DISCUSSION PAPER SERIES

DP16130
(v. 3)

The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist

Milena Djourelova, Ruben Durante and Gregory Martin

PUBLIC ECONOMICS

ISSN 0265-8003

The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist

Milena Djourelova, Ruben Durante and Gregory Martin

Discussion Paper DP16130
First Published 09 May 2021
This Revision 07 November 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Public Economics

The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist

Abstract

How does competition from online platforms affect the organization, performance, and editorial choices of newspapers? And what are the implications of these changes for the information voters are exposed to and for their political choices? We study these questions using the staggered introduction of Craigslist - the world's largest online platform for classified advertising - across US counties between 1995 and 2009. This setting allows us to separate the effect of competition for classified advertising from other changes brought about by the Internet, and to compare newspapers that relied more or less heavily on classified ads ex ante. We find that, following the entry of Craigslist, local papers experienced a significant decline in the number of newsroom and management staff. Cuts in editorial staff disproportionately affected editors covering politics. These organizational changes led to a reduction in news coverage of politics and resulted in a decline in newspaper readership, which was not compensated by increased news consumption online or in other media. Finally, we document that reduced exposure to political news was associated with more partisan voting and increased support for ideologically extreme candidates.

JEL Classification: L82, L86, D72

Keywords: newspapers, internet, advertising, ideological polarization

Milena Djourelova - mdjourelova@uchicago.edu
University of Chicago

Ruben Durante - ruben.durante@upf.edu
ICREA-Universitat Pompeu Fabra, IPEG, BSE, CESifo and CEPR

Gregory Martin - gjmartin@stanford.edu
Stanford GSB

Acknowledgements
We thank Tim Besley and seminar participants at Berkeley, Bocconi, Bristol, Chicago (Harris), EIEF, the European Commission (ECFIN), LSE, Mannheim, Pittsburgh, and UPF, as well as participants in the Workshop on Elections, Democracy, and Populism at the University of Bologna, the CEPR Virtual IO seminar, SAEe, RIDGE-LACEA and the Monash-Warwick-Zurich Text-as-Data Workshop for helpful comments. David Ampudia Vicente, Marco Lo Faso, and Elliot Motte provided excellent research assistance. Irena Djourelova, Anna Griñó, Adriana Oliveres and Silvia Terzulli provided help with digitizing the Editor and Publisher data. Ruben Durante acknowledges financial support from the European Union's Horizon 2020 research and innovation program [Grant 759885].

The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist*

Milena Djourelova† Ruben Durante‡ Gregory J. Martin§

November 2021

Abstract

How does competition from online platforms affect the organization, performance, and editorial choices of newspapers? And what are the implications of these changes for the information voters are exposed to and for their political choices? We study these questions using the staggered introduction of Craigslist – the world's largest online platform for classified advertising – across US counties between 1995 and 2009. This setting allows us to separate the effect of competition for classified advertising from other changes brought about by the Internet, and to compare newspapers that relied more or less heavily on classified ads *ex ante*. We find that, following the entry of Craigslist, local papers experienced a significant decline in the number of newsroom and management staff. Cuts in editorial staff disproportionately affected editors covering politics. These organizational changes led to a reduction in news coverage of politics and resulted in a decline in newspaper readership, which was not compensated by increased news consumption online or in other media. Finally, we document that reduced exposure to political news was associated with more partisan voting and increased support for ideologically extreme candidates.

Keywords: newspapers, Internet, advertising, ideological polarization
JEL codes: L82, L86, D72

*We thank Tim Besley and seminar participants at Berkeley, Bocconi, Bristol, Chicago (Harris), EIEF, the European Commission (ECFIN), LSE, Mannheim, Pittsburgh, and UPF, as well as participants in the Workshop on Elections, Democracy, and Populism at the University of Bologna, the CEPR Virtual IO seminar, SAEe, RIDGE-LACEA and the Monash-Warwick-Zurich Text-as-Data Workshop for helpful comments. David Ampudia Vicente, Marco Lo Faso, and Elliot Motte provided excellent research assistance. Irena Djourelova, Anna Griñó, Adriana Oliveres and Silvia Terzulli provided help with digitizing the Editor and Publisher data. Ruben Durante acknowledges financial support from the European Union's Horizon 2020 research and innovation program [Grant 759885].
†University of Chicago. Email: mdjourelova@uchicago.edu
‡ICREA-UPF, BSE, IPEG, CESifo and CEPR. Email: ruben.durante@upf.edu
§Stanford GSB. Email: gjmartin@stanford.edu

1 Introduction

The Internet has profoundly changed the environment in which newspapers operate. Competition from online platforms has contributed to the sharp decline in newspapers' advertising revenues over the last two decades, forcing many news outlets to drastically rethink their business model and organization.[1] These changes, some warn, may have detrimental consequences for the quality of news reporting and the provision of political information (McChesney and Nichols, 2011; Starkman, 2014; Peterson, 2021). Given the key role played by newspapers in informing citizens about politics (Gentzkow et al., 2011; Snyder and Strömberg, 2010), they may also have important implications for electoral politics.

Despite the potentially grave consequences of these transformations for the future of journalism, rigorous evidence on the impact of online competition on newspapers' organization and editorial choices is surprisingly scant. One reason for this is the challenge of separating the effect of online competition from other technological and socioeconomic changes brought about by the Internet, which may affect both the demand and the supply side of the newspaper market in other ways.

To overcome this limitation, in this paper we investigate the impact on US newspapers of the introduction of Craigslist (henceforth CL), the world's largest online platform for classified ads, whose entry disrupted the market for classified ads, a formerly lucrative niche for newspapers (Seamans and Zhu, 2014; Kroft and Pope, 2014).[2] Tracking the expansion of CL across US counties between 1995 and 2009, we examine how the entry of a local CL website affected the organization and editorial decisions of local newspapers, and ultimately, the electoral choices of local voters.

The expansion of CL in the U.S. provides an attractive setting for several reasons. First, CL's staggered expansion over a period of 15 years, combined with the limited geographic scope of local CL websites, generates significant variation over time and across space in the degree of online competition for classified ads faced by local newspapers. Second, since CL websites do not feature news content or display advertising, CL's entry represents a specific shock to revenues from classified ads but leaves other market conditions unaffected. In addition, CL's narrow focus on classified ads provides an important source of heterogeneity, since the entry of CL should disproportionately affect newspapers that relied more heavily on classified ads *ex ante*. Finally, with a few exceptions in the biggest cities, ads on CL are free of charge, and most local websites do not generate profit for the company. The lack of a clear profit maximization strategy[3] alleviates

[1] According to data from the News Media Alliance (formerly Newspaper Association of America), US newspapers' advertising revenues fell from US$49 billion in 2000 to US$26 billion in 2010.

[2] As of 2000, classified advertising accounted for 40% of US newspapers' total advertising revenues. In that year advertising revenues amounted to US$49 billion, compared to US$11 billion of circulation revenues.

[3] CL founder Craig Newmark was sued in 2010 by eBay, which held a minority stake in CL, for failing in his fiduciary duty to maximize shareholder returns.

concerns that the timing of CL's entry might have been driven by strategic considerations related to the conditions of local newspaper markets. We document that the timing of CL's entry into a local market is not correlated with the characteristics of local newspapers once population and the quality of the local Internet connection are controlled for.

Our empirical strategy compares the evolution of outcomes of interest between areas with and without access to a local CL website, before and after the website is introduced, controlling for the quality of local broadband Internet. In addition, we exploit variation across newspapers in the reliance on classified ads prior to the entry of CL, proxied by the presence of a dedicated classified ads manager in the newspaper's staff prior to the entry of CL.[4]

To analyze the impact of CL's entry on local newspapers, we exploit comprehensive data on the organization and staff of over 1,500 newspapers, covering the period from 1995 to 2010. We find that while CL's entry does not significantly affect the number of active local newspapers, it leads to substantial downsizing. After the entry of CL into a county, newspapers headquartered there cut about 1.2 jobs on average, or about 6% of the sample mean. This effect is driven by newspapers that relied more heavily on classified ads at baseline, which experience a 14% decline in the number of jobs relative to the mean.[5] Staff cuts affect both managerial and editorial positions, but editorial cuts appear to disproportionately affect editors responsible for the coverage of politics, leaving other areas such as sports and entertainment largely unaffected.

We then test to what extent these organizational changes affected newspapers' editorial priorities, with particular regard to their coverage of politics. First, applying keyword searches to the entire corpus of articles published in over 800 newspapers, we compute the number of mentions of local Congressional representatives and candidates for office. We document that, following the entry of CL in a given area, news coverage of local representatives in affected papers declined significantly, by around 30%. We find no corresponding decline in mentions of *national* party leaders. We complement this approach with a semi-supervised topic model, estimated on a random sample of two million articles. We find evidence of a decline in the prevalence of topics related to politics, while other topics such as sports, entertainment, or crime are not affected.[6]

Next we examine how readers responded to these changes in content. First, we document that, in the years after the entry of CL, local newspapers experienced a sharp decline in circulation. We further explore this readership decline using data from two large-scale surveys on media consumption. The results confirm that, following the entry of CL, local respondents are less likely to report reading a newspaper. Interestingly, this change is driven entirely by readers who are likely

[4] To corroborate the validity of this proxy, we document that the presence of classified ads managers is strongly correlated with the number of pages devoted to classified ads prior to the entry of CL (see Appendix Table A1).

[5] We find no significant change in the number of pages published by affected newspapers, which suggests that this downsizing is likely associated with greater workload per staff member.

[6] This is consistent with recent evidence that staff cuts are detrimental to the coverage of politics, from Peterson (2021).

to be interested in the news section of the newspaper, while we find no heterogeneity by interest in classified ads. We interpret this as consistent with readers responding to changes in news coverage, rather than merely lower demand for print classified ads.[7] Evidence from both survey and browsing data suggests that the decline in newspaper readership is not compensated by increased news consumption online.[8] We also find no evidence of substitution to other news sources (i.e., national papers, radio or TV). The changes induced by CL are hence likely to result in an overall decline in political information among the public.

Finally, we study how reduced news coverage of politics affected the behavior of local voters, with particular regard to electoral participation and ideological polarization. For electoral participation we find that the entry of CL has no significant effect on turnout in either presidential or House elections. Regarding ideological polarization, we find evidence that the entry of CL reduced the probability that voters support different parties in concurrent elections (split-ticket voting) and favored the rise of ideologically extreme candidates. These findings are consistent with a greater tendency to rely on national partisan cues when less information about local politicians is available (Darr et al., 2018; Moskowitz, 2021; Trussler, 2020), and with the role of newspapers in allowing voters to discriminate between extreme and moderate candidates in primaries (Hall and Lim, 2018).

Taken together, our results indicate that the impoverishment of local newspapers due to competition from online platforms can jeopardize their ability to inform citizens about politics, with the effect of reducing the centripetal pressure on candidate ideology that elections provide. This evidence supports the concerns expressed by some regulators that newspapers' financial distress, due to lower advertising revenues, may threaten quality reporting and pluralism (FCC, 2016).

Our paper contributes to several streams of research. First, it relates to previous contributions on the impact of CL on US newspapers. Most notably, the work by Seamans and Zhu (2014) documents that, after the entry of CL, local newspapers experienced lower classified-ad rates and circulation, and higher display-ad rates and subscription prices. Our analysis expands upon these findings by documenting the entire chain of events triggered by the entry of CL, and its profound implications for newspapers' staff, editorial priorities, news content, and, ultimately, political outcomes.[9]

Second, our paper relates to previous work on how the introduction of new media technologies affects incumbent media outlets. For example, Angelucci et al. (2020) argue that the introduction

[7] Lower readership is also unlikely to be due to newspapers charging higher subscription prices, since, unlike Seamans and Zhu (2014), we find no significant effect of CL's entry on that dimension.

[8] In particular, we find that the number of visits to popular national news websites by local users is not affected by CL's entry.

[9] Other work use the expansion of CL across the US, or particular design features of the platform, to study questions related to matching efficiency in labor and housing markets (Kroft and Pope, 2014), and the impact of online personal ads on sexually transmitted diseases and violence against women (Cunningham et al., 2019; Chan and Ghose, 2014).

of commercial TV in the US had a significant negative economic impact on newspapers, leading to less coverage of local politics. Bhuller et al. (2020) document large declines in circulation and shifts in editorial priorities for newspapers in Norway in response to the roll-out of broadband Internet. Our empirical setting is different in that it allows us to clearly separate the effect of a specific shock to the advertising market from the demand-side changes brought about by the Internet.

In this regard, our findings dovetail with previous evidence of how shocks to revenues affect news producers. In a long-term historical perspective, Hamilton (2004) and Petrova (2011) argue that the growth of the print advertising market in the late 19th century was essential to the emergence of an independent (non-partisan) press. Looking at France, Angelucci and Cagé (2019) examine how the introduction of advertising on TV in the 1960s and 1970s affected newspapers, leading to a decline in subscription prices and in the amount of journalistic-intensive content. Our analysis differs in that we study the disruptive effect of online competition, and, crucially, its ramifications for editorial choices and political outcomes.[10]

Our findings also relate to previous work on the effect of the Internet on political participation and policy (Falck et al., 2014; Campante et al., 2018; Gavazza et al., 2019; Larcinese and Miner, 2018; Manacorda and Tesei, 2020). While these studies assess the aggregate effect of the various changes brought about by the Internet on both sides of the political market, our analysis isolates the political impact of the Internet through the disruption of legacy media.

Finally, our paper contributes to the broader literature on the link between media, information, and voting behavior (Besley and Burgess, 2002; Snyder and Strömberg, 2010, e.g.,). In particular, it relates to a set of recent studies on the impact of media consumption on ideological polarization (Hall, 2015; Lelkes et al., 2017; Allcott et al., 2020; Levy, 2021). In this regard, our findings indicate that the impoverishment of local newspapers - and the resulting changes in organization, content, and readership - can reduce voters' ability to recognize and punish extreme candidates, and thus weaken the ideologically moderating force of elections.

The remainder of the paper is organized as follows. Section 2 provides some background information about Craigslist and its expansion. Section 3 describes the data used in the analysis. Section 4 discusses the empirical strategy. Section 5 presents the results and evidence on possible mechanisms. Section 6 concludes.

[10] Another relevant contribution on the impact of competitive pressures on newspapers is George and Waldfogel (2006). The study documents how the diffusion of the New York Times to over 100 US towns where it was previously not distributed affected the circulation and content of local newspapers.

2 Background

Craiglist.org (CL) is the world's largest online platform for classified ads. It was founded in San Francisco in 1995, and served only the Bay Area until 2000, when it began to gradually expand to other US locations. CL initially opened new local websites in big cities such as Boston, New York, and Chicago. Over time, it expanded to smaller markets, covering 115 locations in 2005, 331 in 2008, and 416 today.

Consistently ranked among the top 20 US websites by traffic,[11] the CL platform has a simple layout which has remained largely unchanged over time (see Figure A3). Ads on CL websites are organized into sections including housing, jobs, items for sale, professional services, and personals. CL websites only host classified ads and do not include any display ads or news content. Ads on CL are generally posted free of charge, with a few exceptions for brokered apartment rentals in New York and job posts for employers in some major cities.[12] CL's business model reflects the unconventional views of its founder, Craig Newmark, who prioritized providing a useful service to local communities over profit maximization, and always opposed listing the company on the stock market.[13]

Being cheaper and more efficient than traditional newspaper ads, CL became very popular among users, and rapidly disrupted the lucrative market for classified ads that many local newspapers had relied upon. Aggregate data on newspaper revenues from the Newspaper Association of America suggest that classified revenues started declining in the early 2000s, and the downward trend became steeper following CL's major expansion in the mid 2000s (Figure A1). Similarly to print revenues, the aggregate number of newsroom workers has been on a steep downward trend since the mid 2000s (Figure A2).

To investigate whether the link between CL's expansion and newspapers' performance is causal, our analysis exploits variation in the timing of the introduction of CL across US counties, combined with a newspaper-specific measure of ex-ante reliance on classified ads.

