

DISCUSSION PAPER SERIES

DP16130
(v. 2)

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

Milena Djourelova, Ruben Durante and Gregory
Martin

PUBLIC ECONOMICS



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

Milena Djourelova, Ruben Durante and Gregory Martin

Discussion Paper DP16130
First Published 09 May 2021
This Revision 06 July 2021

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

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

- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Milena Djourelova, Ruben Durante and Gregory Martin

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

Abstract

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

JEL Classification: L82, L86, D72

Keywords: newspapers, internet, advertising, ideological polarization

Milena Djourelova - milena.djourelova@upf.edu
Universitat Pompeu Fabra, IPEG, Barcelona GSE

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

Gregory Martin - gjmartin@stanford.edu
Stanford GSB

Acknowledgements

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

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

Milena Djourelova[†] Ruben Durante[‡] Gregory J. Martin[§]

June 2021

Abstract

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

Keywords: newspapers, Internet, advertising, ideological polarization

JEL codes: L82, L86, D72

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

[†]UPF, BGSE, IPEG. Email: milena.djourelova@upf.edu

[‡]ICREA-UPF, BGSE, IPEG, and CEPR. Email: ruben.durante@upf.edu

[§]Stanford GSB. Email: gjmartin@stanford.edu

1. INTRODUCTION

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

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

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

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

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

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

The lack of a clear profit maximization strategy³ alleviates concerns that the timing of CL's entry might have been driven by strategic considerations related to the conditions of local newspaper markets. Indeed, we document that the timing of CL's entry into a local market is not correlated with the characteristics of local newspapers once population and the quality of the local Internet connection are controlled for.

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

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

We then test to what extent these organizational changes affected newspapers' editorial priorities, with particular regard to their coverage of politics. First, applying keyword searches to the entire corpus of articles published in over 800 newspapers, we compute the number of mentions of local representatives in Congress. We document that, following the entry of CL in a given area, news coverage of local representatives in affected papers declined significantly, by around 30%. We complement this approach with a semi-supervised topic model, estimated on a random sample of two million articles. We find evidence of a decline in the prevalence of topics related to politics, while we find no significant effect on

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

⁴ To corroborate the validity of this proxy, we document that: i) the presence of classified ads managers is strongly correlated with the share of pages devoted to classified ads prior to the entry of CL, and ii) the space devoted to classified ads decreased disproportionately for papers with classified managers after the entry of CL (see Appendix Tables A1 and A2.)

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

other topics such as sports, entertainment, or crime.

Next we examine how readers responded to these changes in content. First, we document that, in the years after the entry of CL, local newspapers experienced a sharp decline in circulation. We further explore this readership decline using data from two large-scale surveys on media consumption. The results confirm that, following the entry of CL, local respondents are less likely to report reading a local newspaper. Interestingly, this trend is primarily driven by readers that are relatively *less* likely to be interested in classifieds, and therefore cannot be merely due to lower demand for print classified ads.⁶ Evidence from both survey and browsing data suggests that the decline in newspaper readership is not compensated by increased news consumption online or through other sources (i.e., national papers, radio or TV), and is hence likely to result in an overall decline in political information among the public.⁷

Finally, we study how reduced news coverage of politics affected the behavior of local voters, with particular regard to electoral participation and ideological polarization. For electoral participation we find that the entry of CL has a negative though not robust effect on turnout in presidential elections and no effect in House elections. Regarding ideological polarization, we find robust evidence that the entry of CL favored the rise of ideologically extreme candidates, and reduced the probability that voters support different parties in concurrent elections (split-ticket voting). These findings are consistent with a greater tendency to rely on national partisan cues when less information about local politicians is available (Darr et al., 2018; Moskowitz, 2021; Trussler, 2020).

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 potentially detrimental effect for ideological polarization. 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 builds upon and expands prior evidence on the impact of CL on US newspapers. A study by Seamans and Zhu (2014) shows that, after the entry of CL, local newspapers experienced lower classified-ad rates and circulation, and higher display-ad rates and subscription prices. Our paper expands these findings by documenting the entire “chain of events” triggered by the entry of CL, and its profound implications for newspapers’ staff, editorial priorities, news content, and,

⁶ The readership decline is also unlikely to be due to newspapers charging higher subscription prices, as we find no significant effect of CL’s entry on that margin.

⁷ In particular, we find that the number of visits to popular national news websites by local users is not affected by CL’s entry.

ultimately, political outcomes.⁸

Second, our results add to previous evidence on the impact of the introduction of new advertising technologies, or new media, on incumbent media outlets. In a long-term historical perspective, Hamilton (2004) and Petrova (2011) argue that the growth of the print advertising market in the late 19th century was essential to the emergence of an independent (non-partisan) press. Our paper studies the flip side of this question: what happens to newspapers' political coverage when advertising profits are competed away by new technology. Regarding the impact of new media, Angelucci et al. (2020) argue that the introduction of commercial TV in the US had a significant negative economic impact on newspapers, leading to lower coverage of local politics and a decline in split-ticket voting between national and local elections. Bhuller et al. (2020) document large declines in circulation and shifts in editorial priorities for newspapers in Norway in response to the roll-out of broadband Internet. Our empirical setting is different in that it allows us to separate the effect of a specific shock to the advertising market from the demand-side changes brought about by new communication technology.

More broadly, our paper relates to previous work on the impact of the Internet on political participation, electoral outcomes, and public policies (Falck et al., 2014; Campante et al., 2018; Gavazza et al., 2019; Larcinese and Miner, 2018; Manacorda and Tesei, 2020). Exploiting variation in access to broadband or wireless technology, these studies gauge the net effect of the various changes brought about by the Internet on both the demand and supply side of the political market. In this respect, the novelty of our paper is to isolate the effect of the Internet – and specifically of digital advertising platforms – on political outcomes through its impact on legacy media.

Finally, our paper is generally related to the literature on the link between media, information, voter participation, and polarization. This includes the seminal studies by Besley and Burgess (2002) and Snyder and Strömberg (2010) which document that media coverage of local politicians make voters' more informed and participative and politicians more responsive to the demands of their constituents.⁹ Another set of studies (Lelkes et al., 2017; Allcott et al., 2020; Levy, 2021, e.g.,) examines the influence of media consumption on ideological polarization. Our paper contributes to this literature by demonstrating how news media support the functioning of elections as a selection mechanism. The magnitude of the

⁸ Other work exploits CL's expansion across the US or particular design features of the platform to study questions related to matching efficiency in labor and housing markets (Kroft and Pope, 2014), and the impact of online personal/ erotic ads on sexually transmitted diseases and violence against women (Cunningham et al., 2019; Chan and Ghose, 2014).

⁹ Also related to our work is a recent study by Gao et al. (2020) which documents the negative impact of the closure of local newspapers on local public finances. The authors interpret these findings as evidence of the importance of local newspapers in monitoring local government.

electoral penalty that extreme candidates face (Hall, 2015) 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 voters' 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 provides some background information about Craigslist and its expansion. Section 3 describes the data used in the analysis, while section 4 discusses the empirical strategy. Section 5 presents the results. Section 6 concludes.

2. BACKGROUND

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

Consistently ranked among the top 20 U.S. websites by traffic,¹⁰ the CL platform has a simple layout which has remained largely unchanged over time (see Figure A1). Ads on CL websites are organized into sections including housing, jobs, items for sale, professional services, and personals. CL websites only host classified ads and do not include any display ads or news content. Ads on CL are generally posted free of charge, with a few exceptions for brokered apartment rentals in New York and job posts for employers in some major cities.¹¹ CL's business model reflects the unconventional views of its founder, Craig Newmark, who prioritized providing a useful service to local communities over profit maximization, and always opposed listing the company on the stock market.¹²

Being cheaper and more efficient than traditional newspaper ads, CL became very popular among users, rapidly disrupting the lucrative market for classified ads that many local newspapers had relied upon. Aggregate data on newspaper revenues from the Newspaper Association of America suggest that classified revenues started declining in the early 2000s and the downward trend became steeper following CL's major expansion in the mid 2000s (Figure A2).

To investigate whether the link between CL's expansion and newspapers' performance

¹⁰ <https://www.similarweb.com/top-websites/united-states/>

¹¹ A full list of the exceptions as of 2010, the end of our sample period, is available at: https://web.archive.org/web/20100706030043/https://www.craigslist.org/about/help/posting_fees

¹² For a profile of Craig Newmark and his business strategy see <https://www.theguardian.com/technology/2006/feb/19/news.theobserver1>.

is causal, our analysis exploits variation in the timing of the introduction of CL across U.S. counties, combined with a newspaper-specific measure of ex-ante reliance on classified ads.

3. DATA

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

3.1. CRAIGSLIST’S EXPANSION

To construct a measure of the availability of CL in each county, we first collect information on the timing of the entry of each of CL’s current websites. For a subset of these, the information is directly available on the CL’s “about” webpage (<https://www.craigslist.org/about/expansion>). For the others, we assigned the date of the first snapshot recorded by the Internet Archive (<https://archive.org>). Figure 1 shows the evolution of the number of CL websites between 1995 and 2010, along with the evolution of a proxy for Internet quality - the average number of Internet Service providers by zipcode.¹³ Figure A3 depicts the geographic distribution of CL websites in 2000, 2005, and 2010 respectively.

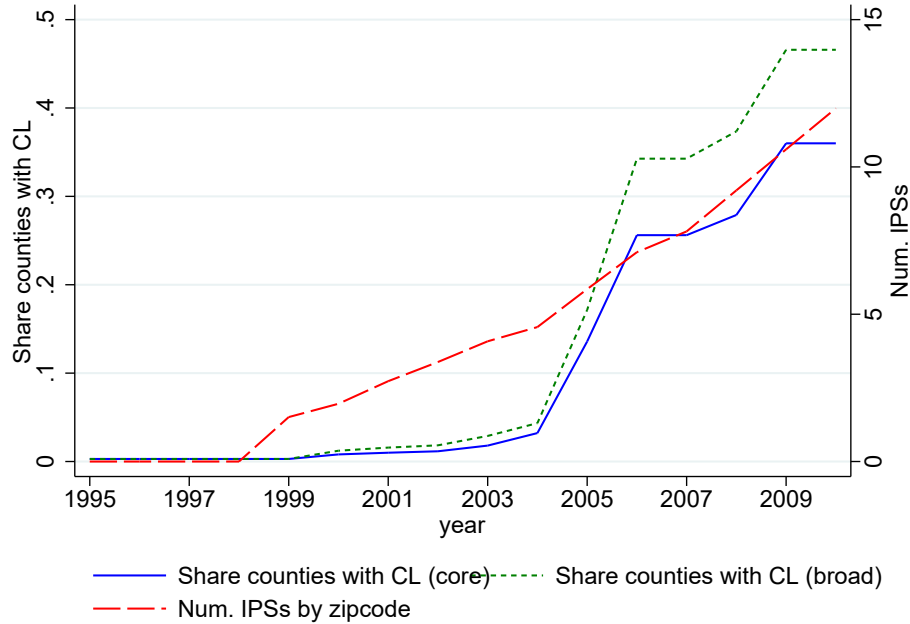
Mapping websites to counties is not straightforward since the area actually served by the website depends on user behavior. To address this issue, we identify the relevant market to a CL website in two ways. In our baseline approach, we assume that CL websites serve primarily the “core” county (or set of counties) containing the place indicated in the headline of the respective local website. The place is usually a single city or town, but can also be a combination of two or three nearby cities, a region, or, in some cases, an entire state. As an alternative, we identify all the counties that account for a non-negligible share of the ads posted on the website. We describe this “broad” definition of CL markets, which we use for purposes of robustness, in Appendix A.2.1.

3.2. NEWSPAPER CHARACTERISTICS AND OUTCOMES

We collect comprehensive data on a range of relevant newspaper characteristics and outcomes from a series of Newspaper Yearbooks published by Editor & Publisher for each year between 1995 and 2010. We accessed print copies the yearbooks and digitized the information contained in them using OCR software. The yearbooks contain detailed information for over 1,500 U.S. daily newspapers, including: address of the headquarters (HQ), circulation,

¹³ For more detail on this proxy, see section 3.5.

FIGURE 1: CRAIGSLIST’S ROLL-OUT OVER TIME



Share of US counties with a local CL website (left axis) and average number of Internet service providers by zipcode (right axis).

subscription prices, number of pages published, as well as the list of staff members with names, broad job categories, and job titles. Figure A4 shows how the information is reported in the yearbooks for two example newspapers.

