Online Competition and News Quality:

Evidence from the Introduction of Craigslist

Milena Djourelova* Ruben Durante† Gregory Martin‡

June 2020

Preliminary and incomplete.

Please click here for the latest version

Abstract

We investigate the effects of online competition on the market outcomes, organization and editorial choices of local newspapers. Our identification is based on the staggered introduction of Craigslist – the largest online platform for classified advertising – across US media markets between 1995 and 2009. This setting allows us to separate the effects of the platform from those of local Internet access. We document that Craigslist's entry caused a reduction in the volume of classified ads, led to a decline in readership and to downsizing of newsrooms and management. These effects are driven by newspapers that relied on classified ads as a source of revenue at baseline, proxied by the presence of a classified ads manager. Looking at content, affected newspapers reduce their political coverage and publish fewer articles covering local Congressional representatives, while we find no change in the volume of content devoted to non-political topics. Finally, we investigate how these changes in coverage affect political participation and the behavior of elected officials.

*UPF, BGSE, IPEG. Email: milena.djourelova@upf.edu
†ICREA, UPF, BGSE, IPEG, and CEPR. E-mail: ruben.durante@upf.edu
‡Stanford GSB. E-mail: gjmartin@stanford.edu

1 Motivation

The Internet has profoundly changed the landscape in which traditional media outlets operate. Due to increased competition from online platforms, newspapers' advertising revenues have sharply declined, forcing many outlets to cut staff and drastically rethink their business model and organization (McChesney and Nichols 2011).[1] These changes are likely to negatively influence the quality of reporting (Starkman 2014), which is particularly alarming given the importance of local newspapers for political participation and government accountability (Gentzkow et al. 2011; Snyder and Strömberg 2010). Despite the potentially grave consequences of these transformations for the future of journalism, rigorous evidence of the impact of online competition on newspapers - and on the quality of information voters are ultimately exposed to - is scant. This question poses an empirical challenge because it requires disentangling the effects of online platforms from the effects of changes in Internet access and quality, which are likely to directly affect the demand for newspapers.

To circumvent this difficulty, our analysis exploits the staggered introduction of a specific online platform, Craigslist - the world's largest online platform for classified advertising - across US media markets between 1995 and 2009. This staggered design allows us to hold local Internet quality constant, and make comparisons between newspapers' outcomes before and after the entry of Craigslist, comparing newspapers more or less exposed to Craigslist competition based on their prior reliance on classified advertising. We examine the effects of the shock to newspapers' revenues from classified advertising on three sets of outcomes: (i) newspapers' market outcomes and internal organization, (ii) news content, and (iii) readers' political awareness and political accountability.

[1] Figure 1 shows the timing and magnitude of this decline, and illustrates that a most of it is due to declines in revenues from *classified* advertising.

2 Data

Our analysis combines data on (1) Craigslist's expansion across the US; (2) Characteristics, market outcomes and content of daily newspapers; (3) Survey data on media consumption; (4) Data on readers' political engagement and congressmen behavior (in progress).

2.1 Craigslist

Expansion timeline. We obtain the timeline of CL's expansion from 2 sources. First, the entry dates for a part of the local websites are listed in CL's "expansion" page (`https://www.craigslist.org/about/expansion`). For the rest of local websites that exist today and are not listed individually in this page, we use the Internet Archive (`https://archive.org//`) to and collect the date of the first snapshot recorded by the archive.

In our baseline regressions, we assume that CL websites serve primarily the county (or 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 2 to 3 nearby cities, or a larger geographic area such as a state.

CL markets. To get a more precise measure of the local reach of CL markets, we also collect the exact locations indicated in ads posted on each local website. Specifically, we recover all archived snapshorts of each location-speicifc url saved by `https://archive.org//` for the first and second year after CL's entry into that location. We then collect the first pages (i.e. first 100 results) for the "housing", "jobs" and "sales" categories. We extract the content of the "location" field of each ad, and match the resulting strings to a comprehensive list of towns and cities, or in case the string contains the word "county", to a comprehensive list of county names. To limit potential false positives, we restrict the search to towns/ cities/ counties in the state corresponding to the

place mentioned in the website's headline or any neighboring state.

Figure 5 illustrates the geographic distribution of ads for the case of `http://brainerd.craigslist.org`. In this case, we find ads in 5 counties, but about 80% of them come from the county in the center of the figure containing the town of Brainerd. Hence, we consider the county (or counties) containing the place mentioned in the CL headline to be a good approximation of Craigslist markets. In robustness checks, we use the set of counties with share of ads above 10%, or alternatively, all counties contained in a radius of 30 / 60 miles from the boundaries of the central county (or counties).

2.2 Newspapers' Outcomes

Editor and Publisher yearbooks. Our main source of data on US daily newspapers is the series of *Editor & Publisher Yearbooks* for the years 1995 to 2010, which we digitized using OCR software. The books contain detailed information for the universe of US daily newspapers, including: headquarters address, circulation, and a list of staff with broad job categories, job titles and names. Figure 4 shows an example of how this data is presented in the books. Using the address field, we assign each newspaper to the county where it is headquartered. This serves as our baseline definition of a newspaper market.

Geographic distribution of circulation. We obtain data on newspapers' circulation disaggregated by zipcode from the Alliance for Audited Media. This data is available for about 300 newspapers. Using the distribution of circulation for the year 2002, we define an alternative measure of newspaper markets as the set of all counties with positive circulation for a given newspapers. For newspapers missing from the AAM data (mostly smaller ones), we keep the definition of county of headquarters as the main newspaper market.

Figure 6 presents an example of the geographic scope of circulation for the case of the *Brainerd Dispatch*. This newspaper circulates in 6 counties in total, but about 80% of its circulation comes from the HQ-county in the center of the figure. With the exception of the biggest national newspapers, this is the typical pattern found in the data. Therefore, we take the HQ-county to be a good approximation of the core newspaper market.

Classified Ads. For a fraction of newspapers we are able to get data on the volume of classified ad pages from *Newspapers.com*. This website makes archived entire copies of newspapers available and allows us to conduct a keyword search for "classified", and extract the respective number of matching pages in each copy. This data covers about 300 newspapers from our sample.

Content. Our data on newspapers' content is from *Newsbank* – an online archive that covers news articles for about 800 newspapers from our sample. We use the archive in 2 ways. First, we obtain the frequency of articles mentioning specific keywords – e.g. the name of a Congressional representative – by newspaper and year. Second, in order to examine the broader coverage of different topics, we construct a corpus consisting of all articles published on 10 randomly sampled dates in each year between 1995 and 2010. This results in about 2 million articles. We extract their first paragraph, headline and date of publication.

2.3 Survey Data

We obtain supplementary data on media consumption from the Annenberg Electoral survey. We use the questions on whether the respondent has consumed a particular type of media in the previous week, and if yes, how often. Three Annenberg waves were conducted during our sample period: in 2000 , 2004 and 2008.

2.4 County Characteristics

Number of Internet Service Providers. To proxy the local quality of broadband Internet we collect data on the number of Internet service providers by zip-code from the FCC. The data is available for the period 1998-2008. We assign years before 1998 zero ISPs for all zip-codes, and use linear interpolation to fill missing data for years after 2008. Finally, we aggregate the data to the county level by taking the population-weighted average across zip-codes within the county. While the number of providers is an imperfect proxy for Internet access, previous research (Lelkes et al. 2015; Larcinese and Miner 2018) has shown that it is highly correlated with number of subscribers.

Other controls. Finally, we collect data on population from the ICHS, data on share urban population, income per capita, share college educated, rental share, from the 2000 Census, unemployment rate from BLS, and turnout and republican vote share in the 2000 presidential election from the David Leip election atlas.

