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

Milena Djourelova† Ruben Durante‡ Gregory J. Martin§

April 2024

Abstract

How does competition from online platforms affect the organization, performance, and editorial choices of newspapers? 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 newspapers reliant on classified ads experienced a significant decline in the number of management and newsroom staff, including in the number of editors covering politics. These organizational changes led to a reduction in news coverage of politics and resulted in a decline in newspaper readership, particularly among readers with high political interest. Finally, we document that reduced exposure to local political news was associated with an increase in partisan voting and increased entry and success of ideologically extreme candidates in congressional elections. Taken together, our findings shed light on the determinants of the decline of print media and on its broader implications for democratic politics.

Keywords: newspapers, Internet, advertising, ideological polarization

JEL codes: L82, L86, D72

*We thank the editor, four anonymous referees and numerous seminar and conference participants for helpful comments. David Ampudia Vicente, Marco Lo Faso, Lorena Mita, and Aaron Schade provided excellent research assistance. We also thank Irena Djourelova, Anna Griñó, Anastasiya Nebolsina, Adriana Oliveres and Silvia Terzulli for their help with digitizing the Editor & Publisher data. Ruben Durante acknowledges financial support from the European Union's Horizon 2020 research and innovation program [Grant 759885].

†Cornell University and CESifo. Email: mnd43@cornell.edu

‡National University of Singapore, BSE, CEPR, CESifo and IZA. Email: rubendurante@gmail.com

§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 a sharp decline in newspapers' revenues over the last two decades, forcing many news outlets to drastically rethink their business model and organization. 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 their representatives (Besley and Burgess, 2002; Snyder and Strömberg, 2010), they may also have important political implications.

Despite the potentially grave consequences of these transformations for the future of journalism, rigorous evidence on the impact of online platforms 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.

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. CL's entry disrupted the market for classified ads, formerly a lucrative niche for newspapers (Seamans and Zhu, 2014; Kroft and Pope, 2014), which accounted for 40% of newspapers' advertising revenues and about 30% of total revenues in the year 2000. 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, editorial decisions, and content production of local newspapers. We then trace the impact of these changes on local representation in the US Congress.

The expansion of CL in the US 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. CL's narrow focus on classified ads provides an additional source of variation, since the entry of CL should disproportionately affect newspapers that relied more heavily on classified ads *ex ante*. Finally, with only 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[1] alleviates 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

[1] 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.

correlated with the characteristics of local newspapers once population and the quality of the local Internet connection are controlled for.

To analyze the impact of CL on local newspapers, we collect data on the date of entry of each local CL website, and digitize comprehensive data on the organization and market outcomes of more than 1,500 newspapers (the universe of daily newspapers), covering the period from 1995 to 2010. Our empirical strategy uses two sources of variation. The first source of variation is CL's staggered introduction: we compare the evolution of outcomes of interest between areas with and without access to a local CL website, before and after the website is introduced. To make sure we separate the effect of CL from that of Internet penetration or other correlates of CL entry, we control for the quality of local broadband Internet, log population, and baseline county characteristics interacted with time fixed effects. While this variation identifies the average effects of CL entry, only newspapers in the classified ad market are plausibly exposed to competition from CL. Therefore, the second source of variation comes from differences 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. We validate this proxy by documenting that it is strongly correlated with both the share of a newspaper's total pages devoted to classified ads and with the rates charged by newspapers for classified ads before the entry of CL.

We begin our analysis by investigating the "first-stage" impact of the local entry of a CL website on the use of the platform and on newspapers' classified ad business. We verify that local take-up of the CL platform increases after the entry of a local CL website, estimating a significant increase in local visits to the craigslist.org domain. We also find evidence that this take-up leads to a significant decline in the volume of classified ads in local newspapers. Following CL entry, the share of classified pages in local newspapers declines by about 3 percentage points on average, or 10% relative to the sample mean. This effect is driven entirely by newspapers that relied more heavily on classified ads at baseline, i.e. those that had a dedicated classified manager. The magnitude of the decline for such newspapers is 4.2 percentage points, or 15% of the mean.[2] Since 25% of newspapers experience the entry of CL and have a classified manager at baseline, this estimate implies that CL led to an overall reduction in the share of classified pages of about 1 percentage point, or 12% of the overall decline observed between 1995 and 2010 (9 percentage points).

Second, we turn to the impact of CL on newspapers' organization, finding substantial downsizing. After the entry of CL into a county, newspapers headquartered in that county cut 1.1 jobs on average, or about 5% relative to the sample mean. This effect is driven entirely by newspapers that relied more heavily on classified ads at baseline, for which we estimate a decline of about 3 jobs

[2] We detect no changes in classified ad rates, implying that the quantity loss should translate proportionately into revenue loss. We can hence roughly approximate the resulting revenue loss using typical levels of classified pages and prices in our dataset. Our estimates of page and price changes imply a median revenue loss among classified-manager papers of at least \$137K-\$298K per year.

(14% of the sample mean). This accounts for 10% of the total decline in newspaper jobs observed between 1995 and 2010 (a 35% decline).

Staff cuts are not limited to advertising managers, but extend to managerial and editorial positions. For editorial staff, a detailed analysis of job titles indicates that the cuts affected political editors disproportionately, while editors covering other areas, such as sports and entertainment, were not significantly affected.

Third, we examine how these organizational changes affected newspapers' editorial priorities, with particular regard to the coverage of politics. To do so, we fit a topic model to a corpus of two million randomly drawn articles (covering ∼850 newspapers in our sample) and examine the distribution of newspapers' coverage across topics. Consistent with the notion that staff cuts are detrimental to the coverage of politics (Peterson, 2021), we find that CL entry is associated with an 8.3% decline in the prevalence of political topics, while "soft news" topics such as sports and entertainment are not affected. We complement this approach with full-text keyword searches for mentions of politicians' names in the universe of articles archived by NewsBank (covering ∼900 newspapers and more than 100M articles). We document that, following the entry of CL, news coverage of local congressional representatives and candidates for congressional offices declines by about 12%.[3] This decline is particularly pronounced for articles published prior to primary elections.

Fourth, we examine how readers respond to these changes in content. We document that, in the years after the entry of CL, affected newspapers experienced a decline in circulation per capita of about 4%, a finding we confirm using self-reported newspaper readership from two independent large-scale surveys of media consumption. The survey data further suggests that the drop in readership is not driven by readers that are likely to read the classified ads section of newspapers, but by readers interested in general news (among whom the decline in readership is as large as 15%). This finding is consistent with readers responding to changes in news content triggered by the entry of CL, rather than merely lower demand for print classified ads.[4] Evidence from both survey and browsing data suggests that the decline in newspaper readership is unlikely to be compensated by increased news consumption online.

Finally, we study how reduced news coverage of politics — and in particular, reduced coverage of local congressional candidates — affected local electoral outcomes, focusing on turnout and ideological polarization in congressional elections. We find that the entry of CL has no significant effect on voter turnout. In contrast, we document that it increased the correlation between results in the local House or Senate race and the national presidential race, reflecting a decline in split-ticket

[3] We document similar declines in the number of mentions of national politicians, such as the president or party leaders in Congress, as well as the number of mentions of titles of local- and state-level politicians.

[4] Lower readership is unlikely to be due to newspapers raising subscription prices, since we find if anything a negative effect of CL entry on that dimension.

voting. This finding is consistent with a greater tendency of voters to rely on national partisan cues when less information about local candidates for office is available.

We also find suggestive evidence that CL entry is associated with an increase in the entry and electoral performance of ideologically extreme candidates in House elections. This result is consistent with the role of newspapers in allowing voters to discriminate between extreme and moderate candidates, particularly in primary elections where voters cannot rely on party labels (Hall and Lim, 2018).[5]

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 literature. First, it relates to prior work by Seamans and Zhu (2014) on the impact of CL on newspapers' business strategies. Our analysis expands upon their findings by documenting the broader implications of the Craigslist shock on newspapers' organization, editorial priorities, news content, and, ultimately, political outcomes.[6]

Second, our paper relates to studies on the effects of new media technologies on incumbent media. Bhuller et al. (2024) and Gavazza et al. (2019) study the roll-out of broadband Internet in Norway and in the UK respectively, and document large declines in newspapers' print circulation. Gentzkow (2006) and Angelucci et al. (2024) study the introduction of television in the US, documenting that it presented a significant negative shock to newspapers' readership and advertising. While these papers study broad technology shocks that may simultaneously affect readers' demand, advertising markets and the production of news, our setting allows us to separate the effect of a specific shock to the advertising market.

Our findings also dovetail with previous evidence of how shocks to advertising revenues affect news producers. In a historical perspective, Hamilton (2004) and Petrova (2011) argue that the growth of the print advertising market in the late 19$^{\text{th}}$ century was essential to the emergence of an independent (non-partisan) press. Our analysis differs in that we study the reverse phenomenon, i.e. the decline in newspapers' advertising business due to online competition and its negative

[5] The classic Calvert-Wittman model of electoral competition with policy-motivated candidates (Calvert, 1985; Wittman, 1983) generates the prediction that greater uncertainty over the preference of the median voter leads to greater divergence in equilibrium platforms. Matějka and Tabellini (2021) microfound this result in a model of rational inattention; when the cost to voters to acquire information rises, voters with more extreme preferences are more willing to keep paying, leading politicians to cater to them more. "That is, as the cost of attention rises, the equilibrium moves closer to the bliss point of the group with more extreme preferences (p. 1920)."

[6] Other studies have used the expansion of CL across the US, or particular design features of the platform, to investigate 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).

implications for political coverage. This complements evidence from France by Angelucci and Cagé (2019), who find that the introduction of advertising on TV in the 1960s and 1970s reduced newspapers' production of journalistic-intensive content.[7]

More broadly, our analysis relates to previous work on the effect of the Internet on electoral politics (Falck et al., 2014; Campante et al., 2018; Gavazza et al., 2019; Larcinese and Miner, 2018). While these studies assess the aggregate effect of the various changes brought about by the Internet, our analysis isolates the political impact of the Internet through the disruption of legacy media.

Finally, we contribute new causal evidence on the link between the media environment and the nationalization of politics (Moskowitz, 2021; Trussler, 2020) and success of extremist candidates (Hall and Lim, 2018). Our results demonstrate how news media support the functioning of elections as a selection mechanism. The magnitude of the electoral penalty that extreme candidates face, and voters' ability to punish extremists electorally (Hall, 2015; Hall and Thompson, 2018), hinges on the quality of the information environment to which voters have access. Our results show how the impoverishment of local newspapers and the resulting changes in organization, content, and readership can limit the electorate's ability to discriminate between candidates and thus weaken the ideologically moderating force of elections.

The remainder of the paper is organized as follows. Section 2 briefly describes our conceptual framework. Section 3 provides background information on Craigslist and its expansion and 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.

2 Conceptual Framework

Our conceptual framework starts from the assumption that the coverage of politics is costlier than the coverage of soft news, such as sports and entertainment, since it requires greater investments in reporting staff with expertise and more time for investigative activities (Hamilton, 2016).[8] Prior to CL, newspapers' profits from classified advertising cross-subsidized the production of news, in general, and political news, in particular. The entry of a superior online competitor unbundled the two products and reduced newspapers' revenues from classified ads, forcing them to cut staff and

[7] 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.

[8] Hamilton argues that this cost differential arises from the combination of two factors. First, the ease of discovering primary information: "press agents are eager to promote stars," while public affairs stories can "require determination to pursue information that government actively tries to keep hidden." Second, the certainty that reporting effort will produce a good story: in sports coverage, "a game's unfolding and outcomes are guaranteed to provide a story," while "accountability coverage is akin to drilling for oil, since tips and suspicions may not pan out." (p. 16)

limit more costly activities such as political reporting. The reduction in political coverage in turn reduced demand among readers interested in this type of content.

This framework is consistent with a model of a two-sided market, such as that proposed by Angelucci and Cagé (2019), in which advertisers pay for readers' attention, (some) readers value political coverage, and producing political coverage involves paying higher fixed costs as it requires a larger newsroom. In such a model, a reduction in advertisers' willingness to pay induces newspapers to reduce the size of the newsroom and produce less political coverage.[9] Intuitively, newspapers have weaker incentives to attract readers interested in political news through quality political coverage when advertisers are less willing to pay for readers' attention. This framework predicts that newspapers affected by CL should: (1) reduce the size of their newsroom (particularly in the area of politics); (2) reduce their coverage of politics; (3) experience a decline in readership (particularly among readers interested in political news).

Given the role of local newspapers as the primary mainstream source of information on local political candidates,[10] a reduction in the volume of newspaper coverage in this domain can have important implications for electoral politics. First, voters may be less likely to participate in elections.[11] Second, with less specific information available on local candidates, voters may rely more on partisan cues when casting their vote, leading to declines in split-ticket voting in general elections.[12] Third, a rise in the cost of accessing information about candidates in mainstream sources should advantage more ideologically extreme candidates. The reason is that politicians cater their platforms to more informed voters (Strömberg, 2001). When political coverage in mainstream sources falls, moderate voters become differentially less informed compared to more extreme voters, because they are on average less willing to pay the costs of seeking out substitutes from less accessible sources (Matějka and Tabellini, 2021). Moderates will thus have a harder time differentiating ideologically extreme candidates from moderate ones. Together with the second implication above, this effect makes it more likely that relatively extreme candidates will be willing to enter electoral contests, and achieve better electoral returns when they do.

[9] This model's prediction for subscription prices, on the other hand, is ambiguous.

[10] E.g., Mahone et al. (2019) show that newspapers produce a disproportionate share of what the FCC calls "Critical Information Needs" (CIN) coverage. This category includes "information on local, regional, and county candidates at all levels of governance (p. 26)." In contrast, local television news tends to have a more national focus — with political coverage heavily tilted towards the presidential level — and provide less substantive information about politicians' policy stances (Fowler, 2020).

[11] Evidence from a variety of settings suggests that substitution into media with less political coverage leads to declines in participation; see Gentzkow (2006); Ellingsen and Hernæs (2018); Campante et al. (2018).

[12] Other work that examines the relationship of the media environment to split-ticket voting includes Moskowitz (2021); Angelucci et al. (2024); Darr et al. (2018).

3 Background and Data

3.1 Background

Craiglist.org (CL) is the world's largest online platform for classified ads and is consistently ranked among the top 20 websites by traffic in the US. The company was created in San Francisco in 1995 and served only the Bay Area until 2000, when it began expanding to other US locations. The expansion started with large cities such as Boston, New York, and Chicago and continued to smaller markets over time, covering 115 locations in 2005, 331 in 2008, and 416 today.

The CL platform is characterized by a simple layout which has remained largely unchanged over time (Figure B1). CL websites only host classified ads and do not include any display ads or news content. The most popular ad categories are housing, jobs, and items for sale.

Posting, browsing and responding to CL ads in the vast majority of locations is completely free of charge. The only exceptions are brokered apartment rentals in New York City and job posts for employers in some major cities, with fees ranging from \$3 to \$75.[13] Fees for these few special categories constitute CL's only source of revenue. This unconventional business model reflects the views of CL's founder, Craig Newmark, who originally founded CL as a non-profit and prioritized providing a useful service to local communities over profit maximization even after the company was incorporated in 1999.[14]

The period of CL's expansion and growth in popularity coincides with a collapse in newspapers' lucrative classified ad business. According to data from the Newspaper Association of America, in the year 2000 US newspapers' classified advertising revenues amounted to \$20 billion, relative to \$49 billion in total advertising revenues and \$11 billion in circulation revenues. By 2012, classified advertising revenues had fallen to \$4.6 billion — a 77% decline. Panel (a) of Figure B2 shows that this decline was steeper and occurred earlier than that in other revenue sources. Similarly, the aggregate number of newsroom workers has been on a steep downward trend since the mid-2000s (panel b).

3.2 Data

Our analysis combines data on: i) Craigslist's expansion across the US; ii) characteristics, organization and market outcomes of daily newspapers; iii) newspapers' content; iv) survey data on media consumption; v) political outcomes including turnout, voting choices and candidates' ideology, and vi) additional covariates, including a proxy for local Internet penetration.

[13] A full list of the exceptions as of 2010 is available at: https://web.archive.org/web/20100706030043/https://www.craigslist.org/about/help/posting_fees.

[14] For a profile of Craig Newmark and his business strategy see https://www.theguardian.com/technology/2006/feb/19/news.theobserver1 and https://www.wired.com/2009/08/ff-craigslist/.

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 local websites. For a subset of 197 websites, this information is directly available on CL's "about" webpage (https://www.craigslist.org/about/expansion). For the remaining 219 websites, we assign the date of the first snapshot recorded by the Internet Archive Wayback Machine (https://archive.org), cross-checking this date against past snapshots of CL's list of websites. The maps in Figure 1 show the geographic distribution of local CL websites in 2000, 2005, and 2010, respectively.

To define the area served by a given CL website, we consider the county (or counties) included in the location reported in the website's URL. This is usually a single city, but can also be a combination of nearby cities, a region, or, in a few cases, an entire state. For example, we assume that the website chicago.craigslist.org (Chicago) serves Cook and DuPage counties, and that the website westernmass.craigslist.org (Western Massachusetts) serves Hampden, Hampshire, Worcester, Berkshire and Franklin counties.[15]

Figure 2 illustrates the timeline of CL's expansion and take-up. Panel (a) plots the evolution of the share of US counties served by a CL website, along with the evolution of a proxy for Internet quality: the average number of Internet service providers by zip code. Panel (b) plots the evolution of CL's traffic, measured by the share of visits to the domain craigslist.org recorded by Comscore, a dataset that tracks browsing behavior for a representative sample of US online users (more information below). For comparison, we also plot the equivalent measure of traffic for CL's major competitors in the three main categories of classified ads: monster.com for job ads, realtor.com for housing ads, and ebay.com for items for sale. The figure shows that the use of CL evolved differently from that of its competitors and from the general trend in Internet penetration.

Newspaper characteristics and outcomes. We collect comprehensive data on relevant newspaper characteristics and outcomes from industry yearbooks published by the company Editor & Publisher (E&P, 2010). We accessed print copies of the yearbooks for the period 1995-2010 and digitized them using OCR software.[16] The yearbooks contain detailed information for over 1,500 US daily newspapers, including: address of the headquarters (HQ), ownership group, circulation, subscription price, as well as the complete list of editorial and managerial staff members with names, broad job categories, and job titles. Figure B3 presents examples of the raw E&P data.

To define newspaper markets, we assume that newspapers' readership is concentrated in the county where the newspaper HQ is located. This approach is common in the literature (Gentzkow and Shapiro, 2010; Seamans and Zhu, 2014) and a good approximation for most newspapers:

[15] As an alternative approach, we consider a broader set of counties: any counties that account for a non-negligible share of the ads posted on the website in year two after its entry (retrieved from the Wayback Machine). We discuss this approach in Appendix B.1.1.