We identify newspaper markets in two alternative ways. Our baseline approach is to assume that the newspaper market coincides with the county where the newspaper HQ is located. Since this a good approximation for the median newspaper,¹⁴ this approach is common in the literature (Gentzkow and Shapiro, 2010; Seamans and Zhu, 2014). We obtain the HQ county for each newspaper from the address reported in the E&P data. In Appendix A.2.2 we describe an alternative approach based on zipcode-level circulation data available for a subset of newspapers from the Alliance of Audited Media (AAM), which we use for purposes of robustness.

Whether and how much a newspaper is affected by the entry of CL depends on how heavily it initially relied on revenues from classified ads. To measure this baseline difference across papers, we use information on the presence in a newspaper’s staff of one or more classified ads managers prior to CL’s entry (i.e. between 1995 and 2000), available from

¹⁴ Our disaggregated subscription data confirms that the median paper in our sample has about 85% of its subscribers in the HQ county; see Appendix A.2.2.

E&P. To validate this measure, we collect information from <https://www.newspapers.com>, an online newspaper archive, on the share of pages devoted to classified ads in a subset of 262 newspapers. The results, shown in Appendix A.2.3, indicate that, prior to the entry of CL into their market, newspapers with classified ad managers devoted a significantly larger page share to classified ads than other papers, about 20% more than non-classified manager papers with similar circulation.

3.3. NEWSPAPER CONTENT

We are also interested in examining how changes in newspapers’ organization triggered by the entry of CL influenced editorial decisions and news content. To this end, we use information from NewsBank, a database containing the text and metadata of articles published in over 800 newspapers in our sample, beginning in 1999. We use the data in two ways. First, we extract a random sample of 2 million articles and apply a semi-supervised topic model to the text of the lead paragraph to assess changes in the coverage devoted to various topics. Second, 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 complementary measure of the prominence of political issues, both national and local. Details of the procedures used to construct these variables are reported in Appendix A.2.5.

3.4. POLITICAL OUTCOMES

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

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

Second, we examine a set of outcomes capturing the electoral performance of ideologically extreme candidates. We expect the quality of political information available to voters should have strong influence on their ability to distinguish candidates on the ideological dimension. Following Hall (2015) and Autor et al. (2020), we classify candidates based on their position in the distribution of campaign-finance-based ideological scores (CFScores) developed by Bonica (2014). Based on the 25th and 75th percentiles of the distribution of scores for all House candidates in 2000, we define thresholds that we use to separate “extremists” from “moderates.” Our outcomes of interest are the presence of an “extremist”

in the general election (related to the outcome studied by Hall (2015)), the shares of general election votes and individual contributions going to “extremist” candidates (as in Autor et al. (2020)), and the absolute value of the CFscore of the winning candidate. Data on contributions and CFscores are from Bonica’s Database on Ideology, Money in Politics, and Elections (DIME, 2016).

Third, we examine split-ticket voting, a measure of partisanship in voting that captures individual candidates’ ability to differentiate themselves from the national party brand. Following Darr et al. (2018), we measure split-ticket voting as the absolute value of the difference between the Republican candidate vote share in the presidential election and the Republican candidate vote share in House elections in the same county and year.¹⁵

3.5. ADDITIONAL DATA SOURCES

Browsing data. To measure CL take-up over time, as well as the downstream impact of CL on online news consumption, we draw on data from *Comscore*¹⁶. These data track the browsing behavior of a large sample of US Internet users, shown to be representative of US online buyers (De los Santos et al., 2012). The data cover the following years in our sample period: 2002, 2004, 2006 and 2007-2010. We aggregate the number of visits of the domain [craigslist.org](https://www.craigslist.org) as well as total visits recorded by Comscore by county and year. To capture online news consumption, we rely on Comscore’s classification of website categories available in the 2002 wave. We aggregate the number of visits by county for the 100 domains classified as news-related, and repeat this procedure for the following waves for the same set of domains.

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

Our second source is the Survey of the American Consumer conducted by GfK Mediamark Research & Intelligence (GfK-MRI). The survey includes an extensive battery of questions about media consumption. In particular, we use information on respondents’ self-reported readership of newspapers (in print and online, national and local), and news con-

¹⁵ Darr et al. (2018) use the Senate-President difference, but since only a third of Senate seats are contested in each election cycle, using House races expands the number of observations available. We measure vote shares for each office at the county level; House shares are computed by aggregating across all House votes cast by voters in the county (which for larger counties may be split across multiple representatives).

¹⁶ <https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/comscore/>

sumption on radio, TV, and on the Internet. We also exploit a question on what sections of the newspaper respondents usually read.¹⁷ We use these data for the period 1999-2010.

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

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

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

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

4. EMPIRICAL STRATEGY

4.1. DETERMINANTS OF CL ENTRY

To implement our empirical strategy, it is necessary to first understand what factors drove the timing of CL's staggered rollout. Anecdotal evidence suggests that CL prioritized larger markets and areas with good access to broadband Internet, which was crucial for the user to take full advantage of the platform. Importantly for our purposes, CL is not in the news business, and thus is not likely to have considered demand-side factors in the news market in determining where to enter. Also importantly, CL is privately held and has always operated

¹⁷ Figure A5 shows the fractions of GfK respondents reporting that they read each section.

as a mixture of profit-making business and community service. The fact that CL did not maximize profits¹⁸ allowed more flexibility for the idiosyncracies of its influential founder and its early user base to determine the timing of the rollout, rather than systematic factors in the underlying local news markets.

To validate the assumption of conditional exogeneity of CL’s timing of entry, in Figure 2 we examine the set of counties with headquartered newspapers in 2000 and plot the correlations of the year of CL’s entry into a given county with its cross-sectional characteristics. The first two coefficients confirm anecdotal accounts that population and the quality of the local Internet connection, as measured by number of ISPs, were important considerations in CL’s entry decisions. The magnitudes of both coefficients are sizable: a one-standard deviation higher log population (number of ISPs) is associated with CL entering a county about 15 months (16 months) earlier. Together, these two variables account for 37% of the variation in year of entry. Given this strong relationship, our analysis will control for log population and number of ISPs throughout.

The other coefficients are from separate, univariate regressions on the time of CL’s entry residualized by log population and number of ISPs. Conditional on these two variables, several demographic and political variables, including share urban population, racial composition, unemployment rate, age, electoral turnout and split-vote, appear unrelated to the timing of CL’s entry. There are three exceptions: consistent with the profile of CL’s early user base, counties with higher share of college educated population, higher rental share of housing and higher income per capita experience significantly earlier entry (3 to 5 months per standard deviation). We therefore also consider specifications controlling for each of the variables listed above.

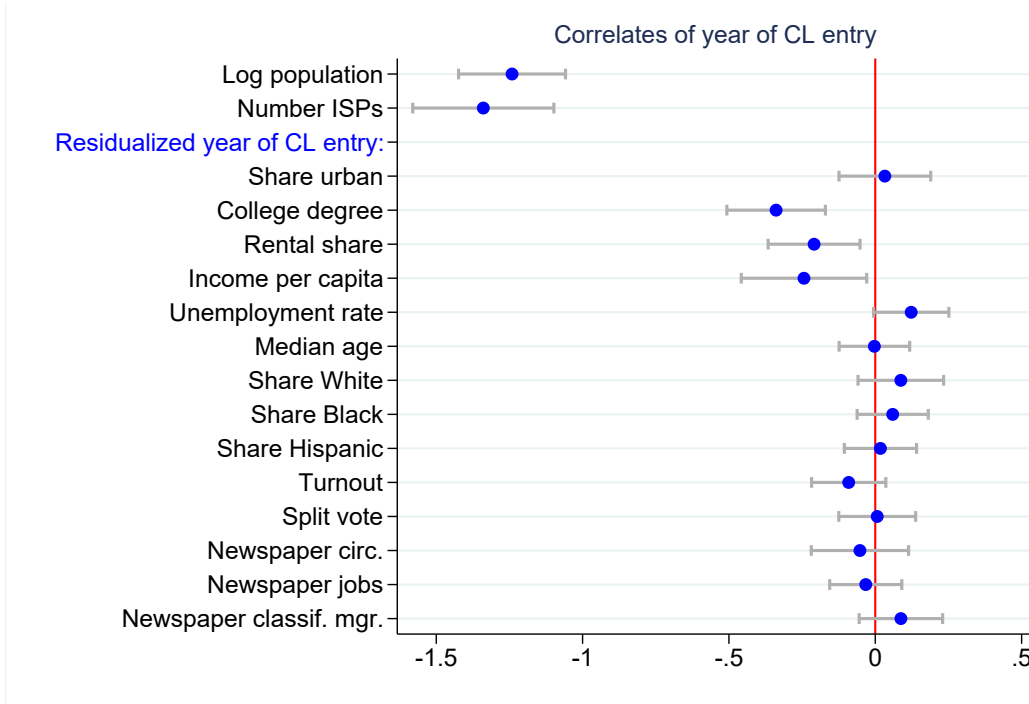
Finally, and crucially for the purpose of our analysis, we find no relationship between timing of entry and the state of local newspapers as measured by circulation, number of jobs or the presence of classified managers. In other words, we find no evidence of CL targeting particular newspaper characteristics in its entry decisions.

4.2. NEWSPAPER-LEVEL REGRESSIONS

To estimate the effect of CL entry on newspaper-level outcomes, we employ a difference-in-differences strategy exploiting CL’s staggered introduction across US counties, combined

¹⁸ The judge in a 2010 civil action against Newmark by eBay concluded that “Craigslisat 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.

FIGURE 2: CORRELATES OF CL ENTRY



Coefficients and 95% confidence intervals representing the univariate correlations of (residualized) year of CL entry with various county characteristics. All variables are measured in the year 2000 and standardized to facilitate the comparison of magnitudes. The sample consists of all counties with newspaper HQs in the year 2000.

with differences across newspapers in ex-ante reliance on classified advertising. The sample consists of the all newspapers covered by E&P, excluding ones with national circulation - i.e. the New York Times, USA Today and the Wall Street Journal.

The following equations summarize our approach:

$$Outcome_{nct} = \beta PostCL_{ct} + \phi_n + \psi_t + \rho' X_{ct} + v_t' Z_{c0} + \varepsilon_{nct}, \quad (1)$$

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

$Outcome_{nct}$ is one of the outcomes of interest for newspaper n , headquartered in county c , at time t . $PostCL_{ct}$ is a an indicator variable equal to one for years after the entry of CL in county c and zero otherwise. In the baseline analysis we follow the “core” definition of CL markets (based on location indicated in the website name) and newspaper markets (based on county of newspaper HQ). ϕ_n and ψ_t are newspaper and year fixed effects, respectively. The vector X_{ct} includes controls for log population and number of ISPs, and the vector Z_{c0} includes additional county-level controls from the 2000 census which we interact with year fixed effects. In alternative specifications we also control for state \times year fixed effects or for

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

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

4.3. COUNTY-LEVEL REGRESSIONS

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

$$Outcome_{ct} = \beta PostCL_{ct} + \phi_c + \psi_t + \rho' X_{ct} + v_t' Z_{c0} + \varepsilon_{ct}, \quad (3)$$

$$Outcome_{ct} = \beta PostCL_{ct} + \gamma PostCL_{ct} \times ClassifMgr_{c0} + \phi_c + \psi_t + \rho' X_{ct} + v_t' Z_{c0} + \varepsilon_{ct} \quad (4)$$

When estimating this specification we focus on the sample of all counties where at least one newspaper was located at baseline, and compute $ClassifMgr_{c0}$ as the circulation-weighted average across newspapers headquartered in county c .

4.4. CONGRESSIONAL DISTRICT-LEVEL REGRESSIONS

For outcomes measured at the level of congressional districts, we estimate equations 3 and 4 at the level of county \times district cells. In other words, we duplicate outcome observations and assign one duplicate to each county contained in the district. We weight observations by the share of the voting-age population of the district accounted for by the respective county, and cluster standard errors by district. To absorb variation due to changing congressional district boundaries, we include district by redistricting regime fixed effects in all regressions.¹⁹

¹⁹ 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 fixed district boundaries.

5. RESULTS

5.1. CRAIGSLIST TAKE-UP

We first confirm that Internet users in a given area were more likely to visit CL’s URL following the opening of a local website. To this end, we use the data on web browsing behavior from Comscore, described in section 3. In table 1 we estimate versions of our baseline specification (equation 1) using as dependent variable the IHS-transformed number of visits of the domain craigslist.org in a county in a given year. In all regressions we control for the total number of visits to any website recorded by Comscore. The results indicate that, after the entry of a local CL website, the number of visits to CL’s URL increase significantly, by between 16 and 40%. The result is robust to expanding the set of county controls (col. 2), to controlling for state \times year fixed effects (col. 3) or DMA \times year fixed effects (col. 4), restricting the comparison to different counties within the same media market.

