

# Political Effects of Newspaper Paywalls\*

Julian Streyczek\*\*

February 4, 2025

## Abstract

I examine how the introduction of paywalls on newspaper websites in the early 2010s affected political knowledge and electoral participation in the United States. Exploiting the staggered adoption of paywalls across newspapers and counties, I document a 25–30 percent decline in website visits following the introduction of a paywall. Moreover, I show that in counties where a larger share of online news consumption was affected by paywalls, survey respondents are 2–3 percentage points less likely to correctly answer political knowledge questions. This effect is driven by declines in knowledge about regional politics and most pronounced among individuals with lower income and education. These findings suggest that these groups may respond to paywalls by substituting towards less costly news sources with less intensive coverage of regional politics. Additionally, affected counties exhibit a suggestive 2 percentage point decline in congressional primary election participation. My results underscore the importance of easy access to high-quality news for democratic processes.

**Keywords:** Paywalls, News, Information, Politics

**JEL:** D72, L82, Z13

\*I express deep gratitude to my advisors Guido Tabellini, Carlo Schwarz, and Luca Braghieri for their outstanding support. I also thank Alberto Manconi, Edoardo Teso, Elliott Ash, Eugen Dimant, Ho Kim, Jaime Marques Pereira, Jan Bakker, Jim Snyder, Julia Cagé, Nico Voigtländer, Rafael Jiménez-Durán, Ro'ee Levy, and Sarah Eichmeyer, as well as seminar and conference participants at Bocconi and outside for valuable discussions and feedback. I gratefully acknowledge financial funding by Bocconi University, the Fondazione Invernizzi, and the German Academic Scholarship Foundation. All errors remain my own.

\*\*Università Bocconi, Department of Economics. [julian.streyczek@phd.unibocconi.it](mailto:julian.streyczek@phd.unibocconi.it)

## 1 Introduction

Over the past 15 years, paywalls have transformed much of newspapers' online content from a public good into a club good: Throughout the 2000s, most newspapers offered free online access to their articles, relying on advertising to generate revenue. However, as the shift to online readership undermined traditional print revenue streams, many publishers implemented paywalls on their websites - subscription models that require a monthly fee for full access. By 2020, more than 80 percent of the most popular US newspapers had implemented such paywalls. Although these models help publishers remain profitable, they also restrict the availability of information, especially for those unwilling or unable to pay.

Access to information can shape voter knowledge and participation, which in turn can influence the accountability of elected representatives and the allocation of public funds ([Besley and Burgess, 2002](#); [Strömberg, 2004](#); [Ferraz and Finan, 2008](#); [Snyder Jr and Strömberg, 2010](#)). Past shifts in digital media availability have prompted consumers to move away from news-intensive outlets, with significant consequences for voting patterns ([Gentzkow, 2006](#); [Drago et al., 2014](#); [Falck et al., 2014](#); [Campante et al., 2018](#); [Gavazza et al., 2019](#)).

In this paper, I study how the emergence of newspaper paywalls affected political outcomes, exploiting the staggered introduction of paywalls across time and geographic regions for causal identification. My analysis consists of two parts: First, I show that paywalls reduce consumption of newspapers' online content. Second, I develop a measure of counties' exposure to paywalls and link it to large-scale survey data to examine effects on political knowledge and electoral participation.

I focus on the 80 largest daily newspapers in the United States based on website page views in 2010. For each newspaper, I measure online consumption using the daily number of page views between 2010 and 2017, collected by [Alexa Internet](#), and offline consumption using county-level print circulation in 2011, provided by the [Alliance for Audited Media](#). I assemble paywall introduction dates by checking mentions in news articles or historical website snapshots in the [Internet Archive](#). I measure individual-level knowledge and voting using the [Cooperative Election Study](#) (CES)<sup>1</sup>, a yearly, nationally representative survey of the adult US population. I construct a political knowledge index using questions that ask participants to specify the party affiliation of politicians or to name the majority party in US legislative bodies. I also use validated data on participation in elections.

For the first part of my analysis, I identify the causal effect of paywalls on page views using the sequential adoption of paywalls across websites. I compare newspapers that introduced a paywall between 2010 and 2017 to those that introduced paywalls later, (much) earlier, or never. I control for newspapers' market characteristics to account for trends among different audiences, and use estimators that are robust to heterogeneous treatment effects across newspapers or time. To corroborate the identifying assumption that trends in page views would be parallel in

---

<sup>1</sup>Formerly called *Cooperative Congressional Election Study* (CCES).

the absence of paywalls, I conduct two sets of robustness checks. First, I verify that selection of newspapers into paywalls is unlikely to drive the results, as pre-paywall trends in page views are parallel, inclusion of controls has minimal impact, and the result is robust to varying the composition of the control group. Second, I show that potential spillover biases, arising from readers' substitution of paywalled with non-paywalled outlets, are negligible, by conducting a series of robustness checks excluding newspapers that are expected to generate the highest spillovers.

I find that paywalls reduce the number of page views for paywalled newspapers by 25–30 percent, on average. These effects materialize in the first two months after the paywall and are persistent over the following years, suggesting that paywalls considerably displaced readers from mainstream US newspapers.

For the second part of my analysis, I start by quantifying the relevance of paywalls for different survey respondents based on their county of residence. For this purpose, I construct a measure for counties' exposure to paywalls over time. Specifically, I define paywall exposure as the share of initial online newspaper consumption that was affected by paywalls. To approximate county-level online readership, I assign newspaper's national page views to counties in proportion to the newspaper's local print circulation, and validate this measure of regional demand for newspapers' online content by comparing it to regional Google search interest. Then, I exploit variation in paywall exposure in a staggered difference-in-differences design, where counties are considered treated once paywall exposure exceeds the 2017-median, which is around 72 percent. While this approach maximizes power, I show that the results hold when varying this threshold. I estimate effects using robust estimators.

Identification relies on the assumption that no omitted variable is correlated with both survey outcomes and paywall exposure, conditional on fixed effects and controls. I provide extensive evidence in support of this assumption. First, all specifications include DMA-by-year fixed effects that capture any time-varying shocks to media markets, such as newspaper exits or the diffusion of social media. Second, I include survey respondent characteristics interacted with year fixed effects that capture changes in outcomes of certain demographics unrelated to paywalls, such as preferences for digital versus analog platforms. Third, I show that adding baseline county characteristics interacted with year fixed effects has little impact on the results, suggesting that observable heterogeneity between counties does not introduce meaningful bias, neither through trends unrelated to paywalls nor through measurement error in paywall exposure.

I find that in counties with above-median paywall exposure, individuals are on average 2–3 percentage points less likely to correctly answer political knowledge questions. The event study rejects the hypothesis that prior trends in outcomes are driving the results. The declines in knowledge pertain to politicians elected at state and district levels, as well as majorities in state rather than national legislative bodies. Furthermore, heterogeneity analyses reveal that the effects are driven by lower-income and lower-education individuals. Finally, I present suggestive

evidence that higher paywall exposure reduces participation in congressional *primary* elections, but not in congressional elections.

My results are consistent with the interpretation that paywalls shift news consumption of those unable or unwilling to pay for news towards outlets with less intensive coverage of regional politics. While the absence of individual-level news consumption data makes it difficult to assess which individuals reduce consumption of paywalled newspapers and which outlets they substitute to, the following arguments support this interpretation: First, substitution to newspapers' print versions is unlikely to explain the decline in page views because aggregate print readership continued to decrease over the sample period, and case studies find little evidence for substitution to print (e.g., [Pew Research Center, 2023](#); [Pattabhiramaiah et al., 2019](#)). Second, the newspapers in my sample traditionally provide content characterized by a relatively large share of hard news and a focus on regional politics, compared to free online news sources such as social media or TV websites. Therefore, substitution likely represents shifts towards alternatives that either focus less on regional compared to national politics, or cover politics less intensively overall. Third, this idea is confirmed by the larger observed declines in knowledge and participation for regional and lower-level politics. Fourth, the fact that individuals with lower education and lower income are most affected by paywalls may reflect the monetary nature of the paywall barrier. Such individuals might have a lower willingness to purchase a subscription due to relatively higher monetary costs and lower perceived benefits from information.

My findings confirm that easy access to news is important for political knowledge and participation. Moreover, they underscore that in a predominantly profit-driven news market, economic shocks can ultimately affect democratic outcomes. To counteract these effects, modern democracies may benefit from funding independent, high-quality news outlets to ensure broad and equitable access to relevant information.

This paper contributes to several streams of literature in economics and political science. First, it relates to existing work on news, knowledge, and accountability. Notable papers include the diffusion of television ([Gentzkow, 2006](#); [DellaVigna and Kaplan, 2007](#); [Durante et al., 2019](#)) and high-speed internet ([Falck et al., 2014](#); [Campante et al., 2018](#); [Gavazza et al., 2019](#)), which demonstrate that biased media can persuade voters, while substitution away from media with high news content can reduce political engagement. Studies on the entry and exit of newspapers underline that access to news can improve political participation ([Gentzkow et al., 2011](#); [Drago et al., 2014](#); [Cagé, 2020](#); [Gao et al., 2020](#); [Djourelova et al., 2024](#)), which may affect the performance of political representatives and the allocation of public resources ([Besley and Burgess, 2002](#); [Strömberg, 2004](#); [Ferraz and Finan, 2008](#); [Snyder Jr and Strömberg, 2010](#)).

To the best of my knowledge, I provide the first evidence on political effects of newspaper paywalls. Unlike most historical trends in media innovation and digitization, paywalls *increased* barriers to information. Moreover, I show that the *monetary* nature of these barriers disproportionately affected lower-income and less-educated individuals.

Second, my paper relates to the literature studying the news industry. Prior studies have shown that entry and exit of newspapers affect existing newspapers' revenues, prompting adjustments in topics, slant, or journalistic intensiveness (George and Waldfogel, 2006; Gentzkow et al., 2014; Angelucci and Cagé, 2019). Digitization has accelerated these pressures, as newspapers faced declining print advertising revenues and increased online competition (Seamans and Zhu, 2014; Bhuller et al., 2024). In response, many newspapers have introduced paywalls as part of their subscription models. Existing case studies indicate that paywalls may cause newspapers to lose readers (Cook and Attari, 2012; Chiou and Tucker, 2013; Pattabhiramaiah et al., 2019), although some outlets have succeeded in offsetting these losses through increased subscriptions (Chung et al., 2019; Aral and Dhillon, 2021). In a broader study, Kim et al. (2020) find that paywalls on 42 large US-based newspapers reduced page views by 30 percent on average.