3 Empirical Strategy

Newspaper Outcomes. To estimate the effect of Craigslist's entry on newspapers' outcomes, we employ a Diff-in-Diff strategy exploiting its staggered introduction. Specifically, we estimate equations of the form:

$$Outcome_{nct} = \alpha + \beta_1 CL_{ct} + \phi_n + \psi_t + \epsilon_{nct}, \tag{1}$$

Where $Outcome_{nct}$ denotes the outcome of interest for newspaper n, headquartered in county c, at time t, CL_{ct} is a an indicator equal to one after Craigslist's entry into county c, ϕ_n and ψ_t are newspaper and year fixed effects respectively. We cluster standard errors at the level of CL

assignment, which is county or counties in cases of websites that target larger regions.

Additionally, we consider a specification interacting the entry of CL with an indicator for whether the newspaper had a classified manager at baseline – $ClassifMgr_n$.

$$Outcome_{nct} = \alpha + \beta CL_{ct} + \gamma CL_{ct} \times ClassifMgr_n + \phi_n + \psi_t + \epsilon_{nt} \quad (2)$$

To verify the identifying assumption of parallel trends, and to understand the timing of CL's effects, we consider the dynamic versions of these specifications, replacing the indicator for post-CL with a set of indicators for year pre-/ post-CL. In light of recent work showing that the standard diff-in-diff estimation can be biased in settings with treatment effects that are heterogeneous across groups or over time, we also present event-studies based on the time-corrected Wald (Wald-TC) estimator proposed by (Chaisemartin and D'Haultfoeuille 2020).

County level outcomes. To examine county-level outcomes, e.g. survey responses matched to the respondent's county of residence, we aggregate $ClassifMgr$ to the county level weighted by circulation for the case of counties with multiple HQs. We then use the same specifications as above, replacing newspaper FEs with county FEs.

4 Results

Correlates of CL-entry We start off by examining which county characteristics correlate with the timing of CL's entry. In the left hand side panel of figure 7 we plot the coefficients from a regression of year of CL entry on log population, share urban population and number of Internet service providers – the likely main determinants of early entry based on anecdotal evidence. The coefficients suggest that one standard deviation larger population/ number of ISPs are associated

with 0.8/ 0.6 years earlier CL entry respectively, while we find no significant correlation with share urban.

In the right hand side panel we test for any correlations between the residual variation in year of entry, after accounting for log population, share urban and ISPs, with other county characteristics. With the exception of the share black population we find no statistically significant correlations. Crucially, the residualized time of CL entry is uncorrelated with newspapers' characteristics and market outcomes, such as circulation, staff size or the presence of a classified manager.

Volume of classified ads. The premise for studying CL's entry as a shock to online competition is that it significantly affected newspapers' classified advertising revenues. In table 1 we show that this premise is supported by the data on classified ads: we find that for newspapers with classified manager at baseline, the number of classified-ad pages per copy declines by about 4 pages after CL's entry.

Furthermore, this data allows us to verify that that the indicator for newspapers having a classified manager in their staff is a good proxy for reliance on revenues from classified ads. Figure 8 shows that the volume of classified ads in newspapers with a classified manager significantly exceeds that in newspapers without a classified manager. The difference is most significant in Sunday editions, which usually feature more pages devoted to ads.

Newspapers market outcomes and organization. We next turn to the effect of CL entry market outcomes, estimating regressions of the form specified in equations 1 and 2. Our results suggest that following the opening of a local CL website, newspapers serving that area experience a significant reduction in circulation (Table 2) and workforce (Tables 4 and 5), with the latter affecting both editorial and management staff (Table 6).

This effect is especially pronounced for newspapers that, prior to Craigslist's entry, relied more

heavily on classified ads, and which were hence more vulnerable to the entry of new competitors in this previously lucrative niche. For such newspapers, CL's entry is associated with a 9% decline in circulation and a 14% decline in staff and jobs count. Alternatively, expressed in terms of a marginal year of exposure to CL competition (Panel B of the respective tables), our estimates imply a 4.5% decline in circulation and 5% decline in staff per year.

Crucially for our identification strategy, we find no pre-trends in these outcomes – neither in the staggered introduction design, nor when looking at the difference between newspapers with and without classified manager (Figures 9 to 11). Our results are robust to using various alternative definitions of the relevant market for both Craigslist's websites and newspapers, though, as expected, we obtain somewhat more precise estimates for more narrow definitions of the treatment (Figures 12 - 14).[2]

Newspapers content. To understand the impact of CL's entry on the distribution of topics covered by affected newspapers, we estimate a Correlation Explanation topic model (Gallagher et al. 2017) using the text of a random sample of 2 million articles.

This method allows us to capture topics of interest in a semi-supervised way by specifying a minimal set of anchor words. Specifically, 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 15 presents the resulting topics, as described by their most representative words. The 10 resulting topics can be labeled as follows: local politics, congressional politics, national politics, foreign politics, entertainment, health/ family, weather, crime, obituaries.

[2]The results are also robust to keeping a balanced panel of newspapers, i.e. abstracting from any entries or exits in the period of interest, to dropping the control group of newspapers never affected by CL, as well as to alternative definitions of the classified manager dummy.

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.[3] 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 equations 1 and 2, with the average probability for each one of the 10 topics as dependent variable.

Tables 7 and 8 present the results for political topics and for other topics respectively. We find negative and significant effects on coverage of political topics (with the exception of the one associated with Congress), with magnitudes in the order of 8 to 11% relative to the mean. At the same time, we find no change in any of the remaining topics.

The significant reduction in political coverage is confirmed looking at mentions of the names of local elected officials. In this case, we search the entire archive (rather than a random sample) for mentions of the names of all Congressmen representing a district in the same state as the newspapers' HQ. Table 9 shows that CL's entry is associated with a significant decline of such mentions by about 13%.

Survey outcomes. Finally, we use use survey data from the National Annenberg Election Survey to confirm the documented decline in newspaper readership. We find that survey respondents in counties affected by Craigslist's entry and with classified-intensive newspapers become less likely to read a newspaper (Table 3), and this true both at the extensive and at the intensive margins.

In progress. The results described above suggest that the entry of Craigslist was associated with significant losses in local newspapers' readership, with staff cuts, and with a shift of content away from reporting on politics. We are currently extending this analysis in several directions.

[3]Importantly, these probabilities do not necessarily sum to 1 – an article can simultaneously belong to more than one topic, or to none.

First, our rich data on newspapers staff allows us to explore further which topical areas suffered most from downsizing. In particular, we are using the job titles listed by E&P to investigate whether staff cuts were more likely to affect political reporters, which would explain the observed decline in political content. We are also investigating whether newspapers affected by the entry of Craigslist became more likely to copy from news wires rather than produce original content.

Second, we are investigating whether the decline in political coverage translated into lower levels of political awareness and participation by local readers. Indeed, our preliminary results looking at electoral turnout suggest that CL's entry is associated with lower participation. Our final step is to analyse the possible implications for the behavior of local politicians, looking at measures ranging from Congressional roll call voting to local municipal finances.

References

Chaisemartin, C. and X. D'Haultfoeuille (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, forthcoming.

Gallagher, R. J., K. Reing, D. Kale, and G. V. Steeg (2017): "Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge," *Transactions of the Association for Computational Linguistics*, 5, 529–542.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011): "The effect of newspaper entry and exit on electoral politics," *American Economic Review*, 101(7).

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 (2015): "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect," *American Journal of Political Science*, 61(1), 5–20.

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.

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.