To get a sense of the timing of CL’s take-up, in Figure 3 we plot the dynamic effects estimated following the method proposed by de Chaisemartin and D’Haultfoeuille (2020). The graph indicates no clear pre-trend, which alleviates concerns about the (conditional) exogeneity of CL’s entry. It also suggests that the effect was immediate but further intensified over the following years. This pattern is consistent with network effects in the local adoption of the website: a larger number of users (and hence, a higher volume of local ads) likely increases the value of the platform and attracts yet more users.

5.2. NUMBER OF NEWSPAPERS

Next we examine how the entry of CL in a given market affected the number of active newspapers. One possibility is that the drop in revenues due to competition from CL might have been so extreme as to force some local papers to close or merge. In Table 2 we estimate our baseline specifications using as dependent variable the number of newspapers that, according to the E&P data, were operating in a given county in a given year. In this case our unit of analysis is a county-year, and the sample includes all counties where at least one newspaper was headquartered in 2000 (i.e., before CL’s major roll-out). We find no evidence that CL entry affected the number of newspapers in a county. CL does not seem to affect the number of newspapers even in those counties where newspapers relied more heavily on classified ads *ex ante*; indeed, in columns 5 through 8, the coefficient on the interaction between *Post-CL*

TABLE 1: VISITS TO [CRAIGSLIST.ORG](https://craigslist.org)

	<i>Dependent variable: CL visits (IHS)</i>			
	(1)	(2)	(3)	(4)
Post-CL	0.440*** (0.129)	0.155* (0.084)	0.200*** (0.073)	0.163*** (0.060)
Total Comscore visits (IHS)	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes
Observations	19896	19896	19896	19501
Number of counties	3053	3053	3053	2995
R ²	0.80	0.83	0.84	0.85
Mean dependent variable	2.42	2.42	2.42	2.40

Regressions of number of visits of the domain craigslist.org by county and year (IHS-transformed) on an indicator for the availability of a local Craigslist website. All specifications control for total visits recorded in Comscore (IHS-transformed). County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

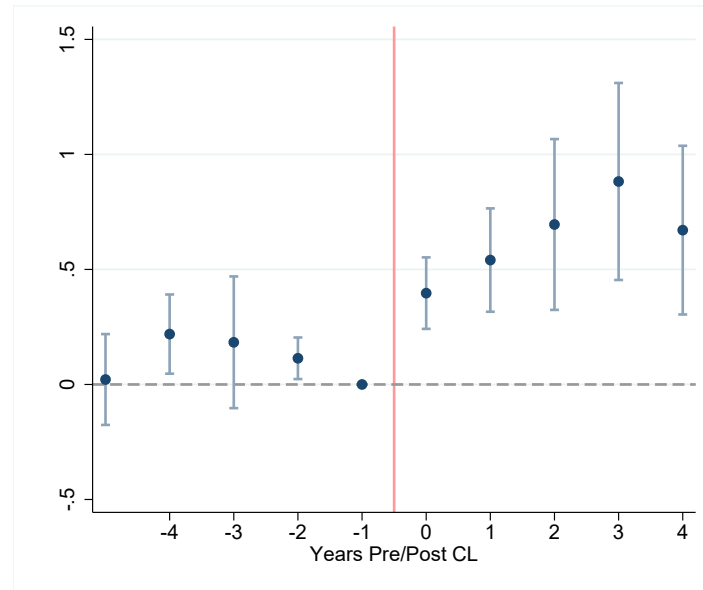
and *Classif. Manager* is small and never statistically significant. The same picture emerges when looking at the event-study graph depicted in Figure 4: all the coefficients for the years after CL entry are close to zero and statistically insignificant.²⁰ Taken together, these findings do not support the view that CL contributed to the disappearance of local newspapers. Yet, it is possible that lower advertising revenues affected the organization, functioning, and content of newspapers in other ways, a hypothesis which we explore below.

5.3. NEWSPAPER STAFF SIZE AND COMPOSITION

In this section we examine to what extent the entry of CL local websites influenced the organization of local newspapers. The first outcome we focus on is staff size. In Table 3 we estimate our baseline specifications using as dependent variable the number of jobs reported in the E&P yearbooks by newspaper and year. In the first four columns we examine the effect

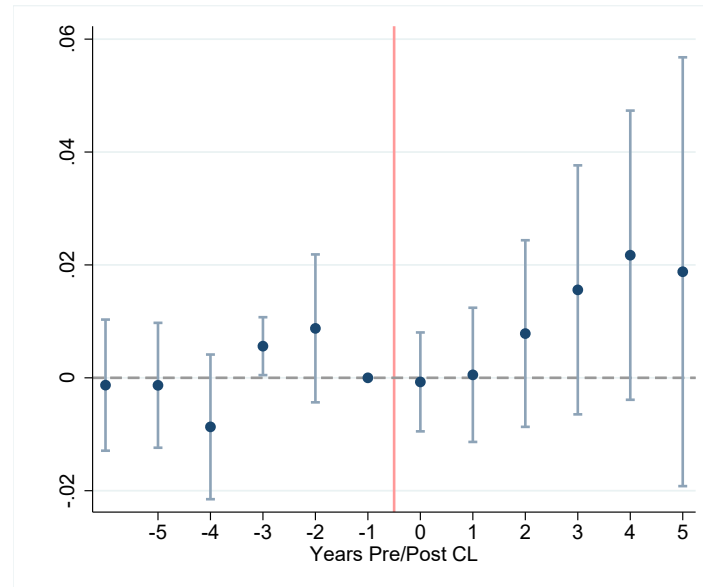
²⁰ Given the lack of evidence that CL affected newspapers' exit and entry, in the remainder of our analysis we use the full unbalanced panel as baseline. We obtain similar results for the balanced panel of newspapers that remained in the sample for the entire period of analysis.

FIGURE 3: VISITS TO [CRAIGSLIST.ORG](https://craigslist.org) – EVENT STUDY



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

FIGURE 4: NUMBER OF NEWSPAPERS – EVENT STUDY



Dynamic effect of the availability of a local CL website on the number of newspapers headquartered by county and year. Coefficients and 95% confidence intervals based on the time-corrected Wald estimator proposed in de Chaisemartin and D’Haultfoeuille (2020). Controls include log population and number of Internet service providers. Standard errors clustered by CL-area.

TABLE 2: NUMBER OF NEWSPAPERS

	<i>Dependent variable: Number of newspapers HQ-ed in county</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.012 (0.011)	0.017 (0.012)	0.009 (0.013)	0.010 (0.013)	0.027** (0.013)	0.020 (0.013)	0.014 (0.014)	0.008 (0.015)
Post-CL \times Classified Mgr.					-0.028 (0.019)	-0.007 (0.019)	-0.010 (0.019)	0.002 (0.020)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	19424	19424	19424	18672	19360	19360	19360	18608
Number of counties	1214	1214	1214	1167	1210	1210	1210	1163
R ²	0.94	0.95	0.95	0.95	0.94	0.95	0.95	0.95
Mean dependent variable	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18

Regressions of the number of newspapers headquartered by county and year on an indicator for the availability of a local Craigslist website and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

of being in the *Post-CL* period. Results in columns 1-4 indicate that following the entry of CL, the number of jobs in local papers decreased significantly. The effect amounts to 1 to 1.2 fewer jobs, or about 5% of the mean, and is robust to the inclusion of additional controls and finer fixed effects. The results in the last four columns confirm that the effect is mainly driven by newspapers that relied more on classified ads prior to CL, for which the effect, looking at the most demanding specification (column 8), is over 12% relative to the mean. Figure 5 displays the associated event-study, which shows the absence of any significant pre-trend, and that the effect, which is significant in the first year, increases considerably over several years.

These result is robust to several additional checks. First, we verify that the indicator for classified manager does not simply capture newspaper size. This may be a concern since in our data newspapers with larger circulation are more likely to have a dedicated classified manager (correlation = 0.26). In Table B1 we show that the results are robust to controlling for baseline circulation interacted with *Post-CL*, and that this interaction has no

significant effect on jobs-count. Second, in Table B2 we verify that the results are robust to defining CL markets based on the locations of ads posted on each website, and to defining newspaper markets based on geographically disaggregated circulation data. Third, in Table B3 re-estimate our regressions on a balanced sample of newspapers, excluding ones that have exited (or entered) the market during our sample period. This does not affect the results and confirms that downsizing, rather than exit, appears to be the major margin of response to the CL shock. Finally, we consider as dependent variable the number of (unique) employees instead of the number of job titles. These two variables can differ if, for example, financial difficulties push a newspaper to assign to the same person multiple jobs that were previously carried out by different people. Table B4 shows that we obtain very similar results.

One important question that our data allow us to examine is what categories of workers were most affected by staff cuts. In fact, based on the job title we can determine whether a worker holds a managerial or an editorial position, and, for editorial staff, we can identify the corresponding topical area (e.g., politics, sports, entertainment, etc.).²¹

In the first two columns of Table 4, we estimate our main specification with the *Post-CL* \times *Classif. Manager* interaction separately for managerial staff and editorial staff. The results indicate that, for the newspapers most affected by the entry of CL, staff cuts concerned both types of positions, with the effect on managerial positions being larger (i.e., 19% vs. 7% of the mean). In the following columns we examine what topic areas were most affected by cuts in editorial staff, looking in particular at politics, sports, and entertainment. The evidence indicates that newspapers most affected by CL were significantly less likely to have dedicated political editors and reporters after CL's entry, while the same was not the case for sports or entertainment. One interpretation of this result is that, when facing financial difficulties, newspapers affected by CL opted to cut staff especially in areas — like local politics — for which producing quality content is more costly.

On the other hand, despite this significant downsizing, we find no effect on the number of pages published by affected newspapers (Appendix Table B5). Together, these results imply increased workload per staff member and may have implications for the distribution of editorial priorities. We examine this issue further in the next section where we look at how the entry of CL affected the evolution of news content.

²¹ Specifically, we classify as managerial positions the jobs listed in the E&P section "Corporate/ General Management", and as editorial positions the jobs listed in the sections "News Executives" and "Editorial Management" (see figure A4). Regarding topic areas, we classify jobs containing the keyword "sports" as sports related, ones containing the keywords "entertainment"/ "lifestyle" / "film" / "music" as entertainment-related, and ones containing the keywords "politics" / "government" / "Washington" / "city" / "local" (excl. "local retail") as politics-related.

TABLE 3: NUMBER OF JOBS

	<i>Dependent variable: Newspaper number of jobs</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-2.030*** (0.398)	-1.084*** (0.390)	-1.242*** (0.418)	-1.262*** (0.404)	-0.103 (0.412)	0.453 (0.411)	0.274 (0.432)	0.163 (0.442)
Post-CL \times Classified Mgr.					-3.623*** (0.575)	-2.975*** (0.548)	-2.911*** (0.549)	-2.802*** (0.594)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22724	22724	22723	21968	22516	22516	22515	21762
Number of newspapers	1540	1540	1540	1492	1505	1505	1505	1459
R ²	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dependent variable	21.31	21.31	21.31	21.14	21.38	21.38	21.39	21.23

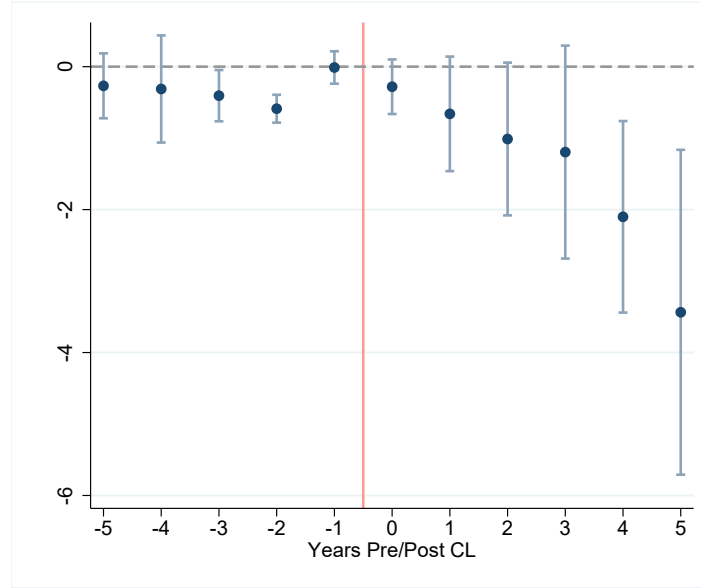
Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.4. NEWSPAPER CONTENT

Next we explore how the transformations in newspapers' organization documented above translate into changes in news content, in general, and news coverage of politics, in particular. This is a question of paramount importance considering that, for most citizens, local newspapers still represent the main source of political information which allows them to monitor and keep politicians accountable (Mahone et al., 2019; Snyder and Strömberg, 2010). To analyse changes in news content we apply text analysis techniques to large corpus of news articles published in a subset of the newspapers in our original sample. Our analysis is composed of three parts.