I add to this literature in three ways. First, I provide the most comprehensive empirical study to date on how paywalls affect news consumption, studying the 80 largest US-based newspaper websites at the time. Second, I identify the causal effect of paywalls using state-of-the-art econometric methods. Third, I demonstrate that the effects on the news industry are both economically significant and persistent.

The remainder of this paper is organized as follows. In Section 2, I introduce my data and provide background on paywalls. In Section 3, I show how paywalls affected news consumption, and in Section 4, I analyze downstream effects on political knowledge and electoral participation.

## 2 Data

### 2.1 Newspapers

My analysis focuses on the largest 80 daily US newspapers by website page views in 2010. The page view data originate from Alexa Internet, a former web traffic analysis company that was discontinued in 2021. It offered a free browser toolbar that provided website statistics to its users, while collecting anonymized data on website visits. For each newspaper website, I obtain the daily number of page views from 2010 through 2017. This figure reflects the total number of clicks to the website and any sub-domain per million toolbar users. According to Alexa Internet, the toolbar had several million users, but precise numbers were never disclosed. Therefore, the variable should be interpreted as a proportional measure of website traffic within the sample of toolbar users. Note that the data is limited to browser users and does not include traffic from mobile devices or tablets. Also, the data does not provide information on individual users.

To supplement information on the regional relevance of each newspaper, I obtain data on newspapers' print circulation from the [Alliance for Audited Media \(AAM\)](#). For each newspaper, the dataset includes the number of physical copies sold per county in 2011, covering all

counties where circulation exceeded 25 copies (2011 is the first year with full coverage of all 80 newspapers in the sample). An exception are the *New York Times*, *USA Today*, and *Wall Street Journal*, which are the only three classified as national newspapers by AAM. For these, print circulation is available on the DMA level.<sup>2</sup>

As a validation measure for regional interest in news websites, I use data from [Google Trends](#) to measure search interest for each newspaper by DMA in 2010. Search interest is defined as the number of Google searches for a newspaper's *topic* divided by the total number of searches.<sup>3</sup> To ensure that these shares are comparable across newspapers and regions, I transform shares into a common scale of 0–100, which reflects proportional differences in newspaper-specific search interest by DMA.<sup>4</sup>

## 2.2 Paywalls

I obtain information on paywall launches through news coverage by the paywall-introducing newspapers themselves or competing outlets. Where ambiguous, I check historical website snapshots for the appearance of a digital subscription button using the [Wayback Machine](#). Figure 1 shows the yearly share of newspapers in my sample that have a paywall on their website. Among the 80 newspapers, only three had introduced a paywall before 2010.<sup>5</sup> 60 newspapers introduced a paywall between 2010 and 2016, and further 13 introduced one by 2021.

These paywalls were designed in a similar way: First, they were usually "metered", allowing free access to the full website up to a certain number of clicks per month (usually 10-20), after which a subscription was necessary to view any additional article.<sup>6</sup> For four newspapers, the paywall had the form of new, dedicated "premium" website.<sup>7</sup> Second, paywalls were "leaky", meaning that they could be circumvented with sufficient effort (often intentionally), for example by clearing one's browser cache, or finding links to specific articles on search engines or social media. Therefore, a paywall increased either the monetary or non-monetary costs of accessing the website. For the purpose of this paper, I focus only on the existence of a paywall, abstracting from design features like the number of free articles, leakiness, and subscription pricing.

---

<sup>2</sup>Designated Market Areas (DMAs) are 210 geographical regions in the United States, defined by the media analytics firm Nielsen.

<sup>3</sup>Google categorizes searches for different keywords that likely refer to the same entity into topics, based on keyword combinations, search history, and clicks. This allows, for example, analyzing the search interest for the company *Apple* as opposed to the fruit.

<sup>4</sup>Search indices for queries to *Google Trends* are scaled proportionally to the maximum in the corresponding set of search terms, time period, and geographies. However, Google allows only queries for at most five keywords at a time. To make indices from different calls comparable, I scale all results against the topic *newspaper*.

<sup>5</sup>The *Wall Street Journal* in 1997, the *Albuquerque Journal* in 2001, and *Newsday* in 2009.

<sup>6</sup>The prominence of metered paywalls was partly due to Google's "First Click Free" policy, which until fall 2017 required news websites to grant each visitor at least three free articles per month, otherwise the website would be penalized in Google's search ranking.

<sup>7</sup>The four newspapers are the *Boston Globe*, *Houston Chronicle*, *Philadelphia Inquirer*, and *San Francisco Chronicle*.

[Figure 1 here]

## 2.3 Survey outcomes

To measure knowledge and voting in the US population, I use the [Cooperative Election Study](#) (CES) for the years 2006–2021 ([Ansolabehere and Schaffner, 2022](#)). The CES is an annual, nationally representative survey of the US population administered by YouGov that uses repeated cross sections of 10,000 to 60,000 respondents.

### Political knowledge

I measure political knowledge using the following two types of questions:

1. *Which party has the majority in [legislative body]?*
2. *Which party does [name of political representative] represent?*

These questions are asked separately for four legislative bodies (US Senate, US House, State Senate, State House) and four political representatives (Governor, House Representative, two Senators), based on the respondent's place of residence.

I compile responses to these eight questions into a political knowledge index by calculating the share of correct answers for each respondent in a given year.<sup>8</sup> To account for possible correlations in the signal of the answers, I use as an alternative measure the first principal component of the eight answers, ensuring that a higher value corresponds to more correct answers.

Table 1 presents summary statistics. On average, respondents answer around 63 percent of knowledge questions correctly, where knowledge is highest for the governor, and lowest for their state's Senate and House majorities. Note that across all questions, respondents may choose "I don't know", which I treat as an incorrect answer.

### Voting behavior

The CES provides self-reported voting data for national-level elections. For a subset of respondents, electoral participation is validated by YouGov, which matches respondents' identifiable information to official voting files, providing information on actual voting in these elections.

[Table 1 here]

## 2.4 Other data sources

I incorporate data on county demographic and economic characteristics for 2010 from the 5-Year American Community Survey ([U.S. Census Bureau, 2010](#)). County partisanship is

---

<sup>8</sup>All eight questions were asked in every year, except for the two questions on state legislative majorities, which were omitted in 2006 and 2009.

measured using the partisanship index in the National Neighborhood Data Archive ([Chenoweth et al., 2020](#)), which is based on Democratic and Republican vote shares in past elections. To link counties with media markets, I use the crosswalk from [Gentzkow and Shapiro \(2008\)](#).

### 3 Effect on news consumption

In this section, I analyze how the launch of paywalls on major US newspaper websites between 2010 and 2017 affected page views. I employ a staggered difference-in-differences design that exploits the sequential timing of paywall introductions.

#### 3.1 Empirical strategy

My analysis is based on the following two-way fixed effects (TWFE) regression model:

$$\log(Pageviews_{n,t}) = \sum_{\tau \neq -1} \beta_\tau D_{n,t}^\tau + \alpha_n + \gamma_t + \delta' \mathbf{X}_{n,t} + \varepsilon_{n,t} \quad (1)$$

$Pageviews_{n,t}$  denotes the number of page views for newspaper  $n$  in year-month  $t$ . I apply a log transformation to account for skewness, enabling interpretation of the coefficients as (approximate) percentage effects on page views.  $D_{n,t}^\tau$  is an indicator equal to one if and only if newspaper  $n$  introduces a paywall  $\tau$  months after  $t$ .  $\alpha_n$  and  $\gamma_t$  denote newspaper and year-month fixed effects, respectively. I cluster standard errors at the newspaper level.

To account for differential trends in newspapers' page views as well as shocks to their respective audiences over time, I include in  $\mathbf{X}_{n,t}$  characteristics of the markets in which the newspapers operated as of 2010, interacted with year-month fixed effects. Specifically, using the counties in which the newspaper is circulated, I compute the average population density, the shares of population with yearly income below 50k and above 100k, the share of college-educated individuals, and the NaNDA partisanship index, weighted by the number of print copies sold in each county.

For this analysis, I restrict my sample as follows: First, I exclude the four newspapers that introduced new premium websites as part of their paywalls. Second, I exclude the newspaper Newsday, which introduced a paywall in October 2009, just before the sample period. Third, I exclude three newspapers with significant data gaps, likely due to data collection errors by Alexa Internet. The final sample comprises the monthly page views for 72 newspapers from 2010 to 2017. Among these newspapers, two introduced paywalls before the sample period, 55 during the sample period, eleven afterward, and four never introduced a paywall (as of 2022). In my preferred specification, I assign the two always-treated and eleven later-treated newspapers as untreated, in addition to the four never-treated. These newspapers arguably act as valid control units because they implemented their paywalls well outside the sample period. I demonstrate that my results are robust to varying the composition of the control group.

### 3.2 Identification

Identification of the monthly treatment effects  $\beta_\tau$  relies on the assumption that page views of paywalled newspaper websites would have evolved parallel to those of non-paywalled websites in the absence of paywalls. In this sub-section, I discuss several challenges to this assumption.

#### Heterogeneous effects

I account for potential biases from heterogeneous treatment effects across units or time, which may arise when estimating Equation (1) via Ordinary Least Squares (for example, see [Baker et al., 2022](#)). I use the robust estimator proposed by [Callaway and Sant'Anna \(2021\)](#), which aggregates doubly-robust estimates for treatment group-time effects using only "clean" control units ([Sant'Anna and Zhao, 2020](#)). Control variables are included in each of those group-time-specific estimates both as a regression adjustment and for the construction of probability weights. Additionally, I employ the robust estimator by [Sun and Abraham \(2021\)](#), which estimates treatment effects by including only treated and never-treated units in the outcome regression and re-weighting the estimates accordingly.

#### Selection

Since the decision to introduce a paywall is endogenous, paywalls may correlate with unobserved factors not accounted for in the regression. To mitigate concerns about selection, I do the following. First, I show that pre-treatment trends in page views are parallel, ensuring that such trends are not biasing the result. Second, I control for market-based factors interacted with month-year fixed effects, which accounts both for different incentives for newspapers to introduce a paywall, and for time-varying shocks to different types of newspaper audiences. Third, to address concerns about selection in the control group, I show that results are statistically and economically significant for different compositions of the control group. Fourth, I explore heterogeneity in treatment effects to ensure that the results are not just driven by a small subset of newspapers.

#### Spillovers

Another potential concern involves biases from spillover effects: Readers displaced by a paywall may switch to a newspaper in the control group, which leads to an upward bias (in absolute terms) in the estimate for the paywall effect. However, in this setting, spillovers are unlikely to be a first-order concern due to the segmented nature of the newspaper market: All but three newspapers are classified as regional by AAM, catering primarily to their respective regional markets. As a result, these newspapers likely act as poor substitutes for one another. For example, it seems unlikely that a reader of the *Chicago Tribune* would respond to a paywall by switching to the *Los Angeles Times*.