5 Figures

Figure 1: Evolution of newspaper revenues by source (1997-2013)

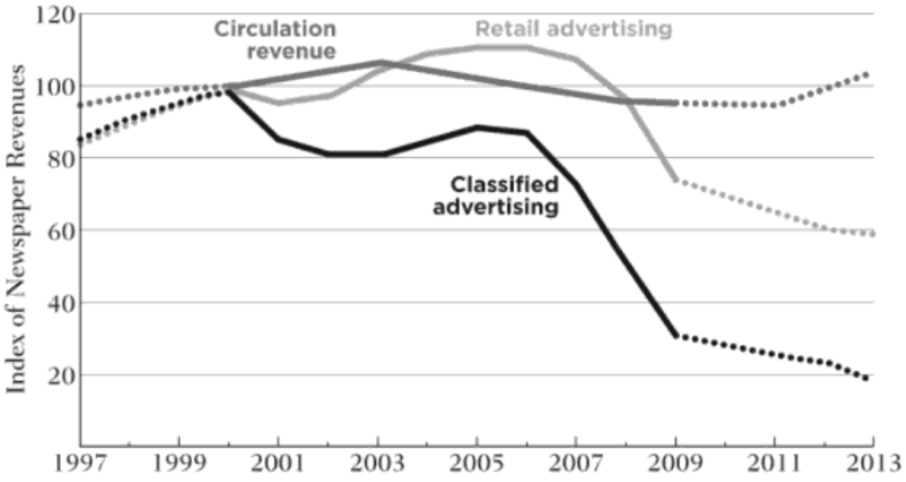


Figure 4: U.S. Newspaper Revenues over Time (Index: year 2000 = 100).

Figure 2: CL expansion across US counties over time

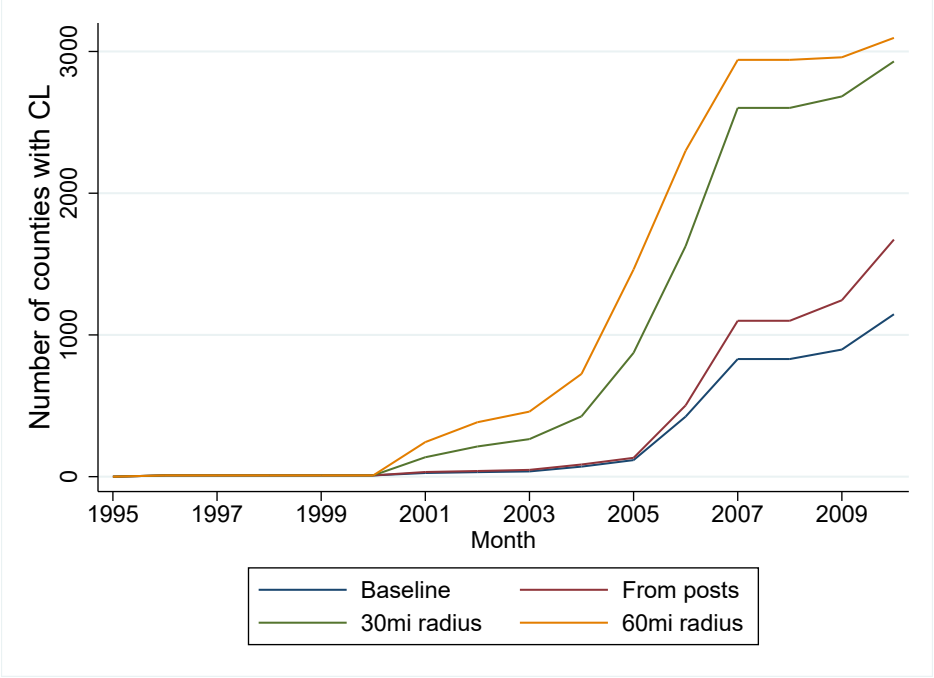
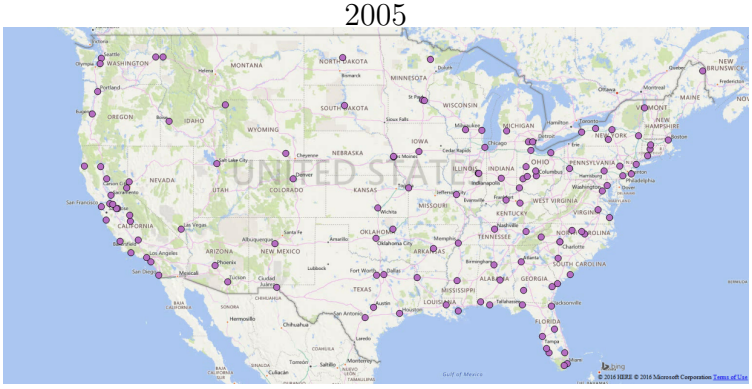


Figure 3: Location of CL websites at 3 points in time

2000

2005



2010

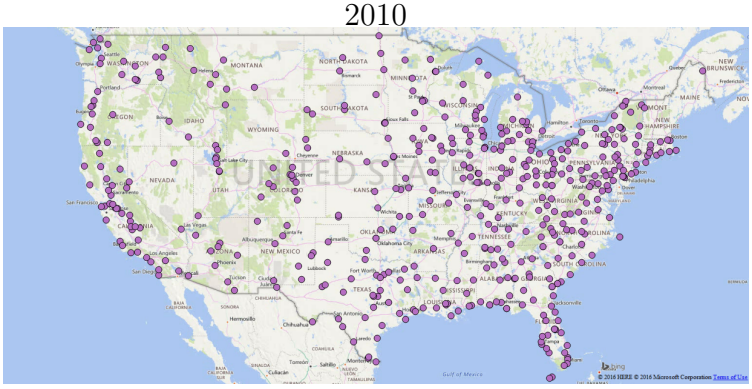
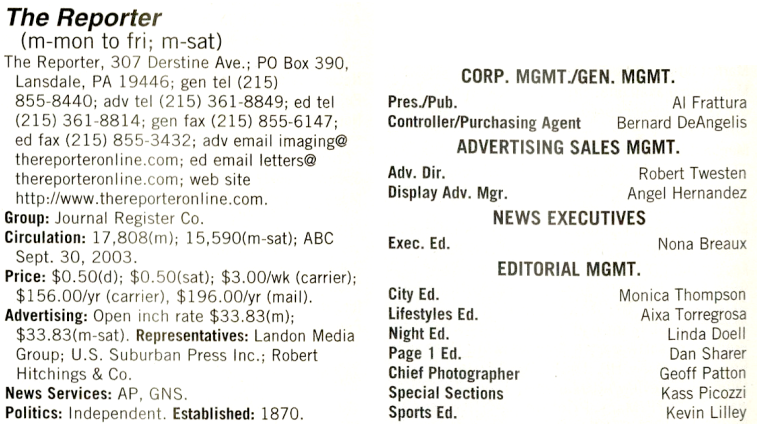


Figure 4: Extracts from the Editor and Publisher Yearbooks



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 Twesten
Display Adv. Mgr.	Angel Hernandez

NEWS EXECUTIVES

Exec. Ed.	Nona Breaux

EDITORIAL MGMT.

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

(a) The Reporter, Landsdale PA

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

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

CORP. MGMT./GEN. MGMT.

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

ADVERTISING SALES MGMT.

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

MARKETING MGMT.

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

CIRCULATION MGMT.

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

NEWS EXECUTIVES

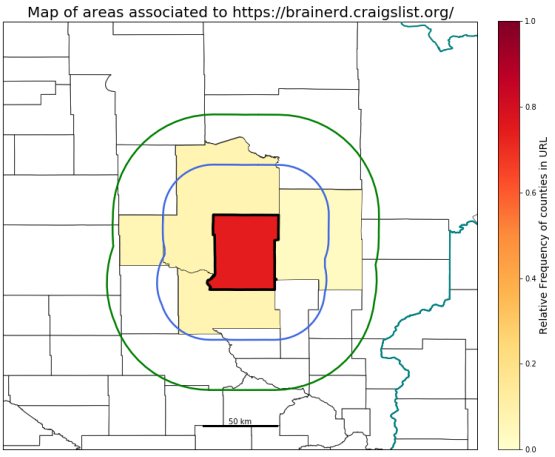
Mng. Ed.	James O'Shea
Public Ed.	Don Wycliff
Deputy Mng. Ed., Features	Jim Warren
Deputy Mng. Ed., News	George de Lama
Deputy Mng. Ed., Opns.	Randy Weissman
Assoc. Mng. Ed., Electronic News	Mark Hinojosa
Assoc. Mng. Ed., Features	Mary Elson
Assoc. Mng. Ed., Financial News	Rob Karwath
Assoc. Mng. Ed.,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

EDITORIAL MGMT.