First, we study how competition from CL affected the volume of coverage devoted to different topics. To do so, we estimate a topic model on a corpus consisting of the first paragraphs of over 2 million randomly drawn articles published in over 800 newspapers between 2000 and 2010, available from NewsBank. We use the Correlation Explanation (CorEx)

FIGURE 5: NUMBER OF JOBS – EVENT STUDY



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

model of Gallagher et al. (2017a). This model has the advantage that it tends to produce coherent topics for corpora consisting of short texts, and also allows us to define anchor words to generate topics corresponding to specific areas. Specifically, we seed the model to produce topics associated with: i) President, ii) Congress, iii) local politics, and iv) foreign policy.²² Figure C1 reports the most frequent words associated with each of the ten topics generated by the algorithm, with the anchor words underlined. For each article in the corpus, we obtain a probability associated with each of the 10 topics, and aggregate the probability weights by newspaper and year. Figure 6 reports the coefficients on the interaction $Post-CL \times Classif. Mgr.$ obtained from regressions of the form specified in equation 4 with dependent variables corresponding to the probability of each topic.²³ The results indicate a general decline in news coverage of politics. For newspapers most affected by CL, the drop is significant for presidential, congressional, and foreign politics, and marginally insignificant for local politics. For the remaining un-anchored topics (i.e., accidents, events, traffic, obituaries, sports, crime) we find mixed results, with insignificant coefficients on the interaction $Post-CL \times Classif. Manager$.

To further explore changes in political coverage, we examine how frequently local

²² Additional details about the procedure are reported in Appendix A.2.5.

²³ Tables B7 and B8 report the corresponding full regression results.

TABLE 4: JOBS BY TYPE AND TOPIC

	(1) Num. Managers	(2) Num. Editors	(3) Editor Politics	(4) Editor Sports	(5) Editor Entertainment
Post-CL	0.050 (0.086)	0.138 (0.306)	0.022 (0.019)	-0.018 (0.018)	-0.010 (0.020)
Post-CL \times Classified Mgr.	-0.667*** (0.122)	-0.781* (0.444)	-0.050** (0.024)	-0.003 (0.021)	-0.028 (0.025)
Log population, share urban, # ISPs	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	Yes	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	21834	22156	21455	21558	21444
Number of newspapers	1503	1504	1504	1504	1504
R ²	0.80	0.89	0.69	0.51	0.75
Mean dependent variable	3.46	10.60	0.40	0.84	0.39

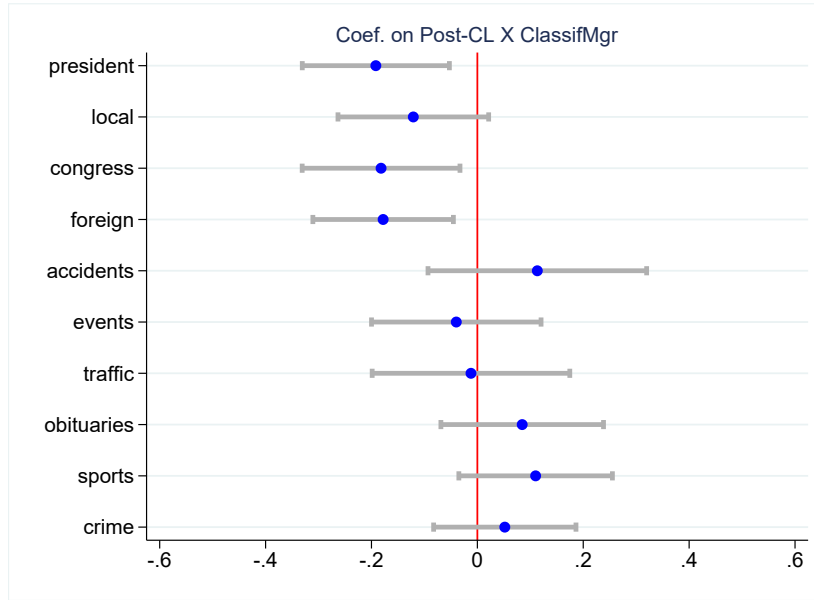
Regressions of number of jobs by type and indicators for the presence of an editor dedicated to a specific topic on an indicator the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

newspapers report about local representatives and national politicians. Using keyword searches in the NewsBank archive, we first identify all articles published in a given newspaper and year that contain the names of the House representative elected in the district corresponding to the newspaper HQ or senate representatives elected in the respective state.²⁴

Table 5 suggests that, following the entry of CL, coverage of local Congressional representatives declines significantly, particularly for newspapers that relied more heavily on classified ads. In that case the differential decline in representatives' coverage amounts to about 30% of the mean. In Table B9 we perform the same exercise for national party leaders: the President and the leadership of both parties in each chamber of Congress. We find no significant effect on coverage of these national politicians. A possible explanation for this pattern is that coverage of national politicians can be sourced from wire services, whereas coverage of the district's representatives is original content produced by in-house reporting staff. It is exactly these staff, per the results in Table 4, whose jobs were most likely to be cut following CL's entry.

²⁴ Details about the keyword searches are reported in Appendix A.2.5. We restrict the sample to the period after decennial redistricting (i.e. post-2003) to maintain stable district boundaries.

FIGURE 6: CONTENT: DISTRIBUTION OF TOPICS



Effect of CL-entry on the (standardized) probability that a randomly sampled article covers a given topic. The graph presents the coefficients and 95% confidence intervals on the interaction of an indicator for the availability of a local Craigslist website with an indicator for the presence of a classified manager at baseline. Controls include log population, number of Internet service providers, and county characteristics in 2000 interacted with time. Standard errors clustered by CL-area.

5.5. NEWSPAPER READERSHIP

How did readers react to the changes newspaper organization and content documented above? One way to tackle this question is by looking at the evolution of the total number of copies sold, i.e., newspaper circulation, available from E&P for all newspapers in our sample. In Table 6 we estimate our standard set of regressions using as dependent variable yearly circulation. The coefficient on Post-CL is negative although it loses significance when including state or DMA fixed effects interacted with year dummies. The effect is instead negative, large, and highly significant for the newspapers most affected by CL. The size of the coefficient is fairly stable across specifications and has a magnitude of about 9% of the sample mean when all county controls are included. The corresponding event study graph, reported in Figure 7, shows the absence of pre-existing trends and a steady increase in the effect within a few years of CL's entry.

We can verify this decline in circulation using self-reported newspaper readership, available from the NAES and GfK-MRI surveys. These data include several questions related to media consumption. While we have only limited information on readership of specific newspapers, we are able to differentiate between respondents who report most frequently reading a national newspaper, i.e. the *New York Times*, *USA Today* or the *Wall Street Jour-*

TABLE 5: MENTIONS OF LOCAL CONGRESSIONAL REPRESENTATIVES

	<i>Dependent variable: Articles mentioning local House/Senate representatives (IHS)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.001 (0.074)	-0.054 (0.079)	-0.028 (0.079)	0.041 (0.099)	0.142 (0.100)	0.097 (0.103)	0.078 (0.104)	0.208 (0.127)
Post-CL \times Classified Mgr.					-0.258** (0.121)	-0.292** (0.122)	-0.218* (0.125)	-0.317** (0.157)
Total articles (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	5453	5453	5426	5036	5410	5410	5383	4985
Number of newspapers	821	821	818	771	814	814	811	763
R ²	0.82	0.82	0.85	0.88	0.82	0.82	0.85	0.88
Mean dependent variable	2.88	2.88	2.87	2.86	2.89	2.89	2.88	2.87

Regressions of the (IHS-transformed) number of articles mentioning the name of a local House or Senate representative by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. All specifications control the (IHS-transformed) total number of articles recorded by Newsbank by newspaper and year. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

nal, and the rest. We rely on residents' county of residence to match them to locally headquartered newspapers, and re-define the variable *Classif. Mgr.* as the circulation-weighted average across newspapers based in that county.

In Table 7 we report the results for self-reported readership of non-national newspapers and its frequency from individual-level regressions. In all specifications we control for a range of respondent characteristics including age, race and education. The results based on NAES data — reported in columns 1-4 — suggest a significant decline readership in areas where newspapers were most exposed to CL-competition. The effect is quite sizeable, corresponding to about 7% of the sample mean for both outcomes. Results are qualitatively similar, though somewhat smaller in size ($\approx 4\%$), when estimated using GfK data. Since in both surveys respondents are asked about reading either the print or the online version of the newspaper, these results indicate that the decline in circulation documented in Table 6 does

TABLE 6: CIRCULATION

	<i>Dependent variable: Newspaper circulation in thousands</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-2.712*** (0.599)	-1.468** (0.645)	-1.194 (0.765)	-0.892 (0.939)	0.012 (0.752)	0.108 (0.874)	0.397 (0.953)	0.700 (1.231)
Post-CL \times Classified Mgr.					-5.148*** (1.297)	-2.991*** (1.096)	-2.896*** (1.055)	-2.921** (1.351)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22950	22950	22948	22195	22633	22633	22631	21880
Number of newspapers	1555	1555	1555	1507	1506	1506	1506	1460
R ²	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99
Mean dependent variable	33.34	33.34	33.34	32.89	33.40	33.40	33.39	32.93

Regressions of circulation by newspaper and year, measured in thousands of copies, on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

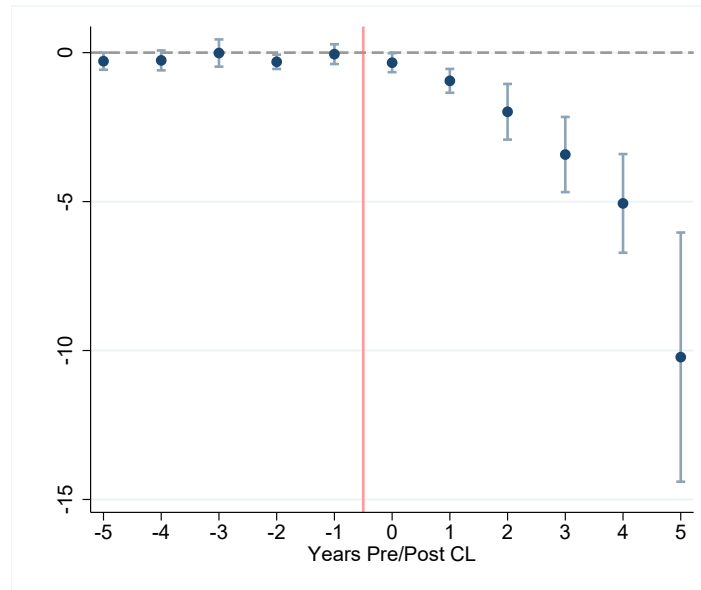
not merely reflect substitution of print editions with online editions.

The decline in readership documented above is consistent with at least two explanations. First, it is possible that newspapers respond to the shock to classified ad revenues by increasing their subscription prices, which in turn would lead to lower demand. We test for this mechanism in Appendix Table B6, looking at the impact of CL's entry on yearly subscription prices reported in the E&P yearbooks. With the exception of the last column, we obtain insignificant coefficients on both the indicator for CL-entry and its interaction with classified manager. There is hence no clear evidence of an increase in subscription prices.²⁵

A second explanation may be that readers respond to the changes in content brought about by CL's entry. One possibility is that the change toward less coverage of politics that we document in Table 5 and Figure 6 alienated readers interested in this type of content.

²⁵ Seamans and Zhu (2014) on the other hand find a significant increase in subscription prices in response to CL entry. Our analysis differs in using more complete data (i.e. covering all newspapers in the period 1995-2010) and a somewhat different empirical strategy, e.g. looking at reliance on classified ads at baseline rather than contemporaneously, to account for its likely endogeneity.

FIGURE 7: CIRCULATION – EVENT STUDY



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

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 1998-2001 waves of the GfK-MRI survey. The distribution of readers’ preferences, depicted in Figure A5, indicates that most readers (63%) report reading the “General News” section (which includes politics), with the Sports and Business sections also being popular (38% and 37%, respectively). The Classified section is not far behind, however, with 34% of respondents reporting Classifieds as one of the sections they frequently read. It is, therefore, possible that a reduction in the value of print classifieds might be a driver of circulation declines.