To address any remaining concerns, I verify that my results hold when removing the largest paywalled newspapers and the smallest control newspapers. Intuitively, paywalls on larger

newspapers displace more readers, whose substitution has a larger relative effect on the control newspapers the fewer readers these newspapers have.

### 3.3 Results

#### Main result

Figure 3 presents estimates for the dynamic effect of paywalls on website page views, using the robust Callaway and Sant'Anna (2021) estimator. On average, this "first wave" of paywalls led to a sharp and persistent reduction in the number of page views for paywalled newspapers. Table 2 reports estimates for the static effect. On average, page views decreased by 25-30 percent, depending on the choice of control variables. Figure 4 and Table 3 demonstrate that these results are robust to using alternative estimators.

[Figure 3 here]

[Table 2 here]

#### Interpretation

The dynamic estimates reject the presence of pre-treatment trends, indicating that newspapers did not select into introducing paywalls based on prior trends in page views. Additionally, the sharp effect following paywall introduction is consistent with the idea that a paywall increases the opportunity cost of accessing the affected website, causing some readers to reduce their visits. Conversely, the paywall might also induce subscribers to use the website more frequently. Therefore, the estimated effect should be interpreted as the net effect of these two opposing mechanisms. However, because the data lack information on individual users, it is not possible to determine whether the results are driven by the displacement of a few highly active users or by many, more casual users.

The timing of the effects is inconsistent with alternative explanations that rely on gradual changes that introduced alongside the paywalls, such as shifts in topic composition or slant. Effects of such changes would likely materialize more slowly. Nonetheless, I cannot entirely rule out the possibility that these factors contributed to the slight downward trend observed in the treatment effects.

However, other long-term trends may also account for this downward trend. For instance, the diffusion of broadband and 3G internet during the sample period likely increased overall interest in online news. If new users entering the online news market disproportionately favored non-paywalled over paywalled websites, the initial effect of paywalls would become more pronounced over time.

How many individuals respond to paywalls by switching to the print version of the paywalled newspaper? While I cannot provide a definite answer based on the available data, this channel is unlikely to explain the observed effects. During the sample period, digital news consumption grew rapidly while print news consumption declined: The growth in the number of newspaper

website visits often exceeded 10% *per year*, while the number of print copies sold declined by around 30% between 2009 and 2017 (Pew Research Center, 2023). Moreover, studying the *New York Times* paywall, Pattabhiramaiah et al. (2019) conclude that spillovers from digital to print were negligible, with a 17 percent decline in website visits accompanied by only a 1-4 percent increase in print circulation. This increase was primarily driven by bundling online subscriptions with the print version, rather than by substitution to standalone print subscriptions.

### Further robustness

In this sub-section, I confirm that my results are robust to changes in the choice of which newspapers constitute the control group, and are not driven by users' substitution between paywalled and non-paywalled newspapers. For a detailed discussion of these potential concerns, see Section 3.2.

Table 4 shows results under different control group compositions. Column (2) shows that excluding the pre-treatment period observations for the 55 not-yet-treated newspapers from the control group has minimal impact on the results. Columns (3) and (4) adopt the most restrictive definition for the control group, comparing newspapers that implemented a paywall during the sample period to those that introduced one just afterwards. If anything, the effects are larger under these specifications, with a 34-36 percent decline in page views compared to the baseline estimate of a 28 percent decrease.

[Table 4 here]

Table 5 shows results when excluding different sets of newspapers from the sample. Columns (1) and (2) indicate that excluding large newspapers, in particular the three national newspapers and the fifteen largest newspapers by page views, increases the estimate slightly. Columns (3) and (4) exclude the fifteen largest paywalled newspapers and the five smallest non-paywalled newspapers. If spillovers were significant, one would expect the absolute magnitude of the estimates to decrease under these restrictions. However the opposite is true: the estimates are either unchanged or increase slightly. Intuitively, while theoretically an issue, spillovers may not play a significant role empirically because of the regional segmentation of the news markets, making most newspapers poor substitutes for one another. Moreover, the control group is sufficiently large to dilute any biases.

[Table 5 here]

## 4 Effect on survey outcomes

In this section I explore how paywalls affected political outcomes, such as voters' knowledge about US politics as well as participation in elections, by leveraging temporal and regional variation in exposure to paywalls.

#### 4.1 Measuring regional exposure to paywalls

A key challenge is the lack of sufficiently detailed data linking web traffic from specific counties to individual websites.<sup>9</sup> As a practical alternative, I assign national page views to counties proportionally based on newspapers' print circulation. Formally, I approximate county-level page views as:

$$\text{Pageviews}_{c,2010}^n = \text{Total Pageviews}_{2010}^n \times \frac{\text{Circulation}_{c,2011}^n}{\text{Total Circulation}_{2011}^n} \quad (2)$$

Here,  $\text{Circulation}_{c,2011}^n$  represents the number of print copies sold for newspaper  $n$  in county  $c$  in the year 2011.<sup>10</sup> The other variables are defined analogously. This approach assumes that at the aggregate (county) level, online and print news consumption are complementary. Intuitively, if a newspaper is popular in print in a given county, it is likely also popular online.

A potential concern is that online and print consumption might also act as substitutes, which would introduce measurement error. For instance, older individuals may consume more print and less online news, leading to overestimation of county-level page views in counties with older populations. If counties' population age correlates with survey outcomes (conditional on controls), this would bias my estimates. To address this concern, I verify that controlling for regional characteristics does not substantially affect the results of my analysis.

To further validate this measure, I compare the approximated regional page views with Google search interest for newspapers. These variables should be positively correlated, as both capture regional interest in the online content of a given newspaper. Figure 5 presents a scatter plot of newspaper-DMA-level page views and Google search interest, along with OLS fitted values. The slope coefficient is highly significant and  $R^2$  is 0.24, confirming that regional page views capture meaningful regional variation of interest in these news websites.

[Figure 5 here]

I use approximated regional page views to construct my main measure of regional paywall exposure. This measure represents the share of online newspaper consumption that has been affected by paywalls. Specifically, I define paywall exposure as the proportion of county  $c$ 's *initial* (2010) newspaper website views that are subject to a paywall by year  $t$ . Formally:

$$\text{Paywall Exposure}_{c,t} = \sum_{n \in \text{Newspapers}} \frac{\text{Pageviews}_{c,2010}^n \times I(\text{Paywall}_t^n)}{\text{Total Pageviews}_{c,2010}^n} \quad (3)$$

---

<sup>9</sup>For example, in the beginning of my sample, *comScore*'s browser panel of 50,000 individuals records less than 100 unique visitors per week to most newspaper websites outside the top 3.

<sup>10</sup>By using circulation numbers from 2011, I include print circulation figures recorded *after* the implementation of a paywall for ten newspapers in my sample. However, two facts should mitigate corresponding concerns: First, substitution from digital to print appears to have been relatively uncommon. Second, such substitution would only matter if it changed the relative *shares* of circulation across counties, rather than the absolute *numbers*.

where  $I(\text{Paywall}_t^n)$  is a binary variable that indicates whether newspaper  $n$  has implemented a paywall by year  $t$ . I fix page views at their 2010 levels, prior to the widespread introduction of paywalls, to avoid endogeneity arising from paywalls themselves affecting page views.<sup>11</sup>

This measure of paywall exposure shares similarities with shift-share variables (see [Bartik, 1991](#); [Blanchard et al., 1992](#)), where initial market shares serve as shares and paywall introductions act as shifts. However, since paywall exposure will be used as a treatment variable within a difference-in-differences framework rather than as an instrumental variable, it does not require the same exogeneity conditions.

To provide further insights into the variation captured by paywall exposure, Figure 6 presents key descriptive statistics. Panel (a) shows that paywall exposure increased significantly between 2010 and 2014, reaching an average of approximately 50 percent. Panel (b) highlights substantial variation across counties, with some experiencing minimal exposure, while others see almost their entire initial consumption paywalled. To account for county population, Panel (c) displays the distribution of paywall exposure among respondents in the CES survey. The results reveal two groups: one with very high paywall exposure, for which 80–90 percent of news consumption is affected by paywalls, and a tail with low to intermediate exposure. Figure 7 shows geographic variation, revealing that paywalls affected all major regions of the US.

[Figure 6 here]

[Figure 7 here]

To gain a deeper understanding of which regions are most likely to be affected by paywalls, Figure 9 presents the results of a regression of paywall exposure in 2017 on county characteristics in 2010. Paywalls disproportionately impacted high-income, Democratic-leaning counties. To ensure that my subsequent analysis does not pick up unrelated trends in these regions that may be correlated with outcomes, I will verify that results are robust to controlling for baseline county characteristics. Interestingly, paywalls do not appear to be correlated with political knowledge conditional on controls, which mitigates concerns about omitted variables or selection based on the outcome.

[Figure 9 here]

---

<sup>11</sup>In principle, I could predict post-2010 page views based on trends in 2010. However, I refrain from this approach for two reasons. First, the training period would be relatively short, as the page view data only begin in 2010, reducing the precision of such predictions. Second, as demonstrated in Section 3.3, paywalls were, on average, not driven by pre-existing trends in page views, so that this extension would likely provide limited additional contribution.

## 4.2 Empirical strategy

I exploit the staggered exposure of counties to paywalls using a difference-in-differences design. For this purpose, I divide counties into a low- and a high-paywall exposure group based on the median paywall exposure among survey respondents in 2017, which is approximately 0.72. This way, survey respondents are equally split between treatment and control group, which helps ensure high power. In robustness checks, I confirm that the results are robust to varying this threshold.

The setting is particularly suited for discretization: In most regions, only few newspapers have large audiences, so if paywall exposure exceeds the threshold in a given year, it is likely due to a major newspaper implementing a paywall. Moreover, in regions with multiple large newspapers, the addition of a later paywall is arguably more influential since fewer non-paywalled substitutes are available.

My analysis is based on the following specification:

$$y_{i,c,t} = \sum_{\tau \neq -1} \beta_\tau D_{c,t}^\tau + \alpha_c + \Gamma(d)_c \times \gamma_t + \delta' \mathbf{X}_{i,c,t} + \varepsilon_{i,c,t} \quad (4)$$

The dependent variable  $y_{i,c,t}$  is an outcome variable for survey respondent  $i$  residing in county  $c$  in year  $t$ , such as the share of correctly answered political knowledge questions.  $D_{c,t}^\tau$  is an indicator equal to one if and only if paywall exposure in county  $c$  first exceeds the population-weighted median of 0.72  $\tau$  years after  $t$ . County fixed effects  $\alpha_c$  account for pre-existing regional differences outcomes. DMA-by-year fixed effects  $\Gamma(d)_c \times \gamma_t$  capture time-varying shocks at the newspaper market level. Note that the specification does not include individual fixed effects because the survey is a repeated cross section in which each individual appears only once.

