total number of possible days is 181. A few major news outlets appear on Google News almost daily (the *New York Times, CNN, Wall Street Journal, Washington Post, Reuters*). Frequency falls for sites below the top 10. The number of links is skewed toward the largest outlets, with a long tail. The top 25 outlets receive 68% percent of links on the Google News page in our sample.

Domains marked with stars in Figure 3 are classified as "soft news" outlets that we exclude from some robustness checks. We define soft news outlets as non-local outlets with more than 75% of Google News links in the categories of health, sports, entertainment and technology.

Table 1 reports distributional details for the top 25 domains illustrated in the figure as well as for the whole sample. The top row in the lower panel shows the total number of scraped links: we scrape 20,038 links to 1,319 outlets in the first half of 2010, 13,585 of which are from the 25 most linked domains. The average number of days on Google News for top outlets is 136 (of 181) and the average number of links is 542. For the full sample, the average number of days on Google News is 6, with an average of 15 links.

Turning to our household sample, figure 4 lists the 25 most-visited outlets in our data. The top bar of the figure shows mean weekly visits to each outlet while the lower bar reports total Google News links to the outlet. The asterisks in figure 4 denote soft news outlets that we exclude from robustness specifications. There is modest overlap between the most-linked and most-visited outlets in figures 3 and 4, especially for hard news.

Distributional information for the most-visited outlets in 4 is reported in the top third of table 2. Information on the full sample is reported in the lower third of the table. We observe a total of 1,309,814 news visits by 36,876 households to 7,742

domains over the 24 weeks of 2010 before the redesign. Outlets in the top 25 receive 779 visits per week on average. Visits are highly skewed toward the most popular outlets (especially *CNN*): the median number of visits among the top 25 is about 500 visits per week. For the full sample, the average domain receives 7 visits (median 0.63) visits per week. Overall, household demand is less concentrated than the supply of links, with the top 25 domains receiving about a third (36%) of all news visits.

The lower bars in figure 4 report the total number of Google News appearances over the 48 weeks for each outlet. There is a positive correlation (0.55) between weekly visits and weekly appearances on Google News, however the figure makes clear that direct visits are important for many outlets, as the relationship is far from monotonic. Broadly speaking, the ratio of outlet visits to Google News links offers a measure of brand awareness or competitiveness, with the ratio higher for *CNN* than for the *Washington Post* or the *New York Times*.

Table 2 also reports the local share of visits to outlets. The middle portion of the table reports visit statistics for local outlets alone, while the lower portion covers the full sample of local and non-local outlets. Average weekly visits are lower for local outlets than non-local ones, with an average of 5 visits per week (median 0.54 visits). There is substantial variation across outlets. The overall share of visits to local outlets from within the MSA is about 40%.

To better illustrate localism in our data, we summarize average weekly visits to the three most visited local outlets in the largest metro areas in our household sample, shown in table 3. The first three data columns report total visits, local visits and the local visit share. The *New York Times* is the least "local" of the local outlets in our sample by a considerable margin, with a share of visitors from the NY metropolitan area of 22%. This is about half of the mean (and median) local share of 40%. In most markets, the outlets with the most local visits are major urban daily newspapers, but in some markets local radio or television station websites make the top three, such as WFAA in Dallas. In general, the local visit share for radio and television stations is considerably higher than for print media, with many broadcast sites reaching shares of over 90 percent.[12]

The last column in table 3 reports a concentration ratio, calculated as the share of all visits in the metro area to one of the top 3 outlets. The average share for the three

---

[12]Some sites in our data host two or more media outlets that span metro areas. This is the case for Alabama.com, which hosts the *Birmingham News* and the *Huntsville Times*. We link the domain to the largest media markets and confirm that our results are robust to exclusion of multi-market domains.

largest outlets in the top ten metro areas is 58%, but is substantially higher in smaller markets. The overall ratio is 88% on average. We consider the effect of geo-targeting on concentration in section 6.

The tables and figures above focus on outlets. Table 4 summarizes household behavior, reporting weekly news consumption patterns for the 36,876 sample households. Households visit 1.89 news outlets per week on average, with 0.48 visits to outlets local to the household. Consumption is skewed toward more active users, with the median household not consuming news in a given week and the 95th percentile household making seven visits on average each week. Households visit an average of 1.45 unique domains per week, again with skew toward more active users. The bottom rows of the table show that individuals are active 43% of the 47 week period of the study.

Our final household table summarizes our treatment variables. The top row of table 5 reports the share of imputed Google News referrals at 7%. The most active users at the 95th percentile have an imputed share of Google News referrals of 33%. The second and third rows report combined news and search referrals from Google and Yahoo, used in alternative treatment and robustness specifications. The average share of news visits referred by Google is 33% (median 25%) and by Yahoo 11% (median 0).

Before turning to the empirical analysis, it is useful to consider time trends in the raw data. Figures 6 and 7 show weekly local news visits and the local visit share for households in the top quintile of imputed Google News use prior to the redesign relative to all other households. The vertical line marks the June 30, 2010 redesign, and we exclude the transition week from the figure. A first point of note on the figure is that Google News households make fewer visits to local news sites than the general population, but that the overall consumption pattern across the year is similar for the two groups. Google News households consume about 30% lower share of local news outlets than other households. Differences in the local visit share after the redesign are more marked, with the local share for heavy Google News households increasing in the second half of the year while declining for other households. The increase in the raw data appears to be modest in magnitude, a small increment off of a low baseline.

We can look more closely at trends by focusing on users likely to be similar in unobserved dimensions but for use of the Google News aggregator. Figures 8 and 9 repeat the analysis above with a comparison group that represents households in the top quintile of referrals from Yahoo. Overall, local news consumption is more closely matched for the two subgroups, but the local news consumption share for *Yahoo*

14

households is higher throughout the year. As in figure 7, the time trend in local news share in the second half of the year remains flat or declining for Yahoo users while increasing modestly for Google News users.

The next section outlines the estimation strategy for testing the patterns suggested by the graphs.

4 Empirical Strategy

The goal of the empirical analysis is to identify the effect of adding geo-targeted local news links to the Google News page on the local news consumption of sample households. We test the hypothesis that households using Google News more intensely *before* the redesign increase the number and share of visits to local news outlets *after* the design change. Our focus is on changes in news consumption within households to best control for unobservable household preferences.

Our baseline specification for estimation is a standard fixed effects formulation:

$$Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 X_i + \beta_3 Post_t * X_i + \gamma_i + \epsilon_{it} \tag{1}$$

where the dependent variable Y captures a local news consumption measure for each household i each week t. $Post$ is a treatment dummy set to 1 after local headlines were added to Google News and X is a treatment measure capturing intensity of Google News use before the redesign. Household fixed effects γ account for fixed, unobserved preferences of households. We cluster standard errors ϵ_{it} by MSA's, which we consider to be the level of the geo-targeting experiment.

In our primary specifications, we estimate equation 1 with a focus on three effects: the number of local news visits each week, the probability of a local news visit, and the share of visits to local outlets. The first two measures address the amount of local news consumed. The choice of modeling approach for count data is not straightforward. We prefer raw counts for ease of interpretation, but also present linear probability models for an alternative behavioral perspective and to account for skew in count data.[13] Our third specification considers the local visit share, which captures overall shifts in attention induced by changes in supply. We supplement our primary specifications with an analysis how geo-targeting affects the number of unique local outlets visited each week and over the entire post-design period of 2010.

---

[13]Results for negative binomial specifications are similar to reported results.

We identify the effects of the redesign from variation in imputed Google News use intensity prior to the redesign. This treatment variable is measured at the household level and does not vary over time. With household fixed effects, we cannot measure the direct effect of Google News use, β_2. However we can measure the interaction term, β_3, which reveals whether changes in local news consumption are greater for households with more intense Google News use before the redesign. The interaction term β_3 is thus the primary coefficient of interest in our estimation. In equation 1, β_0 reflects the household average of our local consumption measures. The coefficient β_1 measures the average change post-redesign.[14]

We also consider an alternate specification that focuses only on households that access news through intermediaries. We divide our household sample into a set of (treated) Google News users and a set of (untreated) Yahoo users, excluding other households. With this approach, we seek to identify the effects of geo-targeting by comparing local news consumption of two populations that are similar but for their choice of news platform. We define treated users to be households in the top quintile of imputed Google News intensity and bottom quintile of Yahoo referral intensity. We set as a control group households in the top quintile of Yahoo referral intensity and bottom quintile of *Google News* intensity. We exclude other users. For this specification, X in equation 1 becomes an indicator variable for the treatment group. The coefficient β_3 remains the estimate of interest, a measure of whether treated households increase local news consumption relative to the control.