To understand what types of readers drive the drop in readership of local papers following the entry of CL, we examine heterogeneity in the readership effect by propensity to read the classified versus general news sections. Using the 1998-2001 waves of the GfK-MRI survey, we estimate an elastic-net penalized regression model to identify the individual

TABLE 7: SELF-REPORTED NEWSPAPER READERSHIP

	NAES				GfK-MRI	
	Read newspaper Dummy		Read newspaper Days per wk		Read newspaper Dummy	
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	-0.023 (0.017)	0.019 (0.015)	-0.108 (0.090)	0.118 (0.108)	-0.016*** (0.005)	-0.003 (0.008)
Post-CL \times Classified Mgr.		-0.053** (0.164)		-0.284* (0.027)		-0.018* (0.009)
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	108332	107503	108332	107503	253513	251442
Number of counties	1207	1203	1207	1203	792	790
R ²	0.06	0.06	0.14	0.14	0.09	0.09
Mean dependent variable	0.75	0.75	3.71	3.71	0.42	0.42

Regressions of self-reported newspaper readership on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

characteristics that are most predictive of reading the general news and the classifieds sections respectively.²⁶ Based on the model estimated in the 1998-2001 data, we then project two propensity scores for respondents in the later years. This procedure allows us to assign to each respondent in the post-CL surveys a probability for reading general news and one for reading classifieds. Projecting based on pre-CL data allows us to focus attention on differential changes among demographic types who would have been likely to read either classifieds or political news prior to CL entry, without the confound of the post-CL changes to newspapers' product. The projected propensity scores have fairly strong negative correlation (with correlation coefficient of -0.4), indicating that the groups that tend to read each section are relatively distinct.

²⁶ The characteristics most strongly associated with general news reading are white race, having a post-graduate degree, income in the 75K-150K range, being retired, being married 25 years or more, and being age 45-49. The characteristics most associated with classifieds reading are being unemployed, living in a small to moderate sized county, having a high school diploma only or "some college", being engaged (to be married), and being 25-29 years old.

We then re-estimate the individual-level readership regressions separately for two groups of respondents: i) those with above-median probability of reading classifieds and below-median probability of reading general news, and ii) those with below-median probability of reading classifieds and above-median probability of reading general news. The results, reported in Table 8, indicate that the decline in readership after the entry of CL is entirely driven by individuals with high news propensity and low classified propensity, and is again more pronounced where newspapers relied more heavily on classifieds ex ante. The effect is not significant for the high-classified-interest, low-news-interest group in any of the specifications.²⁷ Though newspapers which offered the most classifieds were most affected by the CL shock, the readers least interested in classifieds reduced their newspaper reading most.

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

TABLE 8: SELF-REPORTED NEWSPAPER READERSHIP: HETEROGENEITY

	(1) Low news propensity & High classif. propensity	(2) Low news propensity & High classif. propensity	(3) High news propensity & Low classif. propensity	(4) High news propensity & Low classif. propensity
Post-CL	-0.005 (0.007)	-0.004 (0.012)	-0.016** (0.008)	0.009 (0.012)
Post-CL \times Classified Mgr.		0.001 (0.013)		-0.033** (0.013)
Respondent controls	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	84341	83475	84338	83854
Number of counties	784	782	781	779
R ²	0.05	0.05	0.10	0.10
Mean dependent variable	0.32	0.32	0.54	0.54

Regressions of self-reported newspaper readership on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

²⁷ Alternatively, in appendix table B10 we explore the two dimensions of readership propensity separately. We find more pronounced declines in readership for respondents with high rather than low news propensity and for respondents with low rather than high classified propensity.

5.6. SUBSTITUTION TO OTHER NEWS SOURCES

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

Both NAES and the GfK survey include questions about news consumption via sources other than local newspapers, which we use to study substitution patterns. The results are reported in Tables B11, B12 and B13. In columns 1 and 2 we examine how CL affected the readership of national newspapers. In both surveys we find that the entry of CL in an area is associated with an increase in the likelihood that local respondents read national papers. Yet, the effect does not appear to be stronger in areas where newspapers relied more heavily on classified ads. The remaining columns show no significant effect on news consumption on TV, radio, and online, including in areas where newspapers were most affected by CL.²⁸

The data on browsing behavior available from Comscore allows us to perform an alternative test for substitution to online news sources. In Table B14 we examine the effect of CL's entry on the number of visits of 100 domains classified by Comscore as news-related. We detect no significant effect of CL's entry on visits of these domains.

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

5.7. POLITICAL OUTCOMES

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

²⁸ The analysis of online news consumption with NAES data is not possible since the question is not asked consistently across waves.

vote for ideologically extreme candidates (Hall, 2015; Autor et al., 2020), and iii) 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).

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

In Appendix Table B15 we examine the effect of CL on voter turnout. The results provide some evidence of a negative effect of CL on turnout in presidential elections, which is concentrated in counties where newspapers relied more on classified ads. However, the coefficient on $Post-CL \times Classif. Mgr$ becomes smaller when controlling for county-level covariates and, especially, when $State \times Year$ fixed effects are included. In addition, as shown in Appendix Table B16, we find no significant effect when looking at turnout in House elections, for which one might expect the decline in news coverage of local members of Congress to be especially relevant. We conclude that there is at most weak evidence that the changes in newspaper content and readership induced by the entry of CL reduced aggregate electoral participation.

Support for extreme candidates Though changes in the information environment may not impact *whether* people vote, they may affect *how* they vote. In Table 9 we examine whether the reduction in political news coverage and readership brought about by the entry of CL favored the emergence and success of ideologically extreme candidates, focusing on House elections. The hypothesis we test is that a coarser information environment, by making it harder for voters to acquire information about candidates’ ideological positioning, makes the entry of more extreme candidates more likely and improves their electoral prospects. Specifically, we look at the following outcomes: i) the probability that ideologically extreme candidates win a primary election, ii) the vote share they obtain in the general election, iii) the individual campaign contributions they attract, and iv) the ideological extremity²⁹ of the winning candidate.

Outcome variables i), iii) and iv) are defined at the electoral district level, while ii) varies at the county-by-district level. As explained in section 4, all regressions include dis-

²⁹ Measured by the absolute value of the difference between the candidate’s CFScore and the CFScore of the median House candidate in 2000.

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

In column (1) we look at the effect on the probability that a candidate with an extreme CFScore wins a primary election.³⁰ The results indicate that, after the entry of CL, and in districts where newspapers were most affected, this probability increases significantly by about 7 percentage points. In columns (2) and (3) we look at the effect on general election vote share and share of campaign contributions, respectively, won by extreme candidates. Again, we find a positive, significant, and sizeable effect of CL entry on popular support for extreme candidates along both dimensions. Finally, in column (4) we examine the effect on the extremity of the CFScore of the winning candidate relative to the median among House candidates in 2000. Again the effect is positive, significant, and concentrated in areas served by the most CL-affected papers.

Split-ticket voting Finally, in Table 10 we examine the impact of the entry of CL on split-ticket voting, i.e., voters’ tendency to support candidates from different parties in concurrent elections. We focus on presidential election years since split-ticket voting is defined by comparing candidates’ vote shares in presidential vs. legislative elections. The results indicate that, following the entry of CL, voters become significantly less likely to split their vote between candidates of different parties. As for the other outcomes, the effect is driven by areas where newspapers were most vulnerable to CL’s competition, where split-ticket voting drops by 16% of the sample mean. One interpretation of this finding is that, as local media provide less information about local candidates, voters tend to rely more heavily on partisan cues — which are shaped primarily by the national political debate — when deciding on down-ballot races. This substitution of local cues for national ones can lead to higher party alignment between different races, i.e. the “nationalization” of local elections (Moskowitz, 2021; Trussler, 2020), and further incite ideological divisions. This result relates to similar findings by Darr et al. (2018) regarding the impact of newspaper closures on polarization. Our results indicate that closures are not a necessary condition, and that the impoverishment of local newspapers, and the associated cuts in news coverage of local politics, can produce similar consequences.

Taken together, our findings suggest that the profound transformations in the media landscape triggered by the entry of CL had a tangible impact on electoral politics in the US,

³⁰ The variable takes value of 1 if either the Republican or the Democratic candidate who advances to the general election qualifies as an “extremist” according to our definition.

TABLE 9: IDEOLOGICAL POLARIZATION

	Extremist in General	Vote Share of Extremists	Indiv. Contrib. Share of Extremists	Winner CFScore Dev. from 2000 Median
	(1)	(2)	(3)	(4)
Post-CL	-0.005 (0.030)	-0.037* (0.021)	-0.011 (0.024)	0.002 (0.018)
Post-CL \times Classif. Mgr.	0.067** (0.030)	0.055** (0.024)	0.049* (0.025)	0.046** (0.023)
Log pop., pct. urban, #ISPs	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
District FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	12,797	12,797	12,534	11,048
R ²	0.58	0.64	0.74	0.86
Mean dependent variable	0.74	0.38	0.51	0.72

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

and contributed to increase ideological polarization and further divisions across party lines.

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. Counter-acting this unhappy equilibrium to some degree are reporters’ professional norms, which value the production of “hard news” and investigative journalism over cheaper-to-produce and sometimes more popular “soft news.”

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

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

	<i>Dependent variable: House-President Rep. vote share differential</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	-0.011 (0.008)	-0.013* (0.007)	-0.013* (0.007)	0.004 (0.010)	0.004 (0.009)	0.006 (0.010)
Post-CL \times Classified Mgr.				-0.020* (0.011)	-0.022** (0.011)	-0.022* (0.011)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	No	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4807	4807	4807	4791	4791	4791
Number of counties	1214	1214	1214	1210	1210	1210
R ²	0.49	0.51	0.62	0.49	0.51	0.62
Mean dependent variable	0.13	0.13	0.13	0.13	0.13	0.13

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

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

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

The Craigslist effect is not simply a consequence of changes to the demand for news induced by internet availability; rather, it appears to operate by reducing newspapers’ ability to invest in local reporting resources. Papers that were especially reliant on classified advertising in the pre-Craigslist period saw much larger changes on these dimensions than

comparably internet-exposed but less classified-dependent papers. The loss of advertising revenues at these papers seems to have particularly reduced political coverage and especially coverage of local representatives, an area with large positive externalities but also large private costs for newspaper operators.

Consistent with existing work on media effects on political outcomes, we find that there were measurable social consequences of this change in the production of news content. Voters in areas served by papers affected most by the Craigslist shock saw their House 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 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.

Our results have implications for our understanding of the link between advertising market structure and the market for news. They highlight the fragility of compensating the production of a public good — politically relevant information — with proceeds from bundled advertising. Technological innovation that unbundles the two products, as Craigslist did for classified advertising, can have spillover effects on the news market, with significant and lasting consequences for the quality of representation and political polarization.

REFERENCES

- ALLCOTT, H., L. BRAGHIERI, S. EICHMEYER, AND M. GENTZKOW (2020): “The welfare effects of social media,” *American Economic Review*, 110, 629–76.
- ANGELUCCI, C., J. CAGE, , AND M. SINKINSON (2020): “Media Competition and News Diets,” *NBER Working Paper No.26782*.
- AUTOR, D., D. DORN, G. HANSON, AND K. MAJLESI (2020): “Importing political polarization? The electoral consequences of rising trade exposure,” *American Economic Review*, 110, 3139–83.
- BESLEY, T. AND R. BURGESS (2002): “The Political Economy of Government Responsiveness: Evidence from India,” *Quarterly Journal of Economics*, 117, 1415–1451.
- BHULLER, M., T. HAVNES, J. MCCAULEY, AND M. MOGSTAD (2020): “How the Internet Changed the Market for Print Media,” .
- BONICA, A. (2014): “Mapping the Ideological Marketplace,” *American Journal of Political Science*, 58, 367–386.
- (2016): “Database on Ideology, Money in Politics, and Elections,” Stanford, CA: Stanford University Libraries. <https://data.stanford.edu/dime>.
- CAMPANTE, F., R. DURANTE, AND F. SOBBRIO (2018): “Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation,” *Journal of the European Economic Association*, 16, 1094–1136.
- CHAN, J. AND A. GHOSE (2014): “Internet’s dirty secret: assessing the impact of online intermediaries on HIV transmission,” *MIS Quarterly*, 38.
- CUNNINGHAM, S., G. DEANGELO, AND J. TRIPP (2019): “Craigslist’s Effect on Violence Against Women,” .
- DARR, J. P., M. P. HITT, AND J. L. DUNAWAY (2018): “Newspaper closures polarize voting behavior,” *Journal of Communication*, 68, 1007–1028.
- DE CHAISEMARTIN, C. AND X. D’HAULTFOEUILLE (2020): “Two-way fixed effects estimators with heterogeneous treatment effects,” *American Economic Review*, forthcoming.
- DE LOS SANTOS, B., A. HORTAÇSU, AND M. R. WILDENBEEST (2012): “Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior,” *American Economic Review*, 102, 2955–80.
- FALCK, O., R. GOLD, AND S. HEBLICH (2014): “E-lelections: Voting Behavior and the Internet,” *American Economic Review*, 117, 2238–2265.
- FCC (2016): “Federal Communications Commission Report FCC 16-107,” Tech. rep.