[16] To alleviate transcription errors, we flagged observations for which our key variables deviate substantially from past/future values for the same newspaper, and interpolated flagged or missing values.

disaggregated circulation data suggests that the HQ county accounts for 85% of the readership of the median newspaper.[17]

The extent to which a newspaper is affected by the entry of CL depends on how heavily it initially relied on revenues from classified ads. To proxy for pre-CL reliance on classified ads revenue, we consider whether, in the year 2000 (prior to CL's major expansion), the newspaper had a dedicated classified ads manager in its staff. To validate this measure and to estimate the size of the revenue shock to affected newspapers, we collect information from newspapers.com — an online archive of historical newspapers — on the share of pages per issue devoted to classified ads (available for a subset of about 250 newspapers for the period 1995-2010). We also collect information from Standard Rate and Data Service (SRDS, 2006) on classified prices (available for all newspapers for the period 1995-2006).

We correlate these measures of classified reliance with the presence of a classified ads manager prior to CL's entry. The results, presented graphically in Figure 3 and in tabular form in Table B1, indicate that, controlling for a broad set of characteristics of the newspapers and of the county in which they are headquartered, newspapers with a classified ad manager devoted about 7 percentage points more pages to classified ads and had classified ad prices about 9% higher than newspapers without a classified ad manager.

Newspapers' content. We obtain data on newspapers' content from NewsBank (NewsBank, 2010), which contains the full text and metadata of more than 100 million articles published in about 900 newspapers in our sample. 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 amount of coverage it devotes to specific groups of politicians. Second, we extract the text of the lead paragraphs for a random sample of 2 million articles and estimate a topic model on the resulting corpus. Specifically, we estimate a Correlation Explanation (CorEx) model (Gallagher et al., 2017a), which has the advantage of producing coherent topics for corpora consisting of short texts. We obtain a distribution of topic weights for each article in the corpus, which we then aggregate up to the newspaper-year level. Further details about the procedures used to construct these variables are reported in Appendix B.1.4.

Internet penetration and other county controls. To control for local Internet penetration, following Larcinese and Miner (2018), Seamans and Zhu (2014) and Lelkes et al. (2017), we use the number of Internet service providers (ISPs) registered by zip code and year. These data are available for the period 1998-2008 from the Federal Communication Commission (FCC, 2008)

[17] In Appendix B.1.2 we describe an alternative approach based on identifying the broad set of counties where a given newspaper circulates using geographically disaggregated circulation data. These data, however, are only available for a subset of the newspapers in the E&P sample.

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.

While disaggregated data on the number of Internet subscribers is unavailable for the period of interest, prior studies show that the number of ISPs is a strong predictor of the number of subscribers at the state level, as well as at the county level in later periods. As further validation, we examine the correlation of the number of ISPs with self-reported measures of access to the Internet from two large scale electoral and marketing surveys (described in greater detail below). The results, reported in Figure B4, show a strong positive relationship between the number of ISPs and the share of local respondents who report accessing the Internet regularly.

Throughout our analysis we also use data on population by county and year (from the National Center for Health Statistics), and on a series of county-level variables in the baseline census year 2000 including: 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) unemployment rate (from the Bureau of Labor Statistics), and presidential turnout (from the David Leip Election Atlas).

Surveys and browsing data. We use individual survey data on self-reported media consumption from two large-scale surveys. The second source is the Survey of the American Consumer conducted by GfK Mediamark Research & Intelligence (GfK-MRI, 2010), a nationally-representative rolling cross-sectional marketing survey conducted every year in the period 1999-2010. The second source is the National Annenberg Electoral Survey (NAES, 2008), a nationally-representative rolling cross-sectional electoral survey conducted in the lead-up to the 2000, 2004, and 2008 presidential elections.[18] In addition to questions on the frequency and type of media the respondent consumes, GfK-MRI contains information on the newspaper sections read by the respondent, e.g., general news or classified advertising. This allows us to study the characteristics of news- vs. classified-section readers.

We use additional data on online media consumption from WRDS Comscore (Comscore, 2010) – a dataset that tracks the browsing behavior of a large sample of US Internet users representative of online buyers (los Santos et al., 2012). The data cover the years 2002, 2004 and 2006-2010. We aggregate the number of visits to domains of interest as well as total visits recorded by Comscore by county and year.

[18] The 2000 and 2004 waves of the NAES include questions on general media exposure, while the 2008 wave only asks about exposure to campaign-specific information. We therefore focus the analysis of media consumption on the first two waves.

Political outcomes. We obtain county-level electoral data for House, Senate and presidential elections from the David Leip Atlas of American Elections (Leip, 2020). Following Darr et al. (2018), we use these data to construct a measure of split-ticket voting, defined as the absolute value of the difference between the Republican candidate's vote share in the presidential election and the Republican candidate's vote share in House and Senate elections in the same county and year.[19]

We also examine a set of outcomes related to the entry and electoral performance of ideologically extreme candidates. Following Hall (2015) and Autor et al. (2020), we classify candidates using information on contributors to their campaigns. Specifically, we use candidates' position in the distribution of campaign-finance-based ideological scores (CFScores) from the Database on Ideology, Money in Politics, and Elections (Bonica, 2016). Bonica's method is an unsupervised learning technique that recovers candidates' positions in a one-dimensional latent ideological scale under the assumption that donors give "spatially;" e.g. that a contribution from donor i to candidate j is more likely, the closer are i and j in the latent ideological space. Importantly, CFScores are fixed within candidate, meaning that change in the distribution of CFScores of candidates in a given district over time can come only from the entry of new candidates. 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." The main outcomes of interest are the presence of an extremist in a primary election, the presence of an extremist in a general election (which implies that at least one primary election in that district was won by an extremist), and the probability that a general election is won by an extremist.

4 Empirical strategy

4.1 Determinants of CL entry

To implement our empirical strategy, it is necessary first to understand what factors drove the timing of CL's staggered rollout. According to CL founder Craig Newmark, CL's entry into new locations was determined by market demographics and the quality of local broadband Internet, but apart from that was largely idiosyncratic.[20]

[19] 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.

[20] According to Newmark: "We put up a city based on how many people are asking us to do so. It's also based on Jim's [Jim Buckmaster, CL's CEO] perception of a city's demographics and the city's broadband penetration and intuition. We use word of mouth to get the word out, though sometimes the local press is kind...We just wait for things to happen." See "CRAIGSLIST/On the record: Craig Newmark." San Francisco Chronicle, August 15, 2004 (https://www.sfgate.com/business/ontherecord/article/CRAIGSLIST-On-the-record-Craig-Newmark-2733312.php).

Importantly for our purposes, CL is not in the news business, and is thus unlikely to have considered demand-side factors in the news market. Furthermore, various accounts call into question the extent to which CL attempted to maximize profits.[21] The lack of a profit-maximization motive arguably allowed more flexibility for the idiosyncracies of CL's founder and its early user base to determine the timing of the roll-out, rather than systematic economic factors in local news markets.

To test this conjecture, in Table 1 we examine the set of counties where at least one newspaper was headquartered in 2000, and regress the year of CL's entry on county and newspaper characteristics, expressed in levels in panel (a) and in changes between 1995 and 2000 in panel (b). The results confirm that population and the quality of the local Internet connection were important factors in CL's entry decision. The magnitudes 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, we will control for time-varying log population and number of ISPs throughout our analysis. A few other covariates, such as the rental share of housing, median age and the county's racial composition, display weaker correlations with CL's time of entry (see Table A1). To control flexibly for possible secular trends associated with these factors, in our baseline specification we interact a broad set of baseline county characteristics with year dummies.

Importantly for our analysis, we find no relationship between the timing of CL's entry and the state of local newspapers at baseline as measured by circulation, number of jobs, or the presence of a dedicated classified manager, once log population and the number of ISPs are controlled for. These results corroborate the view that CL did not prioritize areas where newspapers were underperforming or had greater dependence on classified ads.

4.2 Main Specifications

To estimate the effect of CL entry, 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 includes the universe of daily newspapers covered by E&P (and respectively, all counties with newspaper HQs), excluding newspapers with national circulation — the New York Times, USA Today, and the Wall Street Journal — that do

[21] In the ruling of a 2010 civil action 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.

not serve a specific local market.[22]

The following equations summarize our econometric strategy:

$$Outcome_{nct} = \alpha_n + \beta_t + \gamma PostCL_{ct} + \theta Controls_{ct} + \varepsilon_{nct}, \tag{1}$$

$$Outcome_{nct} = \alpha_n + \beta_t + \delta PostCL_{ct} + \lambda PostCL_{ct} \times ClassifMgr_n + \theta Controls_{ct} + \varepsilon_{nct} \tag{2}$$

$Outcome_{nct}$ is an outcome of interest for newspaper n, headquartered in county c, at time t; $PostCL_{ct}$ is a an indicator equal to one for years after the entry of CL in county c; $ClassifMgr_n$ is an indicator for the presence of a dedicated classified ad manager at baseline (i.e., in the year 2000). The vector $Controls_{ct}$ includes time-varying log population and number of ISPs, as well as county-level controls from the 2000 census interacted with year fixed effects, including share of urban population, share with college degree, rental share of housing, income per capita, unemployment rate, median age, turnout and racial composition. α_n and δ_t are newspaper and year FEs, and ε_{nct} is the error term. 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.

The identifying assumption of our estimation strategy is that, conditional on controls, the timing of CL entry is uncorrelated with pre-existing trends in the outcomes (specification 1), or with pre-existing differential trends between newspapers with or without a classified manager (specification 2). Table 1 lends support for this assumption by showing that the timing of CL entry is unrelated to levels or trends in newspaper outcomes or to the presence of a classified manager once we condition on Internet penetration and log population.

To further examine the timing of the effects, we estimate dynamic specifications following recent methodological advances in the difference-in-differences literature (Roth et al., 2023). A concern raised by this literature is that, in settings with treatment effect heterogeneity over time or across treated units, dynamic models estimated using the standard two-way fixed effects estimator can estimate non-convex combinations of treatment effects. This concern applies to our setting since network effects in Craigslist's local adoption predict heterogeneity over time, and newspapers' differential ex-ante reliance on classified ads predicts heterogeneity across units. Therefore, we estimate dynamic effects and first-difference pre-treatment placebos based on the DID_M estimator proposed by de Chaisemartin and D'Haultfoeuille (2020) which is robust to treatment effect heterogeneity. We consider two versions: one that defines the treatment as CL entry (akin to specification 1) and one that defines the treatment as CL entry into the market of a newspaper with

[22] The sample includes newspapers in counties where CL never entered, which are about 45% of the total and serve as a control group. The existence of a stable group of "never treated" units is a useful feature for the de Chaisemartin and D'Haultfoeuille (2020) method which we use to estimate dynamic effects robust to treatment-effect heterogeneity. That said, we also present results excluding never-treated newspapers and exploiting only variation in the timing of CL's entry across newspapers that were eventually treated.

a classified manager (akin to specification 2). In other words, in the latter version we consider newspapers without a classified manager as untreated.

An important difference between specifications 1 and 2 is that they identify the effect of CL entry for different groups of newspapers. In specification 1, the coefficient γ captures the average impact across all newspapers in the sample. In specification 2, the coefficient λ captures a more local effect – the differential impact on newspapers that are reliant of classified ads ex ante, and hence more vulnerable to shocks to classified ad revenue. Since we are interested in the effect of exposure to the revenue shock rather than the effect of CL per se, the latter approach comes closer to our estimand of interest. Additionally, this approach allows us to conduct robustness checks that relax the assumption on the conditional exogeneity of CL's entry by restricting the comparison to newspapers in the same location – i.e. subject to the same location-specific trends – but with different ex-ante reliance on classified revenues. Yet, this requires the additional assumption that differences in CL's effect between newspapers with and without a classified manager are indeed attributable to different classified reliance, and not to other characteristics (e.g., newspapers' size). We discuss robustness checks probing the validity of this assumption in section 5.

For outcomes measured at the county level, we estimate specifications equivalent to 1 and 2, but replace newspaper FEs with county FEs and aggregate $ClassifMgr_n$ up to the county level by taking the circulation-weighted average across newspapers headquartered in the county, using circulation as of the year 2000. Note that this aggregation only applies to counties where newspapers are not local monopolies — which is the case for 14% of the counties in our sample.

For outcomes measured at the congressional district level, we adapt the above specifications to the level of county $\times$ congressional district cells, weighting observations by the share of the cell's voting-age population relative to the respective district and clustering standard errors by district.[23] It is worth noting that since outcomes in this case are measured at a higher level of aggregation than the treatment (our sample covers about 1,200 counties in 440 electoral districts), estimates in this specification may be downward biased. Indeed, the misalignment between units of observation inevitably reduces precision as districts contain a mixture of treated and untreated newspapers (i.e., ones that experience CL entry and ones that do not, and ones that are more and less reliant on classified ads ex ante).

[23] To absorb variation due to changing congressional district boundaries, we further include district by redistricting regime fixed effects in all regressions. 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. These fixed effects thus ensure that comparisons in the regressions are within invariant district boundaries.

5 Results

5.1 First stage: Craigslist take-up and effects on the classified ad market

The premise of our conceptual framework is that local CL entry triggers take-up of the platform by local consumers and thereby causes a decline in local newspapers' classified ad revenues.

While in theory CL websites have narrow geographic scope, in practice there are no restrictions on the locations of ads or of users. It is therefore an empirical question whether CL's local entry produces meaningful variation in local take-up. To verify that it does, we use data on web browsing behavior from Comscore (described in section 3.2) and estimate the effect of CL's local entry on the (ihs-transformed) visits to the craigslist.org domain by county and year. Figure 4 shows an event-study for this effect, estimated following the method proposed by de Chaisemartin and D'Haultfoeuille (2020). The figure presents first-difference placebos for the pre-entry period, and cumulative dynamic effects (corresponding to long differences relative to t=-1) for the post-entry period. It suggests a significant increase in visits after the entry of a local website, of about 50% relative to the last pre-entry period. Local take-up increases over time, consistent with network effects in the adoption of the website: a larger number of local users (and hence, a higher volume of local ads) makes the platform more useful and attracts yet more users.

Next, we examine the effect on the newspaper classified business. In Table 2, we estimate effects of CL entry on the quantity (measured in share of pages devoted to classifieds) and prices of classified ads. We find on average a 3 percentage point decline in the share of pages to classifieds (relative to a mean of 28 percent), and no effect on classified prices. This effect is driven entirely by papers with a classified manager at baseline, where the decline is 4.2 percentage points (corresponding to 15% of the mean). We find no effect on the prices of classified ads,[24] implying a proportional impact on revenue. Together with Figure 3, which shows that these papers were more reliant on classified ads at baseline, Table 2 confirms that the entry of CL was a revenue shock with

[24] This result differs from that found by Seamans and Zhu (2014), who document a significant decline in classified rates for newspapers affected by CL. Several factors could explain the discrepancy between our results and theirs. First, our sample covers a larger number of newspapers (about 1400 per year vs. less than 600 per year in Seamans and Zhu (2014)) and a longer time period (1995 to 2006 vs. 1999 to 2006 in Seamans and Zhu (2014)). Second, our empirical strategy is different in that we use the presence of classified managers in the pre-Cragslist period as proxy for newspapers' prior reliance on classifieds, while they use, for each year, the presence of classified managers in that year, which is likely to be endogenous to the entry of Craigslist. Finally, we recognize and try to control for inconsistencies in the unit of measure of classified rates in the SRDS data (i.e., column inch, line or word), an issue that their analysis ignores.

heterogeneous effects depending on baseline classified reliance. [25]

5.2 Main newspaper outcomes

Main results. We present the main results on the impact of CL entry on local newspapers' organization, readership and political coverage in Table 3. The specifications follow equations 1 and 2 in alternating columns. The first dependent variable is the number of jobs listed in E&P's staff section by newspaper and year. We find that CL-entry is associated with a significant decline of 1.1 fewer jobs on average, or 5% relative to the mean of 21.4 jobs (column 1). Turning to our main specification which allows the effect of CL entry to depend on newspapers' baseline reliance on classified ads, we find that this decline is driven entirely by newspapers with a classified manager at baseline (column 2). Comparing newspapers with a classified manager to ones without, the magnitude of the decline reaches 3 fewer jobs, or 14% relative to the mean. Conversely, the coefficient estimated for newspapers without a classified manager is small, insignificant and if anything, positive. The pattern is similar for CL's effect on newspaper readership. We find a significant decline in circulation per capita, with a magnitude of 2% on average, and 4% comparing newspapers with a classified manager at baseline to ones without (columns 3 and 4).

We also find evidence for significant declines in the volume of newspapers' coverage of politics using two alternative measures. One measure is the topic weight on politics obtained from a CorEx model estimated on a random sample of 2 million articles. We find a 4% average decline in the topic weight on politics, and an 8% decline comparing newspapers with a classified manager at baseline to ones without (columns 5 and 6). A complementary measure is the number of articles that mention the name of a congressional representative or candidate running in the state in which the newspaper is headquartered. Looking at the ihs-transformed count of such articles, and controlling for the ihs-transformed count of total articles covered by NewsBank by newspaper and year, we find no significant effect of CL entry on average, but a significant 12% decline comparing newspapers with a classified manager at baseline to ones without (columns 7 and 8).

Figure 5 presents estimates for the dynamic effects of CL entry on these four main outcomes, following the method of de Chaisemartin and D'Haultfoeuille (2020). Placebo estimates for the pre-entry period (corresponding to first differences between consecutive pre-periods) are mostly insignificant, except for the 2nd and 3rd pre-period estimates for jobscount. This suggests some pre-trend in jobscount, though placebos for all further pre-periods are insignificant and close to

[25] The median classified-manager paper in our data in 2000 has about 28 pages per issue, and the median price reported in 2000 among classified-manager papers is \$17.42 per column inch. Classified section layouts range from 6 to 12 columns per page, and broadsheet newspaper page lengths in the US are either 21 or 22.75in. Taking the ends of these ranges yields an approximate loss of between roughly \$2650 and \$5725 per issue; with 52 Sunday issues (the traditionally heaviest day for classifieds) per year, our share-of-pages effect estimate implies a median revenue loss among classified-manager papers of at least \$137K-\$298K per year.

zero.[26] The figure also shows that the estimated post-entry effects (corresponding to long differences relative to t=-1) tend to increase in magnitude with length of exposure – a pattern that is consistent with the dynamics of local CL take-up.

Robustness. These main results are robust to a number of alternative specifications. First, we consider alternative sets of controls. In Table A2 we present results without any controls (panel a), as well as results controlling for log population and number of Internet service providers but excluding the baseline controls $\times$ year FEs (panel b). Consistent with the notion that Internet penetration may have independent detrimental effects on newspaper outcomes, omitting the control for number of ISPs results in estimates of somewhat larger magnitude for some outcomes. In panel (c) on the other hand we include an additional control for Internet penetration: the average share of respondents with self-reported Internet access in the (pooled) GfK-MRI and NAES surveys by county and year. Even though this specification is valid only in county-years where a survey was conducted, which restricts the sample size significantly, our estimates become if anything more precise with the inclusion of this control. This supports the interpretation of the estimated effects as specific to CL-entry and separate from the generic effects of Internet penetration.