3 Data

Our analysis combines data on: i) Craigslist's expansion across the U.S.; ii) characteristics, organization and market outcomes of daily newspapers; iii) newspapers' content; iv) survey data on media consumption; v) political behavior outcomes including turnout, vote choices and campaign

[11] https://www.similarweb.com/top-websites/united-states/

[12] A full list of the exceptions as of 2010, the end of our sample period, is available at: https://web.archive.org/web/20100706030043/https://www.craigslist.org/about/help/posting_fees

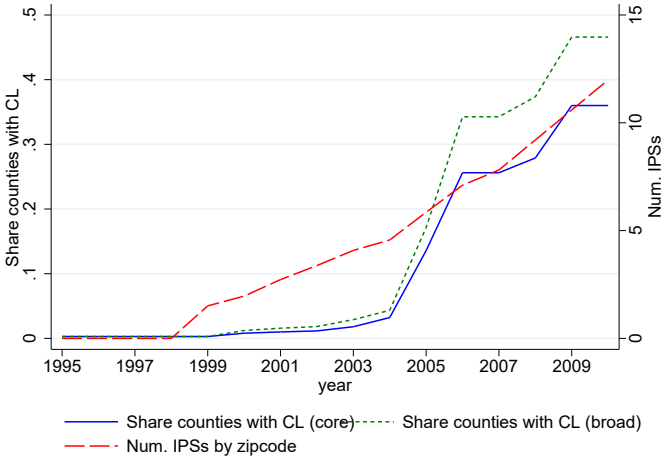
[13] For a profile of Craig Newmark and his business strategy see https://www.theguardian.com/technology/2006/feb/19/news.theobserver1.

contributions, and vi) additional covariates.

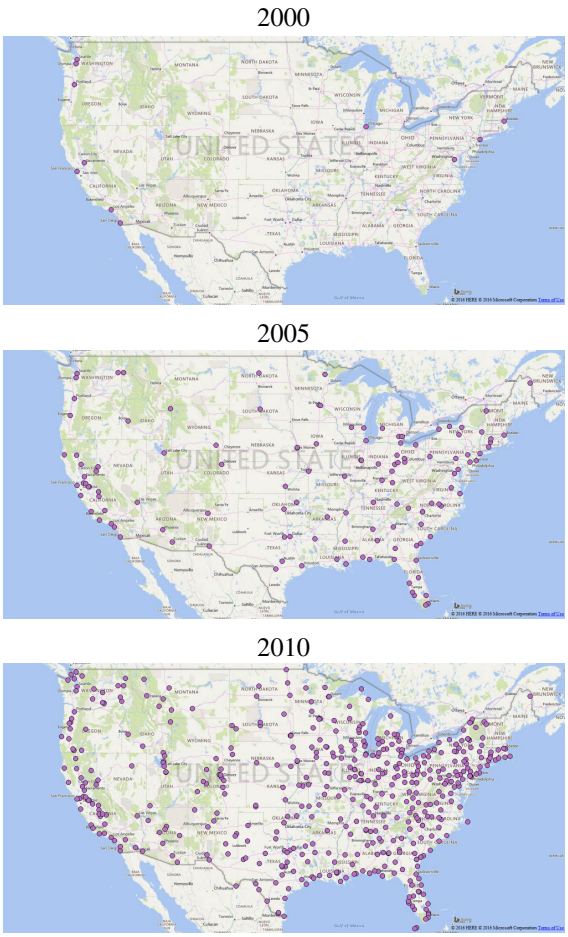
3.1 Craigslist's expansion

To construct a measure of the availability of CL in each county, we first collect information on the timing of the entry of each of CL's current websites. For a subset of these, the information is directly available on the CL's "about" webpage (https://www.craigslist.org/about/expansion). For the others, we assigned the date of the first snapshot recorded by the Internet Archive (https://archive.org). Panel (a) of Figure 1 shows the evolution of the share of US counties served by CL websites between 1995 and 2010, along with the evolution of a proxy for Internet quality - the average number of Internet Service providers by zip code.[14] Panel (b) depicts the geographic distribution of CL websites in 2000, 2005, and 2010, respectively.

FIGURE 1: CRAIGSLIST'S ROLL-OUT OVER TIME



(a) Share of US counties with access to a local CL website (left axis) and average number of Internet service providers by zip-code (right axis). The two definitions of access to CL ("core" and "broad") are described in Appendix A.2.1.



(b) Geographic distribution of CL websites over time.

Mapping websites to counties is not straightforward since the area served by the website de-

[14] For more detail on this proxy, see section 3.5.

pends on user behavior. To address this issue, we identify the relevant market to a CL website in two ways. In our baseline approach, we assume that CL websites serve primarily the "core" county (or set of counties) containing the place indicated in the headline of the respective local website. The place is usually a single city or town, but can also be a combination of two or three nearby cities, a region, or, in some cases, an entire state. As an alternative, we identify all the counties that account for a non-negligible share of the ads posted on the website. We describe this "broad" definition of CL markets, which we use for purposes of robustness, in Appendix A.2.1.

3.2 Newspaper characteristics and outcomes

We collect comprehensive data on a range of relevant newspaper characteristics and outcomes from the Newspaper Yearbooks published by Editor & Publisher for each year between 1995 and 2010. We accessed print copies the yearbooks and digitized the information contained in them using OCR software. The yearbooks contain detailed information for over 1,500 US daily newspapers, including: address of the headquarters (HQ), circulation, subscription prices, number of pages published, as well as the list of staff members with names, broad job categories, and job titles. Figure A4 shows how the information is reported in the yearbooks for two example newspapers. We identify newspaper markets in two alternative ways. Our baseline approach is to assume that the newspaper market coincides with the county where the newspaper HQ is located. Since this a good approximation for the median newspaper,[15] this approach is common in the literature (Gentzkow and Shapiro, 2010; Seamans and Zhu, 2014). In Appendix A.2.2 we describe an alternative approach based on identifying all the counties where a given newspaper circulates using zip code-level circulation data. These data are available from the Alliance of Audited Media (AAM) though only for a subset of the newspapers in our sample. We use this approach for purposes of robustness.

Whether and to what extent a newspaper is affected by the entry of CL depends on how heavily it initially relied on revenues from classified ads. To measure this baseline difference across papers, we use information on the presence in a newspaper's staff of one or more classified ads managers prior to CL's entry (i.e. in 2000), available from E&P. To validate this measure, we collect information from https://www.newspapers.com, an online newspaper archive, on the number of pages devoted to classified ads in a subset of 262 newspapers. The results indicate that, prior to the entry of CL into their market, newspapers with classified ad managers devoted a significantly larger number of pages to classified ads than other papers (see Table A1). These papers devote about 35% more pages to classifieds than papers with no classified manager with similar circulation per capita and similar overall length.

[15] Our disaggregated subscription data confirms that the median paper in our sample has about 85% of its subscribers in the HQ county; see Appendix A.2.2.

3.3 Newspaper content

We are also interested in examining how changes in newspapers' organization triggered by the entry of CL influenced editorial decisions and news content. To this end, we use information from NewsBank, a database containing the text and metadata of articles published in over 800 newspapers in our sample, beginning in 1999. We use the data in two ways. First, we perform keyword searches on the full text of all articles looking for names of specific politicians (e.g., "Rep. Paul Ryan", "Senator Dianne Feinstein", etc.), and use the number of mentions in a given newspaper/year as a measure of the prominence of politicians, both national and local, in newspaper coverage. Second, we extract a random sample of 2 million articles and apply a semi-supervised topic model to the text of the lead paragraph to assess changes in the coverage devoted to various topics. Further details about the procedures used to construct these variables are reported in Appendix A.2.5.

3.4 Political outcomes

Finally, we examine how the entry of CL, and the subsequent changes in newspapers organization and content, influenced citizens' political behavior. To this end, we collect data on a variety of electoral outcomes measured at the county, the Congressional district, or the county-by-Congressional district level.

First, we look at electoral turnout, a standard measure of political participation. We use county-level turnout data from David Leip's Atlas of American Elections, covering both midterm and presidential elections between 1996 and 2010.

Second, we examine split-ticket voting, a measure of partisanship in voting that captures local candidates' ability to differentiate themselves from the national party brand. Following Darr et al. (2018), we measure split-ticket voting as the absolute value of the difference between the Republican candidate vote share in the presidential election and the Republican candidate vote share in House and Senate elections in the same county and year.[16]

Third, we examine a set of outcomes capturing the electoral performance of ideologically extreme candidates, as we expect the quality of the information available to voters to influence their ability to distinguish candidates on the ideological dimension (Hall and Lim, 2018). Following Hall (2015) and Dorn et al. (2020), we classify candidates based on their position in the distribu-

[16] Darr et al. (2018) use the Senate-President difference, but since only a third of Senate seats are contested in each election cycle, also using House races expands the number of observations available. We measure vote shares for each office at the county level; House (Senate) shares are computed by aggregating across all House (Senate) votes cast by voters in the county.

tion of campaign-finance-based ideological scores (CFScores) from Bonica (2016).[17]

3.5 Additional data sources

Browsing data. To measure CL take-up over time, as well as the downstream impact of CL on online news consumption, we use data from *Comscore*[18]. These data track the browsing behavior of a large sample of US Internet users, shown to be representative of US online buyers (De los Santos et al., 2012). The data cover the following years in our sample period: 2002, 2004, 2006 and 2007-2010. We aggregate the number of visits of the domain craigslist.org as well as total visits recorded by Comscore by county and year. To capture online news consumption, we rely on Comscore's classification of website categories available in the 2002 wave. We aggregate the number of visits by county for the 100 domains classified as news-related, and repeat this procedure for the following waves for the same set of domains.

Survey data. To explore the impact of changes in newspapers' content on readers' news consumption habits, we use individual data from two large-scale surveys. Our first source is the National Annenberg Electoral Survey (NAES), a nationally-representative rolling cross-sectional survey that was conducted in the lead-up to the 2000, 2004, and 2008 presidential elections. In particular, we use information on respondents' Internet access and self-reported media consumption in the week prior to the interview.

Our second source is the Survey of the American Consumer conducted by GfK Mediamark Research & Intelligence (GfK-MRI). The survey includes an extensive battery of questions about media consumption. In particular, we use information on respondents' self-reported readership of newspapers (in print and online, national and local), and news consumption on radio, TV, and on the Internet. We also exploit a question on what sections of the newspaper respondents usually read.[19] We use these data for the period 1999-2010.

Number of Internet service providers. To separate the impact of CL entry from a generic Internet effect, we control for a measure of the quality of local broadband Internet. In the absence of disaggregated data on Internet subscribers for this period, we follow Larcinese and Miner (2018) and Seamans and Zhu (2014) and use the number of Internet service providers (ISPs) registered by zipcode as a proxy.

[17] We use the 25th and 75th percentiles of the distribution of scores for all House candidates in 2000 as the thresholds separating "extremists" from "moderates." Our outcomes are the presence of an "extremist" in the general election, the shares of general election votes and individual contributions going to "extremist" candidates, and the absolute value of the CFscore of the winning candidate.

[18] https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/comscore/

[19] Figure A5 shows the fractions of GfK respondents reporting that they read each section.

These data are available for the period 1998-2008 from the Federal Communication Commission (FCC) and cover all providers with more than 250 high-speed lines in a state and transfer speed greater than 200 kilobits per second. We assign zero ISPs to all zip codes for the years before 1998, and use linear interpolation to fill missing data for years after 2008. We then aggregate the number of ISPs at the county level by taking the population-weighted average across all zip codes in a county.

This measure has been shown to be a strong predictor of the number of broadband subscribers at the state level, as well as at the county level in later periods when such disaggregated data are available (Larcinese and Miner, 2018). To further validate the number of ISPs as a proxy of local Internet penetration, we examine its correlation with self-reported Internet access from both the NAES and GfK-MRI surveys. Figure A8 confirms a strong positive relationship.

Other county characteristics. Throughout our analysis we also use data on the following county-level variables: population (from the National Center for Health Statistics), income per capita, share of the population in urban areas, share of the population with college education, share of the population who rent housing, racial composition and median age (all from the 2000 Census), as well as unemployment rate (from the Bureau of Labor Statistics).

4 Empirical strategy

4.1 Determinants of CL entry

To implement our empirical strategy, it is necessary to first understand what factors drove the timing of CL's staggered rollout. Anecdotal evidence suggests that CL prioritized larger markets and areas with good access to broadband Internet, which was crucial for the user to take full advantage of the platform. Importantly for our purposes, CL is not in the news business, and thus is not likely to have considered demand-side factors in the news market in determining where to enter. Also importantly, CL is privately held and has always operated as a mixture of profit-making business and community service. The fact that CL did not maximize profits,[20] allowed more flexibility for the idiosyncracies of its influential founder and its early user base to determine the timing of the rollout, rather than systematic factors in the underlying local news markets.

To validate the assumption of conditional exogeneity of CL's timing of entry, in Table 1 we

[20] In the ruling of a 2010 civil action case against Newmark by eBay, the judge concluded that "Craigslist does not expend any great effort seeking to maximize its profits or to monitor its competition or its market share." (eBay Domestic Holdings Inc. v. Newmark, Delaware Court of Chancery Civil Action No. 3705-CC, decision dated 2010-09-09, https://h2o.law.harvard.edu/cases/3472.) The absence of profit maximization motivated the suit, as eBay (which held a substantial stake in CL) argued that founder Craig Newmark had failed his fiduciary duty to maximize returns to shareholders.

examine the set of counties with headquartered newspapers in 2000, and regress the year of CL's entry into a given county on county characteristics. The coefficients on log population and number of ISPs confirm anecdotal accounts that population and the quality of the local Internet connection, as measured by number of ISPs, were important considerations in CL's entry decisions. The magnitudes of both coefficients are sizable: a one-standard deviation higher log population (number of ISPs) is associated with CL entering a county about 15 months (16 months) earlier. Together, these two variables account for 37% of the variation in year of entry. Given this strong relationship, our analysis will control for log population and number of ISPs throughout.

A few other demographic covariates, namely the rental share and the racial composition of the county in the 2000 census, have weaker correlations with entry time. These will be interacted with year dummies in our main regressions, to allow for the possibility of differential trends along these dimensions.

Crucially for the purpose of our analysis, we find no relationship between timing of entry and the state of local newspapers as measured by circulation, number of jobs or the presence of a dedicated classified manager, once log population and number of ISPs are conditioned on. In other words, we find no evidence of CL targeting particular newspaper characteristics in its entry decisions.

4.2 Newspaper-level regressions

To estimate the effect of CL entry on newspaper-level outcomes, we employ a difference-in-differences strategy exploiting CL's staggered introduction across US counties, combined with differences across newspapers in ex-ante reliance on classified advertising. The sample consists of the all newspapers covered by E&P, excluding ones with national circulation - i.e. the New York Times, USA Today and the Wall Street Journal.[21]

The following equations summarize our approach:

$$Outcome_{nct} = \beta PostCL_{ct} + \phi_n + \psi_t + \rho' X_{ct} + \upsilon_t' Z_{c0} + \varepsilon_{nct}, \tag{1}$$

$$Outcome_{nct} = \beta PostCL_{ct} + \gamma PostCL_{ct} \times ClassifMgr_{n0} + \phi_n + \psi_t + \rho' X_{ct} + \upsilon_t' Z_{c0} + \varepsilon_{nct} \tag{2}$$

$Outcome_{nct}$ is one of the outcomes of interest for newspaper n, headquartered in county c, at time t. $PostCL_{ct}$ is a an indicator variable equal to one for years after the entry of CL in county c and zero otherwise. In the baseline analysis we follow the "core" definition of CL markets (based

[21] The sample includes newspapers that never experience CL entry in their local market (which is the case for 45% of newspapers) – these "never-treated" newspapers serve as a control group. Yet, we obtain similar results excluding this group and relying on "timing-only" comparisons of eventually-treated newspapers (see Table B6).