5   Results

5.1   Local News Consumption

Table 6 reports our primary estimates of equation 1. The first column shows results for the number of local news visits, the second shows effects for the probability of a local news visit, and the final column reports the effect on the local share. Each specification includes household fixed effects, so that estimated coefficients are identified

---

[14]We have estimated all of our specifications with time as well as household fixed effects. Including weekly fixed effects has the advantage of controlling for time-varying patterns in local news consumption, but the disadvantage that we cannot estimate β_1. Since the coefficients of primary interest are essentially unchanged with inclusion of weekly fixed effects, we choose to report results without them so as to allow estimation and interpretation of β_1. All models are estimated in STATA with the *reghdfe* package (Correia, 2016) that produces correct standard errors in the case of fixed effects nested within cluster.

from changes within households. Since Google News use intensity is a static household characteristic, its overall impact is not identified. For compactness in presentation we exclude it from the table and focus on the interaction term. We present results without week dummy variables so that the post-re-design indicator is identified, having confirmed that coefficients of interest are essentially unchanged with their inclusion.

The coefficient estimate in the first row indicates that the level and share of local news consumption fell slightly in the second half of the year on average. The interaction term in the second row is the primary coefficient of interest, and the positive and statistically significant results across all columns indicates that households using Google News more intensely prior to geo-targeting see larger increases (or lower decreases) in local news consumption after the redesign. Magnitudes indicate that an increase in the share of intermediated visits of one standard deviation (12 percentage points) is associated with 0.031 additional local news visits per week after the redesign, or about 6.5%. One standard deviation in intensity is associated with an increase in the likelihood of a local news visit by 0.005, or 3.5%, and the local news share by 0.011, or 4%. A more substantial difference between the 5th percentile (no Google News referrals) and the 95th percentile (33%) is associated with local news consumption about three times this level, 18%, 9% and 14%, respectively. Stated more intuitively, a user at the 95th percentile of imputed Google News use of 33% would visit an additional 0.09 local news outlets each week after the redesign, or one local outlet every eleven weeks.

Table 7 reports results for our treatment-control specification. The top row indicates that local news consumption remains effectively unchanged for the control Yahoo group after geo-targeting, with small, negative and statistically insignificant coefficient estimates. The primary coefficient of interest is again the interaction term, which represents the change in consumption for the Google group relative to the Yahoo group after geo-targeting. The coefficients are positive and statistically significant in all columns. After the redesign, heavy Google News users visit 0.038 more local news outlets and are 0.013 more likely to visit a local news outlet each week. The share of local news consumption likewise increases by 0.024. Although the magnitudes are substantial in percentage terms (about 25% for local news visits, 12% for the probability of a local news visit and 15% for the local news share), baseline consumption for the treatment group is low, 0.158 local visits per week, so again the economic significance is modest. Since local news consumption is higher for the Yahoo group (about double the Google group), geo-targeting works to narrow the consumption gap between the two types of users.

17

The results above indicate that aggregation of local news links increased local news consumption overall and shifted the share of attention from non-local to local content. The result is consistent with theory that models aggregators as reducing search costs in consumption. In the next section we examine the impacts of these lower search costs on consumption diversity.

5.2 Source Diversity

To test whether aggregation impacts consumption diversity, we construct two new dependent variables measuring the number of unique news outlets and unique local news outlets visited each week for each household. We also measure aggregate variety as the total number of unique (local and total) news outlets visited by each household in 2010 in the six months before and six months after the implementation of geo-targeting. We interpret the weekly variables as measuring variety at the article level with the aggregate measure more closely approximating product discovery.

We re-estimate equation 1 with the weekly diversity measure. Table 8 reports results of our baseline specification with all households and table 9 reports the treatment-control specification. Coefficient patterns in both tables are similar. In particular, the interaction term in the second column in table 8 indicates that the number of unique local outlets visited per week after geo-targeting increases with Google News intensity, and the second column in table 9 indicates that weekly variety is higher after geo-targeting for Google News relative to Yahoo users. The baseline model shows that a one standard deviation increase in Google News intensity is associated with consuming 0.028 (4%) more local news sources per week. In the treatment-control specification, heavy Google News users increase the number of local outlets consulted by 18%. Results in the first column in both tables are not statistically significant, indicating that geo-targeting increases in the number of unique local news sources consulted each week without a net increase in diversity.

Columns (3) and (4) in both tables show results for aggregate diversity, the total number of news sources consulted January-June and July-December. Total diversity in column (3) declines after geo-targeting in both specifications, suggesting that aggregation overall might depress diversity. Results in column (4) show mixed evidence on aggregate local diversity, with positive but statistically insignificant (or borderline) effects of geo-targeting. Taken together, results indicate that aggregation of local news lowered search costs and increased variety at the article level, inducing more more fre-

quent visits to local outlets, but did not lead to product discovery. Aggregators appear to increase competition across major outlets at the article level rather than fostering interest in the long tail of media sources.

5.3 Robustness

To check the validity of our results, we consider several robustness and placebo specifications.

We first consider the sensitivity of our results to influential markets, to the time period of study, and to the classification of news outlets. For these checks, we focus on our primary interaction specification of equation 1. We again report results with only household fixed effects, but coefficients are essentially unchanged with addition of week fixed effects.

To consider influential markets, we estimate the model excluding the five markets with the most and five markets with the least news visits in our household sample. The dominant media outlets in the largest markets such as the *New York Times* and *New York Post* have a larger national audience (lower local visit share) than outlets in small markets, and tend to appear more often on Google News. The reverse tends to be true in smaller markets. Excluding the largest and smallest markets provides a check on whether shifts in local consumption are driven by potentially influential regions. Results are presented in table 10. Estimated coefficients are close to and slightly larger than estimates in table 6, with similar standard errors.

As a second check, we restrict attention to the 12 weeks before and after the implementation of geo-targeting. Narrowing the time period of study helps control for unobservable changes in search engines, aggregators or news platforms that might affect local news consumption. Results for short-run effects are shown in table 11. These results are about half the magnitude of our primary results, again with small standard errors. The magnitudes comport with intuition that immediate impacts are strongest but continue over time.

A third robustness check restricts attention to major hard news outlets. Specifically, we exclude visits to non-local outlets that have never appeared on Google News and exclude visits to outlets with more than 75% of Google News links listed in the categories of entertainment, sports, technology and health. (We do not adjust the treatment variables). This screening reduces the sample of news domains from 7,956 to 2,100. We report results in table 12. Again, coefficient estimates and standard errors

19

are similar to those in table 6.

The checks above indicate that our results are robust to the time period of study, exclusion of influential markets and to the classification of news outlets. We also want to consider the robustness of our empirical strategy by considering alternative approaches for identifying the effects of geo-targeting on local news consumption.

We consider first a specification that can better distinguish search effects from aggregator effects. In our primary specifications above we measure Google News intensity as the share of news visits before the redesign that are both referred by Google and made to outlets appearing on Google News the day of the visit. We can separately measure the share of news visits before the redesign that are referred by Google but linked to outlets that that do *not* appear in our scrapes. Adding this term to the basic specification allows us to control for search intensity. This measure is not a true placebo in the sense that Google News and search algorithms might be related in some ways. However, we would expect changes in household news consumption after geo-targeting to respond more strongly to imputed Google News intensity than to search intensity.

To test this, we include the measure of search intensity in our estimate of equation 1. As above, we cannot measure the direct effects of search intensity or Google News intensity in a specification with household fixed effects, but we can examine how these measures affect changes in consumption after the redesign. Results are reported in table 13. Coefficient estimates for Google News intensity after the start of geo-targeting are very close to those reported in table 6. For the number of local news visits and the probability of a news visit in columns (1) and (2), greater search intensity is also associated with greater local news consumption. The magnitude of the coefficient, however, is smaller than the effect of Google News intensity. For the local news share, higher search intensity is associated with a lower local news share after the re-design, though we cannot statistically reject a hypothesis of zero effect. Differences in the coefficients, reported in the last row of the table, are statistically significant.

As a second check of our identification strategy, we include the share of Yahoo referrals as a placebo in our primary specification. This can be seen as an extension of our treatment-control strategy in table 7 to the whole sample of households. Unlike in our search specification in table 13, we do not expect any relationship between Yahoo use intensity and local consumption after Google's introduction of geo-targeting. We report results in table 14. Coefficient estimates for Google News intensity are again very close to those in our original specification. Coefficient estimates for Yahoo use

intensity are not statistically different from zero in any column.