Second, we address the concern that CL entry may be correlated with other location-specific shocks or trends by introducing location $\times$ year fixed effects. In Table A3 we introduce fixed effects at the level of state $\times$ year (panel a) and fixed-effects at the level of Designated Market Area (DMA) $\times$ year (panel b).[27] In both cases the main estimates remain very similar in magnitude and precision. The strongest test for a possible confounding role of location-specific factors is the inclusion of county $\times$ year FEs, which absorb any county-level trends or shocks.[28] This specification is identified only from variation in counties that contain at least one newspaper with a classified manager at baseline and at least one without, so that newspapers that are local monopolies in a county (the vast majority of US newspapers) drop out of the sample. This leaves us with 168 counties and 412 newspapers, for which we can make *within-county* comparisons before and after the entry of CL, comparing newspapers with a classified ad manager to their neighbors without. Despite the demanding specification and the smaller sample size, the resulting estimates, presented in panel (c), suggest similar effect sizes for three out of the four main outcomes of interest.

Third, we consider specifications that model more precisely the dynamics of the effect of CL on newspaper outcomes. In panel (a) of Table A4 we replace the post-CL indicator with the number of years since CL entry, obtaining effects expressed in terms of an additional year of exposure to

[26] This method allows us to estimate up to 10 pre-treatment placebos.

[27] DMA is the definition of media market adopted by the FCC in regulations governing cross-ownership of newspapers and television stations. They typically contain multiple counties, including core metropolitan counties plus surrounding rural counties. There are 210 DMAs in the US.

[28] County $\times$ year FEs also absorb the main effect of the post-CL indicator, as well as log population, number of ISPs and all baseline controls interacted with year FEs.

CL competition: 3.5% fewer jobs, 1% lower circulation per capita, 2.4% lower weight on political coverage, and 3.1% fewer articles covering congressional representatives or candidates. In panel (b) we instead split the post-CL indicator into an indicator for the 0 to 2 years post CL entry and an indicator for more than 2 years post CL entry. Consistent with the event-studies in Figure 5, the estimates comparing newspapers with a classified manager at baseline to ones without are up to twice as large in the long-run than in the short run.

Fourth, we address the concern that the presence of a classified manager may be correlated with newspaper characteristics such as size, which could confound the effects of CL entry. We address this issue in three ways. Our first approach is to consider a subsample of newspapers that is relatively homogeneous in size by excluding newspapers in the first and fourth quartile of the distribution of baseline readership. This narrows the range of baseline circulation per capita to between 0.1 and 0.2. Our main estimates are robust to this restriction (Table A5). Another approach is to explicitly allow for the impact of CL entry to vary by newspaper size by including as a control the interaction of baseline size and post-CL in addition to the interaction of classified manager and post-CL.[29] The estimates are robust to this inclusion (Table A6). Finally, we consider an alternative measure of classified reliance for which a mechanical correlation with newspaper size is unlikely: the average share of pages per issue devoted to classified ads. This measure is available from newspapers.com for a subsample of ~ 250 newspapers.[30] In Table A7 we replace the indicator for the presence of a classified manager with an indicator equal to one if the newspapers' average share of classified pages exceeds the sample median of 0.3. Even though we lose some precision in this significantly smaller sample, the estimates suggest the same direction and similar magnitudes as those in our baseline specification.

Fifth, in Table A8 we probe the robustness of our estimates to different sample and treatment definitions. While our main sample consists of all newspapers covered by E&P, including ones that exit or enter the market, in panel (a) we instead restrict the sample to a balanced panel of 1,300 newspapers that operate throughout the entire sample period of 1995 to 2010. In panel (b) we exclude the control group of newspapers that do not experience CL entry and only use variation in the timing of entry across newspapers that were eventually treated. In panel (c) we adopt a broader definition of CL and newspaper markets based on the geographic distribution of CL ads and newspapers' circulation (see Appendix sections B.1.1 and B.1.2).

Finally, in Table A9 we show that the baseline estimates are robust to alternative levels of clustering of the standard errors– at the higher geographic level of state (panel a), or at the level of

[29] To avoid issues of mean reversion, we use circulation per capita as a measure of size when the dependent variable is number of jobs, and number of jobs when the dependent variable is circulation per capita.

[30] We use the average number of pages of classifieds per issue from the newspapers.com archive as the numerator (see Appendix B.1.3 for details) and the average pages per issue reported by E&P as the denominator in constructing this share.

newspaper groups (panel b).

Spillovers. In Table A10 (a) we examine the impact on papers not themselves exposed to CL, but located in a media market (DMA) in which other papers are exposed. To do so, we augment our baseline specification with the circulation-weighted leave-one-out share of newspapers in the same market that have experienced CL entry, that have a classified manager at baseline, and the interaction of the two. We find evidence of a small positive spillover in circulation, and no evidence of significant spillovers in terms of jobs or content.

Other outcomes. We also investigate whether CL affected newspapers in other ways. First, we test whether the entry of CL in an area affected the number of local newspapers. In the first two columns of Table A11 we estimate our two baseline specifications using as dependent variable the number of active newspapers headquartered in a county and year. The entry of CL does not appear to have any tangible effect on the number of newspapers, including in counties where newspapers had more to lose from increased competition for classified ads. Another possibility is that the revenue shock due to the entry of CL weakened newspapers' financial conditions, making them an easier target for acquisition. Using information from the E&P yearbooks on the group associated with each newspaper, in columns 3 and 4 of Table A11 we document that the entry of CL did not make local newspapers more likely to experience a change in ownership.[31] Finally, in Table A12, we explore whether, in response to the entry of CL, newspapers reduce the total number of pages, to cut costs, or increased subscription prices, to boost revenues. We find no decline in pages and a small negative effect on subscription prices with a magnitude of about 2%.[32]

5.3 Mechanism

In this section we present additional results on the impact of CL entry on newspaper's organization, readership and coverage and discuss our interpretation of the mechanisms that may be at play.

Jobs. An important question that the detailed data from E&P allow us to examine is what categories of workers were most affected by staff cuts. We can use the staff categories reported by

[31] A limitation of this analysis is that we use newspaper groups as listed by E&P and do not trace the ultimate cross-ownership across these groups.

[32] The finding on subscription prices differs from the results in Seamans and Zhu (2014) who document a significant *increase* for newspapers more reliant on classified ads. This discrepancy could be due to the larger number of newspapers and years that our data cover (all daily newspapers for the entire period of 1995 to 2010), or to the fact that we use the pre-CL rather than contemporaneous presence of a classified manager as proxy for newspapers' reliance on classifieds. Another difference may be in the treatment of the subscription period the listed price refers to. We use E&P's yearly subscription price, which is available for most newspapers, and convert weekly or daily prices to the year level if the former is not available.

E&P to 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 or entertainment).

In panel (a) of Table 4 we estimate our main specification 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 technology. The results indicate that CL entry leads to cuts in all types of positions, with magnitudes ranging from 23% for advertising management (which includes classified ad managers) to 7% for other staff. Importantly, staff cuts are not limited to management positions but also affect news executives and editors, whose number declines by about 9%. Since we find that CL entry does not reduce the total number of pages published per issue (columns 1 and 2 of Table A12), this result is suggestive of a net increase in editorial workload.

In panel (b) we focus on the category of news executives and editors and use their individual job titles to gauge the news topics they are likely to cover — an approach similar to Fan (2013) and George and Waldfogel (2006). We code keywords related to three common topics: politics, sports, and entertainment.[33] The results indicate a significant decline in the number of dedicated political editors after CL's entry, while we find no significant impact for sports or entertainment. In line with the conceptual framework described in section 2, this result suggests that, when facing financial difficulties, newspapers affected by CL opted to cut staff especially in areas like politics, for which producing quality content is more costly (Hamilton, 2016).[34]

Content. In Table 3 we have shown that CL entry is associated with a decline in the coverage of politics, measured as either the topic weight on politics or the number of articles covering local representatives and candidates for office. In panel (a) of Table 5 we present the complete set of results from the CorEx model used to estimate the first measure, which produces a probability distribution over five topics. The topics can be labeled as "politics," "sports," "entertainment," "obituaries," and "crime" (Table B2 presents the most representative words by topic). Consistent with the analysis of editors' job titles, we find no significant effect of CL entry on the weights associated with sports or entertainment coverage. We also find no effect on crime coverage. On

[33] We classify job titles containing the keyword "sport[s]" as sports related, ones containing the keywords "entertainment," "lifestyle," "women," "travel," "film," and "music" as entertainment related, and ones containing the keywords "politics / political," "government," "Washington," "city," and "local" as politics related.

[34] An alternative interpretation is that readers' elasticity with respect to the volume of political news is lower than the elasticity with respect to the volume of soft news, and that the risk of an even larger decline in readership motivated newspapers to cut politics disproportionately. Though later in this section we show that readers interested in politics did react to a decline in political coverage, we cannot test how the elasticity with respect to this type of content compares to the elasticity with respect to soft news.

the other hand, we find a sizable (though only marginally significant) positive effect on the weight of obituaries, arguably the only type of classified advertising that local newspapers continue to monopolize.[35] Appendix Table A15 probes the robustness and heterogeneity of these results to a version of the CorEx model that is anchored to separate political topics into 4 subgroups: national, congressional, local and foreign (see Table B3 for representative words). The results suggest significant declines in the topic weights associated with all of the political categories, and again, no significant change in other topics besides a marginally significant increase in obituaries.

Appendix Table A17 confirms the finding that CL entry appears to have reduced coverage of all levels of politics. In this table we focus on the count of articles covering either national or local- and state-level politicians. We measure articles that include the names of the President, the names of leadership of both parties in each chamber of Congress, or titles commonly held by state and local elected officials. We find substantial negative effects on coverage of all of these groups.[36]

In panel (b) of Table 5 we break out the effect of CL entry on congressional coverage, exploring its heterogeneity over the electoral cycle. We consider the number of articles mentioning the names of local representatives and candidates for congressional office published during the general election period (between the primary election date and the end of the election year) in column (1), and the number of articles published in all other times in column (2). Interestingly, the decline in congressional coverage is entirely concentrated outside of general election periods.[37] Appendix Table A18 shows a very similar result using an alternative measure of congressional coverage: the decline in the number of articles mentioning generic Congress-related keywords ("Congress", "Senate" or "U.S. House") is roughly twice as large among such articles which were published in a primary election period and which also contain the keywords "primary" or "nomination."

This temporal pattern implies we should expect to find larger effects on candidate selection in primaries than on general election turnout. The results on political outcomes presented in Section 5.4 will align with this expectation.

Readership. We verify and extend the analysis of newspaper readership using data from the NAES and GfK-MRI surveys, which include questions related to media consumption.[38] In column (1) of Table 6, panels (a) and (b), we report the results for self-reported readership of non-national

[35] Obituaries are generally paid for and written by families of the deceased, similarly to classified ads.

[36] We also find declines in the number of articles using one of the keywords indicative of investigative journalism from Hamilton (2016): see Table A19. The effect on indicators of investigative journalism is consistent with the idea that politics coverage declined due to reductions in investment in costly investigative activities.

[37] Appendix Table A16 presents additional heterogeneity tests distinguishing between coverage of incumbents vs. challengers, and House vs. Senate representatives and candidates. We find no significant differences along these dimensions.

[38] While the two surveys contain 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 ones who most frequently read other newspapers.

newspapers from individual-level regressions controlling for respondents' age, race and education. We find significant reductions of newspaper readership in both surveys, with a magnitude of 3% to 6% relative to the respective sample mean – comparable to the estimates for circulation per capita.

The decline in readership 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. Yet, as mentioned above, we find no evidence of an increase in subscription prices for newspapers affected by Craigslist.

A second explanation is 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 Table 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 B5, 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 reading this section. It is, therefore, possible that a reduction in the value of print classifieds may drive the decline in circulation.

To understand what types of readers drove 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.[39] Based on the model estimated in the 1999-2001 data, we then project two propensity scores for respondents in the following years of GfK-MRI, as well as respondents in the NAES survey. This procedure allows us to assign to each respondent 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

[39] Regularized regression is useful here because GfK-MRI records more than 600 demographic features of respondents. The characteristics most strongly associated with general news reading are being white, having a post-graduate degree, having an income in the 75K-150K range, being retired, being married 25 years or more, and being aged 45-49. The characteristics most associated with reading classifieds 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.

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 in the GfK-MRI dataset and -0.2 in the NAES dataset, indicating that the groups that tend to read each section are relatively distinct. In particular, younger, less educated, less wealthy, and unemployed respondents report high interest in classifieds, whereas older, wealthier, more educated, and retired respondents report high interest in general news.

With these propensity scores at hand, we 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 columns (2) and (3) of Table 6, indicate that the decline in readership after the entry of CL is entirely driven by individuals with high news propensity and low classified propensity. Hence, though newspapers which at baseline offered the most classifieds were most affected by the CL shock, we find that readers of those papers interested in news rather than classifieds are the ones decreasing their readership.[40]

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

Substitution to other news sources. In Appendix A.2 we examine the possibility that readers may have substituted into other forms of news media. Overall, we find little evidence of such a substitution. Tables A13 and A14 show minimal effects of CL entry on visits to major news websites in the Comscore data or on self-reported consumption of other media. For example, looking at the number of visits to the top 100 news websites classified by Comscore, we can rule out an increase in response to CL entry larger than 2.5%.

5.4 Political outcomes

In the previous sections we documented that newspapers affected by the entry of CL reported less about politics, in general, and candidates for congressional office, in particular. We also found that individuals in areas affected by CL reduced their readership of local newspapers. In this section, we examine the electoral implications of these changes. 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

[40] It is also important to note that the profile of readers who reduce their readership in response to the CL shock stands in contrast with the profile of readers who may be more likely to migrate online, i.e., the young and the less-educated (Gavazza et al., 2019).

downstream political outcomes.

We focus on three groups of congressional election outcomes, motivated by the prior literature: i) voter turnout (Gentzkow et al., 2011), ii) voters' propensity to rely on national partisan cues when voting for local candidates, measured by the incidence of split-ticket voting (Darr et al., 2018; Moskowitz, 2021; Trussler, 2020), and iii) the entry and electoral performance of ideologically extreme candidates (Hall, 2015; Hall and Lim, 2018; Autor et al., 2020).

Turnout. Since electoral data for House and Senate elections are available at the county level, we estimate the county-level specification discussed in section 4. 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. We stack House and Senate elections controlling for office $\times$ year fixed effects and, given the considerable differences in the number of voters across counties, weight observations by the county's total voting-age population.

Columns (1) and (2) of Table 7 (a) present the results. We find small and insignificant estimates for the impact of CL on turnout in congressional elections. We can rule out a negative effect of magnitude larger than 1.3 percentage points.[41] One explanation of this null result is that, as we show in Table 6, 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. This contrasts with existing work (Gentzkow, 2006; Ellingsen and Hernæs, 2018; Gavazza et al., 2019) which finds effects on turnout of substitution from newspapers into television or Internet; there the marginal readers tend to be younger, less educated, and lower income.

Split-ticket voting. Though changes in the information environment induced by CL may not impact *whether* people vote, they may affect *how* they vote. In particular, less exposure to coverage of candidates for office could reduce voters' ability to evaluate their specific platforms and valence attributes, and hence increase reliance on party labels.

Following the literature, we measure voters' tendency to deviate from party labels, or split-ticket voting, as the absolute difference in the Republican vote share in presidential vs. concurrent congressional elections. We measure vote shares at the county level and we follow the exact same specification as for turnout, again stacking House (vs. presidential) and Senate (vs. presidential) elections.

The results, presented in columns (3) and (4) of Table 7 (a) indicate that, following the entry

[41] We also find null effects for turnout in presidential elections.

of CL, voters become significantly less likely to split their vote between candidates of different parties. The effect is again driven by areas where newspapers were most vulnerable to CL's competition. The magnitude of the decline in split-ticket voting is about 13% of the sample mean.[42]

In Appendix Figure A1 we present the event-study corresponding to column 4 of Table 7. Since split-ticket voting is only defined for elections held in presidential years, we consider three election cycles before the entry of CL and two after. We define as "treated" counties that experience CL entry and in which more than 50% of newspapers (circulation-weighted) had a classified manager in the year 2000. We find no evidence of pre-existing trends, and a significant decline in split-ticket voting after the entry of CL in counties where newspapers were more reliant on classified ads.

Support for extreme candidates. Finally, we study whether the entry of CL favored the emergence and success of ideologically extreme candidates. The hypothesis 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.[43] As a concrete example, Skocpol and Williamson's (2016) study of the Tea Party movement argues that coverage of the movement by national media outlets focused on national polling data, promoting the "wrong-headed notion that the Tea Party appealed to centrist independents (pp. 147)." The thinness of on-the-ground media coverage of Congressional primaries — in which Tea Party-affiliated candidates consistently challenged Republican incumbents from the right — allowed the movement to maintain an electorally beneficial ambiguity about its ideological aims.

To test this hypothesis, we use a measure of candidate ideology (CFScore) available for all candidates in House elections from Bonica (2016) (see section 3.2 for details of the measure).[44] We examine three outcomes: i) the probability that an ideologically extreme candidate runs in a primary election; ii) the probability that an ideologically extreme candidate *wins* a primary election and thus runs in the general election, and iii) the probability that ideologically extreme candidates win the general election.

Since these outcomes are defined at the electoral district level, we use the version of our baseline specifications defined at the level of county $\times$ district cells (see section 4). All regressions

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

[43] This is in line with correlational findings by Hall and Lim (2018) who document that the advantage of extreme candidates in Congressional primaries is concentrated in areas with low news coverage, and consistent with the theoretical model of Matějka and Tabellini (2021), which predicts that increases in the cost of acquiring political information will induce candidates to cater more to relatively extreme segments of the population, because these extreme segments are willing to pay higher costs of acquiring information than are moderates.

[44] We focus on the House because we can define outcomes at the finer level of electoral district, rather than at the coarse level of state as would be the case for Senate elections.

include district $\times$ redistricting regime fixed effects to absorb the effect of changes in district boundaries, so that we exploit variation over time within a constant district boundary. It is important to stress that the geographic level of electoral districts, at which the outcomes in this analysis are observed, is much coarser than the level of the treatment (i.e., county). To the extent that the assignment of the local treatment to an aggregated outcome introduces measurement error, it should bias our coefficient toward zero.