TABLE 1: CORRELATES OF YEAR OF CL ENTRY

	Dependent variable: Year of CL entry					
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper jobs	-0.002	-0.004				
	(0.007)	(0.005)				
Newspaper circulation			-0.062	0.082		
			(0.576)	(0.529)		
Newspaper classif. mgr.					0.207	0.176
					(0.179)	(0.153)
Log population	-0.507***	-0.486***	-0.524***	-0.521**	-0.540***	-0.526***
	(0.158)	(0.180)	(0.181)	(0.240)	(0.156)	(0.189)
Number ISPs	-0.659**	-0.392*	-0.669**	-0.401*	-0.685**	-0.413*
	(0.307)	(0.213)	(0.304)	(0.212)	(0.296)	(0.210)
Share urban		0.014		0.013		0.012
		(0.009)		(0.010)		(0.009)
College degree		-0.035		-0.035		-0.036
		(0.024)		(0.025)		(0.024)
Rental share		-0.035		-0.037*		-0.033
		(0.021)		(0.020)		(0.021)
Income per capita		-1.113		-0.972		-1.262
		(1.274)		(1.273)		(1.338)
Unemployment rate		0.016		0.017		0.007
		(0.093)		(0.094)		(0.092)
Median age		-0.064		-0.068		-0.063
		(0.044)		(0.044)		(0.045)
Share White		0.081*		0.080*		0.077*
		(0.047)		(0.047)		(0.045)
Share Black		0.078*		0.076*		0.073*
		(0.044)		(0.043)		(0.042)
Share Hispanic		-0.010		-0.009		-0.009
		(0.012)		(0.012)		(0.012)
Turnout		0.877		0.717		0.858
		(2.559)		(2.588)		(2.626)
Observations	614	614	616	616	611	611
R^2	0.37	0.44	0.38	0.45	0.38	0.45

Notes: Regressions of year of CL entry on county characteristics in the year 2000. The sample consists of all counties with newspaper HQ in the year 2000. Standard errors clustered by state. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

on location indicated in the website name) and newspaper markets (based on county of newspaper HQ). ϕ_n and ψ_t are newspaper and year fixed effects, respectively. The vector X_{ct} includes time-varying controls for log population and number of ISPs, and the vector Z_{c0} includes addi-

tional county-level controls from the 2000 census interacted with year fixed effects. In alternative specifications, we also control for state $\times$ year fixed effects or for DMA $\times$ year fixed effects, thus restricting the comparison to newspapers operating in the same state or in the same media market. Finally, $ClassifMgr_{n0}$ is an indicator variable equal to one if the newspaper had a classified manager at baseline. We cluster standard errors by the area affected by the entry of a given CL website (i.e., a single county or group of counties), or, for newspapers never affected by CL, by county.

Equation 1 estimates the impact of CL's entry under the assumption that the timing of entry is conditionally uncorrelated with pre-existing trends in these outcomes. To evaluate the plausibility of this assumption, we also estimate a dynamic version of specification 1 for our main outcomes of interest. In light of recent work showing that two-way fixed effect estimates can be biased in settings where treatments effects are heterogeneous over time or across groups, we present event-studies based on the DID_M estimator proposed by de Chaisemartin and D'Haultfoeuille (2020).

4.3 County-level regressions

To estimate the impact of CL entry on outcomes measured at the county level, we estimate versions of equations 1 and 2 aggregated by county, i.e.:

$$Outcome_{ct} = \beta PostCL_{ct} + \phi_c + \psi_t + \rho' X_{ct} + \upsilon_t' Z_{c0} + \varepsilon_{ct}, \tag{3}$$

$$Outcome_{ct} = \beta PostCL_{ct} + \gamma PostCL_{ct} \times ClassifMgr_{c0} + \phi_c + \psi_t + \rho' X_{ct} + \upsilon_t' Z_{c0} + \varepsilon_{ct} \tag{4}$$

In this specification we focus on the sample of all counties where at least one newspaper was located at baseline, and compute $ClassifMgr_{c0}$ as the circulation-weighted average across newspapers headquartered in county c. We also exclude the counties corresponding to the headquarters of the three national-circulation newspapers – New York Times, USA Today and Wall Street Journal.

4.4 Congressional district-level regressions

For outcomes measured at the level of congressional districts, we estimate equations 3 and 4 at the level of county $\times$ congressional district cells. In other words, we duplicate outcome observations and assign one duplicate to each county contained in the district. We weight observations by the share of the voting-age population of the district accounted for by the respective county, and cluster standard errors by district. To absorb variation due to changing congressional district boundaries,

we include district by redistricting regime fixed effects in all regressions.[22] These fixed effects thus ensure that comparisons in the regressions are within fixed district boundaries.

5 Results

5.1 Craigslist take-up

We first confirm that Internet users in a given area were more likely to visit CL's URL following the opening of a local website. To this end, we use the data on web browsing behavior from Comscore, described in section 3. The dependent variable of interest is the IHS-transformed number of visits of the domain craigslist.org in a county in a given year, controlling for IHS-transformed total visits to any website recorded by Comscore.

In Figure 2 we plot an event-study for the effect of CL entry on CL take-up, estimated following the method proposed by de Chaisemartin and D'Haultfoeuille (2020). In addition to total Comscore visits we control for contemporaneous log population and number of ISPs. The graph indicates no pre-trend, which alleviates concerns about the (conditional) exogeneity of CL's entry. It suggests that take-up was immediate but further intensified over the following years. This pattern is consistent with network effects in the local adoption of the website: a larger number of users (and hence, a higher volume of local ads) likely increases the value of the platform and attracts yet more users.

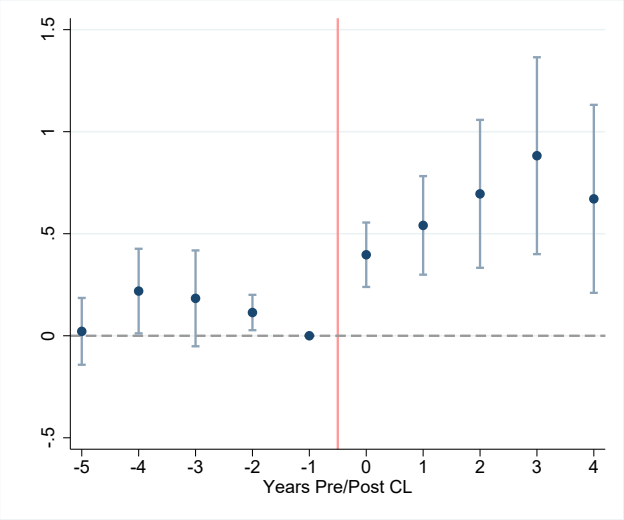
In table 2 we estimate versions of our baseline specification (equation 1) gradually including additional controls and fixed effects in the following columns. The results indicate that, after the entry of a local CL website, the number of visits to CL URLs increased significantly, by a magnitude between 16 and 44%. The estimated coefficient is lower but still significant conditional on controls for baseline county characteristics interacted with time (column 2), and is stable to controlling for state $\times$ year fixed effects (column 3) or DMA $\times$ year fixed effects (column 4), which restrict the comparison to counties within the same state or media market.

5.2 Number of newspapers

How did the entry of CL in a given market affect the number of active newspapers? One possibility is that the drop in revenues due to competition from CL might have been so extreme as to force some local papers to close or merge. In Table B8 we estimate our baseline specifications using as

[22] The major redistricting event in our sample period occurs following the decennial census in 2000, after which all states redrew district boundaries. A handful of states (North Carolina and Virginia in 1997, Texas in 2003, and Georgia in 2005) had additional significant district boundary changes, which we include as well. An example district-redistricting regime fixed effect would be GA-04-2005, which is treated as distinct from GA-04-2000.

FIGURE 2: VISITS TO CRAIGSLIST.ORG – EVENT STUDY



Notes: Dynamic effect of the availability of a local CL website on the number of craigslist.org visits (IHS-transformed) by county and year. Coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include total Comscore visits (IHS transformed), log population and number of Internet service providers. Standard errors clustered by CL-area.

TABLE 2: VISITS TO CRAIGSLIST.ORG

	Dependent variable: CL visits (IHS)			
	(1)	(2)	(3)	(4)
Post-CL	0.440***	0.155*	0.200***	0.171***
	(0.129)	(0.085)	(0.073)	(0.059)
Total Comscore visits (IHS)	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes
Observations	19896	19896	19896	19820
Number of counties	3053	3053	3053	3042
R^2	0.80	0.83	0.84	0.85
Mean dependent variable	2.42	2.42	2.42	2.41

Notes: Regressions of number of visits of the domain craigslist.org (IHS-transformed) by county and year on an indicator for the availability of a local Craigslist website. All specifications control for total visits recorded in Comscore (IHS-transformed). Baseline county controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

dependent variable the number of newspapers that, according to the E&P data, were operating in a given county in a given year. In this case our unit of analysis is a county-year, and the sample includes all counties where at least one newspaper was headquartered in 2000 (i.e., before CL's major roll-out). We find no evidence that CL entry affected the number of newspapers in a county, either looking at CL entry itself or at its interaction with ex-ante reliance on classified ads.[23] These results do not support the view that CL contributed to the disappearance of local newspapers. Yet, it is possible that lower advertising revenues affected the organization, functioning, and content of newspapers in other ways, a hypothesis which we explore below.

5.3 Newspaper staff size and composition

Next we turn to the question of how the entry of CL influenced the organization of newspapers. The first outcome we focus on is staff size.

Figure 3 displays an event-study for the impact of CL-entry on the number of jobs reported in the E&P yearbooks by newspaper and year, controlling for log population and number of ISPs. The figure suggests no significant pre-trend and shows a negative effect that intensifies over time.

In Table 3 we estimate our baseline specifications with number of jobs as the dependent variable.[24] Results in columns 1-4 indicate that following the entry of CL, the number of jobs in local papers decreased significantly. The effect amounts to about 1.2 fewer jobs, or about 6% relative to the mean, and is robust to the inclusion of additional controls and finer fixed effects. The results in the last four columns confirm that the effect is driven by newspapers that relied more on classified ads prior to CL. For such papers the magnitude of the effect is about 14% relative to the mean.

Robustness checks. This finding is robust to a number of additional checks. First, although newspapers with a classified manager are significantly larger, our results are unaffected by allowing for heterogeneity in the effect of CL entry by newspaper size. In fact, an interaction of *Post-CL* with baseline circulation leaves the estimates of the interaction with classified manager unchanged (Table B2), and the relative effects on newspapers with classified manager are if anything somewhat stronger in the sample of newspapers with below-median circulation at baseline (Table B3). Second, the results do not depend on our simplifying assumptions on the geographic scope of CL and newspaper markets. We find similar effects when we define CL markets based on the locations of ads posted on each website, and define newspaper markets based on geographically disaggre-

[23] Given the lack of evidence that CL affected newspapers' exit and entry, in the remainder of our analysis we use the full unbalanced panel as baseline. We obtain similar results for the balanced panel of newspapers that remained in the sample for the entire period of analysis.

[24] Alternatively, we examine the number of unique employees instead of the number of job titles. These two variables can differ if, for example, financial difficulties push a newspaper to assign to the same person multiple jobs that were previously carried out by different people. Table B1 shows that we obtain very similar results.

gated circulation data (Table B4). Third, newspaper entry, exit, or mergers, are not driving these results, as we obtain similar estimates when we focus on a balanced panel of newspapers operating continuously between 1995 and 2010 (Table B5). Fourth, rather than using newspapers never affected by CL-entry as a control group, we can restrict the comparison to newspapers that are affected by CL-entry at some point in time (Table B6). Finally, in addition to ruling out pre-trends in the entire sample of newspapers, in Figure B1 we show that there are no discernible pre-trends within the group of newspapers with classified managers, as well as within the group of newspaper without a classified manager.

Job Types. One important question that our data allow us to examine is what categories of workers were most affected by staff cuts. Based on the staff category and indicated in the E&P data, we can determine whether a staff member holds a managerial or an editorial position, and, for editorial staff, we can in some cases identify the corresponding topical area (e.g., politics, sports, entertainment, etc.).

In the left hand-side panel of Table 4 we estimate specification 2 separately for the number of jobs reported by E&P in each staff category. We differentiate between corporate /general management, advertising management, news executives and editorial management, and a residual category of other staff which captures mostly production and tech. The results indicate that in newspapers reliant on classified ads at baseline, CL entry leads to cuts in all types of positions, with magnitudes ranging from 23% for advertising management (which includes classified ads managers) to 7% for other staff, expressed relative to the mean. Importantly, staff cuts are not limited to management positions and also affect news executives and editors – their number number declines by about 9%.

In the right hand-side panel of Table 4 we focus on news executives and editors and use their individual job titles to gauge the news topics they likely to cover.[25] We code keywords related to three common topics: politics, sports, and entertainment.[26] The results indicate that newspapers exposed to competition from CL show a significant decline in the number of dedicated political editors after CL's entry, while we find no significant impact for sports or entertainment. One interpretation of this result is that, when facing financial difficulties, newspapers affected by CL opted to cut staff especially in areas — like local politics — for which producing quality content is more costly.

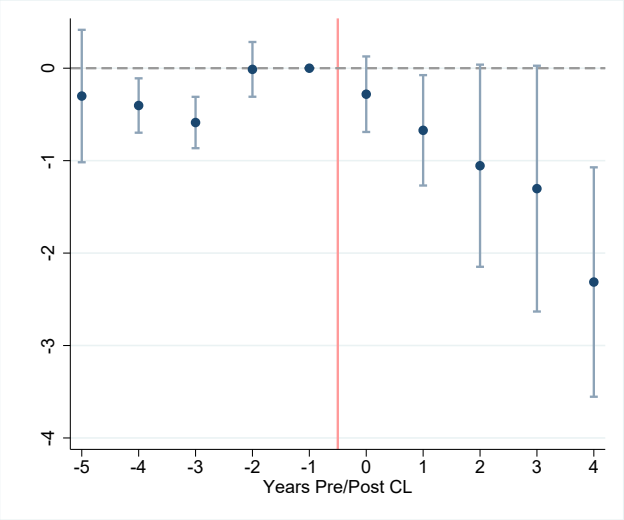
Despite the significant downsizing of editorial staff, we find no effect on the number of pages published in an average issue reported by E&P (Table B9). Together, these results imply increased workload per editor and may have implications for the distribution of editorial priorities. We

[25] This is a similar approach to Fang (2013) and **?**.

[26] We classify editorial jobs containing the keyword "sports" as sports related, ones containing the keywords "entertainment", "lifestyle", "women", "travel", "film", "music" as entertainment related, and ones containing the keywords "politics/ political", "government", "Washington", "city", "local" as politics related.

examine this issue further in the next section, where we look at how the entry of CL affected the evolution of news content.

FIGURE 3: NUMBER OF JOBS – EVENT STUDY



Notes: Dynamic effect of the availability of a local CL website on number of jobs by newspaper and year. Coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

5.4 Newspaper content

Next we explore how the transformations in newspapers' organization documented above translate into changes in news content, in general, and news coverage of politics, in particular. This is a question of paramount importance considering that, for most citizens, local newspapers still represent the main source of political information which allows them to monitor and keep politicians accountable (Mahone et al., 2019; Snyder and Strömberg, 2010).

Our first approach is based on a topic model applied to a random sample of articles. We estimate a Correlation Explanation (CorEx) topic model (Gallagher et al., 2017a) on a corpus consisting of the first paragraphs of over 2 million randomly drawn articles published in over 800 newspapers between 2000 and 2010. This model has two advantages: first, it tends to produce coherent topics for corpora consisting of short texts, and second, it allows us to define anchor words to generate topics corresponding to specific areas. Specifically, we seed the model to produce topics associated with: i) President, ii) Congress, iii) local politics, and iv) foreign policy.[27,28]

[27] Additional details about the procedure are reported in Appendix A.2.5.

[28] The resulting topics are not mutually exclusive, and do not sharply distinguish between different political sub-topics, e.g. local versus national politics. For instance, the Congress topic may include discussion of the local member of Congress, or discussion of the institution.