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

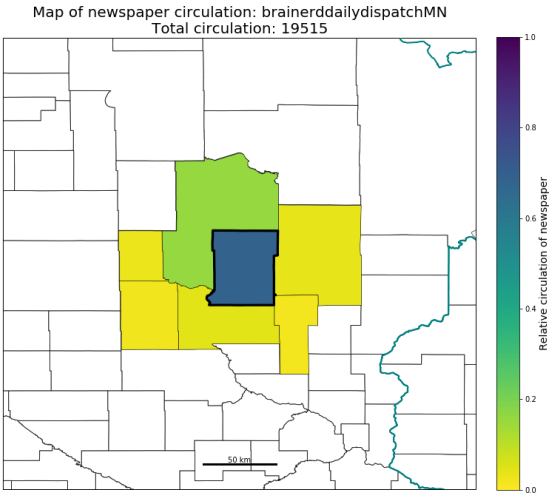
(b) The Chicago Tribune

Figure 5: Geographic scope of Craigslist websites



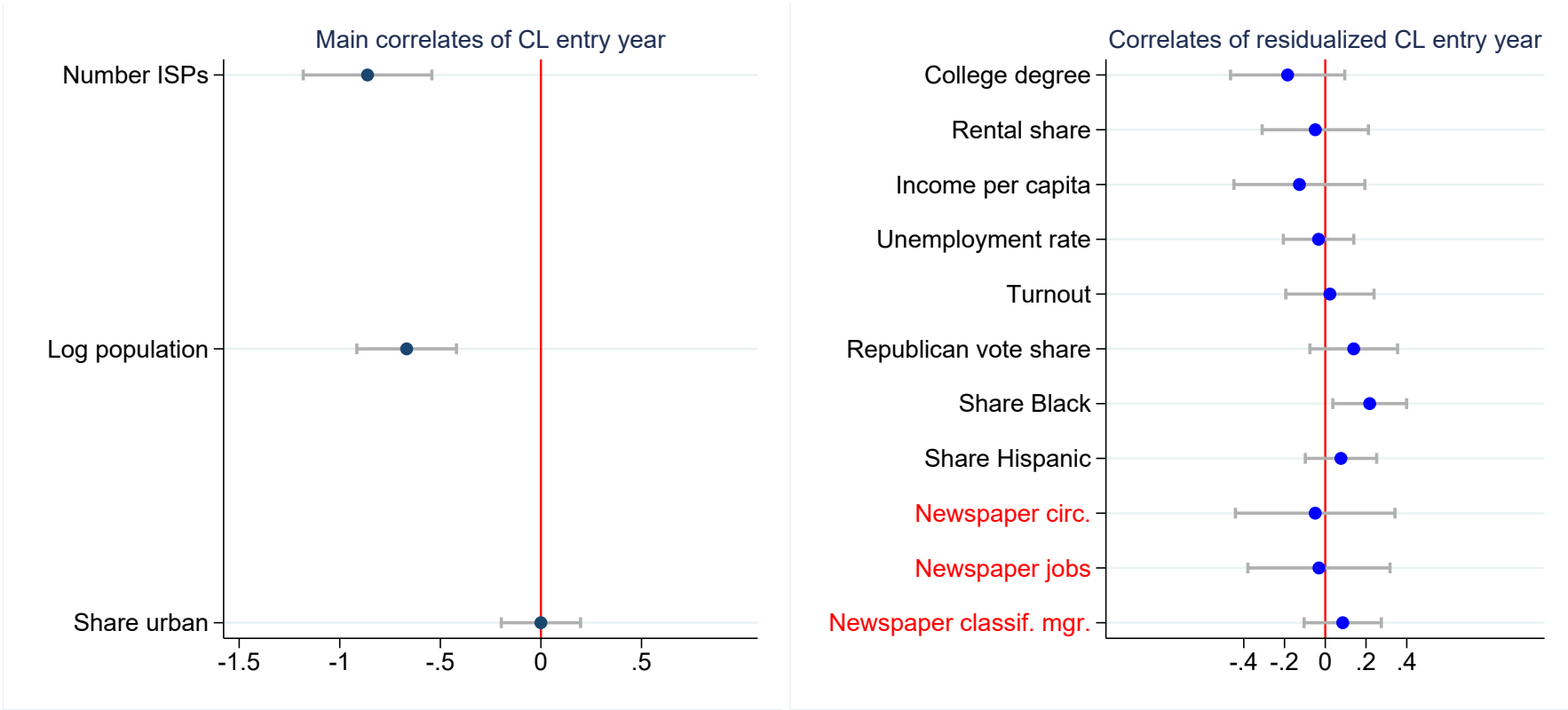
Notes: Example for the case of the Brainerd, MN (`https://brainerd.craigslist.org/`)

Figure 6: Geographic scope of newspaper markets



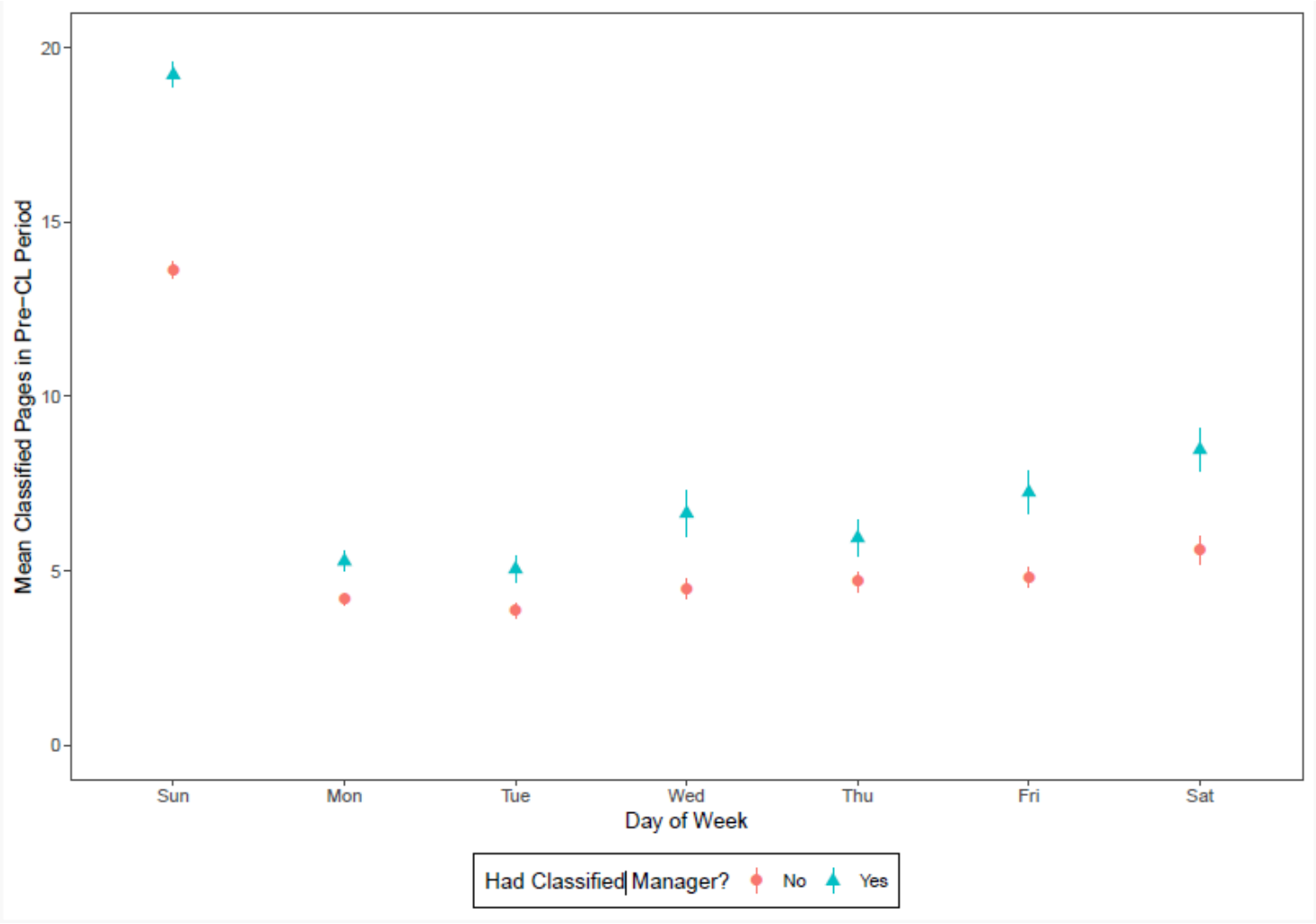
Notes: Example for the case of the Brainerd Dispatch, MN.