$\mathbf{X}_{i,c,t}$  denotes time-varying control variables. In the baseline specification, I include individual demographic variables interacted with year fixed effects that account for changes in outcomes of different demographic groups over time.<sup>12</sup> In robustness checks, I also add county-level demographics<sup>13</sup> as well as partisanship at their 2010-levels interacted with year fixed effects, which capture shocks to counties with specific characteristics. I cluster standard errors at the state level.

I estimate Equation (4) using the robust [Sun and Abraham \(2021\)](#) estimator. Its regression-based nature allows for the straightforward inclusion of the control variables and fixed effects specified above. Consequently, it is better suited to this context than the estimator by [Callaway and Sant'Anna \(2021\)](#). The control group is comprised of survey respondents in counties that never become treated, and in counties that become treated after 2017. Among the latter, I exclude counties from my sample once their paywall exposure exceeds the threshold. Conse-

---

<sup>12</sup>The variables, categorized into equally-sized buckets with the number of levels in parentheses, are sex (2), age (3), ethnicity (2), education (4), and income (3).

<sup>13</sup>The variables are log population, log population density, male share, white share, median age, log median income, and college-educated share.

quently, my estimates should be interpreted as effects of the first wave of paywalls that were introduced between 2011 and 2017.

### 4.3 Identification

Identification requires that, in the absence of paywalls, outcomes in low-exposure and high-exposure regions would have evolved the same. As before, I employ a robust estimator which ensures that comparisons are made between appropriate units and that group-time treatment effects are aggregated correctly.

The key identification assumption would be violated in the presence of trends in outcomes that are unrelated to paywalls but correlated with paywall exposure. For instance, individuals with higher income might be more likely to substitute from newspapers to social media, even in the absence of paywalls, and the corresponding shift from hard news to soft news may reduce political knowledge. Since county-level income is positively correlated with paywall exposure, failure to adequately control for income would introduce downward bias in the estimated effect of paywalls on knowledge.

The following remedies address these concerns. First, I interact individual control variables with year fixed effects to allow the correlation between individual characteristics and outcomes to vary by year. This accounts for shocks to specific subpopulations, such as different propensities to adopt social media or to develop sufficient willingness to pay for news. Similarly, interactions between baseline county characteristics and year fixed effects address both shocks to counties with higher paywall exposure and potential non-random measurement error in paywall exposure that may correlate with observable variables. While the individual controls are expected to have explanatory power for the reasons mentioned above, I verify that additionally including the county controls does not significantly alter the results. Intuitively, if the residual variation in the outcome shows little correlation with *observed* regional variables, it is less likely that there is an *omitted* regional variable with sufficiently strong correlation with both the outcome and paywall exposure to drive the results.

Second, I include DMA-by-year fixed effects to capture time-varying shocks to media markets. These fixed effects account both for observable shocks, such as newspaper exits, and for unobservable shocks, such as shifts in aggregate preferences for social media platforms.

Any remaining threats to identification would need to originate from variation within DMA-year that is correlated with the outcome and paywall exposure, but uncorrelated with both individual- and county-level controls.

### 4.4 Results

#### Main result

Figure 10 shows the dynamic effect of paywalls on the share of correctly answered political knowledge questions, comparing survey respondents in counties with high versus low paywall

exposure. The effect begins to materialize in the year a county becomes treated and remains persistent in subsequent years. Importantly, the results are not driven by pre-treatment trends in political knowledge. If anything, the prior trends are slightly positive, in which case the estimated effect would constitute a lower bound on the true impact of paywalls.

Table 6 demonstrates that this result is robust to the inclusion of a wide range of time-varying controls. Column (1) includes only fixed effects, while column (2) represents the preferred specification that includes time-varying individual demographics. Columns (3) through (5) add county demographics, county partisanship, and state-year fixed effects, respectively. After controlling for survey respondent characteristics, adding regional controls does not significantly alter the results. This suggest that the selection of paywalls into specific regions, as well as non-random measurement error in paywall exposure based on county characteristics, are of limited importance.

[Figure 10 here]

[Table 6 here]

### Further robustness

The findings are robust to varying the threshold that defines whether counties are classified as high- or low-paywall exposure. Figure 11 presents estimates for the baseline specification under different threshold levels. For easier comparison, the figure also reports the associated shares of survey respondents in the control group. The effects remain quantitatively similar and at least marginally significant for a wide range of control group sizes, ranging from approximately 35 to 70 percent.

[Figure 11 here]

Moreover, the results are not driven by particularly small or large counties. Table 8 shows that the estimates remain similar when excluding up to 30 percent of either the smallest or the largest counties in the sample. Importantly, this insight also provides evidence that neither imprecisely measured circulation in small counties nor the omission of small newspapers introduce serious measurement error in my analysis.

[Table 8 here]

### Index decomposition

To better understand which types of political knowledge are affected by paywalls, I decompose the knowledge index by estimating eight separate regressions, each using as the outcome one of the index components. Figure 12 shows results. Estimates in black indicate conventional 95 percent confidence intervals, while those in red apply Bonferroni adjustments that account for multiple hypothesis testing. The four upper rows correspond to questions regarding party

affiliation given the name of a political representative, while the four lower rows pertain to questions about the majority party in state or federal legislative bodies.

The results show that the effects on political knowledge are primarily driven by reduced knowledge of political representatives. To a smaller extent, reduced knowledge about state-level legislative majorities also play a role, while knowledge of federal majorities has no observable effect.

[Figure 12 here]

These findings can be interpreted in two ways. On one hand, they may reflect the different *geographical* focus of these questions. Paywalls may have shifted news consumption from regional to national news sources, resulting in reduced awareness of politicians elected on the state level (governor, senators, state majorities) and on the congressional district level (US House representatives), but not knowledge on national party majorities. This explanation aligns with the observation that many digital substitutes for newspaper websites, such as the websites by ABC, CNN, and Fox, as well as social media platforms, tend to prioritize coverage of national politics.

On the other hand, the results may also reflect differences in the *depth* of knowledge necessary to answer the questions. Questions about national politics may require less detailed knowledge and may be answered correctly even with a lower understanding of (and interest in) political affairs. As such, popular substitutes of paywalled websites may not provide sufficiently deep information to correctly answer the more difficult questions, but enough for the easier ones.

### Distributional effects

Who is most affected by paywalls? To answer this question, Figure 13 shows results from estimating the baseline specification separately for subgroups of the survey respondents, defined either by individual or county characteristics.

Individuals with low income and low education levels experience larger declines in political knowledge. This finding is consistent with the notion that these individuals face relatively higher monetary costs of purchasing a paywall subscription, and may derive lower subjective benefits from high-quality information, either due to less interest in or a lower perceived value of such content.

At the county level, Democratic, urban, and suburban counties exhibit larger effects. These characteristics align with typical consumers of mainstream news. Since paywalls primarily affected mainstream news outlets, it is reassuring that the groups most likely to consume such news experience the largest effects.

[Figure 13 here]

## Magnitudes

In this section I discuss the magnitude of the estimated effect of paywalls on political knowledge. To facilitate comparisons with other studies in the media literature, I present a back-of-envelope calculation for the "persuasion rate" of paywalls, defined as the share of individuals who switched from the correct to the incorrect answer because of paywalls, among those exposed to paywalls.<sup>14</sup> Following the notation of [DellaVigna and Gentzkow \(2010\)](#), the persuasion rate for this setting is calculated as:

$$f = \frac{y_T - y_C}{e_T - e_C} \frac{1}{1 - y_0} \quad (5)$$

$$= \frac{0.023}{0.5 * 0.5} \frac{1}{0.745} = 0.123 \quad (6)$$

Here,  $y_T - y_C$  is the change in the outcome between the treated and the control group in response to the treatment. To provide a conservative figure, I use the estimate from my most restrictive specification, which is 0.023.  $e_T - e_C$  denotes the difference in the share of the population exposed to the treatment. I assume that around 25 percent of the voting-age population is affected, which accounts for the fact that around 50 percent of the US voting population read newspapers online during the sample period, and that the difference in average paywall exposure between treated and control group is roughly 50 percentage points.  $1 - y_0$  represents the share of the population left to be persuaded. In this context, I use the average share of correct questions in the population at baseline, which is 0.745 (calculated using the years 2006 through 2010).

The resulting magnitude is 12.3 percent. This means that approximately 12 percent among individuals who read a newspaper that implemented a paywall lost some political knowledge of the type tested in the survey.

While informative, this calculation has some limitations. First, since I measure knowledge as an index comprised of several outcomes, the estimates reflect the effect of paywalls on the share of answers in the population, rather than on the number of individuals who lose knowledge. Therefore, the calculation implicitly assumes that the effect of paywalls on knowledge is homogeneous across individuals. Consequently, if reductions knowledge are concentrated among fewer individuals, the true persuasion rate is higher; if many individuals lose a smaller amount of knowledge, the rate is lower.

Second, the calculation should be interpreted as the intent-to-treat effect of being exposed to a paywall. It represents the net effect of two opposing mechanisms: increases in consumption by some individuals who purchase a subscription, and decreases in consumption by others who substitute away. Without more detailed data to disentangle these mechanisms, it is not possible to compute a persuasion rate specifically for the subset of readers who reduce their consumption due to paywalls.

---

<sup>14</sup>In this setting, the term "persuasion rate" applies imperfectly, as individuals do not change their *opinion* in response to the treatment.

The estimated persuasion rate of 12.3 is in the medium range of estimates in the literature on the effects of news availability on political outcomes (DellaVigna and Gentzkow, 2010). For example, Enikolopov et al. (2011) find a persuasion rate of 7.7 percent from the availability of the independent anti-Putin TV station *NTV* on reducing the vote share for pro-Putin parties. Similarly, DellaVigna and Kaplan (2007) estimate an 11.6 percent persuasion rate from the availability of *Fox News* on the Republican vote share in presidential elections. Gerber et al. (2009) find a rate of 19.5 percent for free subscriptions to the *Washington Post* increasing intent to vote Democratic. Gentzkow et al. (2011) find a 12.9 percent persuasion rate for local newspaper readership in the late 1800s and early 1900s on voter turnout in presidential elections, while Gentzkow (2006) reports a 4.4 percent persuasion rate for early television exposure reducing congressional election turnout.