- 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.
- GAO, P., C. LEE, AND D. MURPHY (2020): “Financing Dies in Darkness? The Impact of Newspaper Closures on Public Finance,” *Journal of Financial Economics*, 135, 445–467.
- GAVAZZA, A., M. NARDOTTO, AND T. VALLETTI (2019): “Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies,” *Review of Economic Studies*, 86, 2092–2135.
- GENTZKOW, M. AND J. M. SHAPIRO (2010): “What Drives Media Slant? Evidence from U.S. Daily Newspapers,” *Econometrica*, 78, 35–71.
- GENTZKOW, M., J. M. SHAPIRO, AND M. SINKINSON (2011): “The effect of newspaper entry and exit on electoral politics,” *American Economic Review*, 101(7).
- HALL, A. B. (2015): “What happens when extremists win primaries?” *American Political Science Review*, 109, 18–42.
- HAMILTON, J. (2004): *All the news that’s fit to sell: How the market transforms information into news*, Princeton University Press.
- HOPKINS, D. J. (2018): *The increasingly United States: How and why American political behavior nationalized*, University of Chicago Press.
- KROFT, K. AND D. G. POPE (2014): “Does online search crowd out traditional search? Evidence from Craigslist,” *Journal Labor Economics*, 32(2), 259–303.
- LARCINESE, V. AND L. MINER (2018): “Was Obama Elected by the Internet? Broadband Diffusion and Voters’ Behavior in US Presidential Elections,” .
- LELKES, Y., G. SOOD, AND S. IYENGAR (2017): “The hostile audience: The effect of access to broadband internet on partisan affect,” *American Journal of Political Science*, 61, 5–20, publisher: Wiley Online Library.
- LEVY, R. (2021): “Social media, news consumption, and polarization: Evidence from a field experiment,” *American economic review*, 111, 831–70.
- MAHONE, J., Q. WANG, P. NAPOLI, M. WEBER, AND K. MCCULLOUGH (2019): “Who’s Producing Local Journalism? Assessing Journalistic Output Across Different Outlet Types,” Tech. rep.
- MANACORDA, M. AND A. TESEI (2020): “Liberation Technology: Mobile Phones and Political Mobilization in Africa,” *Econometrica*, 88, 533–567.

- MCCHESNEY, R. W. AND J. NICHOLS (2011): *The death and life of American journalism: The media revolution that will begin the world again*, Bold Type Books.
- MOSKOWITZ, D. J. (2021): “Local news, information, and the nationalization of US Elections,” *American Political Science Review*, 115, 114–129.
- PETERSON, E. (2021): “Paper Cuts: How Reporting Resources Affect Political News Coverage,” *American Journal of Political Science*, 65, 443–459.
- PETROVA, M. (2011): “Newspapers and parties: How advertising revenues created an independent press,” *American Political Science Review*, 790–808.
- SEAMANS, R. AND F. ZHU (2014): “Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers,” *Management Science*, 60(2), 265–540.
- SNYDER, J. M. AND D. STRÖMBERG (2010): “Press coverage and political accountability,” *Journal of Political Economy*, 118(2), 355–408.
- STARKMAN, D. (2014): *The watchdog that didn’t bark: The financial crisis and the disappearance of investigative journalism*, Columbia University Press.
- TRUSSLER, M. (2020): “Get Information or Get in Formation: The Effects of High-Information Environments on Legislative Elections,” *British Journal of Political Science*, 1–21.

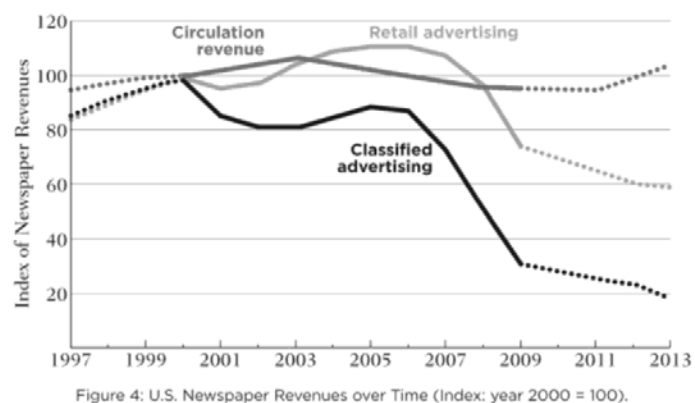
A. BACKGROUND AND DATA

A.1. BACKGROUND

FIGURE A1: CRAIGSLIST: LAYOUT IN 2000 AND 2016

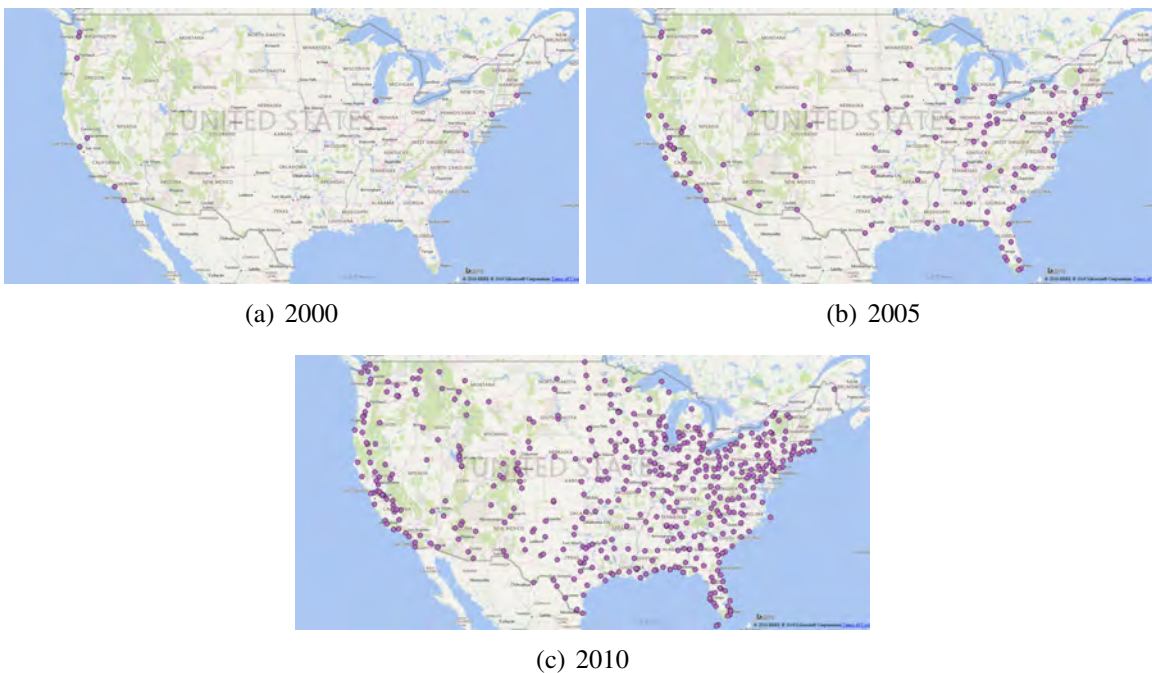


FIGURE A2: EVOLUTION OF NEWSPAPER REVENUES BY SOURCE



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

FIGURE A3: GEOGRAPHIC DISTRIBUTION OF CL WEBSITES (2000-2010)



Distribution of local CL websites in the years 2000, 2005 and 2010.

A.2. DATA

FIGURE A4: EXTRACT FROM THE EDITOR AND PUBLISHER YEARBOOKS

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 Twisten
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, Lansdale 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 Hoyer
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 Solomon

EDITORIAL MGMT.

Books Ed. Elizabeth Taylor
Editorial Page Ed. Bruce Dold
Entertainment Ed. Scott Powers
Foreign Ed. Colin McMahon

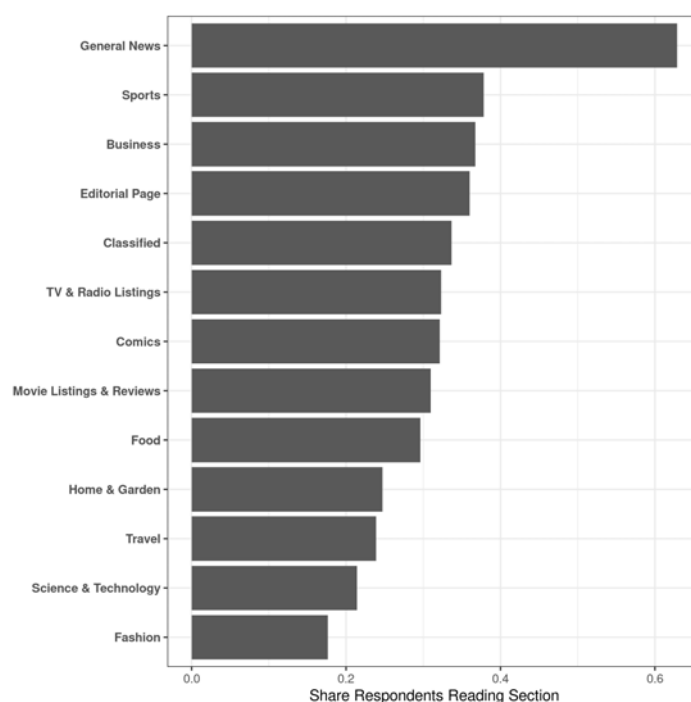
Good Eating Ed. Carol Haddix
Nat'l Ed. Storer Rowley
Special Sections Ed. Janet Franz

Sports Ed. Bill Ade
Sunday Magazine Ed. Elizabeth Taylor
Tempo Ed. Tim Bannon

Travel Ed. Randy Curwen
Womanews Ed. Cassandra West

The Chicago Tribune

FIGURE A5: NEWSPAPER SECTIONS BY READERSHIP (1998-2001)



Source: GfK-MRI Survey

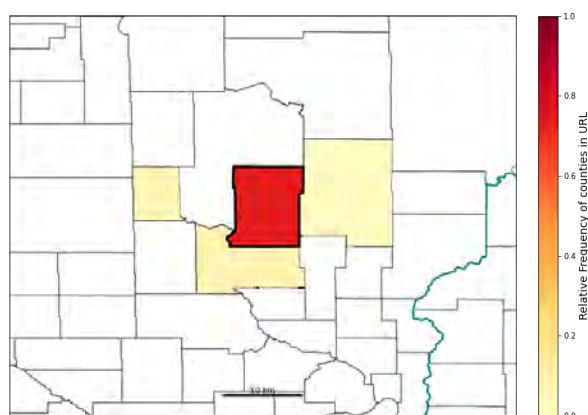
A.2.1. CRAIGSLIST MARKETS

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

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

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

FIGURE A6: DISTRIBUTION OF ADS POSTED ON [HTTPS://BRAINERD.CRAIGSLIST.ORG/](https://brainerd.craigslist.org/)



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

A.2.2. NEWSPAPER MARKETS

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

This method consists in matching a newspaper to all counties where it is read in proportion to circulation. Zip-code-level circulation data are available from the Alliance for Audited Media (AAM) for 2002, which we use to construct a weighted measure of CL availability by newspaper-year. The weights in this “broad” measure of CL availability are the fraction of the paper’s subscribers in each county.³¹ 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.³²

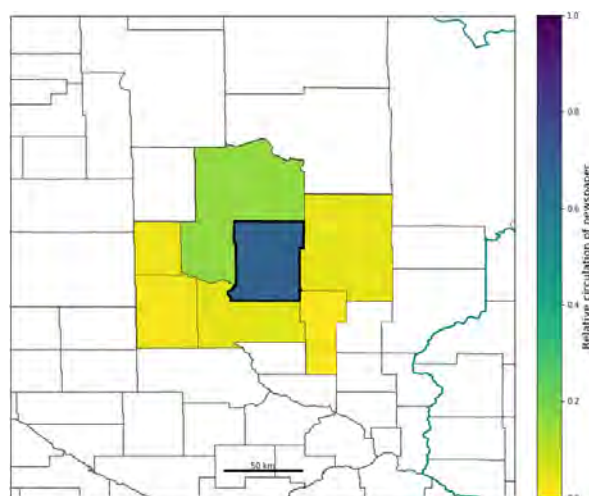
Figure A7 shows the geographic distribution of circulation for a newspaper in our sample, the Brainerd Dispatch, with darker colors representing higher values. Crow Wing County where the newspaper’s HQ is located, shown in blue, accounts for 81% of the total.³³ The median paper in our AAM data has about 85% of its total circulation in the headquarters county once we exclude outliers.