Figure 7: Correlates of CL entry



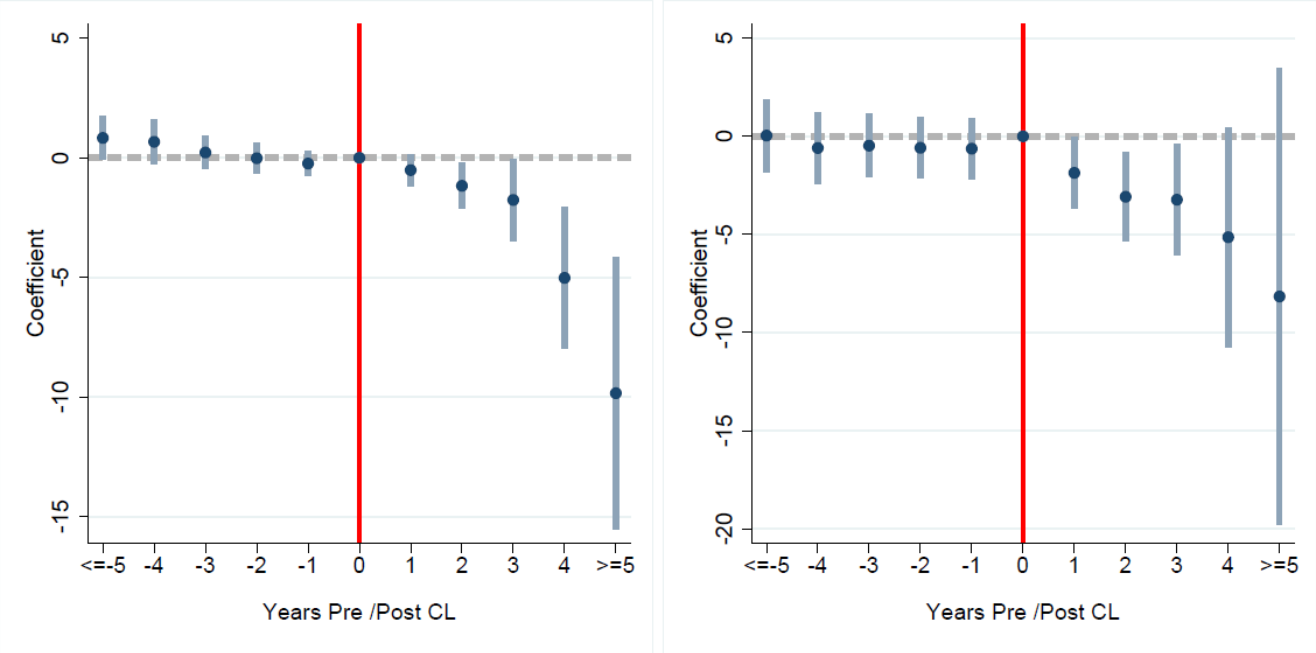
Notes: Left hand side panel: Main county-level correlates of year of CL entry. Right hand side panel: county-level correlates of residualized entry year, after accounting for number of ISPs, log population and share urban population.

Figure 8: *Notes:* Volume of classified ads: comparison of newspapers with/ without classified manager



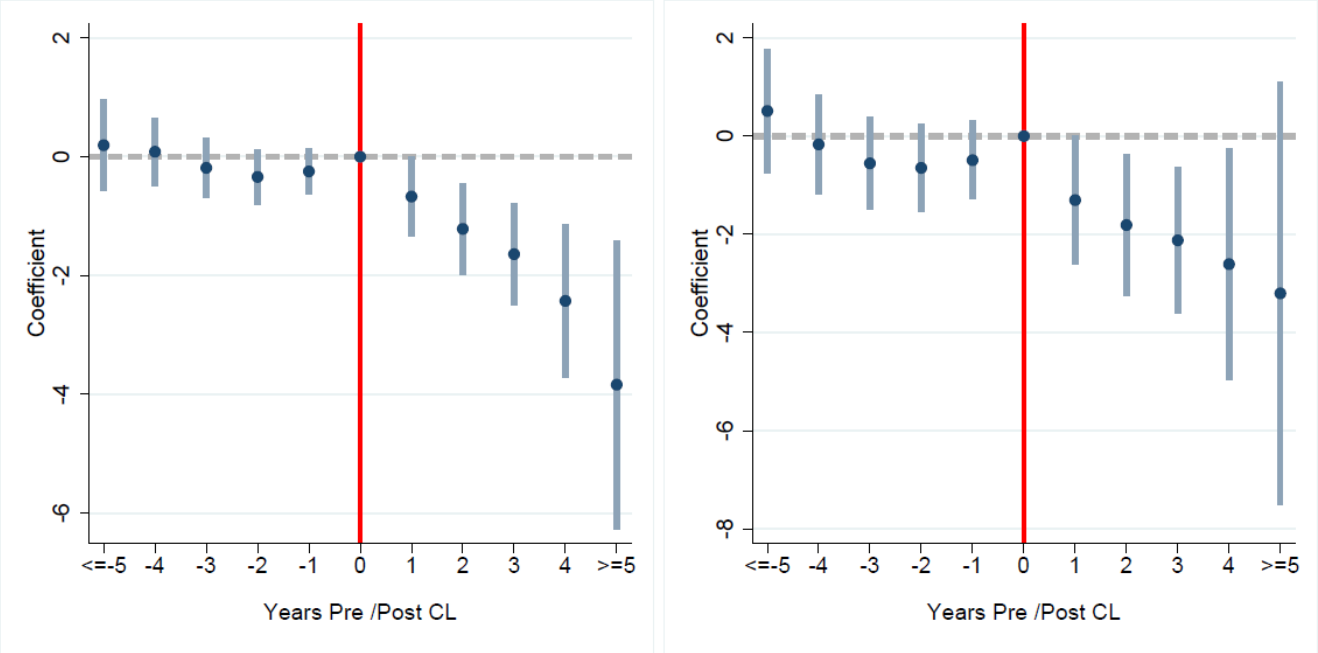
Notes: Average number of classified pages per newspaper copy, split by day-of-week and by presence of a classified manager.