The results are presented in Table 7 (b). Columns 1 and 2 indicate that following the entry of CL and in districts where newspapers were more reliant on classified ads, the probability that a candidate with an extreme CFScore runs in a primary election increases significantly by about 9.5 percentage points, or 12% of the mean. We also find a significant 8.8% increase in the likelihood that such candidates win the primary and run in the general election (columns 3 and 4), as well a significant 13% increase in the likelihood that an ideologically extreme candidate wins the general election (columns 5 and 6).

In Appendix figure A2 we present event-studies corresponding to each of the three outcomes. Since in this case we consider all House elections (held in both presidential and mid-term years), we can estimate dynamic effects with a larger number of leads and lags. The results show no significant pre-trends in the four election cycles prior to CL entry and a clear increase in the probability that extremists participate in and win in primaries in the three election cycles after CL entry, in districts with classified-reliant newspapers.[45] The event-study for the probability of extremists winning general elections is on the other hand noisier and in this case does not match the two-way fixed-effects estimate in Table 7; we therefore interpret this last result with caution.

Taken together, these findings are consistent with the notion that the transformations in the media landscape triggered by the entry of CL had downstream effects on electoral politics. The coarser information on candidates for office may have reduced voters' ability to screen them and thereby increased their equilibrium ideological divergence.

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 under-provision 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

[45] The relatively rapid effect of Craigslist on candidate entry is in line with the short timelines and minimal formal barriers to entry of candidate-centric US House campaigns; according to Herrnson (2008), strategic candidates focus on local circumstances and "respond to opportunities that arise in specific election years (pp. 42)" in making entry decisions.

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 19$^{\text{th}}$-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 political news coverage.

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 reduced political coverage and in particular 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.

In this paper we focus on the implications of the transformations triggered by Craigslist's entry for congressional politics, relying on available measures of election outcomes and candidates' ideology. Yet, it is likely that similar mechanisms apply, potentially to an even greater extent, to political issues at the more local level for which reporting outside of local newspapers is even more scarce. This question presents one possible avenue for future research.

Data Availability Statement: The data and code underlying this research is available on Zenodo with DOI 10.5281/zenodo.10120113.

References

ANGELUCCI, C., J. CAGE, , AND M. SINKINSON (2024): "Media Competition and News Diets," *American Economic Journal: Microeconomics*, forthcoming.

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

AUTOR, D., D. DORN, G. HANSON, AND K. MAJLESI (2020): "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, 110(10), 3139–83.

BESLEY, T. AND R. BURGESS (2002): "The Political Economy of Government Responsiveness: Evidence from India," *Quarterly Journal of Economics*, 117(4), 1415–1451.

BHULLER, M., T. HAVNES, J. MCCAULEY, AND M. MOGSTAD (2024): "How the Internet Changed the Market for Print Media," *American Economic Journal: Applied Economics*, 16(2).

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

CALVERT, R. L. (1985): "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 29(1), 69–95.

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(4), 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(1), 127–140.

CHAN, J. AND A. GHOSE (2014): "Internet's Dirty Secret: Assessing the Impact of Online Intermediaries on HIV Transmission," *MIS Quarterly*, 38(4), 955–976.

COMSCORE (2010): "WRDS ComScore," Wharton Research Data Services. https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/comscore/, accessed April, 2023.

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(6), 1007–1028.

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2020): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110(9), 2964–2996.

ELLINGSEN, S. AND Ø. HERNÆS (2018): "The impact of commercial television on turnout and public policy: Evidence from Norwegian local politics," *Journal of Public Economics*, 159, 1–15.

E&P (2010): *Editor & Publisher International Year Book, 1995-2010*.

FALCK, O., R. GOLD, AND S. HEBLICH (2014): "E-lections: Voting Behavior and the Internet," *American Economic Review*, 104(7), 2238–2265.

Fan, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103(5), 1598–1628.

FCC (2008): "Federal Communications Commission Form 477 Additional Data, 1999-2008," https://www.fcc.gov/general/fcc-form-477-additional-data, accessed April, 2023.

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*.

Fowler, E. F. (2020): "Strategy over Substance and National in Focus?: Local Television Coverage of Politics and Policy in the United States," in *The Routledge Companion to Local Media and Journalism*, Routledge, num Pages: 8.

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(5), 2092–2135.

Gentzkow, M. (2006): "Television and Voter Turnout," *Quarterly Journal of Economics*, 121(3), 931–972.

Gentzkow, M. and J. M. Shapiro (2010): "What Drives Media Slant? Evidence from U.S. Daily Newsapers," *Econometrica*, 78(1), 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), 2980–3018.

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

GfK-MRI (2010): "MRI-Simmons: Survey of the American Consumer 1999-2010," *GfK - Mediamark Research & Intelligence*. https://www.mrisimmons.com/our-data/national-studies/survey-american-consumer/.

Hall, A. B. (2015): "What Happens When Extremists Win Primaries?" *American Political Science Review*, 109(1), 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.

Hall, A. B. and D. M. Thompson (2018): "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in US Elections," *American Political Science Review*, 112(3), 509–524.

Hamilton, J. (2004): *All the News That's Fit to Sell: How the Market Transforms Information into News*, Princeton University Press.

HAMILTON, J. T. (2016): *Democracy's Detectives: The Economics of Investigative Journalism*, Cambridge, Mass.: Harvard University Press.

HERRNSON, P. S. (2008): *Congressional Elections: Campaigning at Home and in Washington*, CQ Press, 5th ed.

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*.

LEIP, D. (2020): "Atlas of American Elections, 1995-2010," https://uselectionatlas.org/, accessed April, 2023.

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(1), 5–20.

LOS SANTOS, B. D., A. HORTAÇSU, AND M. R. WILDENBEEST (2012): "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102(6), 2955–80.

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.

MATĚJKA, F. AND G. TABELLINI (2021): "Electoral Competition with Rationally Inattentive Voters," *Journal of the European Economic Association*, 19(3), 1899–1935.

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(1), 114–129.

NAES (2008): "Annenberg Public Policy Center: National Annenberg Election Survey 2000, 2004, 2008," *University of Pennsylvania*. https://www.annenbergpublicpolicycenter.org/2008-naes-telephone-and-online-data-sets/, accessed April, 2023.

NEWSBANK (2010): "Access World News Database," https://www.newsbank.com/libraries/military/solutions/access-world-news.

PETERSON, E. (2021): "Paper Cuts: How Reporting Resources Affect Political News Coverage," *American Journal of Political Science*, 65(2), 443–459.

PETROVA, M. (2011): "Newspapers and Parties: How Advertising Revenues Created an Independent Press," *American Political Science Review*, 790–808.

ROTH, J., P. H. C. SANT'ANNA, A. BILINSKI, AND J. POE (2023): "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 235(2), 2218–2244.

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.

SKOCPOL, T. AND V. WILLIAMSON (2016): *The Tea Party and the Remaking of Republican Conservatism*, Oxford University Press.

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

SRDS (2006): *Standard Rate and Data Service: Newspaper Advertising Source, 1995-2006*.

STARKMAN, D. (2014): *The Watchdog that Didn't Bark: The Financial Crisis and the Disappearance of Investigative Journalism*, Columbia University Press.

STRÖMBERG, D. (2001): "Mass Media and Public Policy," *European Economic Review*, 45(4), 652–663.

TAUSANOVITCH, C. AND C. WARSHAW (2018): "Does the Ideological Proximity Between Candidates and Voters Affect Voting in U.S. House Elections?" *Political Behavior*, 40(1), 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.

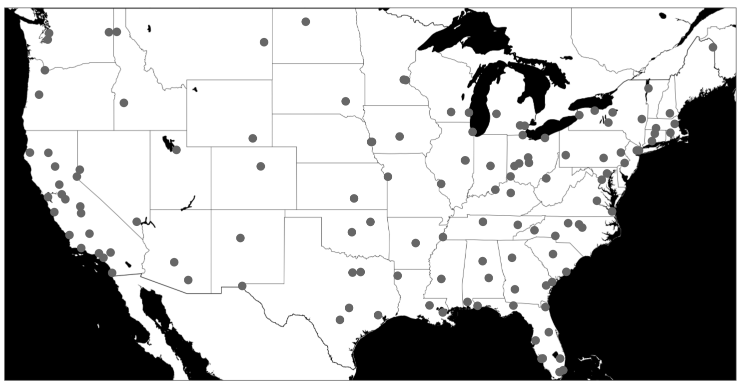
WITTMAN, D. (1983): "Candidate Motivation: A Synthesis of Alternative Theories," *The American political science review*, 142–157.

7 Figures

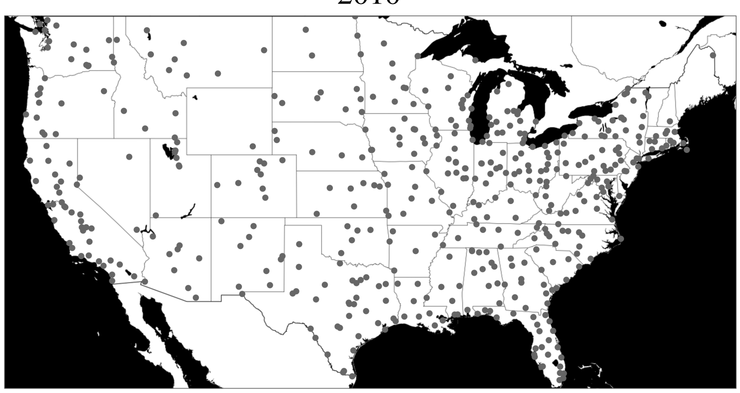
FIGURE 1: CRAIGSLIST LOCATIONS OVER TIME

2000

2005



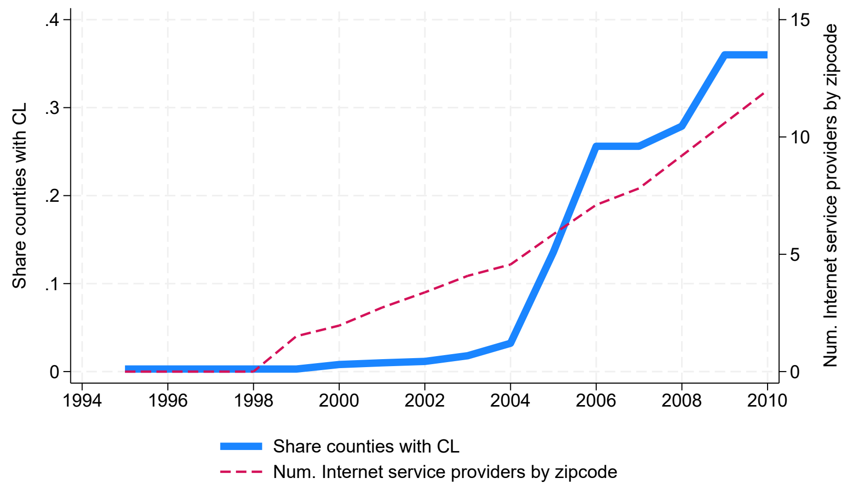
2010



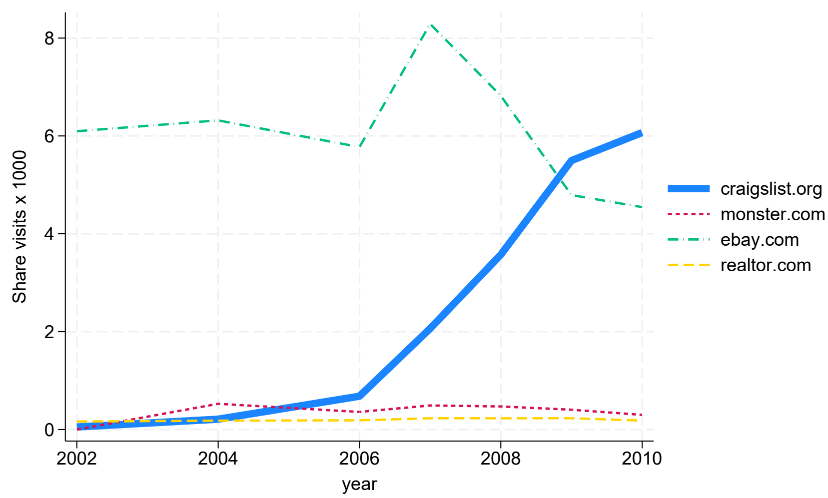
Notes: Geographic distribution of local craigslist.org websites at 3 points in time.

FIGURE 2: CRAIGSLIST'S ROLL-OUT AND TAKE-UP OVER TIME

Panel (a): CL roll-out and Internet penetration



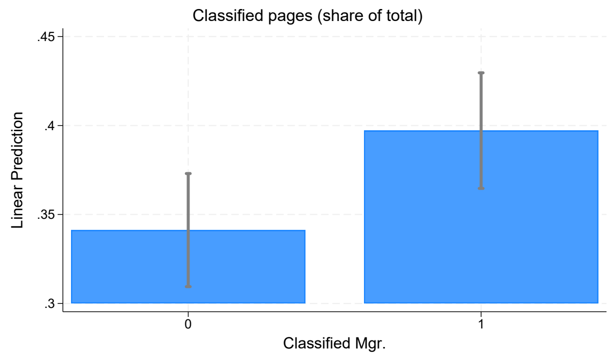
Panel (b): CL take-up relative to other platforms



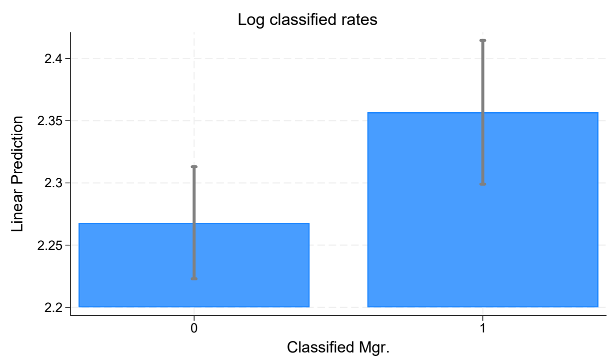
Notes: Panel (a): Share of US counties with access to a local CL website (left axis) and average number of Internet service providers by zipcode (right axis). Panel (b): Number of visits to craigslist.org, monster.com, ebay.com, and realtor.com as a share of total visits by Comscore panelists.

FIGURE 3: CLASSIFIED MANAGER AS A PROXY FOR RELIANCE ON CLASSIFIED ADS

Panel (a): Correlation with classified pages
(conditional on day-of-week FEs, newspaper and county characteristics)

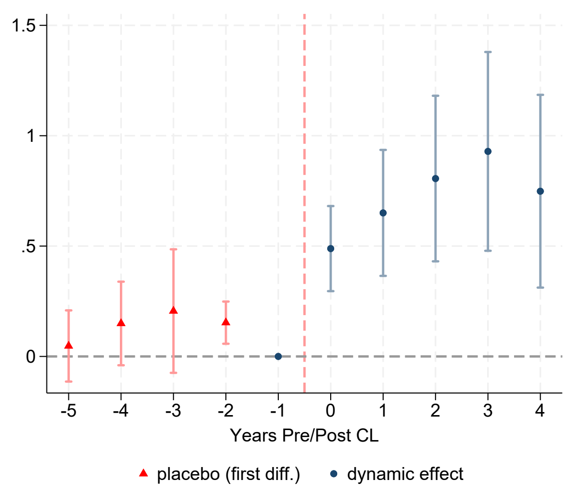


Panel (b): Correlation with classified rates
(conditional on unit FEs, newspaper and county characteristics)



Notes: Linear predictions for the share of classified pages (Panel a) and for log classified rates (Panel b) in newspapers with and without a classified manager at baseline, conditional on covariates. The covariates include circulation, jobs-count, log population, number of Internet service providers, share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout, and share White / Black / Hispanic. Classified pages are pre-2000 newspaper $\times$ day-of-week specific averages, scaled by total pages per issue (recorded by E&P). Classified rates are pre-2000 newspaper $\times$ unit specific averages. Panel (a) additionally controls for day-of-week FEs and is based on regressions weighted by the number of issues sampled. Panel (b) additionally controls for unit FEs.

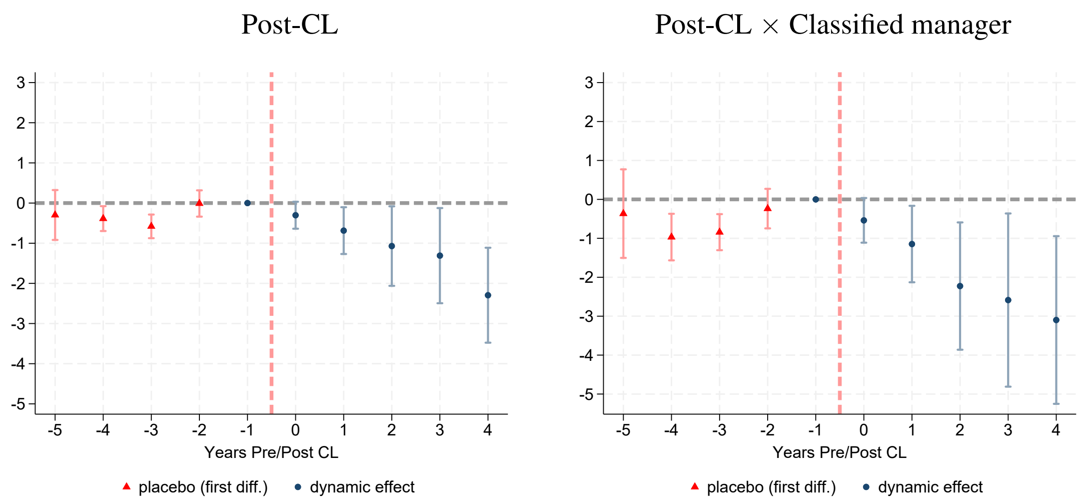
FIGURE 4: VISITS TO CRAIGSLIST.ORG – EVENT STUDY



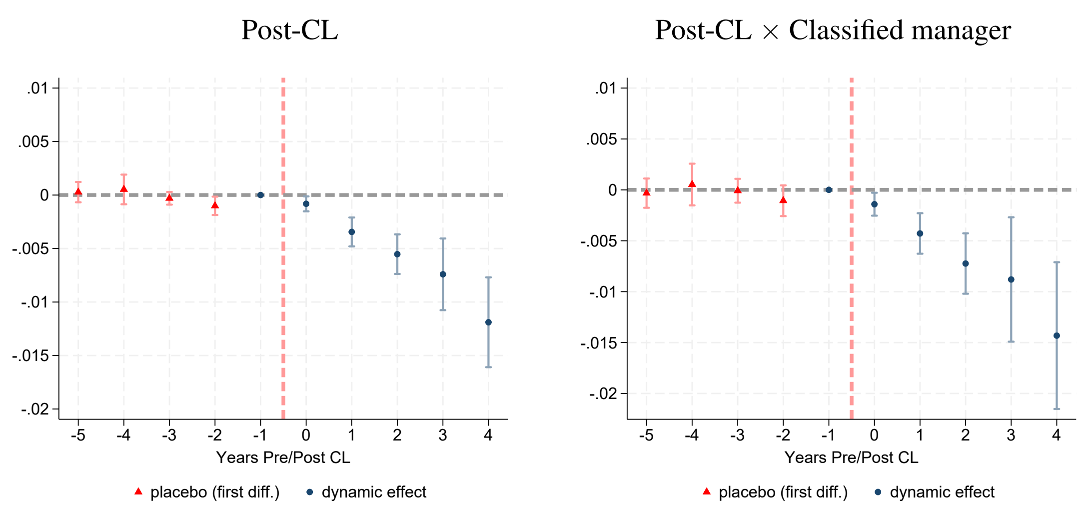
Notes: Pre-treatment placebos and dynamic effects of the entry of a local CL website on the (ihs-transformed) number of visits to craigslist.org. The graph presents coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Pre-treatment coefficients correspond to first-differences between consecutive periods. Post-treatment coefficients correspond to cumulative effects for the respective length of exposure to CL. Controls include total Comscore visits (ihs-transformed), log population and number of Internet service providers. Standard errors clustered by CL-area.