³¹ 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.

³² 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.

³³ Similarly to CL ads, we exclude outlier counties that account for less than 5% of total circulation.

FIGURE A7: DISTRIBUTION OF CIRCULATION OF THE BRAINERD DISPATCH



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

A.2.3. VALIDATING THE CLASSIFIED MANAGER PROXY

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

We measure classified intensity as the number of pages on which the term “Classified” appears, divided by the total number of pages in the issue. We collected this measure for a total of 43,165 issues across the 262 papers available in the Newspapers.com archive. Prior to the entry of Craigslist in a market, the median issue in our sample had 8 pages of classified advertising, or 15% of the issue’s total page count.

We first examine cross-sectional variation in classified intensity prior to Craigslist entry. Table A1 shows the results of regressions where the outcome is the fraction of pages in a newspaper issue that contain classified advertising. The left two columns use all days with weekday fixed effects to control for cyclical variation in classified intensity, while the last two use Sundays only (traditionally the biggest day for classified advertising). Columns (2) and (4) additionally add controls for the newspaper’s circulation and the population of the county in which it operates, to test whether the classified manager dummy merely picks up larger papers.

The table shows that the fraction of pages devoted to classified ads was on average

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

	Classified Page Share		Classified Page Share (Sun.)	
	(1)	(2)	(3)	(4)
Classif. Mgr.	0.029*	0.023	0.034*	0.027
	(0.015)	(0.015)	(0.018)	(0.017)
Pop / Circ controls	No	Yes	No	Yes
Year FEs	Yes	Yes	Yes	Yes
Day-of-week FEs	Yes	Yes	No	No
Observations	31,273	31,273	24,368	24,368
R ²	0.01	0.03	0.01	0.02
Mean dependent variable	0.20	0.20	0.21	0.21
Number of newspapers	262	262	220	220

Regressions of the fraction, by newspaper issue, of pages devoted to classified advertising on an indicator for the presence of a classified manager in 2000, in the pre-CL period. Data from issues published in all years prior to the year of CL entry in the newspaper's market are included. OLS regressions in all columns. Standard errors clustered by newspaper.

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

between 2 and 4 percentage points higher at papers that had a classified manager in 2000, prior to Craigslist entry. Coefficient magnitudes are comparable but fall below conventional significance thresholds when population and circulation controls are added.

TABLE A2: CRAIGSLIST EFFECT ON SHARE OF PAGES DEVOTED TO CLASSIFIED ADS.

	Classified Page Share							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.0010	-0.0008	0.012	0.009	0.018	0.020	0.039**	0.029
	(0.008)	(0.010)	(0.012)	(0.023)	(0.013)	(0.013)	(0.017)	(0.025)
Post-CL × Classif. Mgr.					-0.028*	-0.034**	-0.045**	-0.037
					(0.016)	(0.016)	(0.018)	(0.030)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 Demographics × Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	No	No	Yes	Yes	No	No
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	3,619	3,619	3,619	3,619	3,619	3,619	3,619	3,619
Number of newspapers	262	262	262	262	262	262	262	262
R ²	0.60	0.62	0.71	0.80	0.60	0.62	0.71	0.80
Mean dependent variable	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16

Regressions of the average fraction, by newspaper and year, of pages devoted to classified advertising on an indicator for the availability of a local Craigslist website and interaction with an indicator for the presence of a classified manager in 2000. OLS regressions in all columns. Standard errors clustered by newspaper.

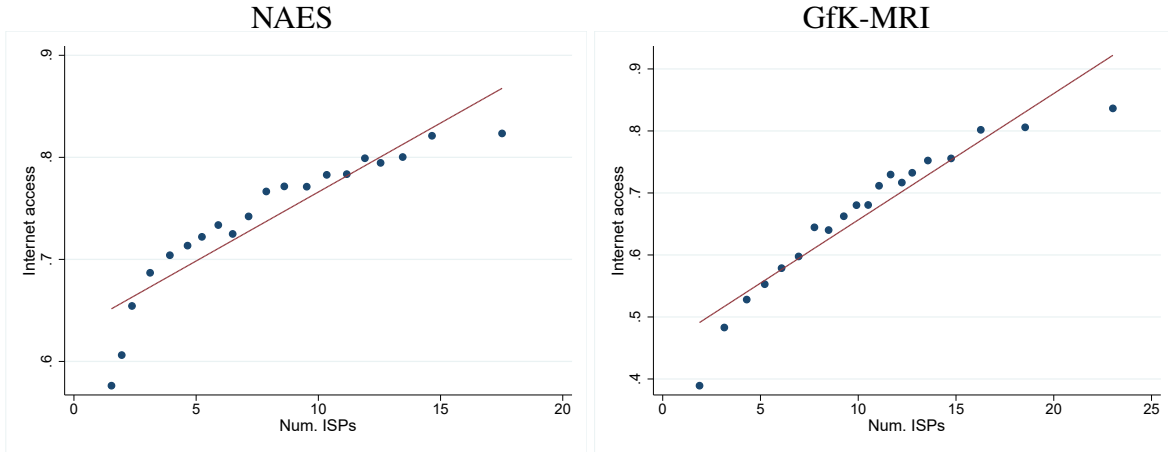
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2 instead uses within-newspaper variation over time related to Craigslist entry. This table shows that papers with classified managers saw their share of pages devoted to classifieds fall substantially after the entry of Craigslist. Papers without dedicated classified managers were unaffected or may even have seen an increase in the fraction of pages devoted to classifieds.

A.2.4. VALIDATING NUMBER OF ISPS AS A MEASURE OF INTERNET QUALITY

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

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



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

A.2.5. DETAILS ON NEWSPAPER CONTENT PROCESSING

This section contains details on the procedures used to process raw text content from the newspapers in our sample to construct lower-dimensional representations of the content. Source data are from the *NewsBank* database. We conduct two main kinds of processing on text data: keyword searches and topic modeling. Keyword searches use the full database

containing more than 100M full-text articles, while our topic model uses a smaller random sample of about 2M articles consisting of all articles published on 10 randomly sampled dates in each newspaper-year between 1999 and 2010. The topic-modeling sample limits to the first paragraph of text, plus the headline.

Politician Names Our first set of keyword searches look for the names of representatives in the US House and Senate. We use a list of representatives from <https://github.com/unitedstates/congress-legislators>, and for each member construct a (case-insensitive) regular expression of the form "(congres.*|rep.*) FIRSTNAME LASTNAME" or "senat.* FIRSTNAME LASTNAME". This expression matches strings like “Rep. Adam Smith” or “Congressman Adam Smith” but not “Adam Smith” alone. We require the inclusion of the title to cut down on false positives, as many members of Congress have common names. This does introduce the possibility of false negatives, but we have found that articles covering a member usually include the title and full name at first mention before switching to a shorter form like “Mr. Smith”.

We count the number of *articles*³⁴ in which the pattern described above appears on each newspaper-day, and then aggregate to the level of congressman by newspaper by year.

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

We use the semi-supervised method rather than the more traditional unsupervised Latent Dirichlet Allocation (LDA) because it allows us to focus on specific topics of interest. We are interested in separating various dimensions of political news coverage: coverage related to local, congressional, national and foreign politics. We seed separate anchors for these 4 topics, and run the CorEx model with 10 topics in total. Figure C1 presents the resulting topics, as described by their most representative words, with our anchor-words highlighted in bold. The 10 resulting topics can be labeled as follows: local politics, congressional politics, national politics, foreign politics, entertainment, health / family, weather, crime, obituaries.

For each of the 2 million articles in the corpus, the CorEx model outputs a set of 10 unconditional probabilities for the article belonging to that a topic. Importantly, these probabilities do not necessarily sum to 1 - an article can simultaneously belong to more than one

³⁴ I.e., each article that mentions the congressperson at all counts as 1, regardless of how many times the congressman is referenced in the article.

FIGURE C1: TOP TEN WORDS ASSOCIATED WITH EACH COREX TOPIC. ANCHOR WORDS ARE UNDERLINED.

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

topic, or to none. To examine the effects of CL's entry, we aggregate the distribution of probabilities by newspaper and year, and estimate the standard diff-in-diff equations specified in section 4, with the average probability for each one of the 10 topics as dependent variable.

B. ADDITIONAL RESULTS

TABLE B1: NUMBER OF JOBS: ROBUSTNESS TO CONTROLLING FOR NEWSPAPER SIZE

	<i>Dependent variable: Newspaper number of jobs</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.049* (0.595)	-0.523 (0.558)	-0.753 (0.596)	-0.893 (0.735)	0.303 (0.482)	0.667 (0.484)	0.451 (0.508)	0.278 (0.599)
Post-CL \times Circ. 2000	-0.016 (0.013)	-0.011 (0.014)	-0.009 (0.014)	-0.007 (0.015)	-0.012 (0.013)	-0.007 (0.014)	-0.006 (0.015)	-0.004 (0.016)
Post-CL \times Classified Mgr.					-3.006*** (0.668)	-2.657*** (0.594)	-2.667*** (0.613)	-2.623*** (0.698)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22516	22516	22515	21762	22499	22499	22498	21745
Number of newspapers	1506	1506	1506	1460	1504	1504	1504	1458
R ²	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dependent variable	21.38	21.38	21.38	21.22	21.39	21.39	21.39	21.24

Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B2: NUMBER OF JOBS: BROAD CL AND NEWSPAPER MARKETS

	<i>Dependent variable: Newspaper number of jobs</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL (broad)	-1.059*** (0.341)	-0.550* (0.283)	-0.697** (0.303)	-0.890*** (0.333)	0.600 (0.365)	0.695** (0.296)	0.520* (0.309)	0.307 (0.345)
Post-CL (broad) \times Classified Mgr.					-3.482*** (0.448)	-2.670*** (0.418)	-2.657*** (0.418)	-2.788*** (0.462)
County controls #1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County controls #2	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22688	22688	22687	21932	22480	22480	22479	21726
Number of newspapers	1534	1534	1534	1486	1499	1499	1499	1453
R ²	0.90	0.91	0.91	0.92	0.90	0.91	0.91	0.92
Mean dependent variable	21.30	21.30	21.30	21.14	21.38	21.38	21.38	21.22

Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Here we define the indicator for CL availability as 1 if a broad-CL market intersects the broad newspaper market (see Appendix sections A.2.1 and A.2.2). County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B3: NUMBER OF JOBS: BALANCED PANEL

	<i>Dependent variable: Newspaper number of jobs</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.992*** (0.414)	-0.991** (0.400)	-1.166*** (0.428)	-1.210*** (0.420)	-0.083 (0.426)	0.493 (0.424)	0.287 (0.445)	0.160 (0.458)
Post-CL × Classified Mgr.					-3.587*** (0.584)	-2.887*** (0.547)	-2.786*** (0.542)	-2.685*** (0.582)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	21148	21148	21147	20381	21138	21138	21137	20371
Number of newspapers	1331	1331	1331	1283	1330	1330	1330	1282
R ²	0.90	0.91	0.91	0.91	0.90	0.91	0.91	0.92
Mean dependent variable	21.72	21.72	21.72	21.55	21.73	21.73	21.73	21.56

Regressions of number of jobs by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. Sample: balanced panel of newspapers in operation throughout 1995 to 2010. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B4: NUMBER OF STAFF

	<i>Dependent variable: Newspaper number of employees</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-1.701*** (0.339)	-1.012*** (0.329)	-1.107*** (0.344)	-1.086*** (0.338)	-0.112 (0.341)	0.260 (0.339)	0.131 (0.345)	0.070 (0.349)
Post-CL \times Classified Mgr.					-3.010*** (0.513)	-2.490*** (0.480)	-2.405*** (0.477)	-2.299*** (0.516)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	22797	22797	22796	22044	22583	22583	22582	21832
Number of newspapers	1540	1540	1540	1492	1505	1505	1505	1459
R ²	0.92	0.92	0.92	0.93	0.92	0.92	0.92	0.93
Mean dependent variable	17.84	17.84	17.84	17.68	17.90	17.90	17.90	17.74