Figure 9: Circulation: Leads and Lags



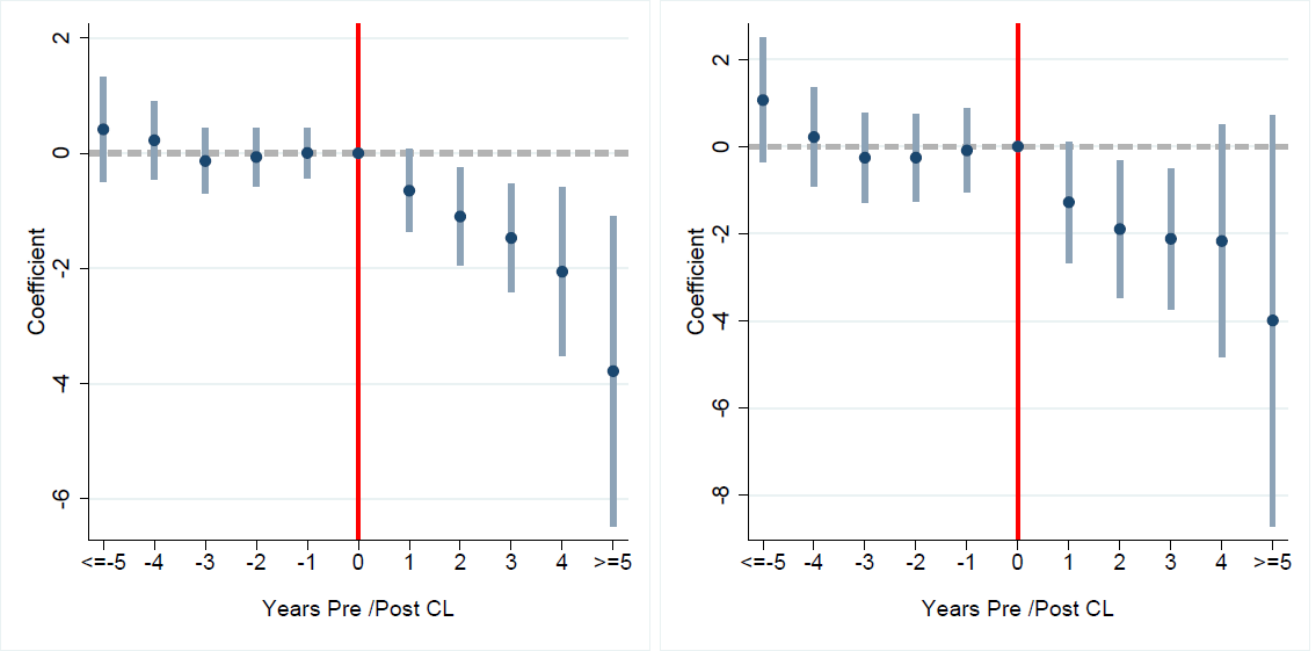
Notes: Evolution of circulation for newspapers in CL-counties relative to non-CL counties (left-hand side), and newspapers with classified manager vs. others (right hand-side). The omitted category contains the year before CL entry, as well as newspapers that never experience CL entry.

Figure 10: Staff count: Leads and Lags



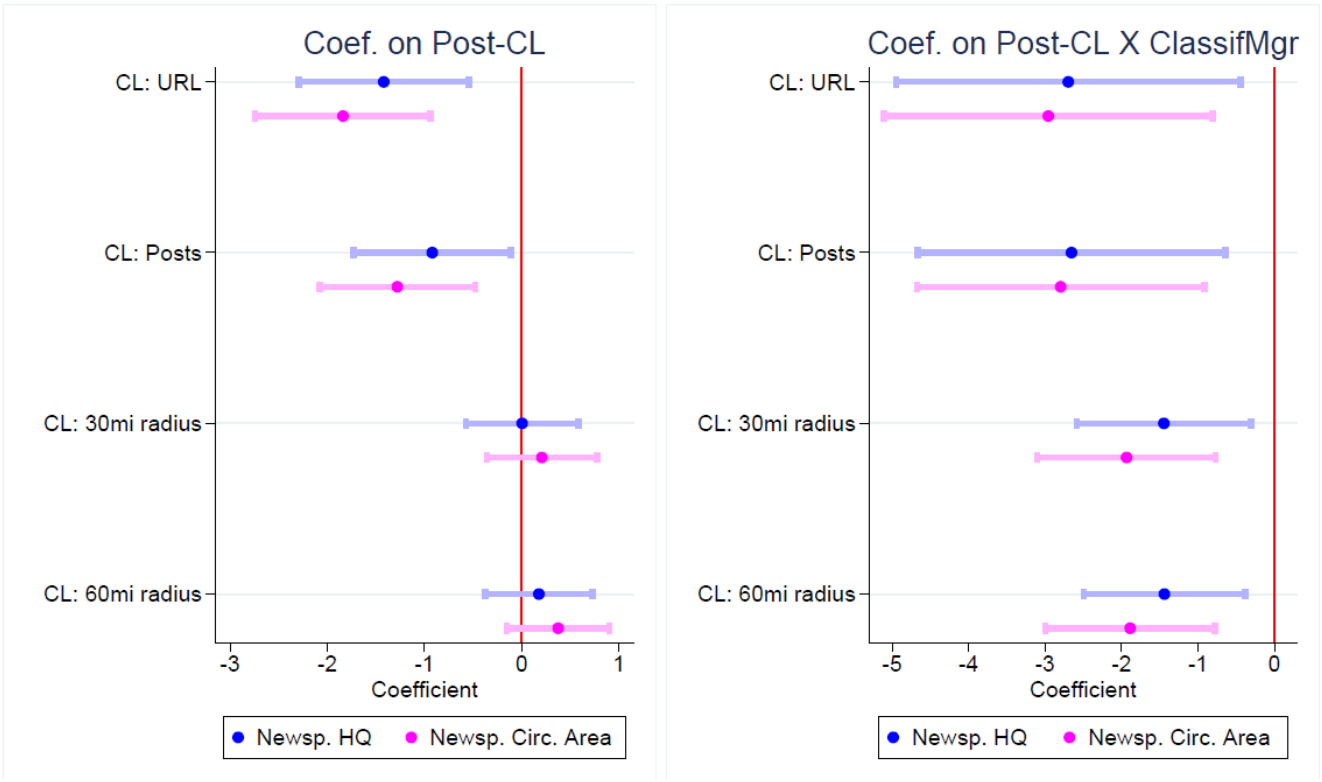
Notes: Evolution of staff count for newspapers in CL-counties relative to non-CL counties (left-hand side), and newspapers with classified manager vs. others (right hand-side). The omitted category contains the year before CL entry, and newspapers that never experience CL entry.