TABLE 3: NUMBER OF JOBS

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-2.031***	-1.126***	-1.198***	-1.235***	-0.252	0.322	0.244	0.146
	(0.398)	(0.392)	(0.420)	(0.405)	(0.465)	(0.472)	(0.513)	(0.494)
Post-CL × Classif. Mgr.					-3.629***	-3.027***	-2.995***	-2.873***
					(0.643)	(0.593)	(0.600)	(0.696)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22733	22733	22733	22411	22177	22177	22177	21830
Num. newspapers	1541	1541	1541	1523	1438	1438	1438	1418
R^2	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dep. var.	21.33	21.33	21.33	21.26	21.44	21.44	21.44	21.37

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4: JOBS BY TYPE AND TOPIC

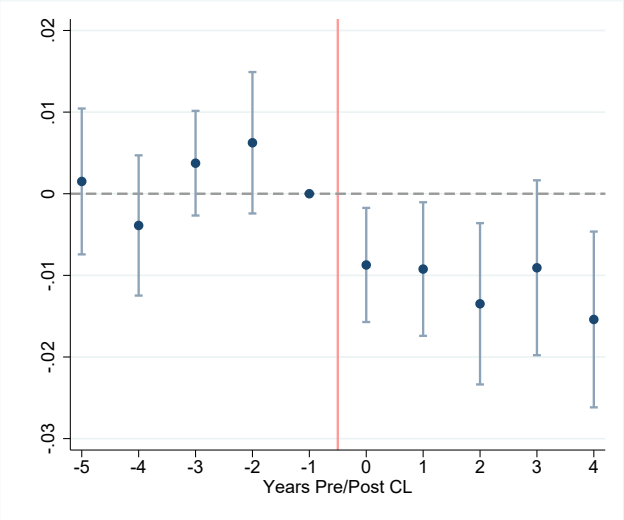
	Job category				Editors by job title		
	(1) Corporate/ General Mgmt.	(2) Advertising Mgmt.	(3) News Exec./ Editorial Mgmt.	(4) Other	(5) Politics	(6) Sports	(7) Entmnt.
Post-CL	0.005	0.079	0.032	-0.046	0.033	0.040	-0.052
	(0.084)	(0.049)	(0.350)	(0.126)	(0.047)	(0.046)	(0.054)
Post-CL × Classif. Mgr.	-0.620***	-0.464***	-1.014**	-0.433**	-0.150**	0.071	-0.090
	(0.145)	(0.080)	(0.459)	(0.212)	(0.063)	(0.085)	(0.092)
Log pop., # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20363	20363	20363	20363	19107	19107	19107
Num. newspapers	1439	1439	1439	1439	1416	1416	1416
R^2	0.80	0.80	0.89	0.76	0.74	0.59	0.76
Mean dep. var.	3.49	1.96	10.58	5.45	0.63	1.01	1.01

Notes: Regressions of number of jobs by category (left hand-side panel) and number of editors by topic specified in the job title (right hand-side panel) on an indicator the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

For each article in the corpus, we obtain a probability associated with each of the 10 topics, and aggregate the probability weights by newspaper and year.

Figure 4 presents an event-study for the impact of CL entry on the probability weight on the topic associated with Congressional politics, which suggests a significant decline. In Table 5 we present the results of specification 4 with dependent variables corresponding to the probability of each one of the 10 topic. The results indicate a general decline in news coverage of politics: newspapers most exposed to competition from CL saw a drop in significant decline in topic prevalence across all the anchored political topics. For the remaining un-anchored topics (i.e., accidents, events, traffic, obituaries, sports, crime) we find mostly insignificant coefficients on the interaction *Post-CL* $\times$ *Classif. Manager*.

FIGURE 4: WEIGHT OF THE TOPIC "CONGRESSIONAL POLITICS" - EVENT STUDY



Notes: Dynamic effect of the availability of a local CL website on the average probability that an article covers Congressional politics by newspaper and year. Coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

To further explore changes in political coverage, we examine how frequently local newspapers report about local representatives and national politicians. Using keyword searches in the NewsBank archive, we identify all articles published in a given newspaper and year that contain the names of the House and Senate representatives elected in the state where the newspaper is headquartered, as well as the names of candidates for these offices.[29]

In Table 6 we estimate our baseline specifications for the IHS-transformed number of articles mentioning the names of Congressional candidates and representatives by newspaper and year, controlling for (IHS-transformed) total number of articles. The results suggest that, following the

[29] Details about the keyword searches are reported in Appendix A.2.5.

TABLE 5: EFFECT OF CL ENTRY ON TOPIC WEIGHTS

Politics-related topics

	(1) presid, feder, govern, compani, tax	(2) council, mayor, board, plan, student	(3) repres, senat, congress, republican, elect	(4) intern, war, foreign, iraq, militari
Post-CL	0.000	-0.001	-0.002	0.004
	(0.005)	(0.007)	(0.003)	(0.003)
Post-CL × Classif. Mgr.	-0.020***	-0.014*	-0.009**	-0.011**
	(0.006)	(0.008)	(0.004)	(0.005)
Log population, # ISPs	Yes	Yes	Yes	Yes
County char. × Year FEs	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	7069	7069	7069	7069
Number of newspapers	855	855	855	855
R^2	0.52	0.47	0.40	0.54
Mean dependent variable	0.21	0.31	0.10	0.10

Other topics

	(1) man, kill, injuri, injur, accid	(2) music, art, food, festival, featur	(3) car, vehicl, driver, road, truck	(4) di, born, funer, son, daughter	(5) game, team, coach, win, season	(6) polic, charg, court, arrest, judg
Post-CL	-0.001	-0.006	-0.003	-0.002	0.008	-0.000
	(0.004)	(0.006)	(0.004)	(0.009)	(0.007)	(0.003)
Post-CL × Classif. Mgr.	0.002	-0.006	-0.005	0.022*	0.011	-0.000
	(0.006)	(0.007)	(0.006)	(0.012)	(0.008)	(0.004)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
County char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7069	7069	7069	7069	7069	7069
Number of newspapers	855	855	855	855	855	855
R^2	0.43	0.41	0.39	0.56	0.44	0.43
Mean dependent variable	0.11	0.17	0.14	0.15	0.21	0.11

Regressions of the average probability of an article covering a particular topic by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

entry of CL, Congressional coverage declines significantly in newspapers that relied more heavily on classified ads. In that case the differential decline in coverage amounts to about 30% of the mean.

In Table B11 we repeat the same exercise for national party leaders: the President and the leadership of both parties in each chamber of Congress. We find no significant effect on coverage of these national politicians. A possible explanation for this pattern is that coverage of national politicians can be sourced from wire services, whereas coverage of the district's representatives is original content produced by in-house reporting staff. It is exactly these staff, per the results in Table 4, whose jobs were most likely to be cut following CL's entry.

TABLE 6: MENTIONS OF LOCAL CONGRESSIONAL REPRESENTATIVES AND CANDIDATES

	Dependent variable: Articles on congressmen/ candidates from the same state (IHS)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.032	-0.099	-0.070	-0.058	0.115	0.047	0.052	0.108
	(0.077)	(0.083)	(0.084)	(0.090)	(0.101)	(0.104)	(0.102)	(0.105)
Post-CL × Classified Mgr.					-0.280**	-0.287***	-0.236**	-0.331**
					(0.109)	(0.110)	(0.109)	(0.137)
Total articles (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	7462	7462	7446	7005	7407	7407	7391	6934
Number of newspapers	884	884	883	853	878	878	877	845
R^2	0.82	0.83	0.85	0.88	0.82	0.83	0.85	0.88
Mean dependent variable	4.10	4.10	4.10	4.08	4.10	4.10	4.11	4.07

Notes: Regressions of the (IHS-tranformed) number of articles mentioning the name of a candidate in a House or Senate race in the same state by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. All specifications condition on the (IHS-transformed) total number of articles recorded by Newsbank by newspaper and year. Baseline county controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

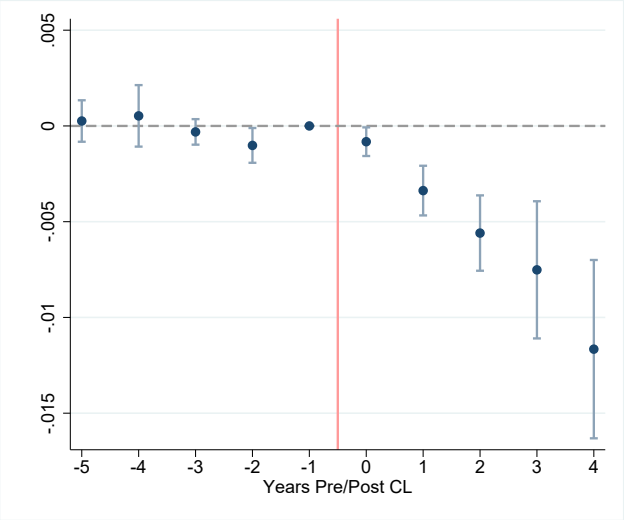
5.5 Newspaper readership

How did readers react to the changes newspaper organization and content documented above? One way to tackle this question is by looking at the evolution of the total number of copies sold, i.e., newspaper circulation, available from E&P for all newspapers in our sample.

The event-study for CL's impact on per capita circulation, presented in Figure 5, suggests a significant negative effect that increases in the years following the entry.

In Table 7 we estimate our standard set of regressions gradually including baseline county controls interacted with time, state × year and DMA × year FEs. The coefficient on *Post-CL* is negative and significant, and fairly stable across specifications. The last four columns again suggest that this is driven by newspapers ex-ante reliant of classified ads, for which the decline in per capita circulation amounts to about 4% relative to the sample mean.

FIGURE 5: CIRCULATION – EVENT STUDY



Notes: Dynamic effect of the availability of a local CL website on circulation per capita by newspaper and year. Coefficients and 95% confidence intervals based on the time-corrected DID_M proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

We can verify this decline in circulation using self-reported newspaper readership, available from the NAES and GfK-MRI surveys. These data include several questions related to media consumption. While we have only limited information on readership of specific newspapers, we are able to differentiate between respondents who report most frequently reading a national newspaper, i.e. the *New York Times*, *USA Today* or *the Wall Street Journal*, and the rest. We rely on residents' county of residence to match them to locally headquartered newspapers, and re-define the variable *Classif. Mgr.* as the circulation-weighted average across newspapers based in that county (following specification 4).

TABLE 7: CIRCULATION

	Dependent variable: Newspaper circulation per capita							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.006***	-0.004***	-0.003*	-0.004**	-0.002	0.000	0.002	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Post-CL × Classif. Mgr.					-0.009***	-0.008***	-0.008***	-0.009***
					(0.003)	(0.003)	(0.002)	(0.003)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22959	22959	22959	22638	22316	22316	22316	21970
Num. newspapers	1556	1556	1556	1538	1439	1439	1439	1419
R^2	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.99
Mean dep. var.	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19

Notes: Regressions of circulation by newspaper and year, measured in thousands of copies, on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In the first two columns of Table 8 we report the results for self-reported readership of non-national newspapers from individual-level regressions. In all specifications we control for a respondent characteristics including age, race and education. In both surveys we find reductions of newspaper readership, with a magnitude of 3 to 6% relate to the respective sample means, which is comparable to our results on circulation. Since in both surveys respondents are asked about reading either the print or the online version of the newspaper, these results indicate that the decline in circulation documented in Table 7 does not merely reflect substitution of print editions with online editions.

The decline in readership documented above is consistent with at least two explanations. First, it is possible that newspapers respond to the shock to classified ad revenues by increasing their subscription prices, which in turn would lead to lower demand. We test for this mechanism in Appendix Table B10, looking at the impact of CL's entry on yearly subscription prices reported in the E&P yearbooks. We obtain insignificant and if anything negative coefficients on both the indicator for CL-entry and its interaction with classified manager. Hence, there is no clear evidence

of an increase in subscription prices.[30]

A second explanation may be that readers respond to the changes in content brought about by CL's entry. One possibility is that the change toward less coverage of politics that we document in Tables 6 and 5 alienated readers interested in this type of content. Alternatively, the fall in circulation may be driven by readers who were primarily interested in classified ads which, after the entry of CL, became relatively less appealing. Though in both cases some readers would ultimately be less exposed to news and political content, understanding which of these scenarios is more plausible can shed light on which segments of the population were most affected by the entry of CL.

To understand this question, it is useful to first get a sense of how many readers were interested in these different newspaper sections at baseline. Information on this is available for a sample of 100,519 respondents from the 1999-2001 waves of the GfK-MRI survey. The distribution of readers' preferences, depicted in Figure A5, indicates that most readers (63%) report reading the "General News" section (which includes politics), with the Sports and Business sections also being popular (38% and 37%, respectively). The Classified section is not far behind, however, with 34% of respondents reporting Classifieds as one of the sections they frequently read. It is, therefore, possible that a reduction in the value of print classifieds might be a driver of circulation declines.

To understand what types of readers drive the drop in readership of local papers following the entry of CL, we examine heterogeneity in the readership effect by propensity to read the classified versus general news sections. Using the 1999-2001 waves of the GfK-MRI survey, we estimate an elastic-net penalized regression model to identify the individual characteristics that are most predictive of reading the general news and the classifieds sections respectively.[31] Based on the model estimated in the 1999-2001 data, we then project two propensity scores for respondents in the following years. This procedure allows us to assign to each respondent in the post-CL surveys a probability for reading general news and one for reading classifieds. Projecting based on pre-CL data allows us to focus attention on differential changes among demographic types who would have been likely to read either classifieds or political news prior to CL entry, without the confound of the post-CL changes to newspapers' product. The projected propensity scores have fairly strong negative correlation (with correlation coefficient of -0.4), indicating that the groups that tend to read each section are relatively distinct.

[30] Seamans and Zhu (2014) on the other hand find a significant increase in subscription prices in response to CL entry. Our analysis differs in using more complete data (i.e. covering all newspapers in the period 1995-2010) and a somewhat different empirical strategy, e.g. looking at reliance on classified ads at baseline rather than contemporaneously, to account for its likely endogeneity.

[31] The characteristics most strongly associated with general news reading are white race, having a post-graduate degree, income in the 75K-150K range, being retired, being married 25 years or more, and being age 45-49. The characteristics most associated with classifieds reading are being unemployed, living in a small to moderate sized county, having a high school diploma only or "some college", being engaged (to be married), and being 25-29 years old.

We then re-estimate the individual-level readership regressions separately for two groups of respondents: i) those with above-median probability of reading classifieds and below-median probability of reading general news, and ii) those with below-median probability of reading classifieds and above-median probability of reading general news. The results, reported in Table 8, indicate that the decline in readership after the entry of CL is entirely driven by individuals with high news propensity, and is again more pronounced where newspapers relied more heavily on classifieds ex ante. We find no such heterogeneity with respect to classifieds propensity. Though newspapers which offered the most classifieds were most affected by the CL shock, we find that readers interested in news rather than classifieds are the ones decreasing their readership.

Taken together, these results support the view that the main driver of circulation reduction was the indirect shift in news content induced by newspapers' revenue loss, rather than the direct effect of the obsolescence and disappearance of print classified ads.

TABLE 8: SELF-REPORTED NEWSPAPER READERSHIP: HETEROGENEITY

	Dependent variable: Read newspaper, dummy					
	NAES	GfK-MRI	GfK-MRI			
	Full sample	Full sample	High news	Low news	High classif.	Low classif.
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	0.015	-0.004	0.008	-0.015*	-0.001	-0.007
	(0.011)	(0.007)	(0.009)	(0.009)	(0.009)	(0.011)
Post-CL × Classified Mgr.	-0.047*	-0.014*	-0.029***	0.004	-0.013	-0.016
	(0.027)	(0.008)	(0.010)	(0.011)	(0.011)	(0.013)
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	105246	243502	121740	121749	122533	120959
Number of counties	1185	775	768	761	774	761
R^2	0.06	0.09	0.08	0.05	0.08	0.11
Mean dependent variable	0.75	0.43	0.53	0.32	0.38	0.47

Notes: Regressions of self-reported newspaper readership on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race indicators. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.6 Substitution to other news sources

To understand the implications of the documented decline in readership, a crucial question is to what extent this decline is offset by consumption of other news sources. If for instance readers merely substitute newspapers for other sources that cover similar content, such mode-switching, though detrimental for newspapers, would not necessarily imply a reduction in overall news consumption.

Both the NAES and the GfK surveys include questions about news consumption via sources other than local newspapers, which we use to study substitution patterns. The results are reported in Tables B12 and B13. In both surveys we find positive but statistically insignificant coefficients on the interaction of *Post-CL* and *ClassifMgr* on the probability that respondents read a national newspaper. The tables also show no significant effect of CL entry on news consumption on TV, radio, and online.[32,33]

The browsing data from Comscore allow us to perform an alternative test for the substitution to online news sources. In Table 9 we examine the effect of CL's entry on the number of visits of 100 domains classified by Comscore as news-related. The results, presented in Figure 6 and Table 9, indicate no significant effect of CL's entry on visits of these domains.