Regressions of number of employees by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B5: NUMBER OF PAGES

	<i>Dependent variable: Number of pages published</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.293 (0.191)	0.250* (0.139)	0.125 (0.171)	0.259* (0.151)	0.140 (0.128)	0.100 (0.126)	-0.026 (0.162)	0.136 (0.213)
Post-CL \times Classified Mgr.					0.286 (0.316)	0.292 (0.286)	0.289 (0.294)	0.238 (0.347)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	21637	21637	21637	20921	21509	21509	21509	20793
Number of newspapers	1471	1471	1471	1426	1450	1450	1450	1406
R ²	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.98
Mean dependent variable	28.50	28.50	28.50	28.31	28.52	28.52	28.52	28.34

Regressions of number of pages published by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B6: SUBSCRIPTION PRICE

	<i>Dependent variable: Yearly subscription price</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.273 (0.896)	-0.872 (0.939)	-0.984 (1.033)	-1.103 (1.259)	-1.372 (1.075)	-2.021* (1.114)	-2.021 (1.352)	-2.781* (1.490)
Post-CL × Classified Mgr.					1.966 (1.384)	2.142 (1.425)	1.911 (1.519)	3.403** (1.628)
Log population, #ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. × year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	No	Yes	No
DMA × Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	19643	19643	19642	18934	19424	19424	19423	18720
Number of newspapers	1431	1431	1431	1387	1393	1393	1393	1349
R ²	0.93	0.94	0.94	0.95	0.93	0.93	0.94	0.94
Mean dependent variable	118.66	118.66	118.67	118.36	118.80	118.80	118.80	118.48

Regressions of yearly subscription price by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B7: TOPIC MODEL: PROBABILITY OF POLITICAL TOPICS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	president	local	congress	foreign	president	local	congress	foreign
Post-CL	-0.010** (0.004)	-0.010* (0.006)	-0.006** (0.003)	-0.001 (0.003)	-0.001 (0.005)	-0.003 (0.007)	-0.000 (0.003)	0.005 (0.004)
Post-CL × Classified Mgr.					-0.017*** (0.006)	-0.013 (0.008)	-0.011** (0.005)	-0.011** (0.004)
Full county controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7178	7178	7178	7178	7129	7129	7129	7129
Number of newspapers	869	869	869	869	862	862	862	862
R ²	0.52	0.47	0.41	0.54	0.52	0.47	0.40	0.53
Mean dependent variable	0.21	0.31	0.10	0.10	0.21	0.31	0.10	0.10

Regressions of the average probability of an article covering a particular topic by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B8: TOPIC MODEL: PROBABILITY OF OTHER TOPICS

	(1) accidents	(2) events	(3) traffic	(4) obituaries	(5) sports	(6) crime	(7) accidents	(8) events	(9) traffic	(10) obituaries	(11) sports	(12) crime
Post-CL	-0.000 (0.003)	-0.009* (0.004)	-0.005* (0.003)	0.008 (0.007)	0.014** (0.006)	-0.000 (0.003)	-0.003 (0.004)	-0.007 (0.006)	-0.004 (0.004)	0.002 (0.010)	0.007 (0.008)	-0.002 (0.003)
Post-CL \times Classified Mgr.							0.006 (0.006)	-0.003 (0.007)	-0.001 (0.006)	0.013 (0.012)	0.012 (0.008)	0.003 (0.004)
County controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7178	7178	7178	7178	7178	7178	7129	7129	7129	7129	7129	7129
Number of newspapers	869	869	869	869	869	869	862	862	862	862	862	862
R ²	0.44	0.41	0.39	0.56	0.44	0.44	0.43	0.41	0.39	0.56	0.44	0.44
Mean dependent variable	0.12	0.17	0.14	0.15	0.21	0.11	0.12	0.17	0.14	0.15	0.21	0.11

Regressions of topic probabilities aggregated by newspaper and year on an indicator for the availability of a local Craigslist website and its interaction with an indicator for the presence of a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B9: MENTIONS OF NATIONAL PARTY LEADERS

	<i>Dependent variable: Articles mentioning the president and party leaders (IHS)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	0.143* (0.074)	0.066 (0.076)	0.078 (0.087)	0.002 (0.115)	0.113 (0.103)	0.080 (0.100)	0.085 (0.108)	0.015 (0.134)
Post-CL \times Classified Mgr.					0.063 (0.129)	-0.037 (0.130)	-0.030 (0.127)	-0.041 (0.171)
Total articles (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	6716	6716	6680	6097	6667	6667	6631	6043
Number of newspapers	825	825	822	775	818	818	815	767
R ²	0.83	0.84	0.86	0.88	0.83	0.84	0.86	0.88
Mean dependent variable	1.79	1.79	1.79	1.84	1.79	1.79	1.79	1.84

Regressions of the (IHS-transformed) number of articles mentioning the names of a set of national party leaders by newspaper and year on an indicator for the availability of a local Craigslist website, and its interaction with an indicator for the presence of a classified manager at baseline. All regressions control for the (IHS-transformed) total number of articles recorded by Newsbank by newspaper and year. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B10: SELF-REPORTED NEWSPAPER READERSHIP: HETEROGENEITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	News propensity				Classified propensity			
	Low		High		Low		High	
Post-CL	-0.011* (0.007)	-0.010 (0.010)	-0.017*** (0.007)	0.007 (0.010)	-0.023*** (0.007)	-0.005 (0.010)	-0.010 (0.007)	-0.000 (0.010)
Post-CL \times Classified Mgr.		-0.002 (0.012)		-0.032*** (0.011)		-0.024** (0.012)		-0.013 (0.012)
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	126747	125255	126752	126173	126747	125637	126756	125795
Number of counties	786	784	790	788	787	785	792	790
R ²	0.05	0.05	0.09	0.09	0.11	0.11	0.07	0.07
Mean dependent variable	0.32	0.32	0.52	0.52	0.47	0.47	0.38	0.38

Regressions of self-reported newspaper readership on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. The sample is split by the median respondent-level propensity to read the news section / classified section. Respondent controls include age bins, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B11: SELF-REPORTED MEDIA CONSUMPTION: GfK

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Newspaper, national		TV news		Radio news		Online news	
Post-CL	0.009** (0.004)	0.007 (0.005)	-0.006 (0.007)	0.004 (0.009)	0.005 (0.006)	0.002 (0.007)	0.003 (0.004)	0.005 (0.007)
Post-CL \times Classified Mgr.		0.003 (0.005)		-0.014 (0.009)		0.004 (0.008)		-0.004 (0.007)
Respondent controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	253513	251442	253513	251442	253513	251442	253513	251442
Number of counties	792	790	792	790	792	790	792	790
R ²	0.14	0.14	0.08	0.08	0.09	0.09	0.16	0.16
Mean dependent variable	0.09	0.09	0.70	0.70	0.16	0.16	0.22	0.22

Regressions of self-reported media consumption on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B12: SELF-REPORTED MEDIA CONSUMPTION: NAES

	(1)	(2)	(3)	(4)	(5)	(6)
	Newspaper, national		TV news		Radio news	
Post-CL	0.030** (0.015)	-0.000 (0.013)	0.002 (0.004)	0.008 (0.009)	0.004 (0.006)	0.000 (0.008)
Post-CL \times Classified Mgr.		0.037 (0.024)		-0.009 (0.009)		0.005 (0.008)
Respondent characteristics	Yes	Yes	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	105989	105173	107778	106981	149111	148083
Number of counties	1207	1203	1207	1203	1208	1204
R ²	0.21	0.20	0.04	0.04	0.05	0.05
Mean dependent variable	0.03	0.03	0.92	0.92	0.37	0.37

Regressions of self-reported media consumption on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B13: SELF-REPORTED MEDIA CONSUMPTION: NAES, CONTINUED

	(1) Newspaper, national days per wk	(2) Newspaper, national days per wk	(3) TV news days per wk	(4) TV news days per wk	(5) Radio news days per wk	(6) Radio news days per wk
Post-CL	0.153** (0.073)	0.004 (0.069)	0.043 (0.038)	0.094 (0.074)	0.024 (0.026)	0.003 (0.039)
Post-CL \times Classified Mgr.		0.186 (0.117)		-0.068 (0.089)		0.030 (0.038)
Respondent characteristics	Yes	Yes	Yes	Yes	Yes	Yes
County controls	Yes	Yes	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	107856	107054	107939	107138	149111	148083
Number of counties	1207	1203	1207	1203	1208	1204
R ²	0.09	0.09	0.10	0.10	0.05	0.05
Mean dependent variable	0.15	0.15	5.04	5.04	1.51	1.51

Regressions of self-reported media consumption on an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Respondent controls include sex, age, an indicator for college degree and race. County controls include contemporaneous log population and number of Internet service providers, as well as share urban population, pct. college educated, pct. rental, median age, share white/black/hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share - all measured in 2000 and interacted with year FEs. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B14: VISITS OF NEWS-RELATED WEB DOMAINS: COMSCORE

	<i>Dependent variable: Visits of news-related web domains (IHS)</i>							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post-CL	-0.021 (0.028)	-0.026 (0.028)	0.001 (0.032)	0.037 (0.035)	-0.035 (0.038)	-0.026 (0.039)	0.010 (0.042)	0.066 (0.047)
Post-CL \times Classified Mgr.					0.032 (0.045)	0.008 (0.047)	-0.010 (0.048)	-0.053 (0.051)
Total Comscore visits (IHS)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	Yes	No	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	No	Yes	No
DMA \times Year FEs	No	No	No	Yes	No	No	No	Yes
Observations	8303	8303	8303	7975	8280	8280	8280	7952
Number of counties	1214	1214	1214	1167	1210	1210	1210	1163
R ²	0.88	0.88	0.89	0.90	0.88	0.88	0.89	0.90
Mean dependent variable	7.89	7.89	7.89	7.89	7.89	7.89	7.89	7.89

Regressions of number of visits of news-related web domains (IHS-transformed) by county and year an indicator for the availability of a local Craigslist website in the county of the respondent, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. OLS regressions in all columns. Standard errors clustered by CL-area. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B15: TURNOUT IN PRESIDENTIAL ELECTIONS

	<i>Dependent variable: Turnout in Presidential elections</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	-0.016*** (0.005)	-0.000 (0.003)	-0.001 (0.002)	-0.003 (0.005)	0.007** (0.003)	0.002 (0.002)
Post-CL \times Classified Mgr.				-0.016** (0.007)	-0.010** (0.004)	-0.004 (0.003)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	No	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4855	4855	4855	4839	4839	4839
Number of counties	1214	1214	1214	1210	1210	1210
R ²	0.97	0.99	0.99	0.97	0.99	0.99
Mean dependent variable	0.48	0.48	0.48	0.48	0.48	0.48

Regressions of midterm turnout county and election year on an indicator for the availability of a local Craigslist website, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Turnout is defined as the ratio of votes cast to Census-estimated voting-age population in the county-year. Years included are 1998, 2002, 2006, and 2010. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. Observations are weighted by voting-age population. OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE B16: TURNOUT IN HOUSE ELECTIONS

	<i>Dependent variable: Turnout in House elections</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-CL	-0.019*** (0.007)	-0.006 (0.006)	0.000 (0.005)	-0.024*** (0.008)	-0.012* (0.007)	-0.006 (0.005)
Post-CL \times Classified Mgr.				0.008 (0.008)	0.008 (0.006)	0.008* (0.005)
Log population, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
2000 county char. \times Year FEs	No	Yes	Yes	No	Yes	Yes
State \times Year FEs	No	No	Yes	No	No	Yes
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4856	4856	4856	4840	4840	4840
Number of counties	1214	1214	1214	1210	1210	1210
R ²	0.79	0.83	0.90	0.78	0.82	0.90
Mean dependent variable	0.37	0.37	0.37	0.37	0.37	0.37

Regressions of midterm turnout county and election year on an indicator for the availability of a local Craigslist website, and its interaction with the circulation-weighted share of newspapers with a classified manager at baseline. Turnout is defined as the ratio of votes cast to Census-estimated voting-age population in the county-year. Years included are 1998, 2002, 2006, and 2010. County characteristics for the year 2000 include pct. college educated, pct. rental, median age, share white/ black/ hispanic, income per capita, unemployment rate, presidential turnout and Republican vote share. Observations are weighted by voting-age population. OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.