FIGURE 5: MAIN NEWSPAPER OUTCOMES – EVENT STUDIES

PANEL (A): JOBS COUNT



PANEL (B): CIRCULATION PER CAPITA

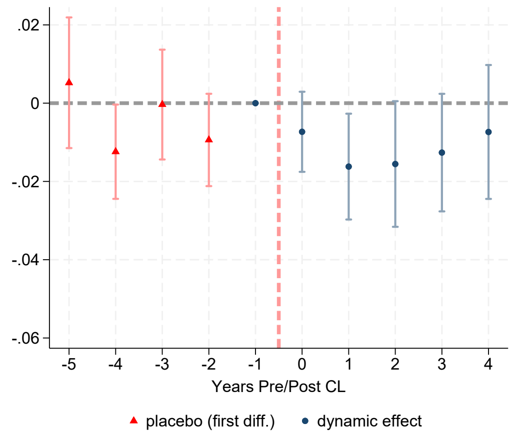


(Continued)

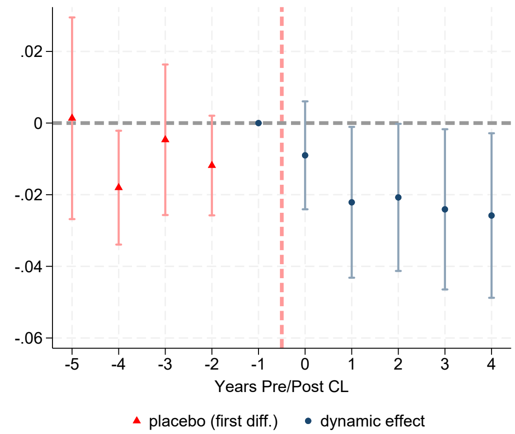
FIGURE 5: MAIN NEWSPAPER OUTCOMES – EVENT STUDIES, CONTINUED

PANEL (C): POLITICS COVERAGE, TOPIC WEIGHT

Post-CL

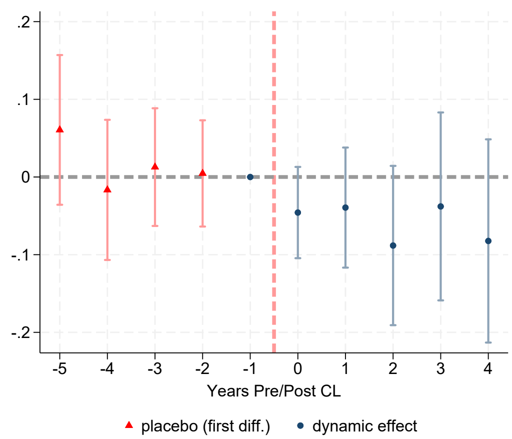


Post-CL × Classified manager

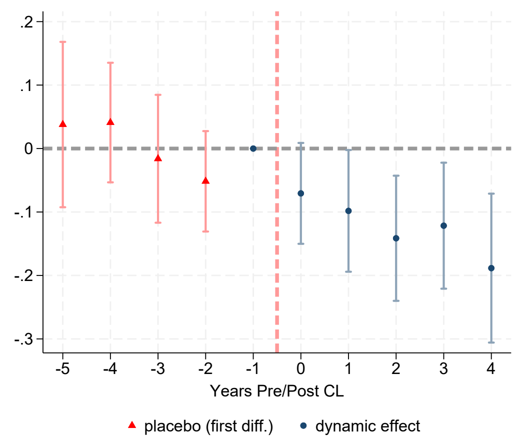


PANEL (D): CONGRESS COVERAGE, NAMES COUNT (IHS)

Post-CL



Post-CL × Classified manager



Notes: Pre-treatment placebos and dynamic effects of the entry of a local CL website on newspaper outcomes. The graphs present coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Pre-treatment coefficients correspond to first-differences between consecutive periods. Post-treatment coefficients correspond to cumulative effects for the respective length of exposure to CL. The dependent variables are number of jobs in panel (a), circulation per capita in panel (b), estimated weight on political topics in panel (c), and the (ihs-transformed) number of articles containing the name of a congressional representative or candidate in panel (d). The left-hand panels define the treatment as any CL entry. The right-hand panels define the treatment as CL entry into the market of a newspaper with a classified manager at baseline. Controls include log population and number of Internet service providers. Panel (d) additionally controls for the (ihs-transformed) total number of articles in relevant sections recorded by Newsbank. Standard errors clustered by CL-area.

8 Tables

TABLE 1: CORRELATES OF YEAR OF CL ENTRY

Panel (a): Correlates in levels

	Dependent variable: Year of CL entry					
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper jobs	-0.004	-0.003				
	(0.003)	(0.003)				
Newspaper circulation per capita			-0.224	0.205		
			(0.512)	(0.687)		
Newspaper classified manager					-0.041	-0.027
					(0.122)	(0.121)
Log population	-0.443***	-0.466**	-0.534***	-0.510**	-0.508***	-0.523***
	(0.150)	(0.184)	(0.176)	(0.247)	(0.150)	(0.190)
Internet service providers	-0.647**	-0.389**	-0.659**	-0.403**	-0.670**	-0.401**
	(0.289)	(0.174)	(0.299)	(0.173)	(0.278)	(0.168)
Other county characteristics	No	Yes	No	Yes	No	Yes
Observations	616	616	616	616	616	616
R^2	0.38	0.45	0.38	0.45	0.38	0.45

Panel (b): Correlates in changes

	Dependent variable: Year of CL entry					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Newspaper jobs	-0.003	-0.000				
	(0.005)	(0.005)				
Δ Newspaper circulation per capita			0.395	0.488		
			(0.844)	(0.727)		
Δ Newspaper classified manager					-0.017	-0.032
					(0.114)	(0.122)
Δ Log population	-4.176***	-2.759	-4.163**	-2.725	-4.184**	-2.741
	(1.180)	(2.500)	(1.852)	(2.519)	(1.847)	(2.514)
Δ Internet service providers	-0.953***	-0.587***	-0.953***	-0.587***	-0.953***	-0.587***
	(0.089)	(0.119)	(0.210)	(0.120)	(0.209)	(0.120)
Other county characteristics	No	Yes	No	Yes	No	Yes
Observations	615	615	615	615	615	615
R^2	0.35	0.42	0.35	0.42	0.35	0.42

Notes: Regressions of year of CL entry on county characteristics in the year 2000, or changes in county characteristics between 1995 and 2000. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 2: EFFECT OF CL ENTRY ON CLASSIFIED AD QUANTITIES AND PRICES

	Share classified pages		Log classified rates	
	(1)	(2)	(3)	(4)
Post-CL	-0.027**	-0.008	0.009	0.014
	(0.013)	(0.014)	(0.011)	(0.015)
Post-CL × Classified Mgr.		-0.042**		-0.010
		(0.021)		(0.019)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Newspaper × Day-of-Week FEs	Yes	Yes	No	No
Newspaper × Unit FEs	No	No	Yes	Yes
Observations	6,368	6,352	14,980	14,869
Number of newspapers	255	254	1,407	1,395
R^2	0.76	0.76	0.98	0.98
Mean dependent variable	0.28	0.28	2.47	2.47

Notes: Regressions of share of classified pages and log classified rates 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. Share classified pages is defined for the Newspapers.com subsample; classified rates are collected from SRDS but available for all E&P papers. The outcome is the average value per newspaper-weekday-year (for page share) or newspaper-unit of measure-year (for classified rate). Baseline controls include the share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/ Black/ Hispanic, all measured in the year 2000. Columns 1 and 2 additionally control for pages per issue recorded by Newspapers.com. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Main newspaper outcomes

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.114***	0.322	-0.004***	0.000	-0.012**	-0.000	-0.028	0.034
	(0.396)	(0.472)	(0.001)	(0.002)	(0.005)	(0.006)	(0.036)	(0.048)
Post-CL × Classified Mgr.		-3.027***		-0.008***		-0.024***		-0.119**
		(0.593)		(0.003)		(0.008)		(0.055)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Notes: Regressions of selected newspaper-level outcomes 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. The dependent variables are number of jobs (columns 1 and 2), circulation per capita (columns 3 and 4), weight on political coverage estimated from a Corex topic model (columns 5 and 6), and the (ihs-transformed) number of articles containing the name of a congressional representative or candidate from the newspaper's state (columns 7 and 8). Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/ Black/ Hispanic, all measured in the year 2000. Columns 7 and 8 additionally control for the (ihs-transformed) total number of total articles in relevant sections in the Newsbank database; see notes to Table 5 for details. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4: NEWSPAPERS' JOBS BY TYPE

Panel (a): Job categories

Num. jobs in category:	(1) Corporate/ General Mgmt.	(2) Advertising Mgmt.	(3) News Exec./ Editorial Mgmt.	(4) Other
Post-CL	0.025	0.085*	0.021	-0.026
	(0.079)	(0.049)	(0.348)	(0.123)
Post-CL × Classified Mgr.	-0.681***	-0.463***	-0.884**	-0.518***
	(0.133)	(0.080)	(0.444)	(0.191)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	20,327	20,327	20,327	20,207
Number of newspapers	1,436	1,436	1,436	1,436
R^2	0.80	0.80	0.89	0.82
Mean dependent variable	3.47	1.96	10.52	5.46

Panel (b): Editor job titles

Num. editors by job title:	(1) Politics	(2) Sports	(3) Entmnt.
Post-CL	0.036	0.037	-0.055
	(0.048)	(0.047)	(0.054)
Post-CL × Classified Mgr.	-0.158**	0.080	-0.081
	(0.063)	(0.085)	(0.093)
Baseline controls × Year FEs	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes
Observations	19,071	19,071	19,071
Number of newspapers	1,413	1,413	1,413
R^2	0.74	0.59	0.76
Mean dependent variable	0.63	1.01	1.01

Notes: Dependent variables: number of jobs by E&P category (panel a) and number of editors by job-title derived topic (panel b). Regressions 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. "Politics" is the number of editorial job titles that contain the keywords "politics/ political," "government," "Washington," "city" or "local." "Sports" is the number of editorial job titles that contain the keyword "sport/ sports." "Entertainment" is the number of editorial job titles that contain the keywords "entertainment," "lifestyle," "film," "music," "women" or "travel". Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 5: NEWSPAPERS' COVERAGE

Panel (a): Topic model weights

Topic probability weight	(1) Politics	(2) Sports	(3) Entertainment	(4) Obituaries	(5) Crime
Post-CL	-0.000	0.007	-0.007	-0.003	-0.001
	(0.006)	(0.007)	(0.007)	(0.009)	(0.005)
Post-CL × Classified Mgr.	-0.024***	0.012	-0.010	0.022*	0.001
	(0.008)	(0.008)	(0.008)	(0.012)	(0.008)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	7,067	7,067	7,067	7,067	7,067
Number of newspapers	855	855	855	855	855
R^2	0.52	0.44	0.47	0.56	0.45
Mean dependent variable	0.29	0.21	0.22	0.15	0.14

Panel (b): Coverage of congressional representatives and candidates: By time in the election cycle

	Split by timing	
Congress coverage names count (ihs)	(1) After primary election	(2) Prior to primary election
Post-CL	0.005	0.040
	(0.047)	(0.064)
Post-CL × Classified Mgr.	-0.001	-0.144**
	(0.047)	(0.072)
Baseline controls × Year FEs	Yes	Yes
Log population, num. ISPs	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes
Observations	7,375	7,375
Number of newspapers	878	878
R^2	0.92	0.85
Mean dependent variable	2.73	4.95
Test for equality of coefficients (interaction term)	(1) vs (2): p-val = 0.032	

Notes: Regressions of content measures 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. Dependent variables: Corex model topic weights by newspaper and year (Panel a), and (ihs-transformed) number of articles containing the name of a congressional representative or candidate by newspaper and year (Panel b). In column 1 of panel (b), the dependent variable is limited to mentions occurring after the primary election date for the given two-year election cycle. In column 2, the dependent variable is limited to articles published before the primary election date. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/ Black/ Hispanic, all measured in the year 2000. Panel (b) additionally controls for the (ihs-transformed) total number of articles in relevant sections recorded by Newsbank. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 6: SELF-REPORTED NEWSPAPER READERSHIP

Panel (a): GfK-MRI

	(1) Read newspaper dummy: Full sample	(2) News propensity ≥ median, Classif. propensity < median	(3) News propensity < median, Classif. propensity ≥ median
Post-CL	-0.006	0.005	-0.004
	(0.007)	(0.011)	(0.010)
Post-CL × Classified Mgr.	-0.014*	-0.029**	-0.003
	(0.008)	(0.012)	(0.013)
Respondent controls	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes
Observations	248,460	83,494	81,703
Number of counties	781	756	762
R^2	0.09	0.10	0.05
Mean dependent variable	0.43	0.54	0.32

Panel (b): NAES

	(1) Read newspaper dummy: Full sample	(2) News propensity ≥ median, Classif. propensity < median	(3) News propensity < median, Classif. propensity ≥ median
Post-CL	0.016	0.029	0.005
	(0.012)	(0.032)	(0.019)
Post-CL × Classified Mgr.	-0.057**	-0.122***	-0.019
	(0.026)	(0.044)	(0.025)
Respondent controls	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes
Observations	106,348	30,133	34,831
Number of counties	1,192	1,087	1,167
R^2	0.05	0.09	0.06
Mean dependent variable	0.75	0.82	0.71

Notes: Dependent variable: indicator for self-reported newspaper readership (excluding national newspapers) in the GfK-MRI (panel a) and NAES surveys (panel b). Individual-level regressions 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 controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE 7: POLITICAL EFFECTS

Panel (a): Turnout in Congressional elections and split-ticket voting

	Turnout House/Senate		Split-ticket vote	
	(1)	(2)	(3)	(4)
Post-CL	-0.003	0.000	-0.014***	-0.005
	(0.003)	(0.004)	(0.005)	(0.006)
Post-CL × Classified Mgr.		-0.005		-0.013*
		(0.004)		(0.007)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
County FEs, Year-Office FEs	Yes	Yes	Yes	Yes
Observations	15,938	15,938	7,900	7,900
Number of counties	1,201	1,201	1,201	1,201
R^2	0.91	0.91	0.37	0.37
Mean dependent variable	0.45	0.45	0.11	0.11

Panel (b): Entry and performance of extreme candidates in House elections

	Extremist in primary		Extremist in general		Extremist wins general	
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	0.024	-0.032	0.042*	0.003	0.003	-0.026
	(0.023)	(0.027)	(0.024)	(0.029)	(0.019)	(0.018)
Post-CL × Classified Mgr.		0.095***		0.066**		0.049**
		(0.030)		(0.030)		(0.023)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,700	12,700	12,998	12,998	12,868	12,868
Number of counties	1,201	1,201	1,201	1,201	1,191	1,191
Number of districts	439	439	439	439	437	437
R^2	0.56	0.56	0.57	0.57	0.79	0.79
Mean dependent variable	0.79	0.79	0.75	0.75	0.37	0.37

Notes: Regressions of electoral outcomes 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. Panel (a): The level of observation is county × election. Observations are weighted by voting-age population. Standard errors clustered by CL-area. Panel (b): The level of observation is electoral district - county cell × election. Observations are weighted by the share of the voting-population in the district-county cell relative to the district. Standard errors clustered by district. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A Appendix: Additional Results

A.1 Newspaper outcomes: Robustness

TABLE A1: CORRELATES OF YEAR OF CL ENTRY

	Dependent variable: Year of CL entry					
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper jobs	-0.004	-0.003				
	(0.003)	(0.003)				
Newspaper circulation per capita			-0.224	0.205		
			(0.512)	(0.687)		
Newspaper classified manager					-0.041	-0.027
					(0.122)	(0.121)
Log population	-0.443***	-0.466**	-0.534***	-0.510**	-0.508***	-0.523***
	(0.150)	(0.184)	(0.176)	(0.247)	(0.150)	(0.190)
Internet service providers	-0.647**	-0.389**	-0.659**	-0.403**	-0.670**	-0.401**
	(0.289)	(0.174)	(0.299)	(0.173)	(0.278)	(0.168)
Share urban		0.013*		0.013		0.013*
		(0.007)		(0.008)		(0.007)
College degree		-0.038*		-0.035		-0.036*
		(0.021)		(0.022)		(0.021)
Rental share		-0.032*		-0.040**		-0.037*
		(0.018)		(0.020)		(0.019)
Income per capita		-0.971		-1.023		-0.963
		(1.374)		(1.485)		(1.367)
Unemployment rate		0.013		0.017		0.017
		(0.097)		(0.097)		(0.097)
Median age		-0.063*		-0.070*		-0.067*
		(0.038)		(0.039)		(0.038)
Share White		0.080*		0.079*		0.080*
		(0.048)		(0.046)		(0.047)
Share Black		0.075		0.075*		0.076*
		(0.046)		(0.044)		(0.046)
Share Hispanic		-0.008		-0.009		-0.009
		(0.013)		(0.013)		(0.014)
Turnout		0.876		0.717		0.729
		(2.379)		(2.420)		(2.410)
Observations	616	616	616	616	616	616
R^2	0.38	0.45	0.38	0.45	0.38	0.45

Notes: Regressions of year of CL entry on county characteristics in the year 2000. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A2: ALTERNATIVE CONTROLS

Panel (a): No controls

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-2.456***	-0.429	-0.007***	-0.003	-0.013***	0.001	-0.042	0.037
	(0.421)	(0.464)	(0.001)	(0.002)	(0.005)	(0.006)	(0.034)	(0.048)
Post-CL × Classified Mgr.		-4.039***		-0.010***		-0.026***		-0.150***
		(0.661)		(0.003)		(0.008)		(0.052)
Baseline controls × Year FEs	No	No	No	No	No	No	No	No
Log population, num. ISPs	No	No	No	No	No	No	No	No
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.90	0.90	0.98	0.98	0.51	0.51	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Panel (b): No baseline controls