Our final placebo specification extends the intuition in table 7 and replaces our imputed Google News measure entirely. Instead, we include only the share of news visits referred by Google and Yahoo without distinguishing aggregation from search. This specification eliminates the noise associated with imputing Google News intensity and identifies effects based on overall reliance on the two platforms, one of which implemented geo-targeting and one which did not. As above, we expect that local news consumption will rise after implementation of geo-targeting with Google intensity and expect no relationship with Yahoo intensity. Results are shown in table 15. The coefficients indicate that both Google and Yahoo use is associated with a larger number of local visits after the redesign, but only Google use intensity is associated with greater likelihood of a local news visit or a greater local share. Differences are reported in the last row of the table.

Taken together, robustness checks in this section support our primary result that the implementation of geo-targeting increased the amount and share of local news consumption among platform users. The magnitude of effects are modest, especially given low initial local consumption.

6   Outlets

With evidence above that geo-targeting raises local news consumption, we continue with an investigation of how aggregation might affect outlets.

As a first step, it is useful to quantify the average effect of a Google News appearance on outlet visits before the advent of geo-targeting. To estimate this, we regress measures of outlet traffic each day on a dummy variable indicating whether or not the outlet appeared that day on Google News, including outlet and day fixed effects so that the indicator measures the impact of a Google News link within outlet. We study total outlet traffic, local traffic, and the local traffic share, transforming the traffic measures to logs because we expect that a Google News appearance impacts visits in proportion to outlet size.[15] We restrict attention to the months before the the redesign and to outlets appearing at least once on the Google News platform. We do not expect that estimated effects are entirely causal, since outlet appearance on Google News is related

---

[15]For households, variation in visits tended to be of a zero-one "visit or not" nature. Since there are many fewer outlets than households, outlet traffic data is more skewed, so we prefer a log specification for this reason as well. Results are very similar in magnitude and precision with a linear specification.

not just to fixed but also to time-varying outlet quality attributes that are not observed. Outlets breaking a major story, for example, are more likely to appear on Google News as well as other platforms and news feeds. But given how little is known about aggregator influence, the estimate offers a useful upper bound on potential aggregation effects.[16]

We report results of the preliminary specification in table 16. A link on Google News is associated with a 3.34% increase in total visits relative to days without an appearance.[17] There is no associated increase on the number of local visits or the local visit share before geo-targeting. The result suggests that aggregation expands the geographic reach of local news outlets.

To investigate the effect of geo-targeting on local news outlets, we consider how non-local traffic, local traffic and the local traffic share change after the redesign. Recognizing that a single difference specification might be biased by unobserved changes in consumption patterns not fully captured by time fixed effects, we model traffic in a way similar to our placebo regressions in table 15. Specifically, we study the relationship between traffic to local outlets and the share of outlet traffic referred through Google and Yahoo before the redesign. In other words, we hypothesize that outlets with a greater reliance on Google will see larger shifts in local traffic than outlets more reliant on Yahoo referrals, all else equal. Specifically, we estimate:

$$Y_{jt} = \beta_0 + \beta_1 Post_t + \beta_2 GR_j + \beta_3 YR_j + \beta_4 Post_t * GR_j + \beta_5 Post_t * YR_j + \gamma_j + \epsilon_{jt} \quad (2)$$

where Y_{jt} is a measure of (log) traffic to outlet j in week t. Independent variables GR_j and YR_j measure the share of outlet traffic referred through Google and Yahoo prior to the redesign. We specify outlet fixed effects γ_j and error ϵ_{jt}.

As with equation 1, we cannot estimate the overall effect of Google and Yahoo referrals β_2 and β_3 because these measures do not vary over time, but we can estimate the interaction terms of interest β_4 and β_5. Following our approach to the household estimates we present results without week fixed effects so that we can estimate β_1, but verify that coefficients of interest do not vary with their inclusion.

Results are presented in table 17. The top row show an overall decline in traffic after the redesign, but no change in the local share of outlet traffic. The second and

---

[16]Our measure puts an upper bound on average effects, but the data suggest heterogeneity in the impact of general versus specific categories (for example, "top news" versus "business").

[17]Calculated as $p = exp(b - 0.5 * V(b)) - 1$, following Kennedy (1981) and implemented in Stata using the SELDUM package (Ries, 2011).

third rows report estimates of β_4 and β_5. A higher Google referral intensity before the redesign is associated with larger increases in local traffic and larger increases in the local share of visitors after the redesign. Coefficients are statistically significant in all columns. Results show no relationship between Yahoo referral intensity and traffic or the local share.

An interesting result in table 17 is that non-local traffic as well as local traffic increases after the redesign, suggesting that changes to the Google News page might have improved click-through-rates overall. But, as with the household results, the magnitude of the estimated effects is modest, with the local share of traffic increasing by about 1 % for an average local outlet with 25% Google referral share.

As a last step, we offer some suggestive evidence on the competitive effects of geo-targeting. Given theoretical results in George and Hogendorn (2012) and Jeon and Nasr (2016) that aggregation can increase competition at the article level, we would like to know if aggregation of local news increases or decreases the concentration of attention in local news markets. We have very limited direct information on geo-targeted outlets because most of our Google News scrapes originate in San Francisco. However we directly observe geo-targeted links in San Francisco on 44 dates. We also observe local scrapes on 7 days in Indianapolis and 6 days in Philadelphia. We summarize linked outlets in table 18, which shows the outlets linked in the geo-targeted portion of Google News on these dates. The second column reports the outlet share of scraped links in each market. In San Francisco, geo-targeted links are highly concentrated among the *San Jose Mercury News* and the SFGate portal that hosts the *San Francisco Chronicle*. Philadelphia links are likewise concentrated in the Philly.com portal for the *Philadelphia Inquirer* and *Philadelphia Daily News*. Links in Indianapolis are less concentrated, with Google News links to six different newspapers and radio stations in the metro area.

To investigate competitive effects more systematically, we calculate a concentration ratio for the top 1, 2, and 3 most visited local news outlets in each metro area each week. We calculate the measure before and after the start of geo-targeting and regress it on an indicator for the geo-targeting period and market fixed effects. Results are shown in table 19. Coefficients are negative and statistically significant in all columns. Although we cannot rule out unobserved heterogeneity with single-difference specifications, these results do suggest that aggregation reduces concentration in local news markets. This result is consistent with our demand specification showing increases in consumption variety on a weekly basis.

The results in this section highlight that aggregation affects both household behavior and the composition of the audience for local outlets. Although the shifts recorded here are modest, over an entire market a small change in the local share can be enough to alter competition for both readers and advertisers.

7   Conclusion

This paper exploits a rare opportunity to study the effect of aggregation on the market for news. We find that adding geo-targeted links to the Google News site increases household visits to local outlets and the share of local content in household news consumption. The magnitude of estimated effects are modest, with increases in local news consumption for the heaviest Google News users of 25% from a very low baseline. Although magnitudes are not large, they are robust to the alternative classification of news domains, to the time frame considered, and to the exclusion of influential markets. They are also robust to alternative identification strategies that rely on direct rather than imputed measures of intermediation. Our results also likely reflect a lower bound on aggregate effects given continued growth of the Google News platform.

We also find evidence that aggregation increases the diversity of sources consulted each week but does not increase variety over an extended period of time. The result supports theoretical frameworks in which aggregators increase competition at the article level among top outlets. But the result also suggests that aggregators may play less of a role in product discovery than their owners sometimes argue.

Our findings are limited by the nature of available data, which provides only top-level domains. This precludes us from precisely measuring the intensity of Google News use for sample households, or for studying the effects of targeting on Google News use itself. We also consider the early years of the platform, whose early users might have different news preferences than the general population. Although our results are robust to alternative identification approaches, much could be learned from further analysis with data from technology firms. Many interesting avenues for future study remain.

References

Alaoui, L., and Germano, F. (2015). "Time Scarcity and the Market for News." Working Papers halshs-01251522, HAL.

24

Athey, S., and Mobius, M. (2012). "The impact of news aggregators on internet news consumption: The case of localization." *Stanford GSB Working Paper.*

Athey, S., and Mobius, M. (2017). "The impact of news aggregators on internet news consumption: The case of localization." *Stanford GSB Working Paper.*

Chiou, L., and Tucker, C. (2017). "Content aggregation by platforms: The case of the news media." *Journal of Economics & Management Strategy.*

Correia, S. (2016). "Linear models with high-dimensional fixed effects: An efficient and feasible estimator." Tech. rep., working Paper.

de Corniere, A., and Sarvary, M. (2017). "Social media and the news industry." *SSRN Working Paper 3049358.*

De Los Santos, B., Hortasu, A., and Wildenbeest, M. R. (2012). "Testing models of consumer search using data on web browsing and purchasing behavior." *American Economic Review*, *102*(6), 2955–80.