### **Effect on electoral participation**

To estimate the effects of paywalls on electoral participation, I use validated measures of whether a survey respondent voted in a US congressional election - either for Senator or House Representative - or in a congressional primary election.<sup>15</sup> Since elections take place every two years, I collapse the specification from Equation (4) to a bi-yearly frequency. In addition to county and DMA-by-year fixed effects, I control for individual characteristics.<sup>16</sup> Figure 14 present results. Panel (a) suggests that paywalls might have negatively affected participation in congressional *primary* elections. Panel (b) points to a similar but less persistent pattern for the congressional elections. Table 12 shows that for both outcomes, the average effect across all periods is not significant at conventional levels. However, for participation in congressional primaries, the estimated 2.4 percentage point decrease has a p-value of 0.17, which is partly due to the fact that the effect materializes rather slowly. Moreover, the point estimate is sizable given the baseline participation rate of 42 percent.

These findings align with the interpretation that switching away from politics-intensive news reduces attention to political events. The more pronounced effect for congressional primaries might reflect that following these lower-level contests requires deeper engagement with politics, making their turnout more vulnerable to changes in news consumption.

[Figure 14 here]

[Table 12 here]

---

<sup>15</sup>I do not consider presidential and gubernatorial elections because they occur too infrequently in my sample period for reliable inference

<sup>16</sup>Here, I do not interact individual characteristics with year fixed effects, as doing so proves too demanding for these outcomes.

## 5 Conclusion

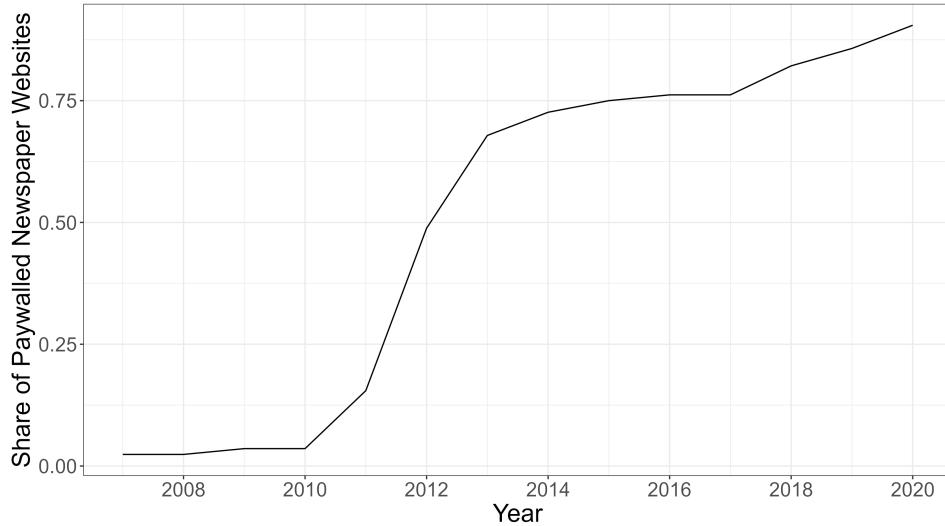
In this paper, I show that paywalls on US newspaper websites reduced online consumption of traditional US newspapers. Regions more affected by paywalls experienced declines in knowledge about regional politics, especially among low-educated and low-income individuals. These results support the idea that these groups are more reluctant or unable to pay for digital news, leaving them disproportionately uninformed. Additionally, paywalls may have reduced participation in already less-frequented elections.

My findings underscore that news media play a critical role for the democratic process. In a media landscape dominated by private, profit-oriented news outlets, economic shocks can induce adjustments that ultimately affect voters' information and participation. Therefore, to counteract the potential political effects of such economic shifts, modern democracies may profit from funding independent, high-quality news outlets.

To better understand both the mechanisms behind my results as well as potential remedies, further research could explore how readers substitute paywalled websites, and examine supply-side changes to topics, slant, or quality that may accompany paywalls. Moreover, one could further investigate downstream effects on political opinions and voter behavior, particularly in regional elections. Finally, it would be valuable to determine whether the displacement of readers from traditional media contributed to the rise of right-wing populism in the following years.

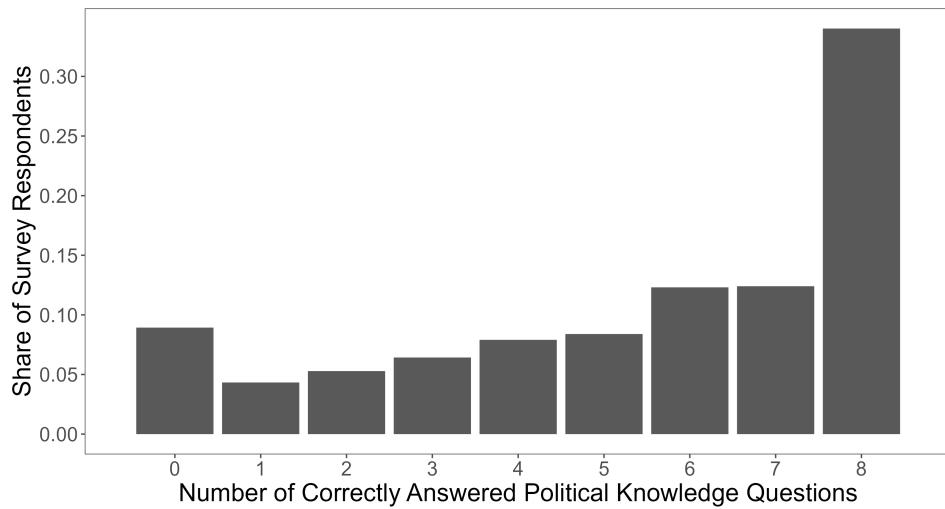
## 6 Figures

Figure 1. Share of newspapers with paywalls



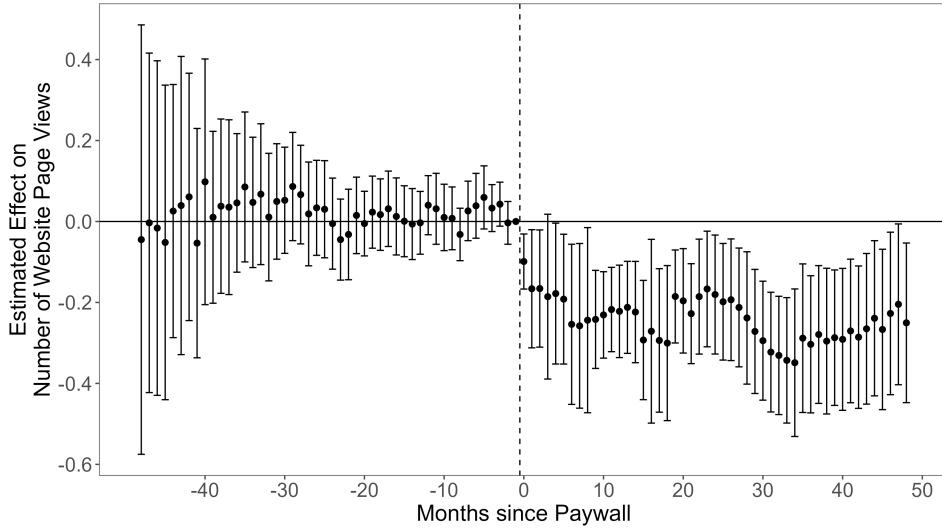
Notes: Share of US newspapers that have ever implemented a paywall on their website, among largest 80 newspapers by number of page views in 2010.

Figure 2. Distribution of Number Correctly Answered Political Knowledge Questions



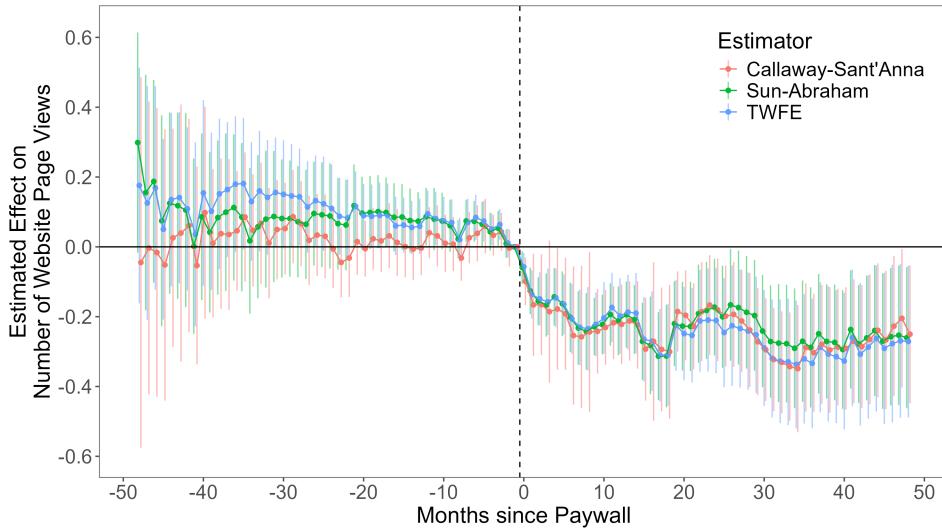
Notes: Distribution of the number of correctly answered political knowledge questions by survey respondent. For the years 2006 and 2009, in which only six of the eight questions were asked, as well as for the 1 percent of respondents who do not answer all questions, I round the share of correctly answered questions, multiply by eight, and round to the closest integer.

Figure 3. Effect of paywalls on website page views



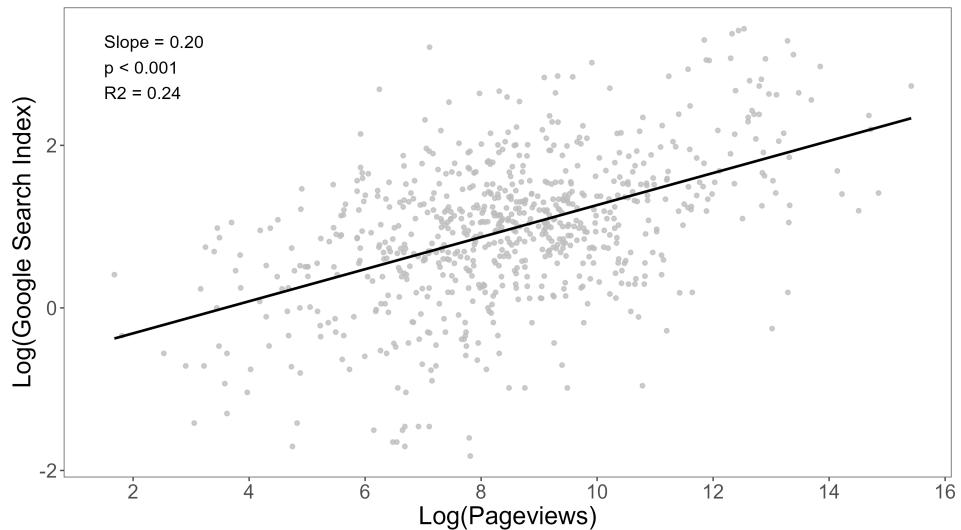
*Notes:* Monthly coefficients for  $\beta_\tau$  from Equation (1): Regression of monthly number of newspaper website views on indicators denoting the number of months since paywall implementation. The omitted category is the month before the paywall. Includes newspaper fixed effects, year-month fixed effects, and log population density, shares of three income buckets, college-educated share, and partisanship index for newspapers' audiences, interacted with year-month fixed effects. Estimated using Callaway-Sant'Anna estimator. Vertical bars denote 95 percent pointwise confidence intervals. Standard errors are clustered by newspaper.