Taken together, these findings suggest that the decline in readership of local newspapers associated with the entry of CL is not fully compensated by increased news consumption online or through other media. These effects are therefore likely to translate in a net decline in exposure to political information.

5.7 Political outcomes

In the previous sections we documented that newspapers affected by the entry of CL reported less about politics, in general, and local politicians, in particular. We also found that individuals in areas affected by CL experienced a decline in readership of local papers not compensated by increased consumption of other news sources. In this section, we examine how these changes affected the behavior of local voters. Given existing evidence on the relationship between exposure to political information and citizens' political decisions, it is plausible that changes in news content and newspaper readership may have ramifications for downstream political outcomes. We focus on outcomes examined in the existing literature on media, political participation, and electoral accountability. Specifically, we investigate how the entry of CL affects: i) voters' propensity to turn out in elections (Gentzkow et al., 2011), ii) to rely on national partisan cues when voting for

[32] The analysis of online news consumption with NAES data is not possible since the question is not asked consistently across waves.

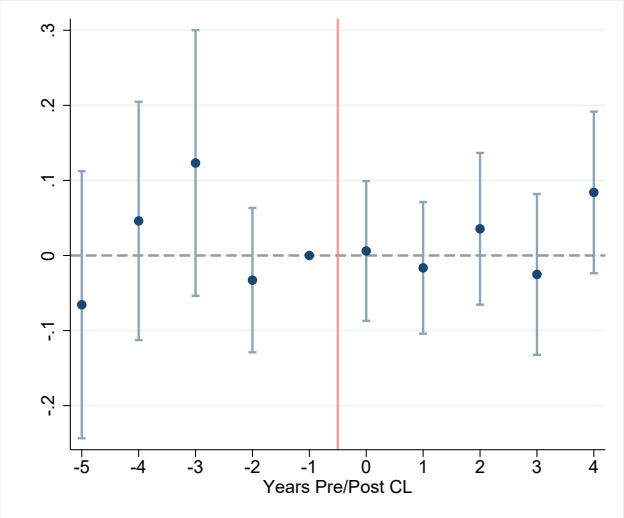
[33] We also find no significant positive effects on consumption of other media when we consider heterogeneity with respect to news propensity or classified propensity.

TABLE 9: ONLINE NEWS CONSUMPTION: VISITS OF NEWS-RELATED WEB DOMAINS

	Dependent variable: Visits top 100 news websites (IHS)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.022	-0.022	0.005	0.040	-0.029	-0.018	0.020	0.063
	(0.028)	(0.028)	(0.032)	(0.035)	(0.037)	(0.037)	(0.041)	(0.046)
Post-CL × Classified Mgr.					0.019	-0.008	-0.028	-0.054
					(0.047)	(0.047)	(0.049)	(0.053)
Total Comscore visits (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	8202	8202	8202	8035	8155	8155	8155	7988
Number of counties	1199	1199	1199	1175	1192	1192	1192	1168
R^2	0.88	0.88	0.89	0.90	0.88	0.88	0.89	0.90
Mean dependent variable	7.89	7.89	7.89	7.88	7.89	7.89	7.89	7.87

Regressions of number of visits of news-related web domains (IHS-transformed) by county and year an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE 6: ONLINE NEWS CONSUMPTION: VISITS OF NEWS-RELATED WEB DOMAINS – EVENT STUDY



Notes: Dynamic effect of the availability of a local CL website on number of jobs by newspaper and year. Coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

local candidates, measured by the incidence of split-ticket voting (Darr et al., 2018; Moskowitz, 2021; Trussler, 2020), iii) to vote for ideologically extreme candidates (Hall, 2015; Hall and Lim, 2018; Dorn et al., 2020).

Turnout. As described in section 3, electoral turnout is measured at the county level, and turnout data are available for both presidential and house elections. We assume that newspapers affect voters' behavior in the county in which they are based, which limits the sample to the 1,234 counties in which, according to the E&P data, at least one newspaper is headquartered. Since turnout is defined at the county-level, we estimate equation 4 as our baseline specification. Given the considerable differences in the number of voters across counties, we weight observations by the county's total voting-age population.

In the first three columns of Table B14 we examine the effect of CL on voter turnout. We find some evidence of a negative effect of CL on turnout in presidential elections, which is concentrated in counties where newspapers relied more on classified ads. However, as shown in the last three columns of Table B14, we find a smaller and non-significant effect when looking at turnout in House elections, for which one might expect the decline in news coverage of local members of Congress to be especially relevant. We conclude that there is suggestive but not conclusive evidence that the changes in newspaper content and readership induced by the entry of CL reduced aggregate electoral participation. One explanation of this result is that the reduction in local

newspaper readership primarily affected individuals with characteristics associated with high electoral turnout (i.e., older, higher income, higher education). These readers were, for the most part, not marginal voters, and even if exposed to less local political news were unlikely to drop into abstention.

Split-ticket voting. Though changes in the information environment may not impact *whether* people vote, they may affect *how* they vote. In particular, less exposure to local political news could reduce voters' ability to evaluate local candidates and their platforms. This could encourage them to rely more on national party labels when deciding on down-ballot races.

To test this hypothesis, in Table 10 we study the impact CL's entry on split-ticket voting, i.e., voters' tendency to support candidates from different parties in concurrent elections. We focus on presidential election years since split-ticket voting is defined by comparing candidates' vote shares in presidential vs. congressional elections. The results indicate that, following the entry of CL, voters become significantly less likely to split their vote between candidates of different parties. As for the other outcomes, the effect is driven by areas where newspapers were most vulnerable to CL's competition, where split-ticket voting drops by 13% of the sample mean. This finding suggests that the impoverishment of local papers, and the associated cuts in news coverage of local politics, can favor the "nationalization" of local elections (Moskowitz, 2021; Trussler, 2020), which, in turn, further incites partisan divisions and ideological polarization.[34]

Support for extreme candidates. Finally, we study whether the entry of CL favored the emergence and success of ideologically extreme candidates. The hypothesis we test is that a coarser information environment, by making it harder for voters to acquire information about candidates' ideological positioning, makes the entry of more extreme candidates more likely and improves their electoral prospects.[35]

We examine this question in Table 11 focusing on House elections and looking at the following outcomes: i) the probability that ideologically extreme candidates win a primary election, ii) the vote share they obtain in the general election, iii) the individual campaign contributions they attract, and iv) the ideological extremity[36] of the winning candidate. Outcome variables i), iii) and iv) are defined at the electoral district level, while ii) varies at the county-by-district level. As explained in section 4, all regressions include district $\times$ redistricting regime fixed effects to absorb

[34] This result relates to similar findings by Darr et al. (2018) regarding the impact of newspaper closures on polarization. Our results indicate that closures are not a necessary condition, and that the impoverishment of local newspapers can produce similar consequences.

[35] This is in line with findings by Hall and Lim (2018) who document that the advantage of extreme candidates is concentrated in areas with low news coverage.

[36] Measured by the absolute value of the difference between the candidate's CFScore and the CFScore of the median House candidate in 2000.

TABLE 10: SPLIT-TICKET VOTING IN HOUSE AND SENATE VS. PRESIDENTIAL ELECTIONS

	House/Senate-Pres R Vote Differential					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL (core)	-0.008*	-0.012**	-0.008*	0.0009	-0.003	0.0010
	(0.005)	(0.005)	(0.004)	(0.007)	(0.006)	(0.005)
Post-CL (core) × Classif. Mgr.				-0.015*	-0.013*	-0.014**
				(0.008)	(0.007)	(0.006)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year-Office FEs	Yes	Yes	No	Yes	Yes	No
State-Year-Office FEs	No	No	Yes	No	No	Yes
Observations	8,076	8,076	8,076	8,076	8,076	8,076
R^2	0.36	0.38	0.62	0.36	0.38	0.62
Mean dependent variable	0.11	0.11	0.11	0.11	0.11	0.11

Regressions of split-ticket voting by county and election year on an indicator for the availability of a local Craigslist website, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Split-ticket voting is defined as the absolute value of the difference between the Republican candidate vote share in the Presidential election and the Republican candidate vote share in the House or Senate election(s) in the same county. The regression is stacked (with one observation corresponding to either a House or Senate election) and we allow the fixed effects for (state-)year and county to vary by office (House or Senate). County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. Observations are weighted by voting-age population. OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the effect of changes in district boundaries. Hence, we exploit variation over time within a fixed district boundary. To simplify the presentation, we report the results from the specification with the interaction term *Post-CL* × *Classif. Mgr.* and the full set of time-varying controls.

The results in column 1 indicate that, following the entry of CL and in districts where newspapers were most affected, the probability that a candidate with an extreme CFScore wins a primary election increases significantly by about 7 percentage points. We also find a positive and significant effect on the vote share of these candidates in the general election, and on the share of total campaign contributions they receive (columns 2 and 3). Finally, in column 4, we find that the entry of CL in areas where newspapers were most vulnerable, is associated with a significant increase in the ideological score of the winning candidate (relative to the median among House candidates in 2000).

Taken together, our findings suggest that the profound transformations in the media landscape triggered by the entry of CL had a tangible impact on electoral politics in the US, and contributed to increase ideological polarization and further divisions across party lines.

TABLE 11: IDEOLOGICAL POLARIZATION

	Extremist in General	Vote Share of Extremists	Contrib. Share of Extremists	Winner CFScore Dev. from 2000 Median
	(1)	(2)	(3)	(4)
Post-CL	-0.006	-0.037*	-0.008	0.002
	(0.029)	(0.020)	(0.024)	(0.017)
Post-CL × Classif. Mgr.	0.070**	0.056**	0.047*	0.047**
	(0.030)	(0.024)	(0.025)	(0.023)
Log population, #ISPs	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	12,862	12,862	12,597	11,104
R^2	0.58	0.64	0.74	0.86
Mean dependent variable	0.74	0.38	0.51	0.72

Regressions of electoral outcomes by county × district cell and year on an indicator for the availability of a local Craigslist website, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Even-numbered years from 1996 to 2010 are included. Vote shares are computed by county × district cell; other outcomes are defined at district level only. The column (1) outcome is an indicator for the presence in the general election of a candidate with CFScore outside the central 50% interval of House candidates in 2000, following the method of Dorn et al. (2020). Columns (2) and (3) are the share of general election votes and contributions from individuals of such candidates. Column (4) is the absolute value of difference between the CFScore (Bonica, 2014) of the candidate who won the election and the median CFScore of all House candidates in the 2000 cycle. "District" means unique combination of state, congressional district number, and redistricting regime (either 1991 or 2001 for all states, plus 1997 for VA and NC, 2003 for TX, and 2005 for GA). All regressions control for log population, number of ISPs, and baseline county controls interacted with year FEs. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. Observations are weighted by the share of the county in the district's voting-age population. OLS regressions in all columns. Standard errors clustered by district. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusion

Hamilton (2004) lays out the basic economics of the news-gathering business: high fixed costs – in the form of reporting staff who must develop expertise in their subjects and form long-term relationships with their sources – combined with a non-excludable product lead generally to underprovision of news production relative to the social optimum. Counteracting this unhappy equilibrium to some degree are reporters' professional norms, which value the production of "hard news" and investigative journalism over cheaper-to-produce and sometimes more popular "soft news."

For a time in the 20th century, local monopoly papers were able to extract sizable profits from the advertising business. Reporters employed by those papers captured some of these rents in the

form of resources dedicated to reporting of local political news and other "hard" topics valued by journalists themselves (rather than readers or advertisers). The growth of advertising profits, in fact, can be directly tied to the emergence of the ideal of an independent press staffed by professional journalists, in contrast to the 19th-century norm of newspapers operated as propaganda organs of local party organizations (Petrova, 2011).

The emergence of competition in the advertising business from new internet-based entrants in the early 2000s upset this tenuous balance, eliminating the economic profits which had supported investments in money-losing but high-prestige reporting. We show that the entry of one particularly important such competitor, the classified advertising platform Craigslist, had severe impacts on newspapers' staffing levels and production of news coverage relating to local politics.

The Craigslist effect is not simply a consequence of changes to the demand for news induced by internet availability; rather, it appears to operate by reducing newspapers' ability to invest in local reporting resources. Papers that were especially reliant on classified advertising in the pre-Craigslist period saw much larger changes on these dimensions than comparably internet-exposed but less classified-dependent papers. The loss of advertising revenues at these papers seems to have particularly reduced political coverage and especially coverage of local representatives, an area with large positive externalities but also large private costs for newspaper operators.

Consistent with existing work on media effects on political outcomes, we find that there were measurable social consequences of this change in the production of news content. Voters in areas served by papers affected most by the Craigslist shock saw their Congressional elections become more nationalized, which we interpret as a consequence of thinner information about the local incumbent's behavior. Changes in the media landscape may thus be an important driver of the overall trend towards nationalization of elections in the United States (Hopkins, 2018).

The change in voters' information about candidates also had consequences for ideological polarization. We show that the reduction of representative-specific information led to greater entry and better electoral performance by relatively extreme candidates at the expense of their more moderate peers. This change provides evidence of newspapers' role in providing information about candidates' ideological and issue positioning. The contraction in coverage generated by Craigslist's entry diminished voters' ability to distinguish between moderate and extreme candidates, reducing the electoral penalty to taking positions "out-of-step" with district preferences (Canes-Wrone et al., 2002). The coarseness of voters' information environment thus provides a plausible explanation for the documented failure of Downsian convergence in candidate positioning in the U.S. House (Fowler and Hall, 2016) and for the relatively weak relationship between candidate positions and vote intentions (Tausanovitch and Warshaw, 2018).

Our results have implications for our understanding of the link between advertising market structure and the market for news. They highlight the fragility of compensating the production of a

public good – politically relevant information – with proceeds from bundled advertising. Technological innovation that unbundles the two products, as Craigslist did for classified advertising, can have spillover effects on the news market, with significant and lasting consequences for the quality of representation and political polarization.

References

ALLCOTT, H., L. BRAGHIERI, S. EICHMEYER, AND M. GENTZKOW (2020): "The welfare effects of social media," *American Economic Review*, 110, 629–76.

ANGELUCCI, C., J. CAGE, , AND M. SINKINSON (2020): "Media Competition and News Diets," *NBER Working Paper No.26782*.

ANGELUCCI, C. AND J. CAGÉ (2019): "Newspapers in Times of Low Advertising Revenues," *American Economic Journal: Microeconomics*, 11(3), 319–364.

BESLEY, T. AND R. BURGESS (2002): "The Political Economy of Government Responsiveness: Evidence from India," *Quarterly Journal of Economics*, 117, 1415–1451.

BHULLER, M., T. HAVNES, J. MCCAULEY, AND M. MOGSTAD (2020): "How the Internet Changed the Market for Print Media," *Working Paper*.

BONICA, A. (2014): "Mapping the Ideological Marketplace," *American Journal of Political Science*, 58, 367–386.

——— (2016): "Database on Ideology, Money in Politics, and Elections," Stanford, CA: Stanford University Libraries. https://data.stanford.edu/dime.

CAMPANTE, F., R. DURANTE, AND F. SOBBRIO (2018): "Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation," *Journal of the European Economic Association*, 16, 1094–1136.

CANES-WRONE, B., D. W. BRADY, AND J. F. COGAN (2002): "Out of step, out of office: Electoral accountability and House members' voting," *American Political Science Review*, 96, 127–140.

CHAN, J. AND A. GHOSE (2014): "Internet's dirty secret: assessing the impact of online intermediaries on HIV transmission," *MIS Quarterly*, 38.

CUNNINGHAM, S., G. DEANGELO, AND J. TRIPP (2019): "Craigslist's Effect on Violence Against Women," *Working Paper*.

DARR, J. P., M. P. HITT, AND J. L. DUNAWAY (2018): "Newspaper closures polarize voting behavior," *Journal of Communication*, 68, 1007–1028.

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, forthcoming.

DE LOS SANTOS, B., A. HORTAÇSU, AND M. R. WILDENBEEST (2012): "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102, 2955–80.

DORN, D., G. HANSON, AND K. MAJLESI (2020): "Importing political polarization? The electoral consequences of rising trade exposure," *American Economic Review*, 110, 3139–83.

FALCK, O., R. GOLD, AND S. HEBLICH (2014): "E-lections: Voting Behavior and the Internet," *American Economic Review*, 117, 2238–2265.

FANG, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103, 1598–1628.