Dellarocas, C., Katona, Z., and Rand, W. (2013). "Media, aggregators, and the link economy: Strategic hyperlink formation in content networks." *Management Science*, *59*(10), 2360–2379.

George, L. M. (2008). "The internet and the market for daily newspapers." *The B.E. Journal of Economic Analysis & Policy*, *8*(1), 1–33.

George, L. M., and Hogendorn, C. (2012). "Aggregators, search and the economics of new media institutions." *Information Economics and Policy*, *22*(1), 40–51.

George, L. M., and Waldfogel, J. (2006). "The new york times and the market for local newspapers." *American Economic Review*, *96*(1), 435–447.

George, L. M., and Waldfogel, J. (2008). "National media and local political participation: The case of the new york times." In R. Islam (Ed.), *Information and public choice : from media markets to policy making*, 33–49, World Bank Washington, D.C.

Heimlich, R. (2011). "Fact tank: Google drives news." Tech. rep., Pew Research Center, available http://www.pewresearch.org/fact-tank/2011/06/02/google-drives-news/ (viewed January 31, 2019).

Hong, S. (2014). "A simple model of news aggregators, information cascades, and online traffic." *SSRN.*

Jeon, D.-S., and Nasr, N. (2016). "News aggregators and competition among newspapers on the internet." *American Economic Journal: Microeconomics*, *8*(4), 91–114.

Kennedy, P. E. (1981). "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations [The Interpretation of Dummy Variables in Semilogarithmic Equations]." *American Economic Review*, *71*(4), 801–801.

Napoli, P. M. (2016). "Requiem for the long tail: Towards a political economy of content aggregation and fragmentation." *International Journal of Media & Cultural Politics*, *12*(3), 341–356.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., and Nielsen, R. K. (2017). "Reuters institute digital news report 2017." Tech. rep., Reuters Institute for the Study of Journalism, available https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf (viewed January 31, 2019).

Ries, J. (2011). "SELDUM: Stata module to transform indicator variables coefficients in semilog model." Statistical Software Components, Boston College Department of Economics.

Robles-Garcia, C. (2019). "Competition and incentives in mortgage markets: The role of brokers." Tech. rep., London School of Economics Working Paper, available https://drive.google.com/file/d/1brHu_96DVphsN5NxWGbDVLSNJcv1ZBt9/view (viewed January 31, 2019).

Rutt, J. (2011). "Aggregators and the news industry: Charging for access to content." *Working paper*.

26

Table 1: Google News Domains, January - June 2010

| | Top 25 Outlets | | | | |
	Mean	SD	p5	p50	p95
Total Scraped Links (13,585)					
Google News Links	542.28	464.30	146.00	284.00	1270.00
Google News Days	135.60	33.62	93.00	129.00	181.00
Soft News Ratio	0.47	0.20	0.11	0.47	0.76
Observations	25				
	Full Sample				
	Mean	SD	p5	p50	p95
Total Scraped Links (20,038)					
Google News Links	15.19	97.64	0.00	1.00	38.00
Google News Days	6.38	21.15	0.00	1.00	30.00
Soft News Ratio	0.60	0.42	0.00	0.74	1.00
Observations	1,319				

Notes: Top panel reports the number average number of days the top 25 outlets appear on the Google News page and the number of links to each outlet over 181 days January-June 2010. The lower panel reports averages for all outlets over this period. See text for details.

27

Figure 1: Google News Front Page Before Redesign, June 28, 2010

Web Images Videos Maps News Shopping Gmail more ▼

Google news

Search News
Advanced news search

U.S.

Top Stories
Starred
World
U.S.
Business
Sci/Tech
Entertainment
Sports
Health
Spotlight
Most Popular

All news
Headlines
Images
YouTube News
Google Fast Flip

Supreme Court rules against Christian group that bars gays ☆
Reuters - James Vicini - 33 minutes ago
WASHINGTON (Reuters) - A university can legally deny recognition to a Christian student group that bars gays and nonbelievers, the US Supreme Court ruled on Monday in a case that pitted anti-discrimination principles against religious freedom.
Supreme Court rules against UC student group that refused to admit gays Los Angeles Times
Court: Christian group can't bar gays, get funding The Associated Press
Wall Street Journal - MiamiHerald.com - CNN - CBS News
all 319 news articles » ✉ Email this story

Supreme Court Rules That Gun Rights Apply to Local Laws ☆
New York Times - Adam Liptak - 55 minutes ago
The Supreme Court ruled on Monday that the Second Amendment, which forbids Congress from infringing the right to keep and bear arms, applies to state and local governments as well.
Supreme Court extends rights of gun owners Los Angeles Times
Supreme Court: Second Amendment rights apply across US Christian Science Monitor
The Associated Press - Newsweek - Atlanta Journal Constitution - Chicago Tribune
all 1,050 news articles » ✉ Email this story

Kagan Senate hearings begin ☆
USA Today - Kathy Kiely - 1 hour ago
By Jack Gruber, USA TODAY By Kathy Kiely, USA TODAY WASHINGTON - For Elena Kagan, adroit Ivy League super-achiever, it's the toughest oral exam she's ever faced.
Kagan Promises Impartiality as Hearings Open New York Times
Elena Kagan Faces Republican Critics as Supreme Court Hearings Begin ABC News
Wall Street Journal - Baltimore Sun - The Atlantic - CBS News
all 3,648 news articles » ✉ Email this story

Recommended for you »
Sign in to add a section with stories recommended for you, by using search history.

World »
Total says it has halted petrol sales to Iran ☆

On Top Magazine

Economist (blog)

msnbc.com

Justices Take Broad View of Business Method Patents
New York Times - 23 minutes ago
Federer and Clijsters Advance on a Busy Day
New York Times - 46 minutes ago

2010 FIFA World Cup South Africa™
FIFA.com - Schedule - Standings - Teams
Round of 16
Brazil   0 : 0   Chile   Live
Latest Matches
Netherlands   2 : 1   Slovakia   Highlights
Argentina   3 : 1   Mexico   Highlights
Upcoming matches
June 29 07:00 (Pacific Time) on ESPN
Paraguay   vs.   Japan
June 29 11:30 (Pacific Time) on ESPN
Spain   vs.   Portugal
In The News
BET Awards   Robert C. Byrd
Chris Brown   Cristie Kerr
Robert Byrd   Michael Jackson
Elena Kagan   Vince Neil
Jennifer Capriati   Bubba Watson

Crisis response: Gulf of Mexico oil spill

Local News »
View stories near: [City, State, or Zip Code] Add

U.S. »
Byrd's absence could be problematic for financial regulation bill ☆

Figure 2: Google News Front Page After Redesign, July 2, 2010

Figure 3: Google News Links, January - June, 2010

Top Google News Domains

Links on Google News

nytimes.com
reuters.com
wsj.com
washingtonpost.com
latimes.com
cnn.com
ap.org
usatoday.com
businessweek.com
bbc.co.uk
abcnews.com
pcworld.com *
foxnews.com
cnet.com *
boston.com
sfgate.com
mtv.com *
mercurynews.com
msnbc.com
miamiherald.com
espn.com *
marketwatch.com
voanews.com
afp.com
ew.com *

Links

Days on Google News

nytimes.com
reuters.com
wsj.com
washingtonpost.com
latimes.com
cnn.com
ap.org
usatoday.com
businessweek.com
bbc.co.uk
abcnews.com
pcworld.com *
foxnews.com
cnet.com *
boston.com
sfgate.com
mtv.com *
mercurynews.com
msnbc.com
miamiherald.com
espn.com *
marketwatch.com
voanews.com
afp.com
ew.com *

Days

Note: Total links scraped from 181 days of archived Google News pages, January - June 2010.
See text for details.