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-2.026***	-0.252	-0.006***	-0.002	-0.013***	0.001	-0.034	0.040
	(0.401)	(0.465)	(0.001)	(0.002)	(0.005)	(0.006)	(0.034)	(0.048)
Post-CL × Classified Mgr.		-3.629***		-0.009***		-0.026***		-0.141***
		(0.643)		(0.003)		(0.008)		(0.053)
Baseline controls × Year FEs	No	No	No	No	No	No	No	No
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.90	0.90	0.98	0.98	0.51	0.51	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

(Continued)

TABLE A2: ALTERNATIVE CONTROLS, CONTINUED

Panel (c): Additional control for Internet access

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.797*	0.686	-0.003**	0.002	-0.023***	-0.005	-0.098***	-0.020
	(0.464)	(0.545)	(0.002)	(0.002)	(0.006)	(0.008)	(0.031)	(0.054)
Post-CL × Classified Mgr.		-2.781***		-0.008***		-0.033***		-0.135**
		(0.695)		(0.002)		(0.010)		(0.055)
Share with self-reported Internet access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,760	8,673	8,822	8,720	3,868	3,838	4,183	4,158
Number of newspapers	1,423	1,410	1,427	1,412	701	694	738	732
R^2	0.91	0.91	0.98	0.98	0.56	0.56	0.96	0.94
Mean dependent variable	25.03	25.09	0.16	0.16	0.30	0.30	6.83	6.00

Notes: Regressions of selected newspaper-level outcomes 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. The table replicates Table 3, modifying the set of control variables. All specifications control for newspaper and year FEs. Panel (a) includes no other controls. Panel (b) includes controls for log population and number of Internet service providers. Panel (c) includes log population, number of Internet service providers and baseline controls interacted with year FEs (as Table 3) and adds a control for the share of survey respondents with self-reported Internet access. Internet access is computed as the average share of respondents reporting Internet access at home or at work by county and year in the pooled GfK-MRI and NAES surveys. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Location × year fixed effects

Panel (a): State × year FEs

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.199***	0.244	-0.003	0.002	-0.010*	0.001	-0.019	0.033
	(0.423)	(0.513)	(0.002)	(0.002)	(0.005)	(0.007)	(0.035)	(0.043)
Post-CL × Classified Mgr.		-2.995***		-0.008***		-0.021***		-0.091*
		(0.600)		(0.002)		(0.008)		(0.050)
State × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,095	7,023	7,414	7,359
Number of newspapers	1,451	1,438	1,454	1,439	860	852	883	877
R^2	0.91	0.91	0.98	0.98	0.57	0.57	0.94	0.94
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Panel (b): DMA × year FEs

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.222***	0.146	-0.004**	0.001	-0.006	0.006	-0.005	0.047
	(0.409)	(0.494)	(0.002)	(0.002)	(0.006)	(0.008)	(0.037)	(0.047)
Post-CL × Classified Mgr.		-2.873***		-0.009***		-0.023**		-0.101*
		(0.696)		(0.003)		(0.009)		(0.057)
DMA × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,040	21,830	22,207	21,970	6,613	6,514	6,971	6,900
Number of newspapers	1,431	1,418	1,434	1,419	835	824	853	845
R^2	0.92	0.92	0.98	0.99	0.66	0.67	0.95	0.95
Mean dependent variable	21.32	21.37	0.19	0.19	0.29	0.29	5.49	5.49

(Continued)

TABLE A3: PLACE × YEAR FIXED EFFECTS, CONTINUED

Panel (c): County × year FEs

	Number of jobs	Circulation per capita	Politics coverage topic weight	Congress coverage names count (ihs)
	(1)	(2)	(3)	(4)
Post-CL × Classified Mgr.	-4.310**	-0.007*	-0.054***	-0.019
	(1.690)	(0.004)	(0.019)	(0.152)
County × Year FEs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	6,105	6,193	1,203	1,413
Number of newspapers	413	413	162	182
R^2	0.94	0.99	0.82	0.97
Mean dependent variable	24.57	0.10	0.31	6.04

Notes: Regressions of selected newspaper-level outcomes 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. The table replicates Table 3, adding: state × year FEs in panel (a), DMA × year FEs in panel (b) and county × year FEs in panel (c). All specifications control for newspaper and year FEs. Note that county × year FEs absorb all controls that vary at the county × year level, as well as the main effect of *Post-CL*. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Dynamic effects of CL entry

Panel (a): Years since CL entry

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years post-CL	-0.488***	-0.115	-0.000	0.000	-0.003*	0.001	-0.012	0.005
	(0.119)	(0.097)	(0.000)	(0.001)	(0.002)	(0.002)	(0.011)	(0.013)
Years post-CL × Classified Mgr.		-0.749***		-0.002**		-0.007***		-0.031***
		(0.170)		(0.001)		(0.002)		(0.011)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

(Continued)

TABLE A4: DYNAMIC EFFECTS OF CL ENTRY, CONTINUED

Panel (b): Short- vs. long-term effects

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL short-term	-0.804**	0.319	-0.004***	-0.001	-0.012**	-0.003	-0.026	0.025
	(0.405)	(0.515)	(0.001)	(0.002)	(0.005)	(0.006)	(0.035)	(0.048)
Post-CL long-term	-2.117***	-0.207	-0.004*	0.001	-0.014**	0.003	-0.041	0.038
	(0.555)	(0.472)	(0.002)	(0.003)	(0.007)	(0.009)	(0.051)	(0.066)
Post-CL short-term × Classified Mgr.		-2.387***		-0.006***		-0.017**		-0.100*
		(0.588)		(0.002)		(0.008)		(0.056)
Post-CL long-term × Classified Mgr.		-3.949***		-0.011***		-0.035***		-0.152**
		(0.931)		(0.004)		(0.010)		(0.064)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Notes: Regressions of selected newspaper-level outcomes on the length of availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. The table replicates Table 3, replacing the indicator for the availability of a local CL website with years since CL entry in panel (a), and indicators for CL short-term (0-2 years) and long-term (>2 years) availability in panel (b). All specifications control for newspaper and year FEs. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Sample restrictions by newspaper size

Panel (a): Excluding top 100 largest newspapers

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.905**	0.403	-0.004***	-0.001	-0.014***	-0.004	-0.015	0.055
	(0.423)	(0.515)	(0.001)	(0.001)	(0.006)	(0.007)	(0.038)	(0.050)
Post-CL × Classified Mgr.		-2.824***		-0.007***		-0.021**		-0.134**
		(0.626)		(0.002)		(0.008)		(0.056)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,782	20,598	20,939	20,722	6,492	6,441	6,753	6,718
Number of newspapers	1,350	1,338	1,352	1,338	791	785	808	804
R^2	0.90	0.90	0.97	0.97	0.52	0.52	0.93	0.93
Mean dependent variable	20.69	20.73	0.16	0.16	0.29	0.29	5.42	5.42

Panel (b): Excluding top 25% and bottom 25% of newspapers

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.698	0.557	-0.002	0.000	-0.013*	-0.000	-0.034	0.059
	(0.683)	(0.831)	(0.001)	(0.002)	(0.007)	(0.009)	(0.050)	(0.067)
Post-CL × Classified Mgr.		-2.553***		-0.006**		-0.028**		-0.181**
		(0.805)		(0.002)		(0.012)		(0.073)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,273	11,165	11,380	11,256	3,726	3,688	3,753	3,728
Number of newspapers	725	718	726	718	445	441	445	442
R^2	0.92	0.92	0.90	0.90	0.50	0.50	0.92	0.93
Mean dependent variable	20.69	20.75	0.17	0.17	0.28	0.28	5.38	5.38

Regressions of selected newspaper-level outcomes 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. The table replicates Table 3, excluding the top 100 largest newspapers in terms of per-capita circulation in the year 2000 in panel (a), and excluding the top 25% and bottom 25% of newspapers in terms of per-capita circulation in the year 2000 in panel (b). Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A6: CONTROLLING FOR HETEROGENEITY BY NEWSPAPER SIZE

	Number of jobs	Circulation per capita	Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	0.268	-0.000	-0.001	-0.000	0.034	0.036
	(0.465)	(0.002)	(0.006)	(0.006)	(0.048)	(0.049)
Post-CL × Classified Mgr.	-2.860***	-0.006**	-0.026***	-0.025***	-0.115**	-0.089
	(0.581)	(0.003)	(0.008)	(0.010)	(0.055)	(0.055)
Post-CL × Baseline circulation	Yes	No	Yes	No	Yes	No
Post-CL × Baseline job-count	No	Yes	No	Yes	No	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,167	22,264	7,067	7,067	7,375	7,375
Number of newspapers	1,437	1,435	855	855	878	878
R^2	0.91	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.44	0.19	0.29	0.29	5.51	5.51

Notes: Regressions of selected newspaper-level outcomes 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. The table replicates Table 3 with the following modifications: Columns 1, 3 and 5 include the interaction of the availability of a local Craigslist website with newspapers' (demeaned) baseline circulation per capita. Columns 2, 4 and 6 include the interaction of the availability of a local Craigslist website with newspapers' (demeaned) baseline jobs-count. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A7: CLASSIFIED RELIANCE BASED ON THE SHARE OF CLASSIFIED PAGES

	Number of jobs	Circulation per capita	Politics coverage topic weight	Congress coverage names count (ihs)
	(1)	(2)	(3)	(4)
Post-CL	0.138	0.004	0.016	0.026
	(0.943)	(0.004)	(0.014)	(0.094)
Post-CL × [Share classif. pages≥ median]	-2.247	-0.013***	-0.059***	-0.131
	(1.486)	(0.004)	(0.017)	(0.106)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	3,852	3,866	1,238	1,335
Number of newspapers	244	244	129	138
R^2	0.92	0.98	0.57	0.94
Mean dependent variable	29.41	0.21	0.29	5.97

Notes: Regressions of selected newspaper-level outcomes on an indicator for the availability of a local Craigslist website and its interaction with baseline reliance on classified ads. The table replicatess Table 3, replacing the indicator for the presence of a classified manager at baseline with an indicator for above-median share of classified pages at baseline. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A8: Alternative samples and treatment definitions

Panel (a): Balanced panel

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.030**	0.362	-0.004**	0.000	-0.012**	-0.000	-0.025	0.035
	(0.402)	(0.485)	(0.001)	(0.002)	(0.005)	(0.006)	(0.036)	(0.049)
Post-CL × Classified Mgr.		-2.938***		-0.008***		-0.024***		-0.116**
		(0.591)		(0.003)		(0.008)		(0.055)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,148	20,973	21,291	21,089	6,995	6,926	7,282	7,227
Number of newspapers	1,331	1,320	1,334	1,321	838	831	856	850
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.72	21.77	0.19	0.19	0.29	0.29	5.52	5.52

Panel (b): Excluding newspapers that do not experience CL entry

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.088	1.222*	0.000	0.004**	-0.012**	0.001	-0.067*	-0.018
	(0.612)	(0.703)	(0.001)	(0.002)	(0.006)	(0.006)	(0.036)	(0.047)
Post-CL × Classified Mgr.		-2.737***		-0.008***		-0.025***		-0.087
		(0.587)		(0.003)		(0.008)		(0.055)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,280	12,173	12,387	12,253	4,405	4,374	4,615	4,585
Number of newspapers	792	785	795	786	493	490	512	509
R^2	0.90	0.90	0.98	0.98	0.54	0.54	0.93	0.93
Mean dependent variable	26.62	26.68	0.20	0.20	0.30	0.30	6.00	6.00

(Continued)

TABLE A8: ALTERNATIVE SAMPLES AND TREATMENT DEFINITIONS, CONTINUED

Panel (c): Broad definition of CL and newspaper markets

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL (broad)	-0.785**	0.461	-0.004***	-0.000	-0.008*	0.001	-0.017	0.029
	(0.336)	(0.391)	(0.001)	(0.001)	(0.005)	(0.006)	(0.032)	(0.043)
Post-CL(broad) × Classified Mgr.		-2.861***		-0.008***		-0.019**		-0.091*
		(0.492)		(0.002)		(0.007)		(0.050)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22377	22177	22543	22316	7139	7067	7430	7375
Number of newspapers	1451	1438	1454	1439	863	855	884	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Notes: Regressions of selected newspaper-level outcomes 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. The table replicates Table 3 with the following modifications. Panel (a) restricts the sample to a balanced panel of newspapers in operation in the entire period of 1995 to 2010. Panel (b) excludes from the sample newspapers that never experience CL entry. Panel (c) uses a broader definition of CL markets (defined based on the location of posted ads) and newspaper markets (defined based on dis-aggregated circulation data) — see Appendix sections B.1.1 and B.1.2. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A9: Alternative clustering of standard errors

Panel (a): Standard errors clustered by state

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.114**	0.322	-0.004**	0.000	-0.012**	-0.000	-0.028	0.034
	(0.470)	(0.466)	(0.001)	(0.002)	(0.005)	(0.007)	(0.041)	(0.052)
Post-CL × Classified Mgr.		-3.027***		-0.008***		-0.024***		-0.119**
		(0.517)		(0.002)		(0.007)		(0.055)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,377	22,177	22,543	22,316	7,139	7,067	7,430	7,375
Number of newspapers	1,451	1,438	1,454	1,439	863	855	884	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.38	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Panel (b): Standard errors clustered by newspaper group

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.210***	0.426	-0.003*	-0.001	-0.015***	-0.002	-0.005	0.051
	(0.331)	(0.466)	(0.002)	(0.002)	(0.006)	(0.007)	(0.041)	(0.051)
Post-CL × Classified Mgr.		-3.561***		-0.006**		-0.028***		-0.110*
		(0.574)		(0.003)		(0.007)		(0.062)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,332	18,147	18,455	18,270	6,283	6,224	6,517	6,474
Number of newspapers	1,296	1,283	1,298	1,285	788	780	805	800
R^2	0.91	0.91	0.98	0.98	0.51	0.52	0.93	0.93
Mean dependent variable	21.27	21.32	0.18	0.18	0.28	0.28	5.46	5.46

Notes: Regressions of selected newspaper-level outcomes 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. The table replicates Table 3 with standard errors clustered by state in panel (a) and by newspaper group in panel (b). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A10: SPILLOVER EFFECTS

	Number of jobs		Circulation per capita		Politics coverage topic weight		Congress coverage names count (ihs)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.312***	0.154	-0.004**	0.000	-0.012**	0.000	-0.031	0.031
	(0.445)	(0.526)	(0.002)	(0.002)	(0.005)	(0.007)	(0.037)	(0.049)
Post-CL DMA	0.355**	0.273	0.000	-0.000	-0.001	-0.001	0.006	0.006
	(0.177)	(0.174)	(0.001)	(0.001)	(0.002)	(0.003)	(0.018)	(0.018)
Post-CL × Classified Mgr.		-2.997***		-0.008***		-0.023***		-0.115**
		(0.598)		(0.003)		(0.008)		(0.054)
Post-CL DMA × Classified Mgr. DMA		0.068		0.001**		0.000		-0.006
		(0.068)		(0.000)		(0.001)		(0.009)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,339	22,135	22,543	22,306	7,110	7,050	7,407	7,363
Number of newspapers	1,451	1,437	1,454	1,438	862	855	883	878
R^2	0.91	0.91	0.98	0.98	0.52	0.52	0.93	0.93
Mean dependent variable	21.39	21.44	0.19	0.19	0.29	0.29	5.51	5.51

Notes: Regressions of selected newspaper-level outcomes on measures of the CL-exposure of other newspapers in the same media market (DMA). *Post-CL DMA* denotes the leave-one-out circulation weighted average number of newspapers in the same DMA that have experienced CL entry. *Classified Mgr. DMA* denotes the leave-one-out circulation weighted average number of newspapers in the same DMA that have a classified manager at baseline. Both variables are standardized to mean equal to 0 and standard deviation equal to 1. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A11: NUMBER OF NEWSPAPERS AND CHANGES IN OWNERSHIP

	Num. newspapers HQ-ed in county		Change in ownership	
	(1)	(2)	(3)	(4)
Post-CL	0.018	0.012	0.001	-0.000
	(0.012)	(0.014)	(0.007)	(0.009)
Post-CL × Classified Mgr.		0.011		0.004
		(0.018)		(0.010)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	No	No
Newspaper FEs	No	No	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	19,216	19,216	21,428	20,857
Number of counties	1,201	1,201		
Number of newspapers			1,541	1,442
R^2	0.95	0.95	0.52	0.52
Mean dependent variable	1.19	1.19	0.13	0.13

Notes: Regressions of the number of local newspapers (columns 1 and 2) and an indicator for a change in newspapers' ownership (columns 3 and 4) 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. The specifications in columns 1 and 2 are at the county-year level, so that *Classified Mgr.* denotes the circulation-weighted share of newspapers with a classified manager at baseline. The specifications in columns 2 and 3 are at the newspaper-year level. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A12: NUMBER OF PAGES PER ISSUE AND SUBSCRIPTION PRICES

	Total pages		Subscription price	
	(1)	(2)	(3)	(4)
Post-CL	0.247*	0.147	-0.020***	-0.011
	(0.139)	(0.144)	(0.007)	(0.009)
Post-CL × Classified Mgr.		0.193		-0.020*
		(0.311)		(0.011)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	21,402	21,211	19,393	19,197
Number of newspapers	1,413	1,400	1,365	1,351
R^2	0.97	0.97	0.94	0.94
Mean dependent variable	28.51	28.53	4.71	4.71

Notes: Regressions of the average number pages per newspaper issue and average yearly subscription price (both reported by E&P) 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 controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share of White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.2 Substitution to other media

To understand the implications of the documented decline in readership, an important 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 Table A13. With the exception of a positive coefficient for the readership of national newspapers reported in the NAES, we find no consistent evidence of substitution to other media, including TV, radio, and online sources.[46][47] To the extent that local newspapers represent a unique source of detailed information about local political affairs that is hard to come across in other media, increased consumption of other sources can hardly compensate for this loss of information.

The browsing data from Comscore allow us to perform an alternative test for the substitution to online news sources. In Table A14 we examine the effect of CL's entry on the number of visits to the websites of the top 3 national newspapers (nytimes.com, wsj.com and usatoday.com), as well as visits to any of the top 100 domains classified by Comscore as news-related. We find no significant effect on visits to these domains.

Taken together, these findings suggest that the decline in readership of local newspapers associated with the entry of CL is unlikely to be fully compensated by increased news consumption online or through other media. These effects are therefore likely to translate into a net decline in exposure to political information.

[46] The NAES does not include consistent questions on general consumption of online media, but does ask whether the respondent has seen online information on the presidential election campaign. We find no effect of CL entry on this measure.

[47] The NAES survey also asks about the number of days in the past week the respondent consumed news through each source. We obtain similar results using the continuous number of days or a dummy for 0 vs. positive number days as a dependent variable.