FCC (2016): "Federal Communications Commission Report FCC 16-107," Tech. rep.

FOWLER, A. AND A. HALL (2016): "The elusive quest for convergence," *Available at SSRN 2262600*.

GALLAGHER, R. J., K. REING, D. KALE, AND G. V. STEEG (2017a): "Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge," *Transactions of the Association for Computational Linguistics*, 5, 529–542.

——— (2017b): "Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge," *Transactions of the Association for Computational Linguistics*, 5, 529–542.

GAVAZZA, A., M. NARDOTTO, AND T. VALLETTI (2019): "Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies," *Review of Economic Studies*, 86, 2092–2135.

GENTZKOW, M. AND J. M. SHAPIRO (2010): "What Drives Media Slant? Evidence from U.S. Daily Newsapers," *Econometrica*, 78, 35–71.

GENTZKOW, M., J. M. SHAPIRO, AND M. SINKINSON (2011): "The effect of newspaper entry and exit on electoral politics," *American Economic Review*, 101(7).

GEORGE, L. M. AND J. WALDFOGEL (2006): "The New York Times and the Market for Local Newspapers," *American Economic Review*, 96(1), 435–447.

HALL, A. B. (2015): "What happens when extremists win primaries?" *American Political Science Review*, 109, 18–42.

HALL, A. B. AND C. LIM (2018): "Ideology and News Content in Contested US House Primaries," Tech. rep., Working Paper.: http://andrewbenjaminhall. com/news. pdf.

HAMILTON, J. (2004): *All the news that's fit to sell: How the market transforms information into news*, Princeton University Press.

HOPKINS, D. J. (2018): *The increasingly United States: How and why American political behavior nationalized*, University of Chicago Press.

KROFT, K. AND D. G. POPE (2014): "Does online search crowd out traditional search? Evidence from Craigslist," *Journal Labor Economics*, 32(2), 259–303.

LARCINESE, V. AND L. MINER (2018): "Was Obama Elected by the Internet? Broadband Diffusion and Voters' Behavior in US Presidential Elections," *CESifo Working Paper No. 6882*.

LELKES, Y., G. SOOD, AND S. IYENGAR (2017): "The hostile audience: The effect of access to broadband internet on partisan affect," *American Journal of Political Science*, 61, 5–20, publisher: Wiley Online Library.

LEVY, R. (2021): "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 111, 831–70.

MAHONE, J., Q. WANG, P. NAPOLI, M. WEBER, AND K. MCCULLOUGH (2019): "Who's Producing Local Journalism? Assessing Journalistic Output Across Different Outlet Types," Tech. rep.

MANACORDA, M. AND A. TESEI (2020): "Liberation Technology: Mobile Phones and Political Mobilization in Africa," *Econometrica*, 88, 533–567.

MCCHESNEY, R. W. AND J. NICHOLS (2011): *The death and life of American journalism: The media revolution that will begin the world again*, Bold Type Books.

MOSKOWITZ, D. J. (2021): "Local news, information, and the nationalization of US Elections," *American Political Science Review*, 115, 114–129.

PETERSON, E. (2021): "Paper Cuts: How Reporting Resources Affect Political News Coverage," *American Journal of Political Science*, 65, 443–459.

PETROVA, M. (2011): "Newspapers and parties: How advertising revenues created an independent press," *American Political Science Review*, 790–808.

SEAMANS, R. AND F. ZHU (2014): "Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers," *Management Science*, 60(2), 265–540.

SNYDER, J. M. AND D. STRÖMBERG (2010): "Press coverage and political accountability," *Journal of Political Economy*, 118(2), 355–408.

STARKMAN, D. (2014): *The watchdog that didn't bark: The financial crisis and the disappearance of investigative journalism*, Columbia University Press.

TAUSANOVITCH, C. AND C. WARSHAW (2018): "Does the Ideological Proximity Between Candidates and Voters Affect Voting in U.S. House Elections?" *Political Behavior*, 40, 223–245.

TRUSSLER, M. (2020): "Get Information or Get in Formation: The Effects of High-Information Environments on Legislative Elections," *British Journal of Political Science*, 1–21.

A Background and Data

A.1 Background

FIGURE A1: EVOLUTION OF NEWSPAPER REVENUES BY SOURCE

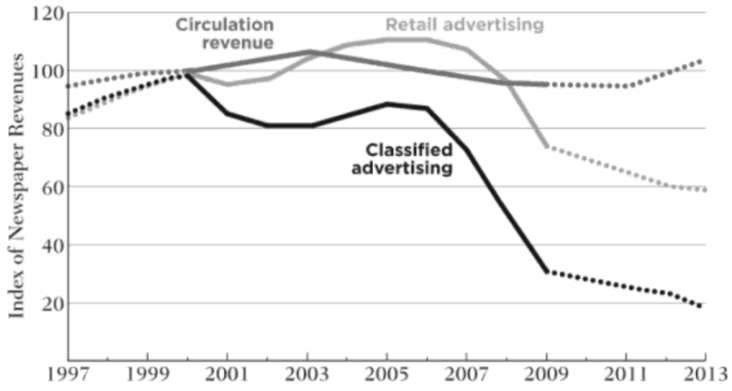
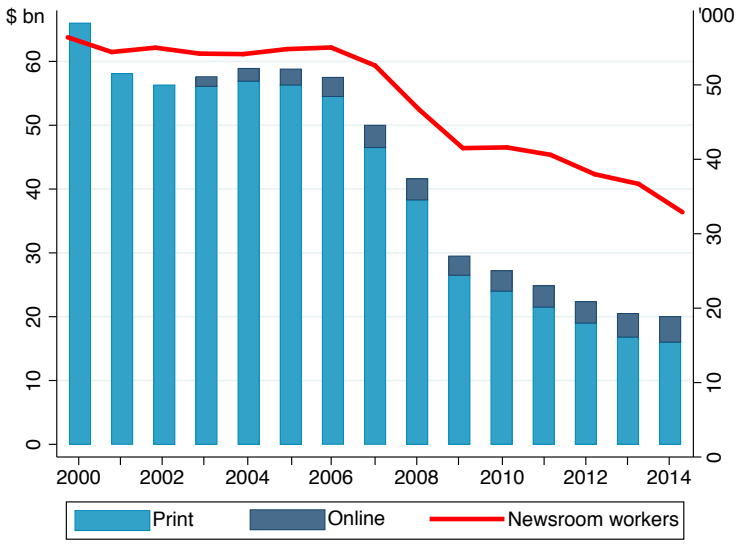


Figure 4: U.S. Newspaper Revenues over Time (Index: year 2000 = 100).

Notes: Index of newspaper revenues from circulation, retail advertising and classified advertising - 1997 to 2013. Source: Newspaper Association of America.

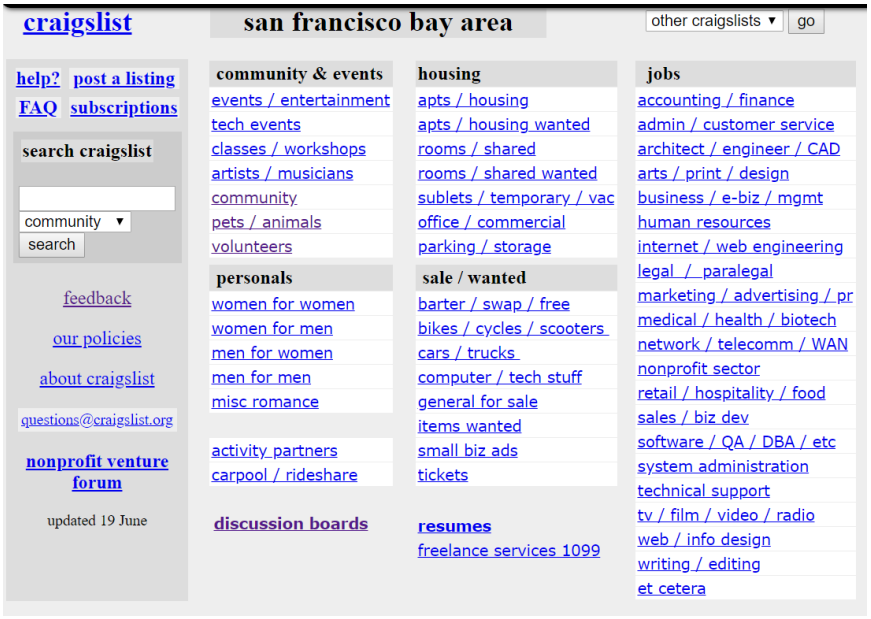
FIGURE A2: EVOLUTION OF NEWSPAPER REVENUES AND NUMBER OF NEWSROOM WORKERS



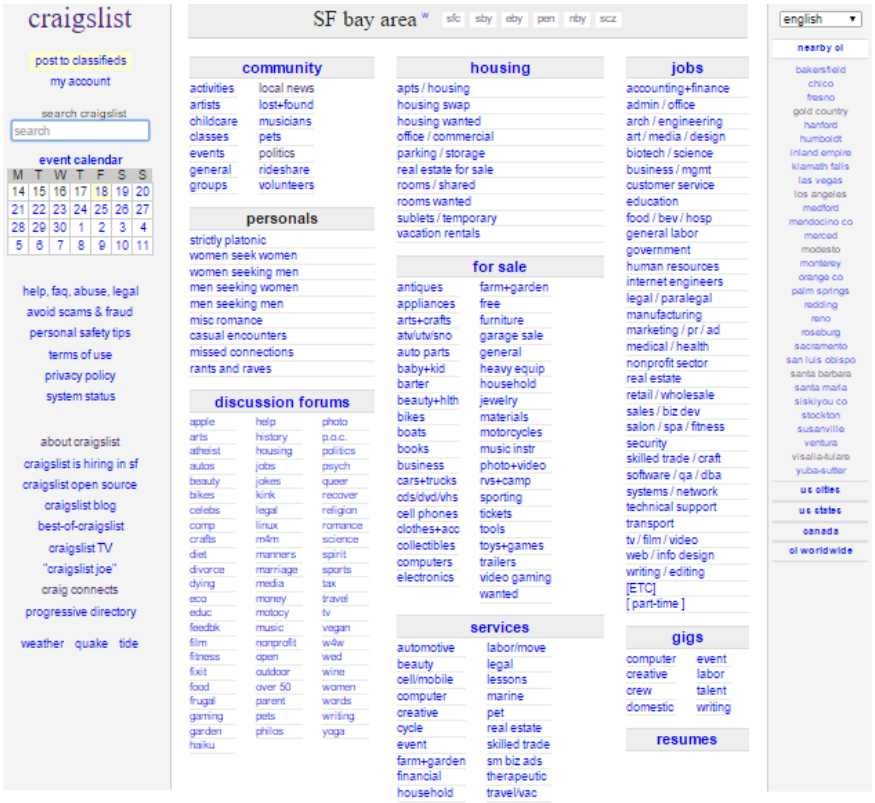
Notes: Newspaper revenues and number of newsroom workers - 2000 to 2014. Sources: Newspaper Association of America and American Society of Newspaper Editors.

FIGURE A3: CRAIGSLIST: LAYOUT IN 2000 AND 2016

2000

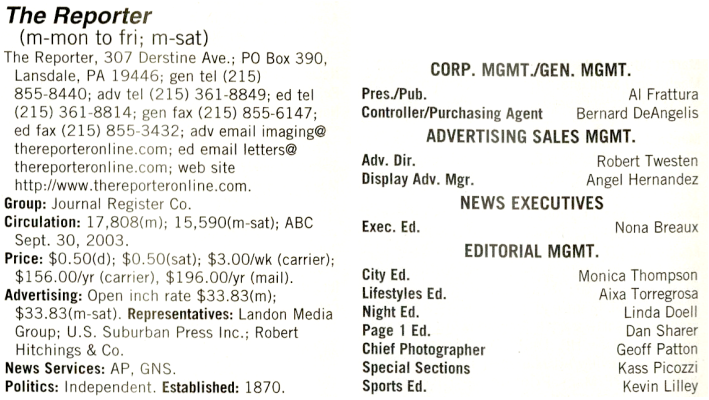


2016



Notes: Layout of CL's San Francisco/ Bay Area page in 2000 (upper panel) and 2016 (lower panel).

FIGURE A4: EXTRACT FROM THE EDITOR AND PUBLISHER YEARBOOK



The Reporter
(m-mon to fri; m-sat)

The Reporter, 307 Derstine Ave.; PO Box 390, Lansdale, PA 19446; gen tel (215) 855-8440; adv tel (215) 361-8849; ed tel (215) 361-8814; gen fax (215) 855-6147; ed fax (215) 855-3432; adv email imaging@thereporteronline.com; ed email letters@thereporteronline.com; web site http://www.thereporteronline.com.
Group: Journal Register Co.
Circulation: 17,808(m); 15,590(m-sat); ABC Sept. 30, 2003.
Price: $0.50(d); $0.50(sat); $3.00/wk (carrier); $156.00/yr (carrier), $196.00/yr (mail).
Advertising: Open inch rate $33.83(m); $33.83(m-sat). **Representatives:** Landon Media Group; U.S. Suburban Press Inc.; Robert Hitchings & Co.
News Services: AP, GNS.
Politics: Independent. **Established:** 1870.

CORP. MGMT./GEN. MGMT.

Pres./Pub.	Al Frattura
Controller/Purchasing Agent	Bernard DeAngelis

ADVERTISING SALES MGMT.

Adv. Dir.	Robert Twesten
Display Adv. Mgr.	Angel Hernandez

NEWS EXECUTIVES

Exec. Ed.	Nona Breaux

EDITORIAL MGMT.

City Ed.	Monica Thompson
Lifestyles Ed.	Aixa Torregrosa
Night Ed.	Linda Doell
Page 1 Ed.	Dan Sharer
Chief Photographer	Geoff Patton
Special Sections	Kass Picozzi
Sports Ed.	Kevin Lilley

The Reporter, Landsdale PA

Chicago Tribune
(m-mon to tues; m-wed to fri; m-sat; S)

Chicago Tribune, 435 N. Michigan Ave., Chicago, IL 60611; gen tel (312) 222-3232; gen fax (312) 222-2595; gen email tribletter@tribune.com; web site http://www.chicagotribune.com.
Group: Tribune Co.
Circulation: 680,879(m); 512,455(m-mon to tues); 571,576(m-sat); 1,002,166(S); ABC Sept. 30, 2003.
Price: $0.50(d); $0.50(sat); $1.79(S); $4.40/wk; $228.80/yr.
Advertising: Open inch rate $580.00(m); $580.00(m-sat); $842.00(S). **Representatives:** Western States Associates Inc.
News Services: AP, RN, NYT, TMS, DJ, KRT.
Politics: Independent. **Established:** 1847.
Advertising not accepted: Handguns, ammunition and tobacco.

CORP. MGMT./GEN. MGMT.

Pres./Pub./CEO	Scott C. Smith
Sr. Vice Pres./Gen. Mgr.	Richard Malone
Sr. Vice Pres./Ed.	Ann Marie Lipinski
Vice Pres., Circ./Consumer Mktg.	Vincent Casanova
Vice Pres./Chief Tech. Officer	Darko Dejanovic
Vice Pres., Adv. Mktg./Sales	Ken DePaola
Vice Pres., Finance	Phil Doherty
Vice Pres., Human Resources	Janice Jacobs
Vice Pres., Devel.	Owen Youngman
Vice Pres./Dir., Opns.	Tony Hunter
Gen. Mgr., Chicago Tribune Interactive	Alison Scholly
Dir., Technical Devel.	Scott Tafelski
Dir., Technical Opns./Help Desk	Robert Trinchet
Dir., Client Servs.	Deepak Agarwal

ADVERTISING SALES MGMT.

Dir., Nat'l Adv.	Dan Dunn
Dir., Network Adv.	Ron Goldberg
Dir., Classified Adv.	Barbara Swanson
Dir., Major Accts.	Douglas Thomas
Dir., Preprint Adv.	John Wollney
Dir., Adv. Planning/Analysis	Margaret Durkin
Dir., Adv. Devel.	Kathy Manilla
Dir., Regl. Accounts	Steve Brooks
Dir., Group Sales/Mktg.	Robert Fleck
Dir., Devel.	Susan Zukrow
Dir., Devel.	Sue Klose

MARKETING MGMT.