30

Figure 4: Weekly Visits to Top News Outlets, January - June, 2010

Mean Weekly Visits and Links
Most Visited Outlets, January-June 2010

Note: Mean weekly news visits from sample households and Google News links, January-June 2010.
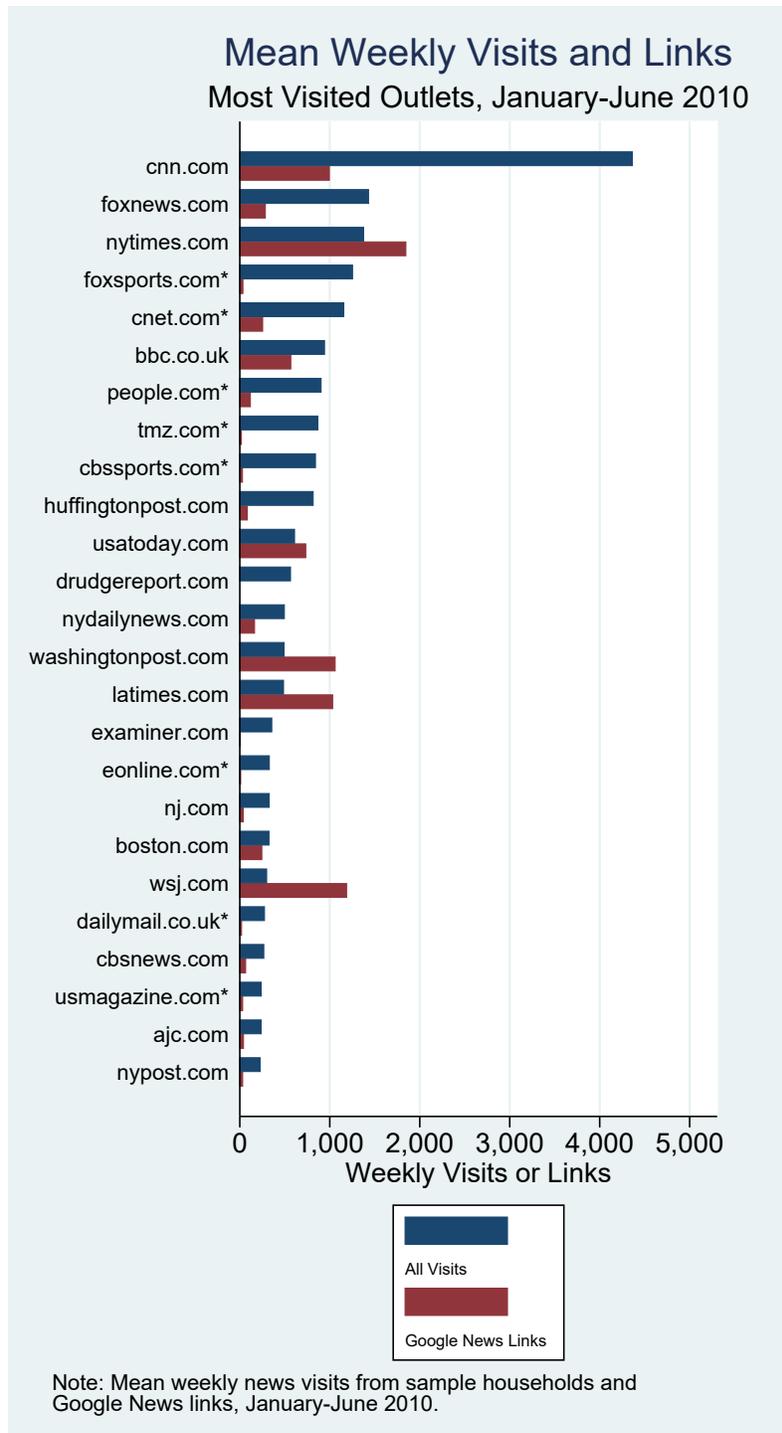
Figure 5: Weekly Visits to Top Local News Outlets, January - June 2010

Weekly News Visits

Most Visited Local Outlets, January - June 2010

Note: Mean weekly total and local new visits to top local
news domains from sample households, January-June 2010.

32
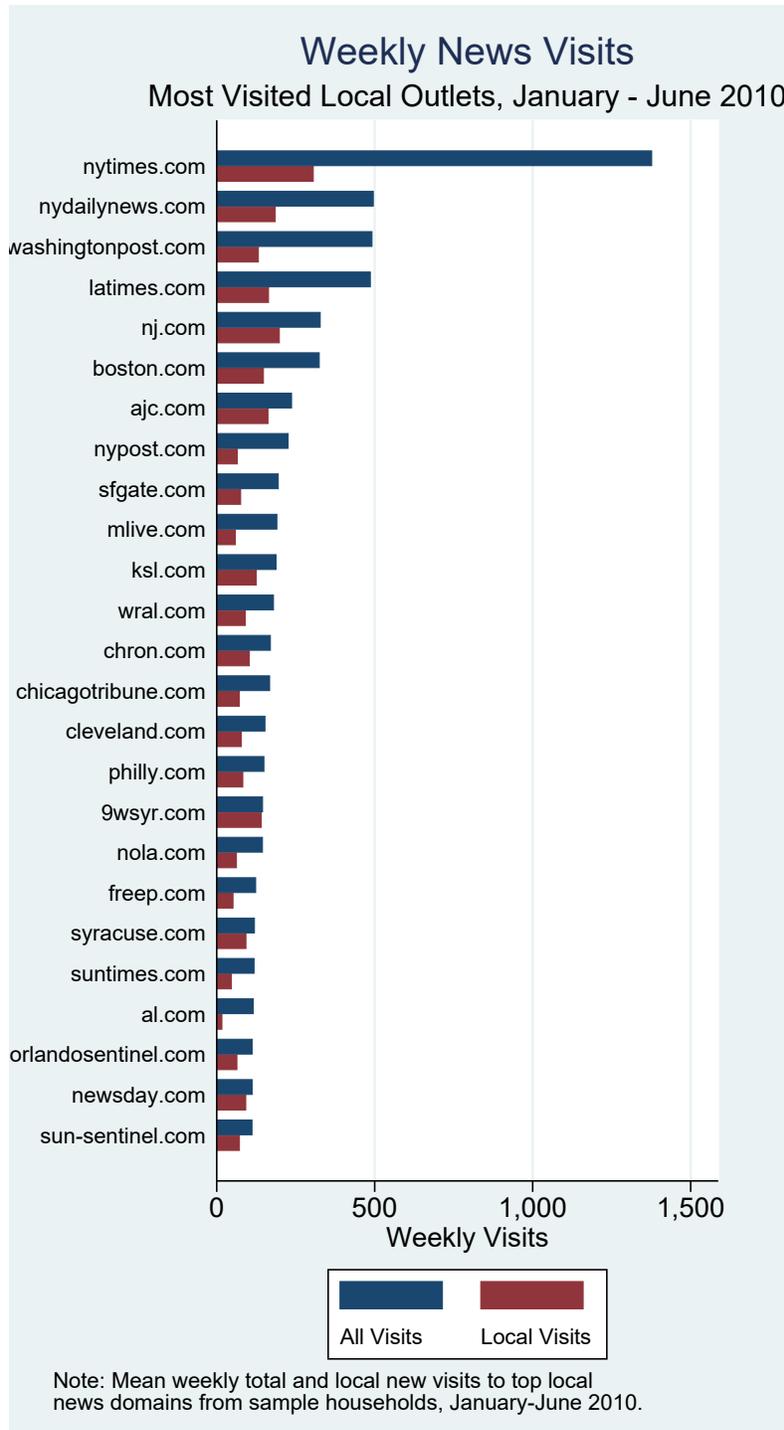
Table 2: Weekly News Visits, January - June 2010

	Top 25 Outlets				
	Mean	SD	p5	p50	p95
Weekly Total Visits	778.83	836.80	236.92	495.58	1430.75
Weekly Local Visits	54.40	88.91	0.00	0.00	198.21
Weekly Google News Links	14.82	20.76	0.04	3.46	49.46
Local Visit Share	0.13	0.21	0.00	0.00	0.61
Observations	25				
	Local Outlets				
	Mean	SD	p5	p50	p95
Weekly Total Visits	5.14	27.49	0.04	0.54	20.88
Weekly Local Visits	2.79	10.85	0.00	0.13	13.33
Weekly Google News Links	0.06	1.44	0.00	0.00	0.04
Local Visit Share	0.41	0.37	0.00	0.40	1.00
Observations	4,940				
	All Outlets				
	Mean	SD	p5	p50	p95
Weekly Total Visits	7.05	65.62	0.04	0.63	21.25
Weekly Local Visits	1.78	8.77	0.00	0.00	8.50
Weekly Google News Links	0.10	1.69	0.00	0.00	0.08
Local Visit Share	0.26	0.36	0.00	0.00	1.00
Observations	7,742				

Notes: The top panel reports average weekly visits from sample households to the 25 most-visited news outlets January - June 2010. The middle panel reports corresponding visit data for all local news outlets. The lower panel reports visit data for the full sample of outlets. See text for details.