TABLE A13: SELF-REPORTED CONSUMPTION OF OTHER MEDIA

Panel (a): GfK-MRI

	(1) Read newspaper, national	(2) Watched TV	(3) Listened radio	(4) Read online
Post-CL	0.007	0.002	0.011	0.006
	(0.005)	(0.009)	(0.008)	(0.007)
Post-CL × Classified Mgr.	0.002	-0.011	-0.002	-0.009
	(0.005)	(0.009)	(0.009)	(0.007)
Respondent controls	Yes	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	248,460	248,460	248,460	248,460
Number of counties	781	781	781	781
R^2	0.14	0.08	0.09	0.16
Mean dependent variable	0.09	0.70	0.17	0.21

Panel (b): NAES

	(1) Read newspaper, national	(2) Watched TV	(3) Listened radio
Post-CL	0.003	0.011	0.009
	(0.010)	(0.008)	(0.008)
Post-CL × Classified Mgr.	0.040*	-0.013	-0.006
	(0.024)	(0.009)	(0.009)
Respondent controls	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes
Observations	106,348	106,432	147,090
Number of counties	1,192	1,192	1,193
R^2	0.09	0.03	0.05
Mean dependent variable	0.03	0.92	0.37

Notes: Dependent variables: indicators for self-reported news consumption via specific media types in the GfK-MRI (panel a) and NAES surveys (panel b). Individual-level regressions 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 controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share of White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A14: ONLINE NEWS CONSUMPTION

	(1) Visits to nytimes.com (ihs)	(2) Visits to wsj.com (ihs)	(3) Visits to usatoday.com (ihs)	(4) Visits to top 3 newspaper websites (ihs)	(5) Visits to top 100 news websites (ihs)
Post-CL	-0.065	0.080	-0.027	-0.084	-0.004
	(0.060)	(0.058)	(0.065)	(0.062)	(0.011)
Post-CL × Classified Mgr.	-0.099	0.080	0.119	0.018	0.004
	(0.075)	(0.077)	(0.078)	(0.077)	(0.012)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	8,331	8,331	8,331	8,331	8,331
Number of counties	1,200	1,200	1,200	1,200	1,200
R^2	0.81	0.73	0.76	0.81	0.99
Mean dependent variable	2.90	1.38	2.75	3.70	10.40

Notes: Regressions of (ihs-transformed) visits to selected news websites on an indicator for the availability of a local Craigslist website and its interaction with the circulation-weighted average of newspapers with a classified manager at baseline. All specifications control for the (ihs-transformed) total visits recorded by Comscore by county and year. "Top 100 news websites" indicates Comscores' classification, available for the year 2002. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share of White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Additional results on content

TABLE A15: TOPIC MODEL WEIGHTS: SEPARATING POLITICAL 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 × Classified Mgr.	-0.020***	-0.014*	-0.009**	-0.011**
	(0.006)	(0.008)	(0.004)	(0.005)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	7,067	7,067	7,067	7,067
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

(Continued)

TABLE A15: TOPIC MODEL WEIGHTS: SEPARATING POLITICAL TOPICS, CONTINUED

	(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 × Classified 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)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,067	7,067	7,067	7,067	7,067	7,067
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

Notes: Dependent variables: Topic weights by newspaper and year, estimated from an anchored Corex model with ten topics. Each topic is labeled by its 5 most representative words. The topic anchors are: ['washington', 'feder', 'govern', 'presid'], ['council', 'mayor'], ['repres', 'congress', 'senat'], ['intern', 'abroad', 'foreign']. Regressions 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 controls include share of urban population, share educated residents, rental share of housing, log income per capita, median age, turnout and share of White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE A16: MENTIONS OF IN-STATE CONGRESSIONAL INCUMBENTS AND CANDIDATES: HETEROGENEITY

	Split by incumbency		Split by office	
	(1)	(2)	(3)	(4)
Congress coverage names count (ihs)	Incumbents	Challengers	Senate	House
Post-CL	0.027	-0.005	0.000	0.039
	(0.048)	(0.101)	(0.059)	(0.071)
Post-CL × Classified Mgr.	-0.089	-0.083	-0.103	-0.122*
	(0.055)	(0.105)	(0.068)	(0.068)
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes	Yes
Observations	7,375	7,375	7,375	7,375
Number of newspapers	878	878	878	878
R^2	0.93	0.72	0.87	0.90
Mean dependent variable	5.31	3.29	4.48	4.99
Test for equality of coefficients (interaction term)	(1) vs (2): p-val = 0.940		(3) vs (4): p-val = 0.752	

Notes: Dependent variables are (ihs-transformed) counts of articles mentioning candidates and current officeholders for Congressional offices (US House and Senate) in the same state of the newspaper's headquarters. Regressions 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 control for the (ihs-transformed) total number of articles in relevant sections recorded by Newsbank. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share of White/Black/Hispanic, 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 A17: MENTIONS OF OTHER POLITICIANS

	(1) Congress. leaders articles count (ihs)	(2) President articles count (ihs)	(3) State and local articles count (ihs)
Post-CL	0.050	-0.008	0.005
	(0.079)	(0.057)	(0.029)
Post-CL × Classified Mgr.	-0.250***	-0.150**	-0.075**
	(0.092)	(0.067)	(0.037)
Baseline controls × Year FEs	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes	Yes
Observations	7,375	7,375	7,375
Number of newspapers	878	878	878
R^2	0.86	0.90	0.95
Mean dependent variable	3.78	5.14	7.58

Notes: The dependent variable is the (ihs-transformed) count of articles containing any of a set of search expressions. Column 1 searches for mentions of presidents in office during the sample period: "`President|Bill) Clinton`", "`President|George W?\.? ?)Bush`", "`(President|Barack) Obama`". Column 2 searches for names of speakers, majority and minority leaders, and party whips in Congress during the sample period (see Appendix B.1.4 for the list). Column 3 searches for a set of regular expressions capturing commonly used titles of elected officials at the local and state levels: "`mayor`", "`council ?person`", "`council ?(wo)?man`", "`council ?member`", "`state (sen(ator\|\.)|rep(resentative|\.))`", "`governor`", "`alderman`", "`commissioner`", "`(city|county) manager`", "`county judge`", "`district attorney`", and "`attorney general`". All searches are case insensitive. Regressions 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 control for the (ihs-transformed) total number of articles in relevant sections recorded by Newsbank. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A18: Generic mentions of Congressional institutions

	(1) Congress keywords articles count (ihs)	(2) Congress. primary keywords articles count (ihs)
Post-CL	0.015	0.061
	(0.038)	(0.053)
Post-CL × Classified Mgr.	-0.128***	-0.215***
	(0.045)	(0.064)
Baseline controls × Year FEs	Yes	Yes
Log population, num. ISPs	Yes	Yes
Newspaper FEs, Year FEs	Yes	Yes
Observations	7,375	7,375
Number of newspapers	878	878
R^2	0.95	0.84
Mean dependent variable	6.39	3.98

Notes: Dependent variables are (ihs-transformed) counts of articles mentioning "US House", "Senate," or "Congress" in a given newspaper-year (cols. 1); and articles that in addition to the previous condition mention the words "primary" or "nomination" and are published prior to the primary election date in the given election cycle (cols 2). Regressions 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 control for the ihs-transformed total number of articles in relevant sections recorded by Newsbank.. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

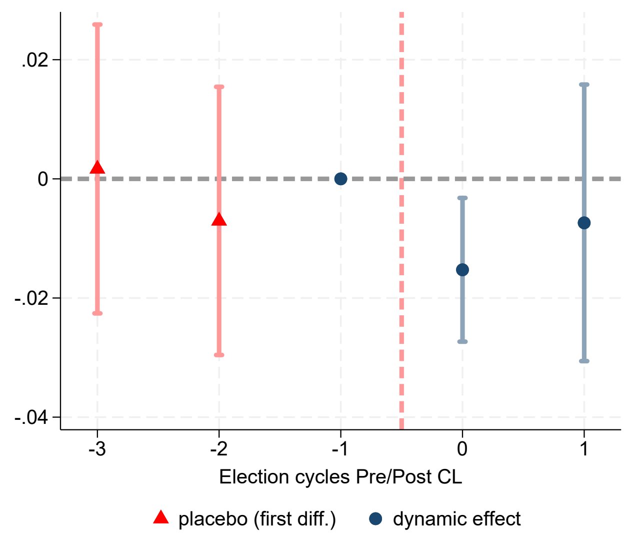
TABLE A19: MENTIONS OF HAMILTON (2016) ACCOUNTABILITY KEYWORDS

	(1) Accountability words
Post-CL	0.009
	(0.034)
Post-CL × Classified Mgr.	-0.086**
	(0.037)
Baseline controls × Year FEs	Yes
Log population, num. ISPs	Yes
Newspaper FEs, Year FEs	Yes
Observations	7,375
Number of newspapers	878
R^2	0.95
Mean dependent variable	6.68

Notes: The dependent variable is the (ihs-transformed) count of articles mentioning a set of keywords used by Hamilton (2016) as indicators of conflicts of interest revealed by investigative journalism: “wasteful”, “mismanagement”, “neglect”, “bribery”, “embezzlement”, “steal”, “corrupt”, “nepotism”, “patronage”, “conflict of interest”, “rent seeking”, “influence peddling”, “favoritism”, “abuse”, “harassment”, “misconduct”, “discrimination”, “misuse”, “fraud”, “deception”, and “mislead.” Regressions 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 control for the (ihs-transformed) total number of articles in relevant sections recorded by Newsbank. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by CL-area. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.4 Additional results on political outcomes

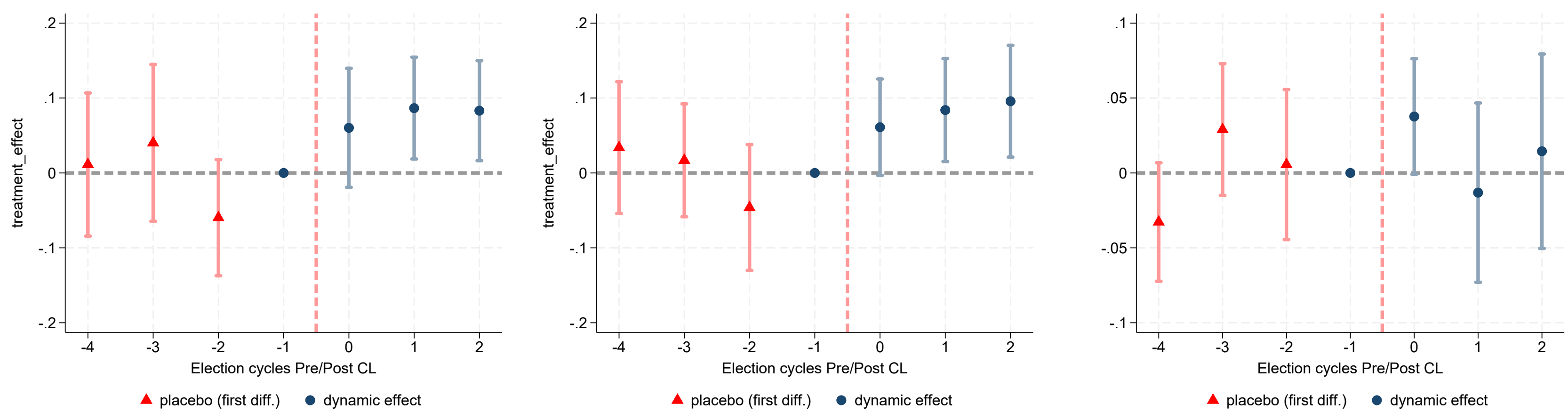
FIGURE A1: SPLIT-TICKET VOTING – EVENT STUDY



Notes: Pre-treatment placebos and dynamic effects of the entry of a local CL website on split-ticket voting. The graph presents coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Pre-treatment coefficients correspond to first-differences between consecutive periods. Post-treatment coefficients correspond to cumulative effects for the respective length of exposure to CL. The dependent variable is the absolute difference in the Republican vote share in congressional elections and concurrent presidential elections. The treatment is defined as CL-entry into a county in which more than 50% of newspapers (circulation-weighted) had a classified manager in the year 2000. The level of observation is a district-county cell $\times$ election and observations are weighted by the share of voting-age population in the district-county cell relative to the district. Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

FIGURE A2: ENTRY AND PERFORMANCE OF EXTREMIST CANDIDATES – EVENT STUDIES

PANEL (A): EXTREMIST IN PRIMARY ELECTION — PANEL (B): EXTREMIST IN GENERAL ELECTION — PANEL (C): EXTREMIST WINS GENERAL ELECTION



Notes: Pre-treatment placebos and dynamic effects of the entry of a local CL website on the presence and performance of extremist candidates in House elections. The graphs present coefficients and 95% confidence intervals based on the DID_M estimator proposed in de Chaisemartin and D'Haultfoeuille (2020). Pre-treatment coefficients correspond to first-differences between consecutive periods. Post-treatment coefficients correspond to cumulative effects for the respective length of exposure to CL. The dependent variables are an indicator for the presence of an extremist candidate in a primary election in panel (a), the presence of an extremist candidate in the general election in panel (b), and a general election win for an extremist candidate in panel (c). The treatment is defined as CL-entry into a county in which more than 50% of newspapers (circulation-weighted) had a classified manager in the year 2000. The level of observation is a district-county cell × election and observations are weighted by the share of voting-age population in the district-county cell relative to the district. Controls include log population and number of Internet service providers. Standard errors clustered by electoral district.

TABLE A20: POLITICAL EFFECTS:
ROBUSTNESS TO THE INCLUSION OF STATE-SPECIFIC TRENDS

Panel (a): Turnout in Congressional elections and split-ticket voting

	Turnout House/Senate		Split-ticket vote	
	(1)	(2)	(3)	(4)
Post-CL	-0.002	0.000	-0.010**	-0.002
	(0.002)	(0.002)	(0.005)	(0.005)
Post-CL × Classified Mgr.		-0.004		-0.013**
		(0.003)		(0.006)
State × Linear time trend	Yes	Yes	Yes	Yes
Baseline controls × Year FEs	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes
County FEs, Year-Office FEs	Yes	Yes	Yes	Yes
Observations	15,938	15,938	7,900	7,900
Number of counties	1,201	1,201	1,201	1,201
R^2	0.92	0.92	0.41	0.41
Mean dependent variable	0.45	0.45	0.11	0.11

Panel (b): Entry and performance of extreme candidates in House elections

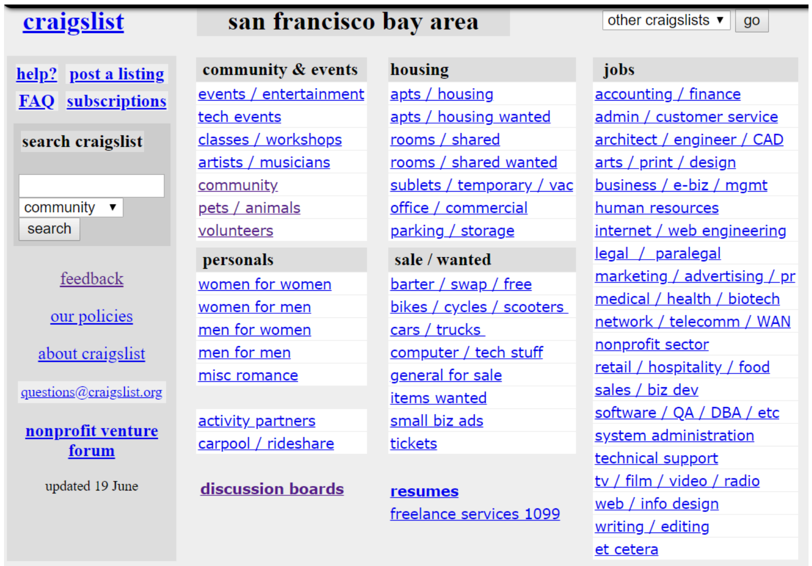
	Extremist in primary		Extremist in general		Extremist wins general	
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	0.015	-0.040	0.043*	0.004	0.009	-0.014
	(0.023)	(0.026)	(0.024)	(0.029)	(0.017)	(0.018)
Post-CL × Classified Mgr.		0.092***		0.066**		0.038*
		(0.028)		(0.029)		(0.021)
State × Linear time trend	Yes	Yes	Yes	Yes		
Baseline controls × Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Log population, num. ISPs	Yes	Yes	Yes	Yes	Yes	Yes
County FEs, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
District - redistricting regime FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,700	12,700	12,998	12,998	12,868	12,868
Number of counties	1,201	1,201	1,201	1,201	1,191	1,191
Number of districts	439	439	439	439	437	437
R^2	0.57	0.57	0.58	0.58	0.80	0.80
Mean dependent variable	0.79	0.79	0.75	0.75	0.37	0.37

Notes: Regressions of electoral outcomes 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. Panel (a): The level of observation is county × election. Observations are weighted by voting-age population. Standard errors clustered by CL-area. Panel (b): The level of observation is electoral district - county cell × election. Observations are weighted by the share of the voting-population in the district-county cell relative to the district. Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/Black/Hispanic, all measured in the year 2000. Standard errors clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01.