Sr. Mgr., Multimedia Mktg.	Tom Garritano
Dir., Community Rel.	Frank Gihan
Dir., Brand Mktg.	Kelly Shannon

CIRCULATION MGMT.

Dir., Distr.	Shelia Davidson
Dir., Consumer Mktg.	Carrie Hoye
Dir., Circ. Planning/Opns.	Becky Brubaker

NEWS EXECUTIVES

Mng. Ed.	James O'Shea
Public Ed.	Don Wycliff
Deputy Mng. Ed., Features	Jim Warren
Deputy Mng. Ed., News	George de Lama
Deputy Mng. Ed., Opns.	Randy Weissman
Assoc. Mng. Ed., Electronic News	Mark Hinojosa
Assoc. Mng. Ed., Features	Mary Elson
Assoc. Mng. Ed., Financial News	Rob Karwath
Assoc. Mng. Ed., Foreign News	Tim McNulty
Assoc. Mng. Ed., Graphics/Design	Stacy Sweat
Assoc. Mng. Ed., Lifestyle	Geoff Brown
Assoc. Mng. Ed., Metropolitan News	Hanke Gratteau
Assoc. Mng. Ed., Nat'l News	Joycelynn Winnecke
Assoc. Mng. Ed., Photography	Bill Parker
Assoc. Mng. Ed., Sports	Dan McGrath
Assoc. Mng. Ed., Washington Bureau	Vicki Walton-James
Sr. Ed.	Tony Majeri
Sr. Ed., Recruiting	Sheila Solomom

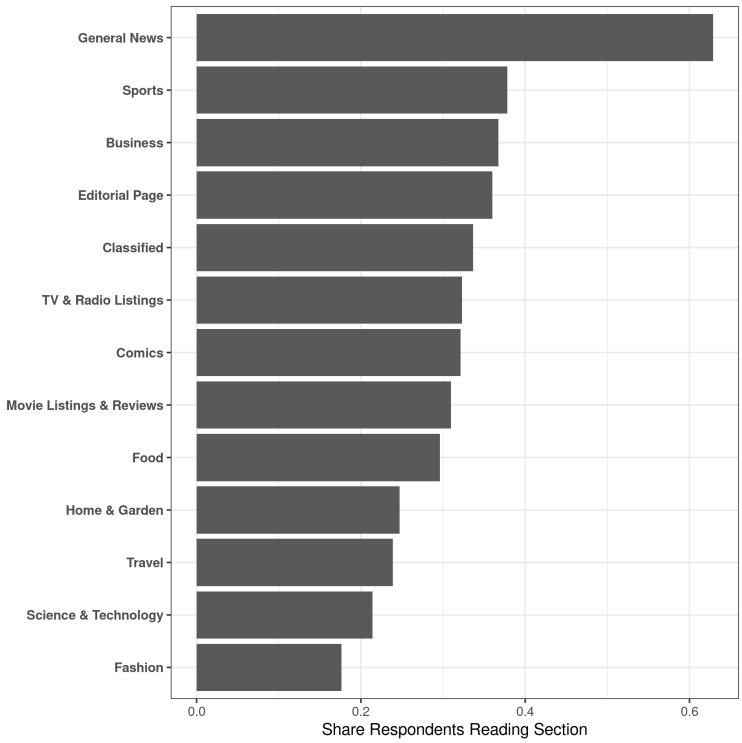
EDITORIAL MGMT.

Books Ed.	Elizabeth Taylor
Editorial Page Ed.	Bruce Dold
Entertainment Ed.	Scott Powers
Foreign Ed.	Colin McMahon
Good Eating Ed.	Carol Haddix
Nat'l Ed.	Storer Rowley
Special Sections Ed.	Janet Franz
Sports Ed.	Bill Adee
Sunday Magazine Ed.	Elizabeth Taylor
Tempo Ed.	Tim Bannon
Travel Ed.	Randy Curwen
Womanews Ed.	Cassandra West

The Chicago Tribune

Notes: Extracts from the print version of the 2003 Editor & Publisher Yearbook for the Landsdale Reporter (upper panel) and the Chicago Tribune (lower panel).

FIGURE A5: NEWSPAPER SECTIONS BY READERSHIP (1999-2001)



Notes: Distribution of readership by newspaper section (self-reported). Based on GfK-MRI survey waves for 1999 to 2001. The categories are not mutually exclusive.

A.2 Details on data construction and validation

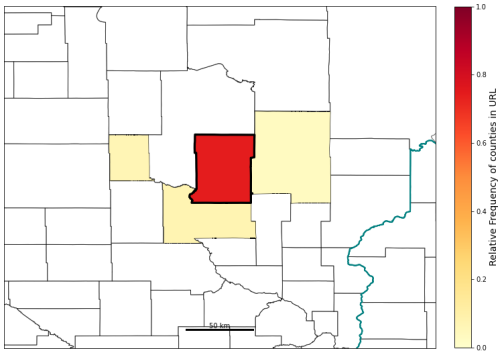
A.2.1 Craigslist markets

In our baseline analysis we assume that Craigslist markets consist of the county (or counties) containing the locality indicated in the website's url. In this section we discuss an alternative approach which relies on the locations indicated in ads posted on the respective websites.

To do so, we retrieve the snapshots of each website available from https://archive.org//, and code the exact location of all the ads posted on the first page of the "housing", "jobs", and "sales" sections. Here we focus on the ads post in the first two years after the entry. We then match the resulting locations to a comprehensive list of towns, cities, and counties (if the location includes the word "county") in the same or a neighboring state. Finally, we consider all counties that account for at least 5% of the ads as part of what we define as the website's "broad" market.

Figure A6 depicts the geographic distribution of ads posted on https://brainerd.craigslist.org/ in the 1st and 2nd year after the opening of the website. In this case, the "core" market is represented by the central county (Crow Wing County) containing the city of Brainerd. This "core" county accounts for over 80% of total ads, while the "broad market" includes five additional neighboring counties. This is a typical pattern in our data: on average the "core" market accounts for 73% (median 76%) of posted ads once we exclude outliers.

FIGURE A6: DISTRIBUTION OF ADS POSTED ON HTTPS://BRAINERD.CRAIGSLIST.ORG/



Notes: Geographic distribution of the location of ads posted in the housing, jobs and sales sections of https://brainerd.craigslist.org/ in years 1 and 2 after the website opening. Source: Internet Archive.

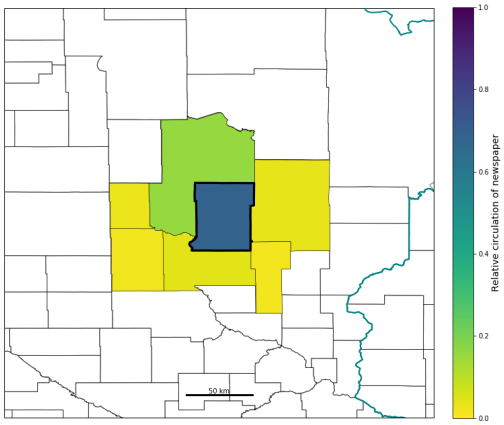
A.2.2 Newspaper markets

In our baseline analysis we assume that local newspaper markets consist of the county in which they are headquartered. In this section we discuss an alternative approach which relies on the availability of geographically dis-aggregated circulation data.

This method consists in matching a newspaper to all counties where it is read in proportion to circulation. Zip-code-level circulation data are available from the Alliance for Audited Media (AAM) for 2002, which we use to construct a weighted measure of CL availability by newspaper-year. The weights in this "broad" measure of CL availability are the fraction of the paper's subscribers in each county.[37] However, AAM data covers only about 40% of the papers in the E&P sample. For the newspapers for which no zip-code circulation data is available, we assign 100% of the circulation to the county where the paper's HQ is located.[38]

Figure A7 shows the geographic distribution of circulation for a newspaper in our sample, the Brainerd Dispatch, with darker colors representing higher values. Crow Wing County where the newspaper's HQ is located, shown in blue, accounts for 81% of the total.[39] The median paper in our AAM data has about 85% of its total circulation in the headquarters county once we exclude outliers.

FIGURE A7: DISTRIBUTION OF CIRCULATION OF THE BRAINERD DISPATCH



Notes: Geographic distribution of the circulation of the *Brainerd Dispatch* in 2002. Source: Alliance for Audited Media.

[37] We measure geographically disaggregated circulation only once, in 2002, and hence year-to-year variation is driven entirely by changes in CL availability and not by changes in circulation patterns.

[38] The papers that are missing from AAM are generally smaller papers and, if anything, less likely to have circulation beyond the county boundaries than the papers that appear in AAM. Papers which appear in AAM had median circulation in 2002 of 67K, compared to 14K for papers not appearing AAM. Hence, we believe that assigning all circulation to the headquarters county is a good approximation for these papers.

[39] Similarly to CL ads, we exclude outlier counties that account for less than 5% of total circulation.

A.2.3 Validating the classified manager proxy

We validate the classified manager indicator as a proxy for classified intensity using data from the website *Newspapers.com*, which archives digitized historical copies of newspapers. We located 262 papers in our dataset which appear in the Newspapers.com archive. For each of these papers, we sampled the edition of the paper published on the first Sunday of each month in all years from 1995 until 2010, substituting another day when the Sunday edition was not available.

We measure classified intensity as the number of pages on which the term "Classified" appears, divided by the total number of pages in the issue. We collected this measure for a total of 43,165 issues across the 262 papers available in the Newspapers.com archive. Prior to the entry of Craigslist in a market, the average paper in our sample averaged a little under 8 pages of classified advertising per issue, or about 15% of the typical issue's total page count.

We first examine cross-sectional variation in classified intensity prior to Craigslist entry. Table A1 shows the results of regressions where the outcome is the average number of pages per issue that contain classified advertising, by newspaper-weekday-year. We split by weekday, and include weekday fixed effects in all specifications, because weekends and especially Sundays tend to have much more classified advertising than other days.

TABLE A1: NUMBER OF PAGES DEVOTED TO CLASSIFIED ADS IN PRE-CL PERIOD, BY PRESENCE OF CLASSIFIED MANAGER IN 2000.

	Average Classified Pages per Issue		
	(1)	(2)	(3)
Classif. Mgr.	5.65***	4.86***	2.82**
	(1.65)	(1.56)	(1.31)
2000 Circ. per Capita		16.4*	9.19*
		(8.86)	(5.40)
Average Total Pages per Issue			0.069***
			(0.009)
Add'l County Characteristics	No	No	No
Day-of-week FEs	Yes	Yes	Yes
Observations	794	794	785
R^2	0.16	0.19	0.47
Mean dependent variable	7.7	7.7	7.7

Notes: Regressions of the average number of pages per issue devoted to classified advertising on an indicator for the presence of a classified manager in 2000, in the pre-CL period. Data from issues published in all years prior to the year of CL entry in the newspaper's market are included. OLS regressions in all columns. Standard errors clustered by newspaper.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

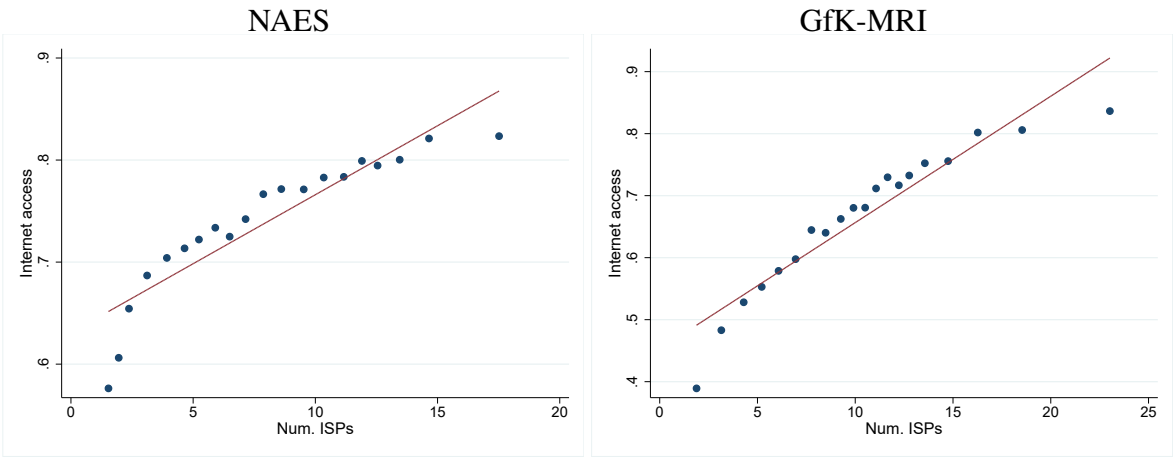
The table shows that amount of space devoted to classified ads was on average a little under three pages higher at papers that had a classified manager in 2000, prior to Craigslist entry. This is roughly 35% of the sample mean level of approximately 8 pages. The predictive power of the

classified manager dummy is preserved when controls for newspaper circulation per capita and the overall size of the paper (in pages) are added.

A.2.4 Validating number of ISPs as a measure of Internet quality

Throughout the paper, we use the number of ISPs registered in a county as a proxy for broadband internet penetration, a measure used in prior work (Larcinese and Miner, 2018; Seamans and Zhu, 2014; Lelkes et al., 2017). To further validate the number of ISPs as a proxy of local Internet penetration, we examine its correlation with self-reported Internet access from both the NAES and GfK-MRI surveys. Figure A8 shows a binned scatter plot comparing the number of ISPs with the fraction of respondents in the county who report having internet access at home (in GfK-MRI) or either at home or at work (in NAES), along with the OLS line of best fit. Both datasets show a strong positive relationship between the two.

FIGURE A8: NUMBER OF ISPS AS A PROXY FOR LOCAL INTERNET PENETRATION



Notes: Binned scatter plot for the relationship between number of ISPs available in the county of the respondent and self-reported Internet access at home (GfK-MRI) and at home or at work (NAES).

A.2.5 Details on Newspaper Content Processing

This section contains details on the procedures used to process raw text content from the newspapers in our sample to construct lower-dimensional representations of the content. Source data are from the *NewsBank* database. We conduct two main kinds of processing on text data: keyword searches and topic modeling. Keyword searches use the full database containing more than 100M full-text articles, while our topic model uses a smaller random sample of about 2M articles consisting of all articles published on 10 randomly sampled dates in each newspaper-year between 1999 and 2010. The topic-modeling sample limits to the first paragraph of text, plus the headline.

Politician Names Our first set of keyword searches look for the names of House and Senate representatives and candidates for office in US House and Senate races from the same state in which the newspaper is headquartered. We use a list of representatives from https://github.com/unitedstates/congress-legislators and a list of candidates who filed campaign finance reports from Bonica's (2016) Database on Ideology, Money in Politics, and Elections (DIME).

For each representative, we construct a (case-insensitive) regular expression of the following form: "(congress.*|rep*) Firstname Lastname" or "(senat.*) Firstname Lastname". This expression matches strings like "Rep. Adam Smith" or "Congressman Adam Smith" but not "Adam Smith" alone. We require the inclusion of the title to cut down on false positives, as many members of Congress have common names. This does introduce the possibility of false negatives, but we have found that articles covering a member usually include the title and full name at first mention before switching to a shorter form like "Mr. Smith".

For each candidate, we search for the number of articles which contain the candidate's standardized full name (in the form "Firstname Lastname"). To cut down on false positives, we exclude articles published in the Obituaries, Sports, Arts and Entertainment, Automotive, or Real Estate sections, and count only articles in which one of the terms "Congress," "Senate," or "US House" appears.

We count the number of *articles*[40] in which any of the list of same-state candidate names appears on each newspaper-day, and then aggregate to the level of newspaper by year.

Topic model Our method for extracting the topical coverage of affected newspapers follows Gallagher et al.'s 2017b Correlation Explanation (CorEx) method. This is a semi-supervised method that allows input of a minimal set of "anchor" words, and then finds topics by searching for groups of words that co-occur with the anchors. We apply this method to the text of a random sample of 2 million articles from the NewsBank corpus.

[40] I.e., each article that mentions the candidate at all counts as 1, regardless of how many times the candidate is referenced in the article.