33

Table 3: Weekly Local News Visits, January - June 2010

Metro Area	Outlet	Visits	Local Visits	Local Share	Top 3 Share
New York	New York Times	1377	306	0.22	0.59
	New York Daily News	496	185	0.37	0.59
	NJ.com	328	198	0.61	0.59
Los Angeles	LA Times	486	164	0.34	0.58
	Orange County Register	93	72	0.78	0.58
	LA Daily News	45	34	0.76	0.58
Washington DC	Washington Post	492	132	0.27	0.61
	Baltimore Sun	94	48	0.51	0.61
	Washington Times	45	21	0.47	0.61
Chicago	Chicago Tribune	167	71	0.43	0.52
	Chicago Sun Times	118	47	0.39	0.52
	NW Illinois Times	27	15	0.56	0.52
Dallas	Dallas Morning News	101	46	0.45	0.54
	Fort Worth Star Telegram	55	31	0.56	0.54
	WFAA	31	24	0.78	0.54
San Francisco	San Francisco Chronicle	195	76	0.39	0.55
	San Jose Mercury News	86	27	0.31	0.55
	Inside Bay Area	44	28	0.63	0.55
Philadelphia	Philly.com	149	83	0.55	0.44
	Courier Post	35	22	0.64	0.44
	NBC Philadelphia	27	18	0.66	0.44
Atlanta	Atlanta Journal Constitution	237	162	0.68	0.66
	WSBTV	81	55	0.68	0.66
	Access Atlanta	34	25	0.73	0.66
Houston	Houston Chronicle	170	103	0.61	0.66
	KHOU	34	28	0.81	0.66
	Click 2 Houston	32	27	0.86	0.66
Miami	Florida Sun-Sentinel	112	72	0.64	0.64
	Miami Herald	93	44	0.48	0.64
	El Spectador	26	9	0.33	0.64

Table 4: Household Sample Statistics

	Mean	SD	p5	p50	p95
All Visits	1.89	6.11	0.00	0.00	8.00
Local Visits	0.48	2.21	0.00	0.00	2.00
Local Visit Share	0.24	0.38	0.00	0.00	1.00
Unique Domains	1.45	3.67	0.00	0.00	7.00
Unique Local Domains	0.62	2.67	0.00	0.00	4.00
Active Week Share	0.43	0.49	0.00	0.00	1.00
Active Weeks	16.90	11.48	3.00	14.00	41.00
Observations	1,207,239				

Note: Weekly total, local and unique outlet visits per week over 47 weeks of 2010 by 36,876 households.

35

Table 5: Household Treatment Measures

	Mean	SD	p5	p50	p95
Google News Intensity	0.07	0.14	0.00	0.00	0.33
All Google Referrals	0.33	0.33	0.00	0.25	1.00
All Yahoo Referrals	0.11	0.22	0.00	0.00	0.67
Observations	36,876				

Note: Google News intensity is the imputed share of visits to outlets referred by Google to outlets appearing on Google News. Google and Yahoo intensity is the observed share of news visits referred by Google and Yahoo, respectively. See text for details.

Figure 6: Weekly Visits to Local News Outlets

Local News Visits by Week

Weekly visits to local news outlets by households in the top quintile of imputed Google News referrals relative to all other households. See text for details.
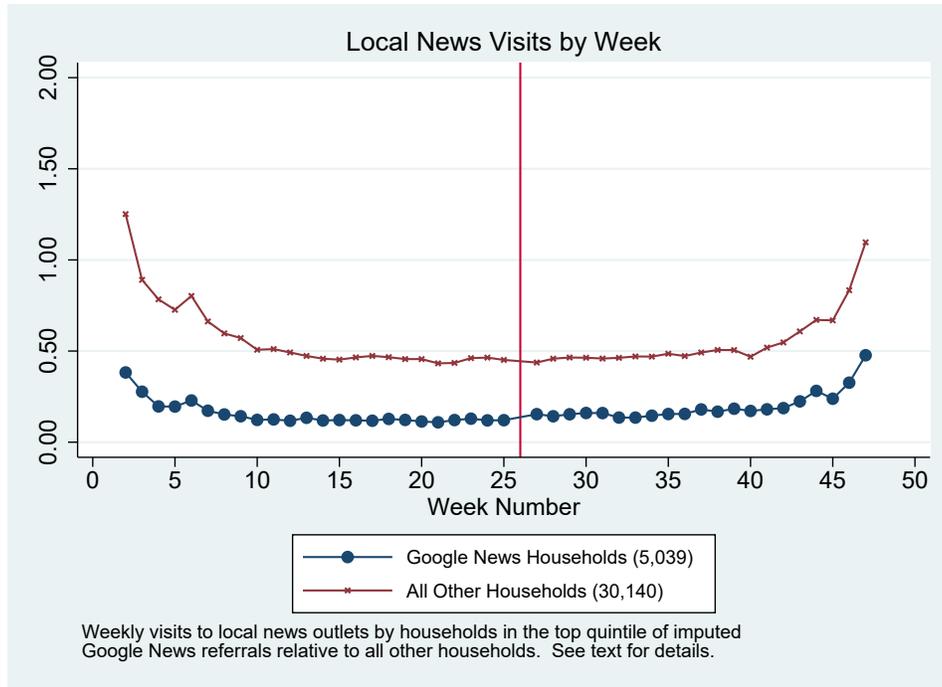
Figure 7: Weekly Local Visit Share

Local Share of News Visits by Week

Weekly share of news visits to local outlets by households in the top quintile of imputed Google News referrals relative to all other households. See text for details.

37
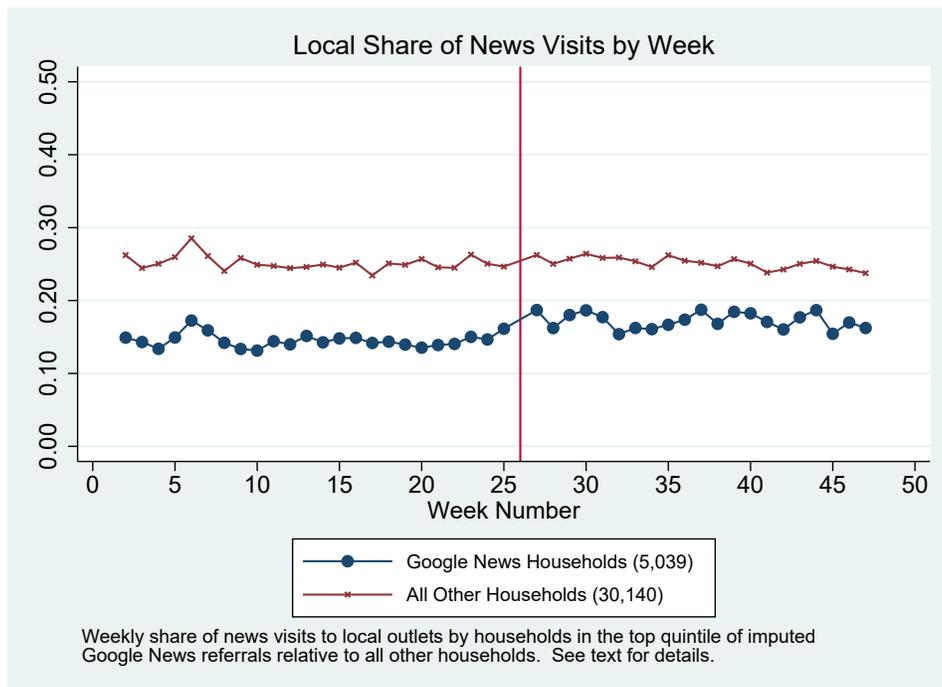
Figure 8: Weekly Visits to Local News Outlets