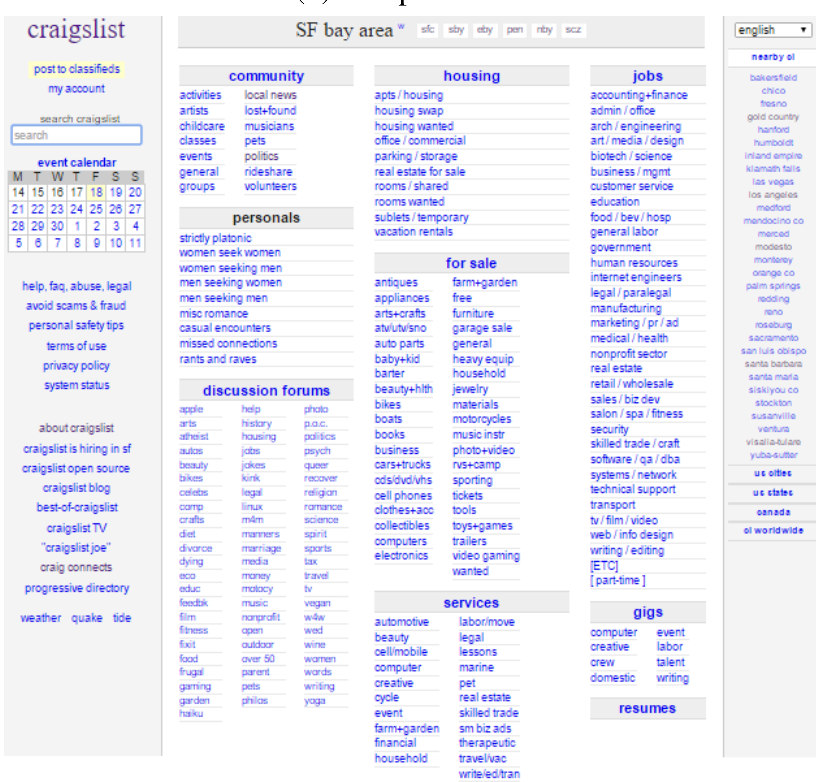
B Appendix: Background and Data

FIGURE B1: CRAIGSLIST: LAYOUT IN 2000 AND 2016

Panel (a): Snapshot from 2000



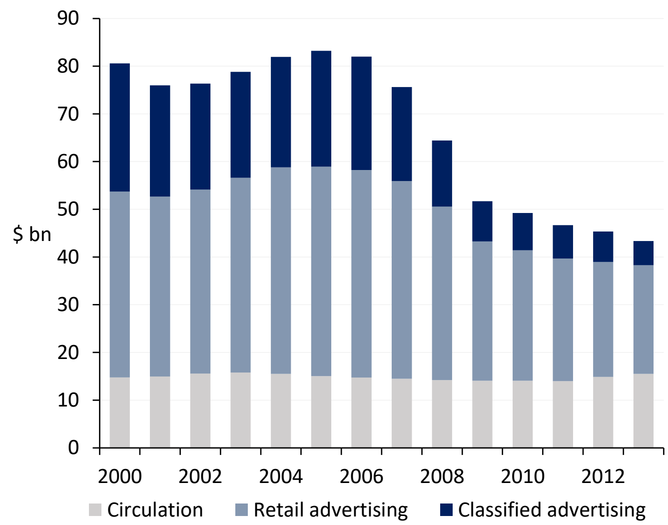
Panel (b): Snapshot from 2016



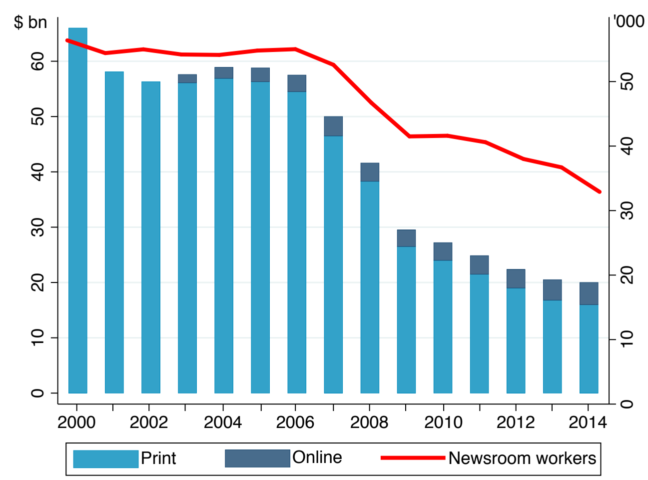
Notes: Layout of CL's San Francisco/ Bay Area page in 2000 (Panel a) and 2016 (Panel b).

FIGURE B2: EVOLUTION OF NEWSPAPER REVENUES AND NEWSROOM WORKERS

Panel (a): Revenues by source



Panel (b): Advertising revenues and newsroom workers



Notes: Panel (a): Newspaper revenues from circulation, retail advertising and classified advertising - 2000 to 2013. Source: Newspaper Association of America. Panel (b): Newspaper advertising revenues and number of newsroom workers - 2000 to 2014. Sources: Newspaper Association of America and American Society of Newspaper Editors. Both series are expressed in constant 2012 US dollars.

Figure B3: 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.

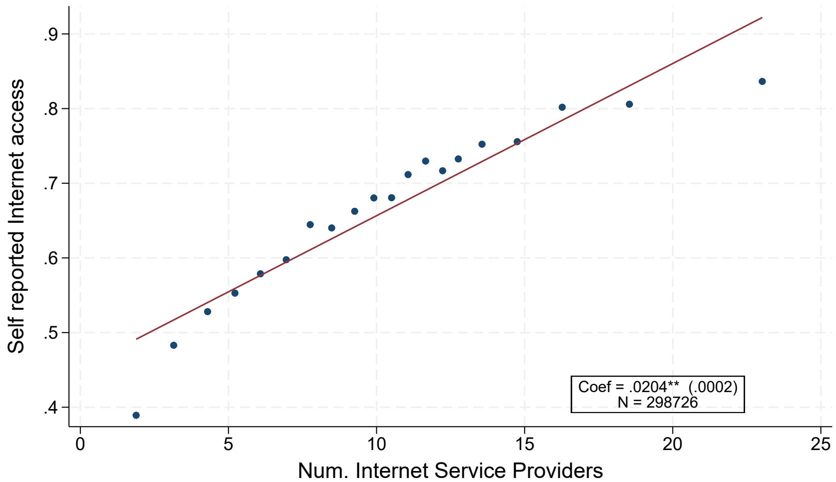
Books Ed.	Elizabeth Taylor
Editorial Page Ed.	Bruce Dold
Entertainment Ed.	Scott Powers
Foreign Ed.	Colin McMahon
Good Eating Ed.	Carol Haddix
Nat'l Ed.	Storer Rowley
Special Sections Ed.	Janet Franz
Sports Ed.	Bill Adee
Sunday Magazine Ed.	Elizabeth Taylor
Tempo Ed.	Tim Bannon
Travel Ed.	Randy Curwen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

The Chicago Tribune

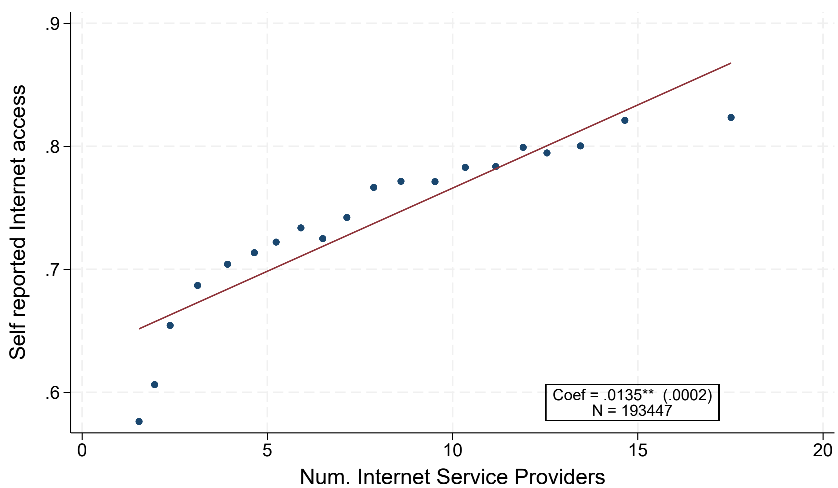
Notes: Extracts from the print version of the 2003 Editor & Publisher Yearbook for the Landsdale Reporter (upper panel) and the Chicago Tribune (lower panel).

FIGURE B4: NUMBER OF ISPS AND SELF-REPORTED INTERNET ACCESS

Panel (a): GfK-MRI

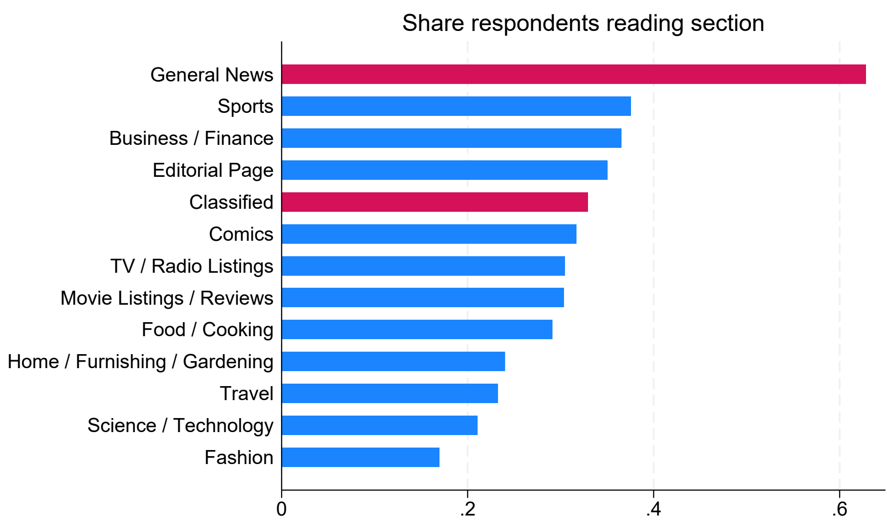


Panel (b): NAES



Notes: Binned scatter plots and best linear fit 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). The figures also report the estimate slope of the regression line.

FIGURE B5: NEWSPAPER READERSHIP BY SECTION



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.

B.1 Details on data construction and validation

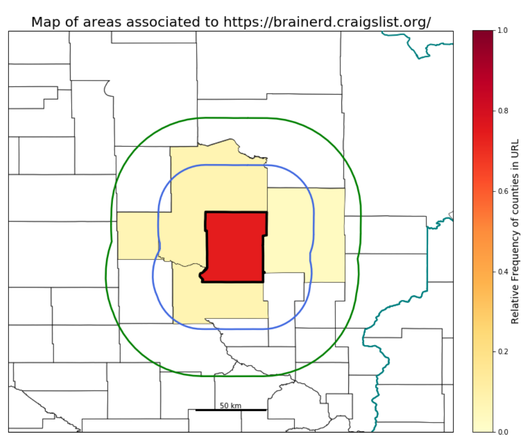
B.1.1 Craigslist markets

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 B6 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 B6: 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.

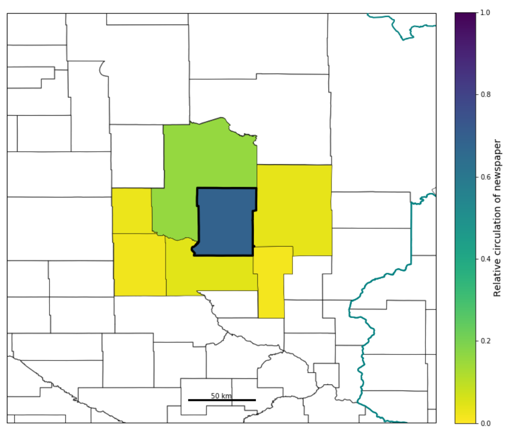
B.1.2 Newspaper markets

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.[48] 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.[49]

Figure B7 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.[50] The median paper in our AAM data has about 85% of its total circulation in the headquarters county once we exclude outliers.

FIGURE B7: DISTRIBUTION OF CIRCULATION OF THE BRAINERD DISPATCH



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

48 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.

49 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.

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

B.1.3 Validating the classified manager proxy

We validate the classified manager indicator as a proxy for classified intensity in two ways. First, 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 with 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.[51] We supplemented this automatic measurement with manual checks for issues with abnormally low detected pages. We collected this measure for a total of 43,165 issues across the 251 papers available in the Newspapers.com archive. Prior to the 2000, classified advertising occupied around 26% of the typical issue's total page count in the average paper in our sample. Because Saturdays and especially Sundays typically had higher classified intensity than other days, we compute averages by newspaper-day of week and include weekday fixed effects in all analyses.

Second, we digitized data on classified rates (prices) from yearbooks published by Standard Rate and Data Service (SRDS). This is the source used in Seamans and Zhu (2014). Following Seamans and Zhu (2014), we measure the log classified rate. Because there is variation in the unit of measurement that newspapers report to SRDS — about half report prices in per-inch terms, with the remainder split roughly evenly between per-column-inch and per-line terms — and prices vary systematically across units, we compute average prices by newspaper-unit and include unit fixed effects in all analyses.

We examine cross-sectional variation in classified intensity prior to 2000, before Craigslist entry. Table B1 shows the results of regressions where the outcome is either the average fraction of pages per issue that contain classified advertising, by newspaper-weekday, or the average log price, by newspaper-unit. We additionally include controls for the newspaper's circulation per capita, the average total pages per issue, and a set of county-level sociodemographic characteristics in both.

The table shows that both the amount of space devoted to classified ads and typical classified prices were higher for newspapers that had a classified manager in 2000, prior to Craigslist entry. The magnitudes imply a roughly 7 percentage point higher share of pages devoted to classifieds and a 9 percent higher classified price for newspapers with classified managers in 2000, even compared to newspapers of similar size, circulation, and in similar counties. These estimates are also reported in graphical form in Figure 3.

[51] We use the number of pages per issue reported in E&P rather than in the newspapers.com scans as the denominator, because the former has less measurement error.

TABLE B1: SHARE PAGES DEVOTED TO CLASSIFIED ADS AND CLASSIFIED RATE IN PRE-CL PERIOD, BY PRESENCE OF CLASSIFIED MANAGER IN 2000.

	Share classified pages	Log classified rate
	(1)	(2)
Classified Mgr.	0.065***	0.093**
	(0.024)	(0.039)
Day-of-week FEs	Yes	No
Unit FEs	No	Yes
Baseline controls	No	No
Log population, num. ISPs	Yes	Yes
Newspaper circulation and jobscount	Yes	Yes
Total newspapers.com pages	Yes	No
Observations	626	2,070
Number of newspapers	243	1,390
R^2	0.41	0.23
Mean dependent variable	0.26	2.30

Notes: Regressions of the average share of pages per issue devoted to classified advertising (column 1) and log classified rate (column 2) on an indicator for the presence of a classified manager in 2000. An observation is a newspaper-weekday (column 1) or a newspaper-unit of measurement (column 2). Baseline controls include share of urban population, share college educated, rental share of housing, log income per capita, median age, turnout and share White/ Black/ Hispanic, all measured in the year 2000. Robust standard errors. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.1.4 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 year between 1999 and 2010. The topic-modeling sample limits to the first paragraph of text, plus the headline.

Politician Names Our keyword searches for politician names 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 GovTrack (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). We harmonized the names of individuals appearing in both datasets, paying particular attention to include both the full name and nicknames where these are commonly used by the representative (e.g., we search for both "Chuck Schumer" and "Charles Schumer").

We pre-filter the set of articles to search by, first, excluding articles with missing main text; second, dropping articles where the section name matches the regular expression

```
Obit|Sport|Art|Entertain|Auto|Estate
```

which excludes articles in sections that are unlikely to discuss candidates for Congress; and third, dropping any articles with duplicated text to another article published by the same source on the same day.

Among the resulting set of articles, we construct two searches. In the first, we construct a (case-insensitive) regular expression for each individual of the form:

```
sen([.]|at[a-z.,]*) firstname lastname
```

for Senators, and

```
(congres[a-z.,]*|rep([.]|resentative)|member( of the house)?) firstname lastname
```

for members of the House, where `firstname` is either the representative's first or nickname. These expressions match 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.

The second search looks for the candidate's standardized full name alone (again in the form "firstname lastname"), but only among articles in which one of the terms "Congress," "Senate," or "US House" appears. The requirement that the article mention a chamber of Congress serves the function of reducing false positive matches.

For sitting representatives, we count an article as mentioning the representative if either the first search OR the second returns a match. For challenger candidates, we use only the second search. When a sitting House member runs for Senate (and thus is both an incumbent representative for the House and a candidate for Senate) we use the result of the incumbent-form search, i.e. the two-criterion version.

We count the number of *articles*[52] in which each of the list of same-state candidate names appears on each newspaper-day, and then aggregate to the level of newspaper by year.

The search for party leaders (in Table 5) looks for the names of Speakers, Majority and Minority Leaders, and party Whips in office during the sample period; these are: "`(Leader|Dick) Gephardt`", "`(Speaker|Nancy) Pelosi`", "`(Speaker|Denn(y|is)) Hastert`", "`(Leader|Dick) Armey`", "`(Leader|Whip|Tom) DeLay`", "`(Whip|Roy) Blunt`", "`(Leader|John) Boehner`",

[52] 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.

“`(Whip|Eric) Cantor`”, “`(Whip|Steny) Hoyer`”, “`(Whip|David) Bonior`”, “`(Jim|James) Clyburn`”, “`(Leader|Whip|Harry) Reid`”, “`(Leader|Dick) Durbin`”, “`(Whip|Tom) Daschle`”, “`(Whip|Don) Nickles`”, “`(Leader|Mitch) McConnell`”, “`(Leader|Bill) Frist`”, “`(Leader|Whip|Trent) Lott`”, and “`(Whip|John) Kyl`”.

Topic model Our method for extracting newspapers' topical coverage follows Gallagher et al.'s 2017b Correlation Explanation (CorEx) method. This is an information-theoretic approach to learning latent topics over documents. Unlike generative models such as the Latent Dirichlet Allocation (LDA), CorEx does not assume a particular data generating model, and instead searches for topics that are "maximally informative" about a set of documents. CorEx has the advantage of extracting coherent and interpretable topics from short texts (in our case: first paragraphs). We apply this method to the text of a random sample of 2 million articles from the NewsBank corpus.

Table B2 presents the resulting topics, as described by their most representative words. The 5 resulting topics can be labeled as follows: politics, sports, entertainment, crime, obituaries.

TABLE B2: TOP TEN REPRESENTATIVE WORDS FOR EACH COREX TOPIC

Sports:	game,team,coach,win,season,victori,plai,score,footbal,player
Obituaries:	di,born,funer,son,daughter,church,surviv,servic,cemeteri,obituari
Politics:	presid,propos,board,vote,plan,approv,govern,elect,tax,republican
Crime:	polic,charg,arrest,man,investig,court,kill,sheriff,suspect,offic
Entertainment:	music,art,movi,featur,festiv,food,event,concert,artist,film

Alternatively, we implement an anchored version of the CorEx model to separate 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']. Table B3 lists the resulting 10 topics.

For each of the 2 million articles in the corpus, the CorEx model outputs a set of unconditional probabilities for the article belonging to a given topic. 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 with the average probability for each one of the topics as dependent variable.

TABLE B3: TOP TEN WORDS REPRESENTATIVE WORDS FOR EACH COREX TOPIC: ANCHORED MODEL

Politics 1 – Presidential:	**presid**, **feder**, **govern**, compani, tax, **washington**, percent, increas, pai, billion
Politics 2 – Local:	**council**, **mayor**, board, plan, student, educ, fund, commun, project, program
Politics 3 – Congressional:	**repres**, **senat**, **congress**, republican, elect, democrat, vote, candid, polit, gov
Politics 4 – Foreign:	**intern**, war, **foreign**, iraq, militari, movi, film, american, soldier, terrorist
Topic 5:	man, kill, injuri, injur, accid, crash, woman, diseas, victim, suffer
Topic 6:	music, art, food, festiv, featur, concert, event, artist, band, holidai
Topic 7:	car, vehicl, driver, road, truck, traffic, highwai, drive, mile, street
Topic 8:	di, born, funer, son, daughter, church, surviv, servic, cemeteri, obituari
Topic 9:	game, team, coach, win, season, plai, victori, footbal, score, player
Topic 10:	polic, charg, court, arrest, judg, investig, attornei, accus, sheriff, suspect