Figure 11: Jobs Count: Leads and Lags



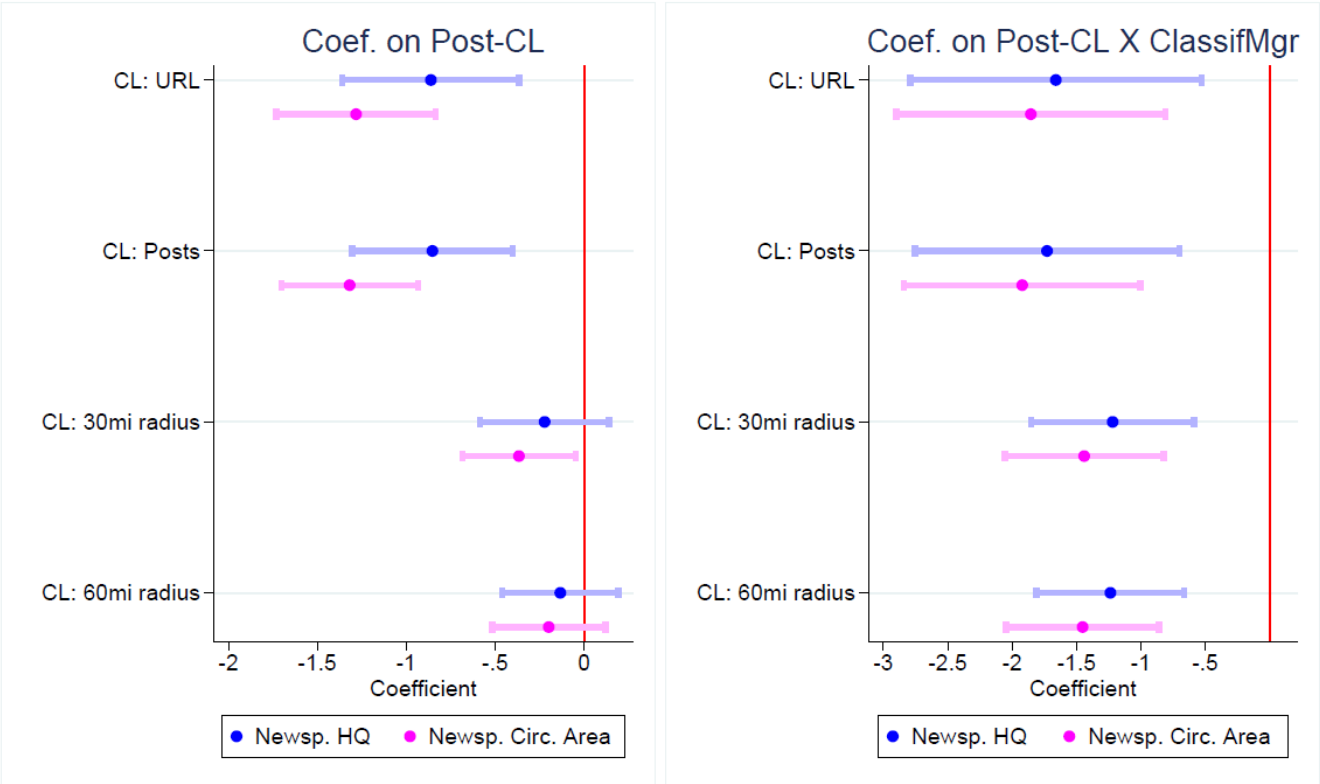
Notes: Evolution of jobs count for newspapers in CL-counties relative to non-CL counties (left-hand side), and newspapers with classified manager vs. others (right hand-side). The omitted category contains the year before CL entry, and newspapers that never experience CL entry.

Figure 12: Newspaper Circulation: Alternative CL / Newspaper Market Definitions



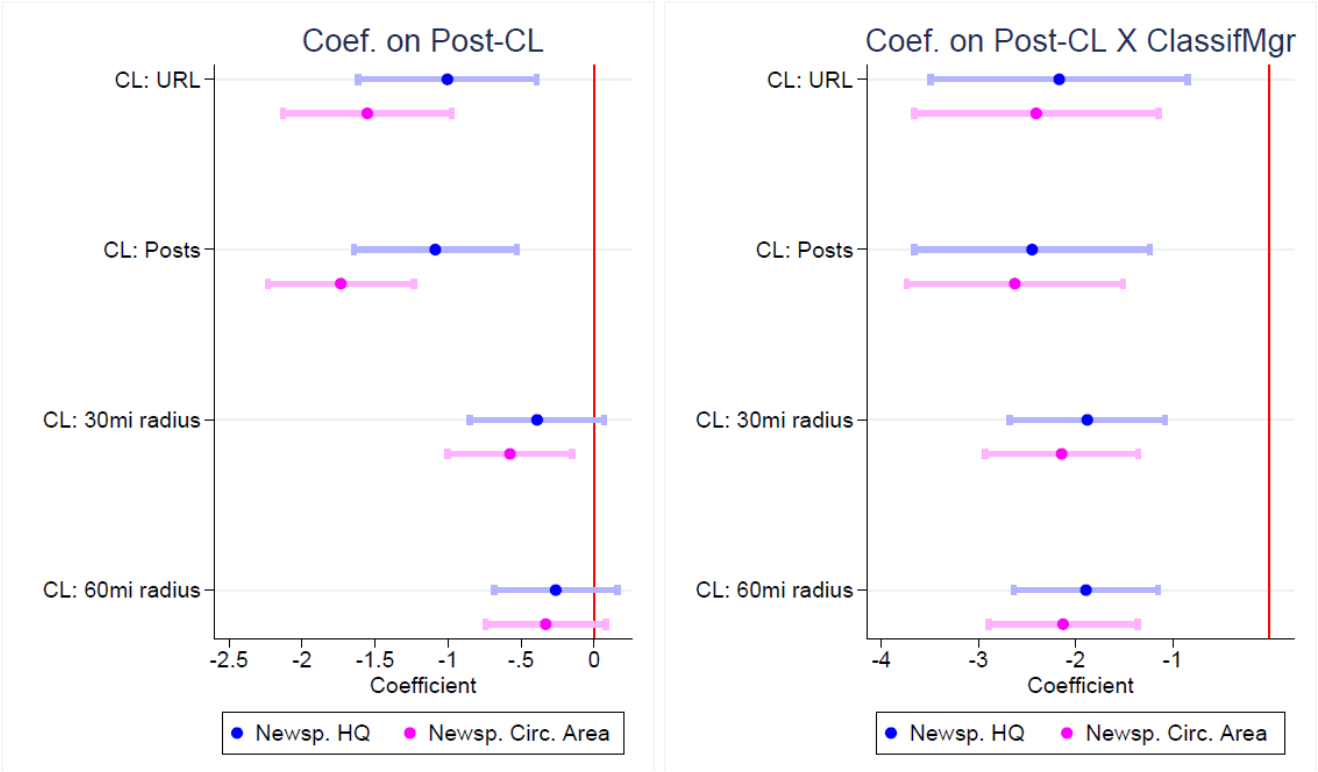
Notes: Robustness of the estimates to alternative definitions of newspaper markets – based on county of HQ or on newspapers' circulation area – and to alternative definitions of CL markets – based on county/-ies corresponding to the website name, on the distribution of posted ads, or on a 30mi or 60mi radius.

Figure 13: Staff Count: Alternative CL / Newspaper Market Definitions



*Notes:*Robustness of the estimates to alternative definitions of newspaper markets – based on county of HQ or on newspapers' circulation area – and to alternative definitions of CL markets – based on county/-ies corresponding to the website name, on the distribution of posted ads, or on a 30mi or 60mi radius.

Figure 14: Jobs Count: Alternative CL / Newspaper Market Definitions



Notes: Robustness of the estimates to alternative definitions of newspaper markets – based on county of HQ or on newspapers' circulation area – and to alternative definitions of CL markets – based on county/-ies corresponding to the website name, on the distribution of posted ads, or on a 30mi or 60mi radius.

Figure 15: CorEx Model Topics

```
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
```

Notes: Topics generated by a CorEx topic model with their most representative unigrams.

6 Tables

Table 1: Volume of classified ads

	Classified Page Count	
	(1)	(2)
Classified Mgr.	4.859***	
	(1.674)	
Post-CL Entry		1.194
		(0.981)
Classified Mgr. × Post-CL Entry		−4.357**
		(1.753)
Years Included:	Pre-CL	All
Number of Papers:	268	268
Year Fixed Effects:	Y	Y
Newspaper Fixed Effects:	N	Y
N	29,047	41,020
R^2	0.126	0.677

*p < .1; **p < .05; ***p < .01
An observation is a newspaper-day. Standard errors (clustered by newspaper) in parentheses.