Figure 4. Effect of paywalls on website page views - Estimators



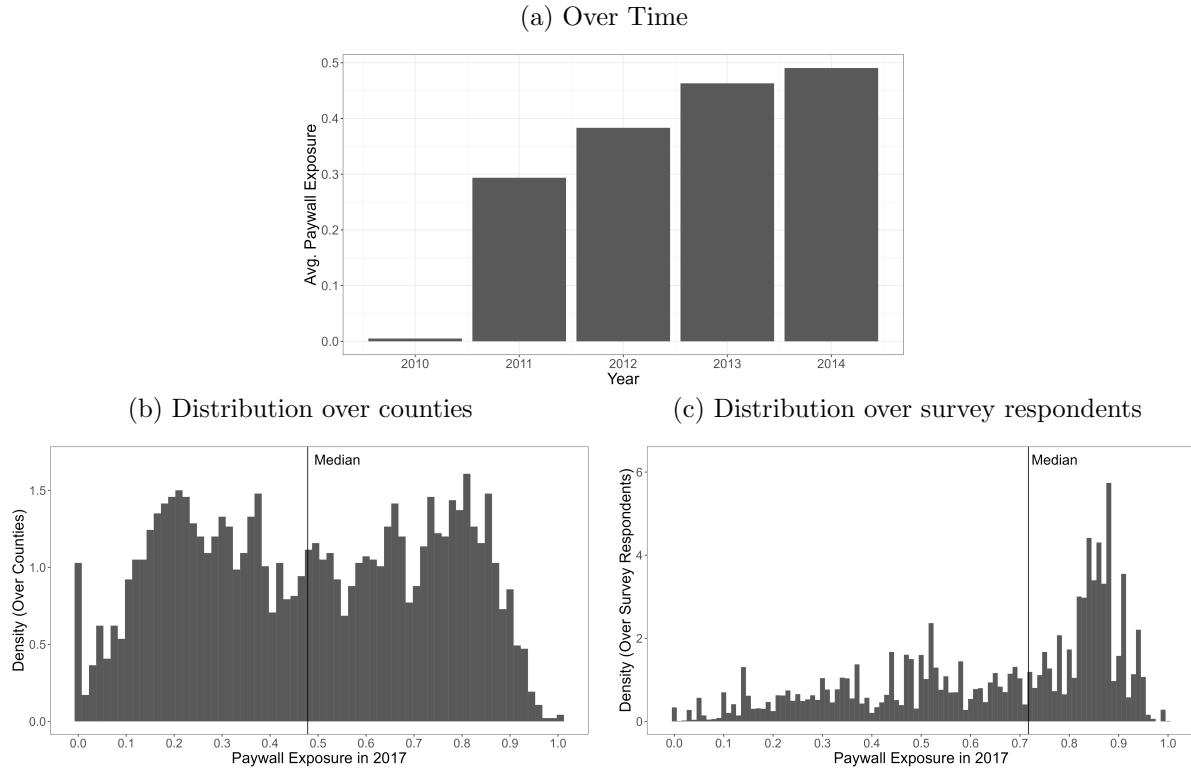
*Notes:* Monthly coefficients for  $\beta_\tau$  from Equation (1), estimated either via Callaway-Sant'Anna, Sun-Abraham, or OLS (TWFE): Regression of monthly number of newspaper website views on indicators denoting the number of months since paywall implementation. The omitted category is the month before the paywall. Includes newspaper fixed effects, year-month fixed effects, and log population density, shares of three income buckets, college-educated share, and partisanship index for newspapers' audiences, interacted with year-month fixed effects. Vertical bars denote 95 percent pointwise confidence intervals. Standard errors are clustered by newspaper.

Figure 5. Correlation of approximated page views and Google search index on newspaper-DMA level



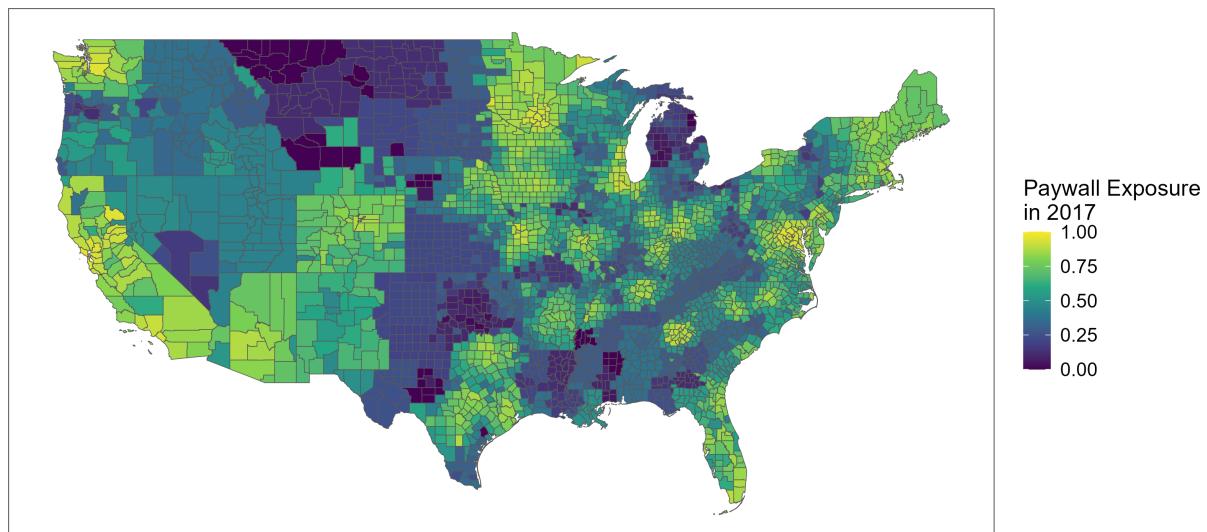
*Notes:* Scatter plot and linear regression for the following variables. X-Axis: Log of number of newspaper page views from DMA in 2010, on DMA-newspaper level. Obtained by assigning national newspaper-specific page views to counties proportionally by newspapers' print circulation in county, as defined in Equation (2), and aggregating to DMA-level. Y-Axis: Log of Google search index for newspaper in DMA in 2010, on DMA-newspaper level. The solid line indicates OLS fitted values. The raw correlation between the variables is 0.49.

Figure 6. Paywall Exposure



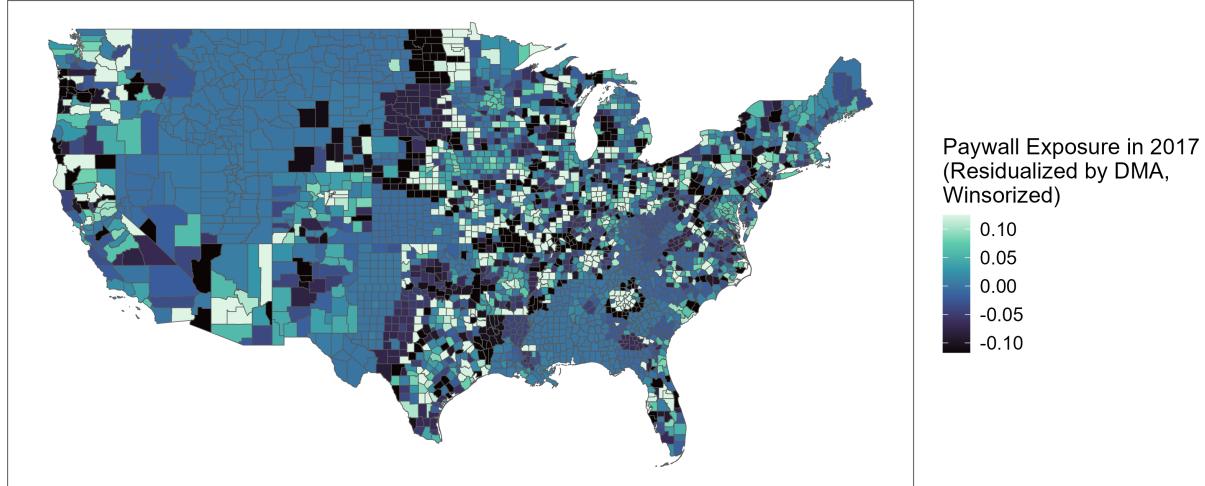
*Notes:* Panel (a): Average paywall exposure across counties by year. Panels (b) and (c): Distribution of paywall exposure over counties and survey respondents, respectively, in 2017.

Figure 7. Geographical variation of paywall exposure



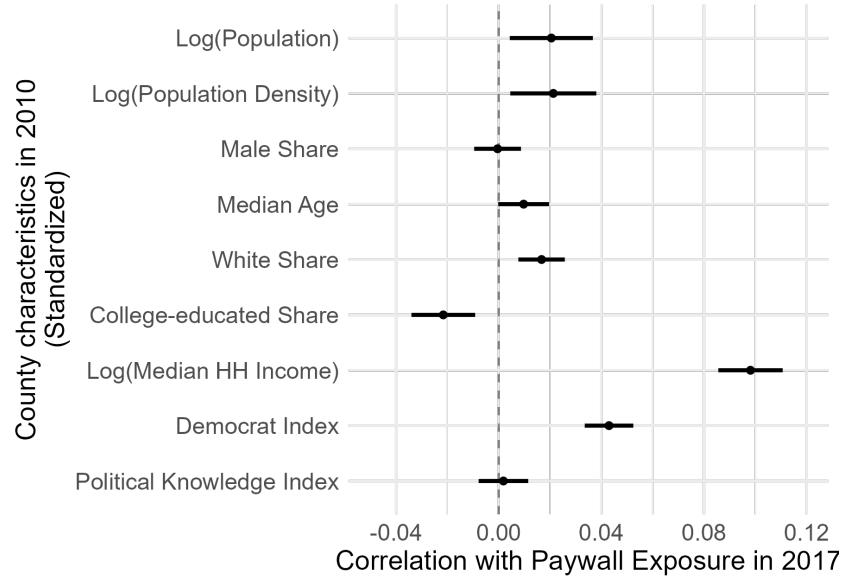
*Notes:* Paywall exposure by county in 2017.