Local News Visits by Week

Note: Average weekly visits to local news outlets by week in 2010 for households in the top quintile of Yahoo referrals and top quintile of imputed Google News referrals. See text for details.
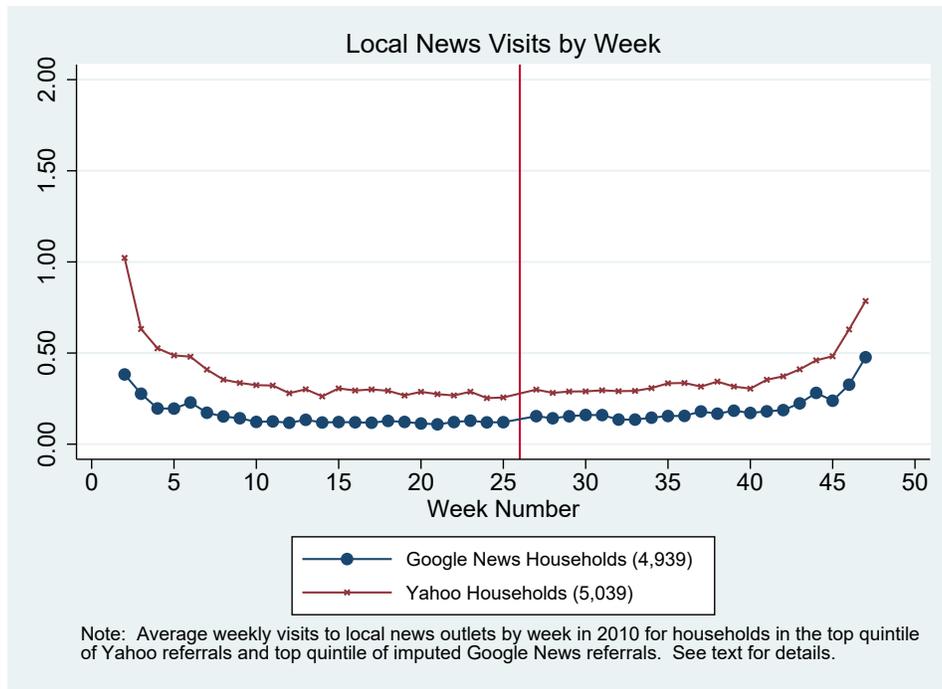
Figure 9: Weekly Local Visit Share

Local Share of News Visits by Week

Note: Average weekly share of news visits to local outlets for households in the top quintile of Yahoo referrals and top quintile of imputed Google News referrals. See text for details.

38
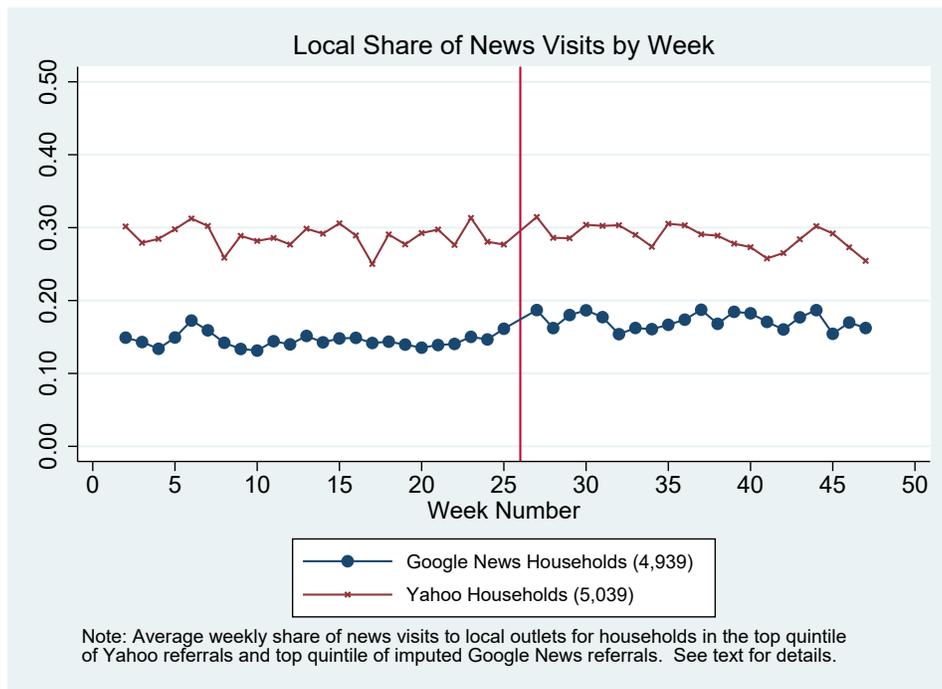
Table 6: Does Geo-Targeting Increase Local News Consumption?

	Local News Visits (1)	Local News Indicator (2)	Local News Share (3)
Post Geo-targeting	-0.07**	-0.01**	-0.01**
	(0.01)	(0.00)	(0.00)
Post x Google News Intensity	0.26**	0.04**	0.10**
	(0.04)	(0.01)	(0.01)
Constant	0.50**	0.16**	0.24**
	(0.00)	(0.00)	(0.00)
Adj. R-squared	0.62	0.30	0.39
Households	36,876	36,876	36,876
Observations	1,207,239	1,207,239	516,834
Y Mean	0.48	0.15	0.24
Y SD	2.21	0.36	0.38
X Mean	0.07	0.07	0.07
X SD	0.12	0.12	0.10

Note: All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 7: Does Geo-Targeting Increase Local News Consumption? (Treatment-Control Approach)

	Local News Visits (1)	Local News Indicator (2)	Local News Share (3)
Post Geo-targeting	-0.012	-0.004	-0.001
	(0.013)	(0.003)	(0.005)
Post x Google News Group	0.038*	0.013**	0.024**
	(0.016)	(0.004)	(0.006)
Constant	0.243**	0.112**	0.217**
	(0.003)	(0.001)	(0.002)
R-squared	0.46	0.22	0.30
Households	10,403	10,403	10,403
Observations	329,613	329,613	118,603
Control Mean	0.339	0.142	0.288
Control SD	1.382	0.349	0.404
Treatment Mean	0.158	0.085	0.157
Treatment SD	0.826	0.279	0.323

Note: All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 8: Does Geo-Targeting Increase Consumption Diversity?

	Weekly Unique Outlets		Aggregate Unique Outlets	
	All Outlets	Local Outlets	All Outlets	Local Outlets
	(1)	(2)	(3)	(4)
Post Geo-targeting	-0.17**	-0.09**	-2.64**	-0.28**
	(0.01)	(0.01)	(0.10)	(0.02)
Post x Google News Intensity	0.18	0.23**	-2.24**	0.33*
	(0.10)	(0.07)	(0.57)	(0.14)
Constant	1.52**	0.65**	12.91**	1.70**
	(0.00)	(0.00)	(0.04)	(0.01)
Adj. R-squared	0.56	0.49	0.62	0.60
Households	36,876	36,876	36,876	36,876
Observations	1,207,239	1,207,239	67,778	67,778
Y Mean	1.45	0.62	11.51	1.57
Y SD	3.67	2.67	15.74	1.93
X Mean	0.07	0.07	0.07	0.07
X SD	0.12	0.12	0.13	0.13

Note: All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 9: Does Geo-Targeting Increase Consumption Diversity? (Treatment-Control Approach)

	Weekly Unique Outlets		Aggregate Unique Outlets	
	All Outlets	Local Outlets	All Outlets	Local Outlets
	(1)	(2)	(3)	(4)
Post Geo-targeting	-0.028	-0.035*	-1.293**	-0.207**
	(0.021)	(0.016)	(0.116)	(0.033)
Post x Google News Group	-0.001	0.045*	-1.296**	0.104
	(0.030)	(0.022)	(0.231)	(0.065)
Constant	0.911**	0.348**	10.383**	1.399**
	(0.006)	(0.005)	(0.061)	(0.021)
R-squared	0.48	0.49	0.69	0.56
Households	10,403	10,403	10,403	10,403
Observations	329,613	329,613	18,939	18,939
Control Mean	0.916	0.431	8.343	1.448
Control SD	2.337	1.851	10.076	1.720
Treatment Mean	0.882	0.259	10.435	1.199
Treatment SD	2.270	1.470	12.772	1.733

Note: All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 10: Robustness Specification: Market Exclusion

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
After Geo-targeting	-0.06**	-0.01**	-0.01**
	(0.01)	(0.00)	(0.00)
After x Google News Intensity	0.29**	0.06**	0.13**
	(0.05)	(0.01)	(0.01)
Constant	0.51**	0.16**	0.24**
	(0.00)	(0.00)	(0.00)
R-squared	0.62	0.31	0.41
Households	27,458	27,458	27,458
Observations	898,527	898,527	385,247
Y Mean	0.49	0.15	0.24
Y SD	2.25	0.36	0.38
X Mean	0.07	0.07	0.06
X SD	0.12	0.12	0.10

Note: Sample excludes ten markets with highest and lowest news visits. All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 11: Robustness Specification: Shortrun Effects