Notes: OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Newspaper Circulation

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Circulation (thousands)					
Panel A						
Post-CL	-0.954*	-1.463**	-1.172	0.998	0.106	0.410
	(0.543)	(0.633)	(0.761)	(0.818)	(0.897)	(0.976)
Post-CL × Classified Mgr.				-3.668***	-2.982***	-2.876***
				(1.206)	(1.138)	(1.096)
Panel B						
Years Post-CL	-1.775***	-1.511***	-1.505***	-0.761**	-0.691	-0.692
	(0.631)	(0.572)	(0.532)	(0.330)	(0.444)	(0.438)
Years Post-CL × Classif. Mgr.				-1.789**	-1.473*	-1.435*
				(0.825)	(0.800)	(0.804)
Log population, share urban, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Additional county characteristics × Year FEs	No	Yes	Yes	No	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	Yes
Observations	23060	22916	22914	22743	22599	22597
Number of newspapers	1558	1549	1549	1509	1500	1500
R^2	0.98	0.98	0.98	0.98	0.98	0.98
Mean dependent variable	33.81	33.38	33.38	33.87	33.43	33.43

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Self-reported frequency of newspaper readership

	(1) Read newsp. days per wk	(2) Read newsp. dummy	(3) Read newsp. days per wk	(4) Read newsp. dummy
Post-CL	0.117	0.019		
	(0.108)	(0.015)		
Post-CL × Classified Mgr.	-0.283*	-0.053**		
	(0.161)	(0.026)		
Years Post-CL			0.089	0.022
			(0.095)	(0.016)
Years Post-CL × Classif. Mgr.			-0.201**	-0.041**
			(0.101)	(0.016)
Respondent characteristics	Yes	Yes	Yes	Yes
Full county controls	Yes	Yes	Yes	Yes
County FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	107503	107503	107503	107503
Number of counties	1203	1203	1203	1203
R^2	0.14	0.06	0.14	0.06
Mean dependent variable	3.71	0.75	3.71	0.75

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Newspaper Staff Count

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Number of employees					
Panel A						
Post-CL	-1.042***	-1.011***	-1.106***	0.198	0.255	0.128
	(0.323)	(0.332)	(0.350)	(0.333)	(0.343)	(0.353)
Post-CL × Classified Mgr.				-2.432***	-2.478***	-2.395***
				(0.498)	(0.476)	(0.473)
Panel B						
Years Post-CL (broad)	-0.404***	-0.382***	-0.413***	0.069	0.093	0.063
	(0.097)	(0.113)	(0.132)	(0.071)	(0.078)	(0.089)
Years Post-CL (broad) × Classif. Mgr.				-0.872***	-0.872***	-0.893***
				(0.151)	(0.146)	(0.143)
Log population, share urban, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Additional county characteristics × Year FEs	No	Yes	Yes	No	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	Yes
Observations	22905	22761	22760	22691	22547	22546
Number of newspapers	1543	1534	1534	1508	1499	1499
R^2	0.92	0.92	0.92	0.92	0.92	0.92
Mean dependent variable	17.90	17.84	17.84	17.96	17.90	17.90

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Newspaper Jobs Count

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Number of jobs					
Panel A						
Post-CL	-1.117***	-1.083***	-1.243***	0.385	0.446	0.269
	(0.401)	(0.393)	(0.425)	(0.414)	(0.413)	(0.436)
Post-CL × Classified Mgr.				-2.914***	-2.959***	-2.900***
				(0.567)	(0.549)	(0.551)
Panel B						
Years Post-CL	-0.483***	-0.514***	-0.551***	0.126	0.124	0.083
	(0.142)	(0.149)	(0.169)	(0.110)	(0.111)	(0.129)
Years Post-CL × Classif. Mgr.				-1.044***	-1.097***	-1.089***
				(0.202)	(0.198)	(0.196)
Log population, share urban, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Additional county characteristics × Year FEs	No	Yes	Yes	No	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	Yes
Observations	22832	22688	22687	22624	22480	22479
Number of newspapers	1543	1534	1534	1508	1499	1499
R^2	0.91	0.91	0.91	0.91	0.91	0.91
Mean dependent variable	21.37	21.30	21.30	21.44	21.38	21.38

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Newspaper Jobs Count: By Job Type

	(1) Num. Managers	(2) Num. Editors	(3) Num. Managers	(4) Num. Editors
Post-CL	0.047	0.136		
	(0.087)	(0.298)		
Post-CL × Classified Mgr.	-0.663***	-0.772*		
	(0.121)	(0.451)		
Years Post-CL			0.044	-0.041
			(0.033)	(0.073)
Years Post-CL × Classif. Mgr.			-0.254***	-0.356**
			(0.039)	(0.144)
Full county controls	Yes	Yes	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	21798	22120	21798	22120
Number of newspapers	1497	1498	1497	1498
R^2	0.80	0.89	0.80	0.89
Mean dependent variable	3.46	10.60	3.46	10.60

Notes: OLS regressions in all columns. Standard errors clustered by CL-area. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Politics Topics

	(1) national	(2) local	(3) congress	(4) foreign	(5) national	(6) local	(7) congress	(8) foreign
Post-CL	-0.001	-0.003	-0.000	0.005				
	(0.005)	(0.007)	(0.003)	(0.004)				
Post-CL × Classified Mgr.	-0.017***	-0.013	-0.011**	-0.011**				
	(0.006)	(0.008)	(0.005)	(0.004)				
Years Post-CL					0.002	0.001	0.001	-0.001
					(0.002)	(0.002)	(0.001)	(0.001)
Years Post-CL × Classif. Mgr.					-0.007***	-0.001	-0.002**	-0.003**
					(0.002)	(0.002)	(0.001)	(0.001)
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	7129	7129	7129	7129	7129	7129	7129	7129
Number of newspapers	862	862	862	862	862	862	862	862
R^2	0.52	0.47	0.40	0.53	0.52	0.47	0.40	0.54
Mean dependent variable	0.21	0.31	0.10	0.10	0.21	0.31	0.10	0.10

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: 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	-0.003	-0.007	-0.004	0.002	0.008	-0.002						
	(0.004)	(0.006)	(0.004)	(0.010)	(0.008)	(0.003)						
Post-CL × Classified Mgr.	0.006	-0.003	-0.001	0.013	0.012	0.003						
	(0.006)	(0.007)	(0.006)	(0.012)	(0.008)	(0.004)						
Years Post-CL							-0.001	-0.003	0.001	0.000	0.002	0.001
							(0.001)	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)
Years Post-CL × Classif. Mgr.							0.001	0.001	0.000	0.004	-0.001	-0.001
							(0.001)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
Full 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	7129	7129	7129	7129	7129	7129	7129	7129	7129	7129	7129	7129
Number of newspapers	862	862	862	862	862	862	862	862	862	862	862	862
R^2	0.43	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

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Mentions of representatives from the same state

	(1)	(2)	(3)	(4)	(5)	(6)
	Articles mentioning Congressmen from state of HQ (ISH)					
Panel A						
Post-CL	-0.189	-0.126	-0.208	0.075	0.181	0.081
	(0.131)	(0.137)	(0.160)	(0.198)	(0.203)	(0.234)
Post-CL × Classified Mgr.				-0.387*	-0.452**	-0.416*
				(0.217)	(0.217)	(0.238)
Panel B						
Years Post-CL (core)	-0.086	-0.055	-0.087	0.001	0.036	-0.008
	(0.052)	(0.055)	(0.067)	(0.067)	(0.068)	(0.081)
Years Post-CL (core) × Classified Mgr.				-0.114**	-0.130**	-0.118**
				(0.052)	(0.052)	(0.057)
Log population, share urban, # ISPs	Yes	Yes	Yes	Yes	Yes	Yes
Additional county characteristics × Year FEs	No	Yes	Yes	No	Yes	Yes
Newspaper FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FEs	No	No	Yes	No	No	Yes
Observations	3731	3676	3665	3714	3659	3648
Number of newspapers	343	338	337	341	336	335
R^2	0.68	0.70	0.76	0.68	0.70	0.76
Mean dependent variable	4.51	4.53	4.53	4.52	4.53	4.54

Notes: OLS regressions in all columns. Standard errors clustered by CL-area.
Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.