We use the semi-supervised method rather than the more traditional unsupervised Latent Dirichlet Allocation (LDA) because it allows us to focus on specific topics of interest. We are interested in separating various dimensions of political news coverage: coverage related to local, congressional, national and foreign politics. We seed separate anchors for these 4 topics, and run the CorEx model with 10 topics in total. The (stemmed) anchor words we use are the following: ['washington', 'feder', 'govern', 'presid'], ['council', 'mayor'], ['repres', 'congress', 'senat'], ['intern', 'abroad', 'foreign'].

Figure A9 presents the resulting topics, as described by their most representative words. The 10 resulting topics can be labeled as follows: local politics, congressional politics, national politics, foreign politics, entertainment, health / family, weather, crime, obituaries.

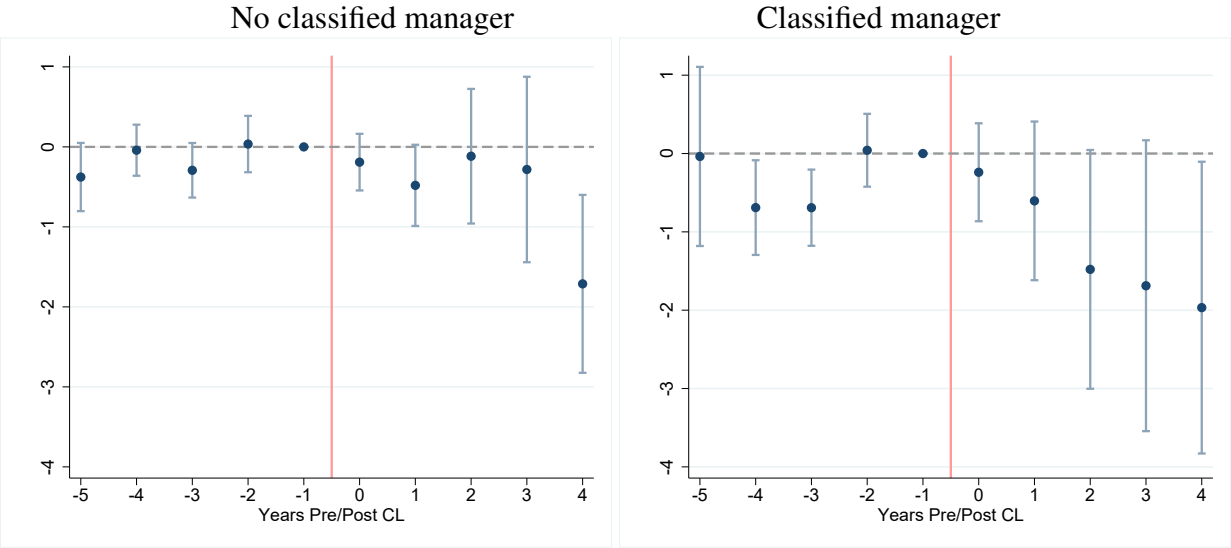
FIGURE A9: TOP TEN WORDS REPRESENTATIVE WORDS FOR EACH COREX TOPIC

```
0: presid,feder,govern,compani,tax,washington,percent,increas,pai,billion
1: council,mayor,board,plan,student,educ,fund,commun,project,program
2: repres,senat,congress,republican,elect,democrat,vote,candid,polit,gov
3: intern,war,foreign,iraq,militari,movi,film,american,soldier,terrorist
4: man,kill,injuri,injur,accid,crash,woman,diseas,victim,suffer
5: music,art,food,festiv,featur,concert,event,artist,band,holidai
6: car,vehicl,driver,road,truck,traffic,highwai,drive,mile,street
7: di,born,funer,son,daughter,church,surviv,servic,cemeteri,obituari
8: game,team,coach,win,season,plai,victori,footbal,score,player
9: polic,charg,court,arrest,judg,investig,attornei,accus,sheriff,suspect
```

For each of the 2 million articles in the corpus, the CorEx model outputs a set of 10 unconditional probabilities for the article belonging to that a topic. Importantly, these probabilities do not necessarily sum to 1 - an article can simultaneously belong to more than one topic, or to none. To examine the effects of CL's entry, we aggregate the distribution of probabilities by newspaper and year, and estimate the standard diff-in-diff equations specified in section 4, with the average probability for each one of the 10 topics as dependent variable.

B Additional Results

FIGURE B1: NUMBER OF JOBS – EVENT STUDY, SEPARATELY BY PRESENCE OF CLASSIFIED MANAGER



Notes: Dynamic effect of the availability of a local CL website on number of jobs by newspaper and year. Left hand-side: sample restricted to newspapers without classified manager at baseline. Right hand-side: sample restricted to newspapers with classified manager at baseline. Coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

TABLE B1: NUMBER OF EMPLOYEES

	Dependent variable: Newspaper number of employees							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.703***	-1.048***	-1.078***	-1.026***	-0.219	0.158	0.103	0.137
	(0.339)	(0.332)	(0.346)	(0.345)	(0.378)	(0.379)	(0.403)	(0.407)
Post-CL × Classif. Mgr.					-3.039***	-2.545***	-2.475***	-2.429***
					(0.566)	(0.530)	(0.530)	(0.610)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22806	22806	22806	22488	22249	22249	22249	21906
Num. newspapers	1541	1541	1541	1523	1438	1438	1438	1418
R^2	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.93
Mean dep. var.	17.86	17.86	17.86	17.80	17.93	17.93	17.93	17.86

Notes: Regressions of number of employees by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B2: NUMBER OF JOBS: CONTROLLING FOR POST-CL × BASELINE CIRCULATION

	Dependent variable: Newspaper number of jobs			
	(1)	(2)	(3)	(4)
Post-CL	0.264	0.930	1.119*	1.020
	(0.592)	(0.639)	(0.648)	(0.646)
Post-CL × Classified Mgr.	-3.481***	-2.848***	-2.687***	-2.545***
	(0.632)	(0.580)	(0.582)	(0.691)
Post-CL × Circ. 2000	-2.894	-3.530	-5.231**	-4.973**
	(2.093)	(2.269)	(2.197)	(2.321)
Log population, #ISPs	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes
Observations	22167	22167	22167	21820
Number of newspapers	1437	1437	1437	1417
R^2	0.90	0.91	0.91	0.92
Mean dependent variable	21.44	21.44	21.44	21.37

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline, and with circulation per capita at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B3: NUMBER OF JOBS: NEWSPAPERS WITH BELOW-MEDIAN BASELINE CIRCULATION

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.086*	-0.336	-0.222	-0.046	0.457	0.963	1.067	1.375*
	(0.593)	(0.626)	(0.679)	(0.667)	(0.691)	(0.732)	(0.783)	(0.807)
Post-CL × Classif. Mgr.					-3.954***	-3.400***	-3.504***	-3.600***
					(1.004)	(0.955)	(1.026)	(1.305)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	11145	11145	11069	10453	11025	11025	10949	10337
Number of newspapers	734	734	729	692	726	726	721	683
R^2	0.91	0.91	0.92	0.92	0.91	0.91	0.92	0.92
Mean dependent variable	18.47	18.47	18.52	18.88	18.51	18.51	18.57	18.93

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Sample restricted to newspapers with below-median circulation per capita in 2000. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B4: NUMBER OF JOBS: BROAD CL AND NEWSPAPER MARKETS

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL (broad)	-1.059***	-0.587**	-0.691**	-0.962***	0.484	0.570*	0.430	0.095
	(0.341)	(0.286)	(0.303)	(0.331)	(0.404)	(0.331)	(0.346)	(0.351)
Post-CL (broad) × Classif. Mgr.					-3.615***	-2.788***	-2.758***	-2.691***
					(0.494)	(0.449)	(0.446)	(0.523)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22688	22688	22688	22366	22153	22153	22153	21806
Num. newspapers	1534	1534	1534	1516	1436	1436	1436	1416
R^2	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dep. var.	21.30	21.30	21.30	21.23	21.44	21.44	21.44	21.37

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. Here we define the indicator for CL availability as 1 if a broad-CL market intersects the broad newspaper market (see Appendix sections A.2.1 and A.2.2). OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B5: NUMBER OF JOBS: BALANCED PANEL

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.991***	-1.030**	-1.120***	-1.151***	-0.234	0.362	0.270	0.157
	(0.414)	(0.402)	(0.429)	(0.421)	(0.482)	(0.485)	(0.525)	(0.510)
Post-CL × Classif. Mgr.					-3.597***	-2.938***	-2.878***	-2.771***
					(0.648)	(0.591)	(0.594)	(0.686)
Log pop., #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	21148	21148	21148	20779	20973	20973	20973	20588
Num. newspapers	1331	1331	1331	1308	1320	1320	1320	1296
R^2	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dep. var.	21.72	21.72	21.72	21.66	21.77	21.77	21.77	21.70

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Sample: balanced panel of newspapers in operation throughout 1995 to 2010. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B6: NUMBER OF JOBS: EXCLUDING "NEVER-TREATED" NEWSPAPERS

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.335	-0.089	-0.396	-0.318	1.389	1.222*	0.849	0.743
	(0.765)	(0.604)	(0.667)	(0.670)	(0.895)	(0.703)	(0.809)	(0.783)
Post-CL × Classif. Mgr.					-3.586***	-2.737***	-2.536***	-2.197***
					(0.645)	(0.587)	(0.619)	(0.775)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	12483	12483	12482	11668	12173	12173	12172	11358
Number of newspapers	844	844	844	795	785	785	785	735
R^2	0.90	0.90	0.91	0.92	0.90	0.90	0.91	0.92
Mean dependent variable	26.53	26.53	26.53	26.53	26.68	26.68	26.68	26.70

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Sample: newspapers that experience CL entry into their local market between 1995 and 2010. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B7: NUMBER OF JOBS: CL TAKE-UP AS CONTINUOUS TREATMENT

	Dependent variable: Newspaper number of jobs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CL	-0.148**	-0.078	-0.104**	-0.078*	0.040	0.053	0.021	0.029
	(0.061)	(0.049)	(0.048)	(0.045)	(0.045)	(0.043)	(0.046)	(0.046)
CL visits (IHS) × Classif. Mgr.					-0.473***	-0.368***	-0.366***	-0.332***
					(0.095)	(0.087)	(0.087)	(0.093)
Total Comscore visits (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	9309	9309	9309	9157	9221	9221	9221	9062
Num. newspapers	1407	1407	1407	1386	1394	1394	1394	1372
R^2	0.92	0.92	0.93	0.93	0.92	0.92	0.93	0.93
Mean dep. var.	18.41	18.41	18.41	18.34	18.45	18.45	18.45	18.38

Notes: Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B8: NUMBER OF NEWSPAPERS

	Dependent variable: Number of newspapers HQ-ed in county							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.009	-0.002	-0.006	-0.006	-0.003	-0.002	-0.007	-0.008
	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.012)
Post-CL × Classified Mgr.					-0.012	-0.001	0.001	0.005
					(0.013)	(0.013)	(0.013)	(0.014)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	18870	18870	18870	18488	18733	18733	18733	18335
Number of counties	1201	1201	1201	1178	1192	1192	1192	1168
R^2	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
Mean dependent variable	1.20	1.20	1.20	1.19	1.20	1.20	1.20	1.20

Notes: Regressions of the number of newspapers headquartered by county and year on an indicator for the availability of a local Craigslist website and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B9: NUMBER OF PAGES

	Dependent variable: Number of pages published							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.293	0.246*	0.130	0.298**	0.174	0.147	0.003	0.164
	(0.191)	(0.138)	(0.171)	(0.151)	(0.147)	(0.144)	(0.185)	(0.238)
Post-CL × Classified Mgr.					0.223	0.193	0.242	0.246
					(0.331)	(0.311)	(0.329)	(0.380)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	21646	21646	21646	21323	21211	21211	21210	20863
Number of newspapers	1472	1472	1472	1454	1400	1400	1400	1380
R^2	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.98
Mean dependent variable	28.51	28.51	28.51	28.44	28.53	28.53	28.53	28.45

Notes: Regressions of number of pages published by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B10: SUBSCRIPTION PRICE

	Dependent variable: Subscription price							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.276	-0.844	-0.963	-1.086	0.330	-0.213	-0.346	-0.681
	(0.896)	(0.934)	(1.038)	(1.254)	(1.223)	(1.222)	(1.413)	(1.551)
Post-CL × Classified Mgr.					-1.429	-1.512	-1.485	-0.873
					(1.524)	(1.542)	(1.607)	(1.605)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	19652	19652	19652	19259	19197	19197	19196	18788
Number of newspapers	1432	1432	1432	1412	1351	1351	1351	1330
R^2	0.93	0.94	0.94	0.95	0.93	0.93	0.94	0.94
Mean dependent variable	118.69	118.69	118.69	118.37	118.80	118.80	118.79	118.43

Notes: Regressions of yearly subscription price by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B11: MENTIONS OF NATIONAL PARTY LEADERS

	Dependent variable: Articles on national politicians							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.038	-0.024	-0.071	-0.024	-0.025	-0.024	-0.063	0.015
	(0.053)	(0.054)	(0.063)	(0.071)	(0.071)	(0.072)	(0.076)	(0.088)
Post-CL × Classified Mgr.					-0.028	-0.006	-0.020	-0.079
					(0.081)	(0.083)	(0.084)	(0.107)
Total articles (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	7462	7462	7446	7005	7407	7407	7391	6934
Number of newspapers	884	884	883	853	878	878	877	845
R^2	0.89	0.90	0.91	0.92	0.89	0.90	0.91	0.92
Mean dependent variable	1.44	1.44	1.44	1.47	1.44	1.44	1.45	1.47

Notes: Regressions of the (IHS-transformed) number of articles mentioning the names of a set of national party leaders by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. All regressions control for the (IHS-transformed) total number of articles recorded by Newsbank by newspaper and year. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B12: SELF-REPORTED MEDIA CONSUMPTION: GFK

	(1) Read online	(2) Read newspaper, national	(3) Watched TV	(4) Listened radio
Post-CL	0.006	0.003	0.005	0.009
	(0.007)	(0.004)	(0.008)	(0.009)
Post-CL × Classified Mgr.	-0.010	0.005	-0.015*	-0.001
	(0.008)	(0.005)	(0.009)	(0.009)
Respondent controls	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	243502	243502	243502	243502
Number of counties	775	775	775	775
R^2	0.16	0.14	0.08	0.09
Mean dependent variable	0.21	0.08	0.70	0.17

Notes: Regressions of self-reported media consumption on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race indicators. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B13: SELF-REPORTED MEDIA CONSUMPTION: NAES

	Read newspaper, national		Watched TV		Listened radio	
	dummy	days per wk	dummy	days per wk	dummy	days per wk
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	0.002	0.032	0.011	0.072	0.008	0.031
	(0.010)	(0.049)	(0.008)	(0.066)	(0.008)	(0.037)
Post-CL × Classified Mgr.	0.034	0.151	-0.014	-0.024	-0.005	-0.008
	(0.025)	(0.120)	(0.009)	(0.084)	(0.009)	(0.040)
Respondent characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	104895	104895	104979	104979	145136	145136
Number of counties	1185	1185	1185	1185	1186	1186
R^2	0.08	0.07	0.03	0.10	0.05	0.04
Mean dependent variable	0.03	0.14	0.92	5.04	0.37	1.51

Notes: Regressions of self-reported media consumption on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. Respondent controls include sex, age, an indicator for college degree and race indicators. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B14: TURNOUT

	Presidential Turnout			House Turnout		
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL (core)	-0.005	0.005*	0.003	-0.009**	-0.002	0.001
	(0.004)	(0.003)	(0.002)	(0.005)	(0.004)	(0.003)
Post-CL (core) × Classif. Mgr.	-0.002	-0.007*	-0.006**	0.002	-0.003	-0.003
	(0.006)	(0.004)	(0.003)	(0.005)	(0.004)	(0.003)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	No	Yes	Yes	No
State-Year FEs	No	No	Yes	No	No	Yes
Observations	4,907	4,907	4,907	9,690	9,690	9,690
R^2	0.95	0.96	0.98	0.90	0.92	0.95
Mean dependent variable	0.54	0.54	0.54	0.44	0.44	0.44

Notes: Regressions of turnout by county and election year on an indicator for the availability of a local Craigslist website, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Observations are weighted by voting-age population. Baseline county controls include share urban population, share college educated, rental share of housing, log income per capita, median age, share White/ Black/ Hispanic and turnout in Presidential elections, all measured in the year 2000. OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.