Figure 8. Geographical variation of paywall exposure - Residualized



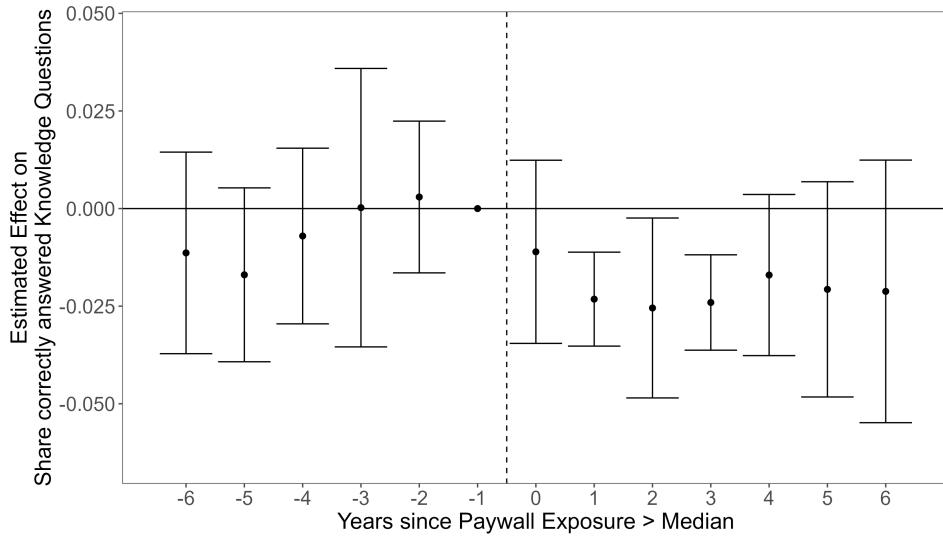
*Notes:* Paywall exposure by county in 2017, residualized by DMA and winsorized at 10 and 90 percent.

Figure 9. County-level predictors of paywall exposure



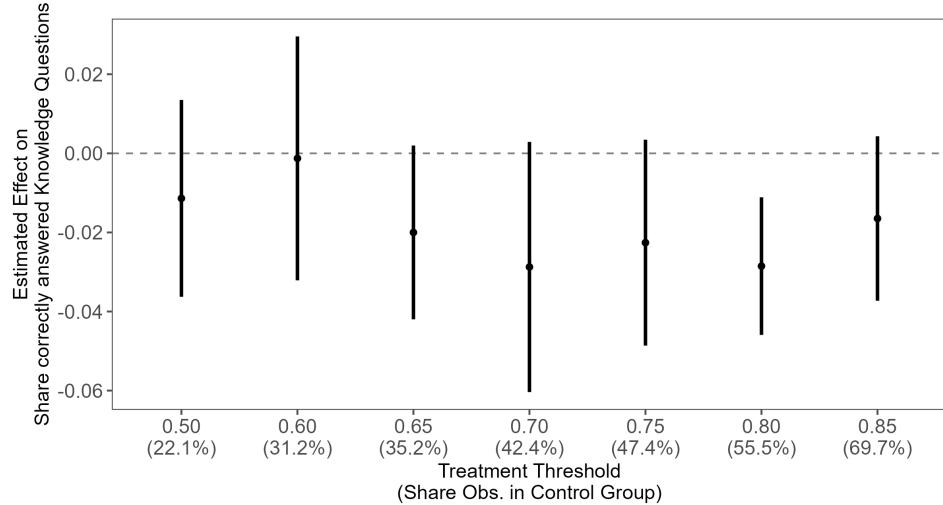
*Notes:* Multivariate regression of county-level paywall exposure in 2017 on county variables in 2010. All variables are standardized. Horizontal bars denote 95 percent confidence intervals.

Figure 10. Effect of paywalls on political knowledge



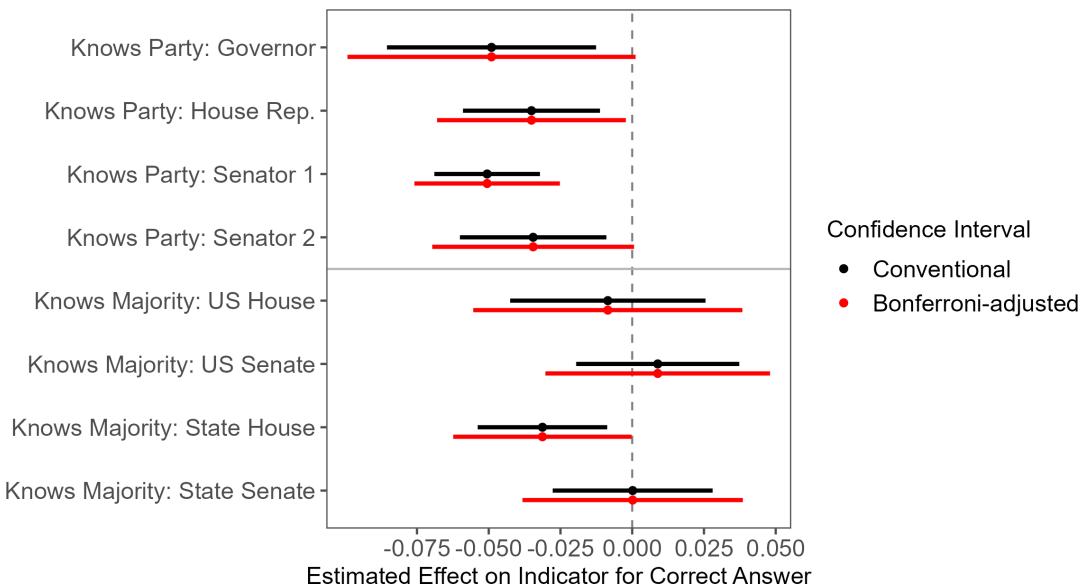
*Notes:* Yearly coefficients for  $\beta_\tau$  from Equation (4): Regression of share correctly answered political knowledge questions by CES survey respondent on indicators denoting the number of years since respondent's county exceeded threshold of paywall exposure of 0.72. The omitted category is the year before exceeding the threshold. Includes county fixed effects, DMA-year fixed effects, and the following individual characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Vertical bars denote 90 percent pointwise confidence intervals. Standard errors are clustered by state.

Figure 11. Effect of paywalls on political knowledge - Thresholds



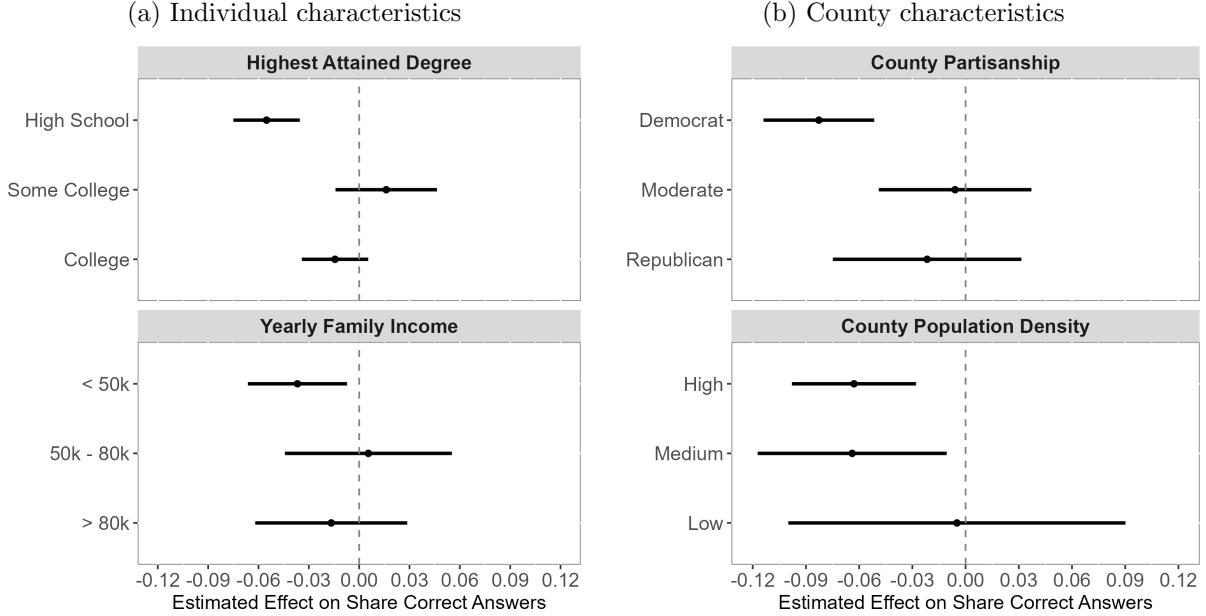
*Notes:* Coefficients for static (pre-post) aggregation of  $\beta_\tau$  from Equation (4), for different thresholds that define whether county switches from low- to high-paywall exposure. Resulting share of survey respondents in never-treated (always low-paywall-exposure) group in parentheses. Regressions of share correctly answered political knowledge questions by CES survey respondent on indicators denoting the number of years since respondent's county exceeded threshold of paywall exposure of 0.72. The omitted category is the year before exceeding the threshold. Includes county fixed effects, DMA-year fixed effects, and the following individual characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Vertical bars denote 95 percent confidence intervals. Standard errors are clustered by state.

Figure 12. Effect of paywalls on political knowledge - Index decomposition



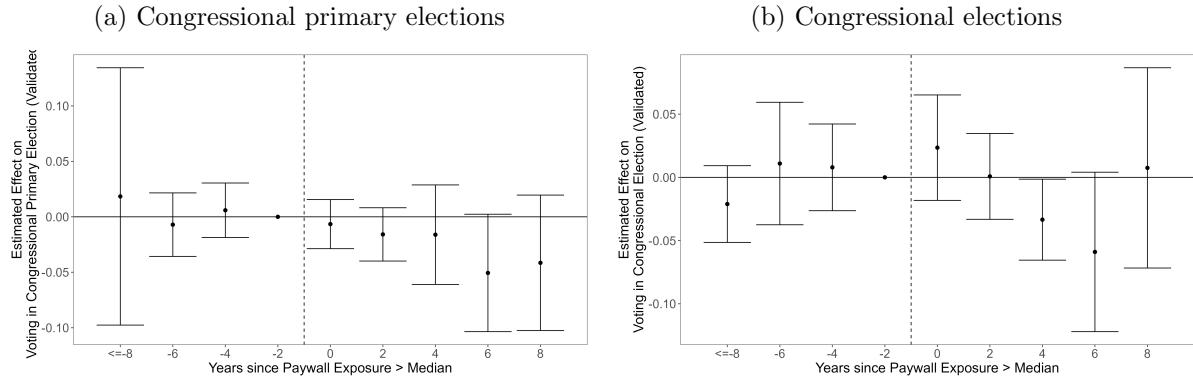
*Notes:* Coefficients for static (pre-post) aggregation of  $\beta_\tau$  from Equation (4), estimated separately for each component of political knowledge index. Regression of indicator for correct answer of CES survey respondent to knowledge question, on indicators denoting the number of years since respondent's county exceeded threshold of paywall exposure of 0.72. The omitted category is the year before exceeding the threshold. Includes county fixed effects, DMA-year fixed effects, and the following individual characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors are clustered by state. Horizontal bars denote 95 percent confidence intervals, either conventional (black) or with Bonferroni adjustment (red).