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
Post Geo-targeting	-0.01	-0.00**	0.00
	(0.01)	(0.00)	(0.00)
Post x Google News Intensity	0.11**	0.02**	0.07**
	(0.03)	(0.01)	(0.01)
Constant	0.42**	0.13**	0.24**
	(0.00)	(0.00)	(0.00)
R-squared	0.67	0.32	0.41
Households	34,786	34,786	33,941
Observations	753,692	753,692	280,236
Y Mean	0.42	0.13	0.24
Y SD	2.05	0.34	0.38
X Mean	0.07	0.07	0.07
X SD	0.12	0.12	0.10

Note: Sample includes news visits 12 weeks before and after the implementation of geo-targeting. All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 12: Robustness Specification: Hard News

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
Post Geo-targeting	-0.07**	-0.01**	-0.00*
	(0.01)	(0.00)	(0.00)
Post x Google News Intensity	0.29**	0.04**	0.12**
	(0.04)	(0.01)	(0.02)
Constant	0.49**	0.16**	0.30**
	(0.01)	(0.00)	(0.00)
R-squared	0.63	0.30	0.43
Households	35,178	35,178	35,178
Observations	1,027,381	1,027,381	390,930
Y Mean	0.47	0.15	0.30
Y SD	2.17	0.36	0.41
X Mean	0.07	0.07	0.07
X SD	0.11	0.11	0.10

Note: Dependent variables calculated based on visits to major hard-news sites only, see text for details. All specifications include household fixed effects. Standard errors clustered by market: * $p<0.05$, ** $p<0.01$.

Table 13: Placebo Regressions: Search

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
Post Geo-targeting	-0.081**	-0.013**	-0.005*
	(0.016)	(0.001)	(0.002)
Post x Google News Intensity (X1)	0.234**	0.043**	0.105**
	(0.036)	(0.010)	(0.016)
Post x Google Search Intensity (X2)	0.129**	0.010*	-0.010
	(0.037)	(0.004)	(0.009)
Constant	0.501**	0.158**	0.241**
	(0.005)	(0.001)	(0.001)
Adj. R-squared	0.62	0.30	0.39
Households	36,876	36,876	36,876
Observations	1,207,239	1,207,239	516,834
Y Mean	0.48	0.15	0.24
Y SD	2.21	0.36	0.38
X1 Mean	0.07	0.07	0.07
X1 SD	0.12	0.12	0.10
X2 Mean	0.13	0.13	0.12
X2 SD	0.18	0.18	0.16
Linear Combination			
Difference	0.105*	0.033**	0.115**
	(0.047)	(0.011)	(0.022)

Note: All specifications include household fixed effects. Standard errors clustered by market: * $p < 0.05$, ** $p < 0.01$.

Table 14: Placebo Regressions: Yahoo

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
Post Geo-targeting	-0.074**	-0.012**	-0.007**
	(0.014)	(0.001)	(0.002)
Post x Google News Intensity (X1)	0.277**	0.045**	0.106**
	(0.042)	(0.010)	(0.015)
Post x Yahoo Referral Intensity (X2)	0.059	0.000	0.011
	(0.032)	(0.004)	(0.007)
Constant	0.501**	0.158**	0.241**
	(0.005)	(0.001)	(0.001)
Adj. R-squared	0.62	0.30	0.39
Households	36,876	36,876	36,876
Observations	1,207,239	1,207,239	516,834
Y Mean	0.48	0.15	0.24
Y SD	2.21	0.36	0.38
X1 Mean	0.07	0.07	0.07
X1 SD	0.12	0.12	0.10
X2 Mean	0.11	0.11	0.11
X2 SD	0.21	0.21	0.21
Linear Combination			
Difference	0.218**	0.045**	0.094**
	(0.040)	(0.009)	(0.014)

Note: All specifications include household fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 15: Placebo Regressions: Referrals

	Local News Visits	Local News Indicator	Local News Share
	(1)	(2)	(3)
Post Geo-targeting	-0.110**	-0.015**	-0.008**
	(0.022)	(0.002)	(0.002)
Post x Google Referral Intensity (X1)	0.157**	0.017**	0.025**
	(0.029)	(0.003)	(0.004)
Post x Yahoo Referral Intensity (X2)	0.102**	0.003	0.012
	(0.038)	(0.005)	(0.007)
Constant	0.501**	0.158**	0.241**
	(0.005)	(0.001)	(0.001)
Adj. R-squared	0.62	0.30	0.39
Households	36,876	36,876	36,876
Observations	1,207,239	1,207,239	516,834
Y Mean	0.48	0.15	0.24
Y SD	2.21	0.36	0.38
X1 Mean	0.33	0.33	0.30
X1 SD	0.31	0.31	0.30
X2 Mean	0.11	0.11	0.11
X2 SD	0.21	0.21	0.21
Linear Combination			
Difference	0.054*	0.013**	0.012
	(0.027)	(0.004)	(0.007)

Note: All specifications include household fixed effects. Standard errors clustered by market: * $p<0.05$, ** $p<0.01$.

Table 16: Do Google News Links Increase Traffic to News Outlets?

	Non-Local Traffic (Log) (1)	Local Traffic (Log) (2)	Local Share (3)
Google News Appearance	0.033**	-0.003	-0.003
	(0.008)	(0.015)	(0.003)
Constant	0.589**	0.639**	0.334**
	(0.000)	(0.000)	(0.000)
R-squared	0.77	0.62	0.68
Outlets	7,178	3,180	7,566
Observations	205,380	110,102	273,913
Y Mean	0.59	0.64	0.33
Y SD	0.92	0.78	0.44
X Mean	0.03	0.02	0.02
X SD	0.17	0.14	0.15

Note: Sample includes outlets appearing on Google News January-June 2010. Dependent variable in column (1) is non-local traffic to news outlets from sample households each day in 2010. Dependent variable in column (2) is traffic from households in the same metropolitan area as the outlet, zero otherwise. Dependent variable in column (3) is the share of outlet traffic from households local to the outlet. All specifications include outlet fixed effects. Standard errors clustered by market: * p<0.05, ** p<0.01.

Table 17: Does Geo-targeting Increase Local Traffic to News Outlets?

	Non-Local Traffic (Log)	Local Traffic (Log)	Local Share
	(1)	(2)	(3)
Post Geo-targeting	-0.123**	-0.138**	-0.006
	(0.016)	(0.019)	(0.005)
Post x Google Ref. Intensity (GR)	0.142**	0.219**	0.041**
	(0.031)	(0.042)	(0.014)
Post x Yahoo Ref. Intensity (YR)	0.072	0.132	-0.002
	(0.043)	(0.068)	(0.020)
Constant	0.988**	1.375**	0.480**
	(0.005)	(0.006)	(0.001)
R-squared	0.73	0.78	0.60
Outlets	4,721	3,403	4,939
Observations	75,918	65,147	102,431
Y Mean	0.95	1.34	0.48
Y SD	1.06	1.21	0.42
GR Mean	0.25	0.23	0.25
GR SD	0.22	0.19	0.22
YR Mean	0.12	0.12	0.12
YR SD	0.15	0.14	0.15

Note: Sample includes local outlets only. Dependent variable in column (1) is (log) non-local traffic each week, dependent variable in column (2) is (log) local traffic each week, and dependent variable in column (3) is the local traffic share. Independent variables are the share of visits referred from Google and Yahoo prior to the re-design, respectively. All specifications include outlet fixed effects. Standard errors clustered by market: * $p < 0.05$, ** $p < 0.01$.

Table 18: Local Google News Links in Three Markets

Local News Outlet	Local Link Share (%)
San Francisco (44 days)	
Bay Citizen	1
San Jose Mercury News	61
NBC Bay Area	1
SFGate	36
Indianapolis (7 days)	
Fox 59	10
Herald Bulletin (Anderson)	10
Indianapolis Star	40
The Indy Channel (WRTV)	20
WISH TV	10
WTHR	10
Philadelphia (6 days)	
Philly.com	78
Press of Atlantic City	11
Times Herald (Norristown)	11

Table 19: Does Aggregation Increase Competition?

	CR1 (1)	CR2 (2)	CR3 (3)
Post Geo-targeting	-0.016*	-0.014**	-0.011**
	(0.006)	(0.004)	(0.003)
Constant	0.627**	0.811**	0.886**
	(0.003)	(0.002)	(0.001)
Adj. R-squared	0.72	0.78	0.81
Markets	280	280	280
Observations	12,168	12,168	12,168

Note: Dependent variable is the weekly share of traffic to the top 1, 2 and 3 local outlets, respectively. All specifications include market fixed effects. Standard errors clustered by market: * $p<0.05$, ** $p<0.01$.