Figure 13. Effect of paywalls on political knowledge - Heterogeneity



*Notes:* Coefficients for static (pre-post) aggregation of  $\beta_\tau$  from Equation (4), estimated separately for different subgroups of respondents. Regression of share correctly answered political knowledge questions by CES survey respondent on indicators denoting the number of years since respondent's county exceeded threshold of paywall exposure of 0.72. The omitted category is the year before exceeding the threshold. Includes county fixed effects, DMA-year fixed effects, and individual characteristics interacted with year fixed effects. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors are clustered by state. Horizontal bars denote 95 percent confidence intervals.

Figure 14. Effect of paywalls on electoral participation



*Notes:* Yearly coefficients for  $\beta_\tau$  from Equation (4) collapsed to bi-yearly level. Outcomes are indicators for whether survey respondent voted in the respective election. Panel (a) denotes voting in any congressional primary election, and Panel (b) refers to voting for either US Senator or US House Representative in the corresponding congressional elections. Regressions on indicators denoting the number of years since respondent's county exceeded threshold of paywall exposure of 0.72. The omitted category is the year before exceeding the threshold. Includes county fixed effects, DMA-year fixed effects, and the following individual characteristics: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Vertical bars denote 95 percent pointwise confidence intervals. Standard errors are clustered by state.

## 7 Tables

Table 1. Summary statistics for variables derived from CES

	Mean	SD	Min	Max	N
<b>Political knowledge</b>					
Knows majority: US Senate	0.63	0.48	0	1	520,495
Knows majority: US House	0.66	0.47	0	1	523,703
Knows majority: State Senate	0.49	0.50	0	1	475,198
Knows majority: State House	0.48	0.50	0	1	474,936
Knows party: Governor	0.75	0.43	0	1	528,258
Knows party: Representative	0.61	0.49	0	1	524,663
Knows party: Senator 1	0.69	0.46	0	1	527,706
Knows party: Senator 2	0.68	0.47	0	1	527,682
Knowledge index: Share correct answers	0.63	0.35	0	1	530,966
<b>Vote participation (validated)</b>					
Voted in congressional election	0.63	0.48	0	1	323,365
Voted in congressional primary election	0.29	0.45	0	1	384,322
<b>Individual Characteristics</b>					
Female	0.52	0.50	0	1	531,087
White	0.72	0.45	0	1	556,746
College-educated	0.37	0.48	0	1	556,746
Family Income > 60k	0.35	0.48	0	1	556,746
Family Income > 100k	0.15	0.36	0	1	556,746
Democrat	0.38	0.49	0	1	506,342
Republican	0.30	0.46	0	1	506,342

Notes: Selected summary statistics of survey respondent-level variables derived from CES, averaged across all years and including survey weights.

Table 2. Effect of paywalls on pageviews

Dependent Variable:	Log(Pageviews)				
	(1)	(2)	(3)	(4)	(5)
Paywall	-0.257*** (0.070)	-0.291*** (0.062)	-0.302*** (0.064)	-0.280*** (0.063)	-0.310*** (0.076)
Newspaper FE	✓	✓	✓	✓	✓
Month-Year FE	✓	✓	✓	✓	✓
<b>Controls x Month-Year FE:</b>					
Log(Population Density)	✓	✓	✓	✓	✓
Sh. HH income $\leq 50k$		✓	✓	✓	✓
Sh. HH income > 100k		✓	✓	✓	✓
Partisanship index			✓	✓	✓
Sh. college-educated				✓	✓
Observations	6,912	6,912	6,912	6,912	6,912

Notes: Coefficients for static (pre-post) version of Equation (1): Regressions of monthly number of newspaper website views on indicator denoting active paywall. All specifications include newspaper fixed effects and year-month fixed effects. Controls represent average county characteristics at 2010-levels where newspaper is active, weighted by number of print copies sold, interacted with year-month fixed effects: Log population density, share of households with yearly income below \$50,000 and above \$100,000, share adults with college degree, and NaNDA partisanship index. Estimated using Callaway-Sant'Anna estimator. Standard errors in parentheses are clustered by newspaper. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 3. Effect of paywalls on pageviews - Estimators

Dependent Variable:	Log(Pageviews)					
	TWFE		Sun-Abraham		Callaway-Sant'Anna	
	(1)	(2)	(3)	(4)	(5)	(6)
Paywall	-0.258*** (0.059)	-0.273*** (0.060)	-0.267*** (0.074)	-0.279*** (0.074)	-0.257*** (0.069)	-0.280*** (0.065)
Newspaper FE	✓	✓	✓	✓	✓	✓
Month-Year FE	✓	✓	✓	✓	✓	✓
Market Controls		✓		✓		✓
Observations	6,912	6,912	6,912	6,912	6,912	6,912
R <sup>2</sup>	0.957	0.960	0.978	0.981		

*Notes:* Coefficients for static (pre-post) version of Equation (1), estimated via Ordinary Least Squares (TWFE), Sun-Abraham, or Callaway-Sant'Anna: Regressions of monthly number of newspaper website views on indicator denoting active paywall, controlling for newspaper fixed effects and year-month fixed effects. Columns (2), (4), and (6) additionally control for average county characteristics at 2010-levels where newspaper is active, weighted by number of print copies sold, interacted with year-month fixed effects: Log population density, share of households with yearly income below \$50,000 and above \$100,000, share adults with college degree, and NaNDA partisanship index. Standard errors in parentheses are clustered by newspaper. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 4. Effect of paywalls on pageviews - Control groups

Dependent Variable:	Log(Pageviews)				
	Control group:		w/o Treated	w/o Earlier/Later	
	All	(1)	(2)	(3)	Only Later (4)
Paywall	-0.280*** (0.067)	-0.277*** (0.070)		-0.342*** (0.068)	-0.355*** (0.068)
Newspaper FE	✓	✓		✓	✓
Month-Year FE	✓	✓		✓	✓
Market Controls	✓	✓		✓	✓
<b>Control group:</b>					
Earlier Treated	✓		✓		
Treated (pre-treatment)	✓			✓	
Later Treated	✓		✓	✓	✓
Never Treated	✓		✓		
Observations	6,912	6,912	6,336	6,336	

*Notes:* Coefficients for static (pre-post) version of Equation (1), for different compositions of the control group: "Earlier treated" are 2 newspapers that implemented paywall before 2010. "Treated (pre-treatment)" are 55 newspapers that implemented paywall between 2010 and 2017. "Later Treated" are 11 newspapers that implemented paywall after 2017. "Never Treated" are 4 newspapers that have not implemented a paywall by 2022. Regressions of monthly number of newspaper website views on indicator denoting active paywall. All specifications include newspaper fixed effects and year-month fixed effects. Controls represent average county characteristics at 2010-levels where newspaper is active, weighted by number of print copies sold, interacted with year-month fixed effects: Log population density, share of households with yearly income below \$50,000 and above \$100,000, share adults with college degree, and NaNDA partisanship index. Estimated using Callaway-Sant'Anna estimator. Standard errors in parentheses are clustered by newspaper. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 5. Effect of paywalls on pageviews - Restrict sample

Dependent Var.:	Log(Pageviews)				
	Exclude Newspapers				Add Weights
	National	Top 15	Top 15 Paywalled	Bottom 5 Control	Weights
	(1)	(2)	(3)	(4)	
Paywall	-0.282*** (0.065)	-0.332*** (0.071)	-0.357*** (0.073)	-0.320*** (0.075)	-0.273*** (0.068)
Newspaper FE	✓	✓	✓	✓	✓
Month-Year FE	✓	✓	✓	✓	✓
Market Controls	✓	✓	✓	✓	✓
Restricted Control Group			✓	✓	
Sample Weights					Log(Pageviews 2010)
Observations	6,624	5,472	5,472	6,432	6,912

*Notes:* Coefficients for static (pre-post) version of Equation (1), excluding different subsets of newspapers from the sample. Column (1) excludes the three national newspapers *New York Times*, *USA Today*, and *Wall Street Journal*. Column (2) excludes the largest 15 newspapers in 2010 by website page views. Columns (3) and (4) exclude the largest 15 newspapers in the treated group, and the bottom 5 newspapers in the control group, respectively. In both columns, treated observations pre-treatment are also excluded from the control group. Column (5) includes the log of 2010 page views as sample weights. Regressions of monthly number of newspaper website views on indicator denoting active paywall. All specifications include newspaper fixed effects and year-month fixed effects. Controls represent average county characteristics at 2010-levels where newspaper is active, weighted by number of print copies sold, interacted with year-month fixed effects: Log population density, share of households with yearly income below \$50,000 and above \$100,000, share adults with college degree, and NaNDA partisanship index. Estimated using Callaway-Sant'Anna estimator. Standard errors in parentheses are clustered by newspaper. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6. Effect of paywalls on political knowledge

Dependent Variable:	Share Correct Answers				
	(1)	(2)	(3)	(4)	(5)
Paywall Exposure > Median	-0.041*** (0.012)	-0.029*** (0.008)	-0.025** (0.010)	-0.025*** (0.008)	-0.023** (0.011)
Dep. Var. Mean	0.745	0.745	0.745	0.745	0.745
County FE	✓	✓	✓	✓	✓
DMA-Year FE	✓	✓	✓	✓	✓
State-Year FE					✓
<b>Controls x Year FE:</b>					
Individual		✓	✓	✓	✓
County Demographics			✓	✓	✓
County Partisanship				✓	✓
Observations	482,757	481,258	481,258	479,979	479,979
R <sup>2</sup>	0.107	0.312	0.284	0.277	0.324

*Notes:* Coefficients for static (pre-post) version of Equation (4): Regressions of share correctly answered political knowledge questions by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects and DMA-year fixed effects. Column (2) adds survey respondent characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Column (3) adds county demographics, and Column (4) adds NaNDA partisanship index, both at 2010-levels, interacted with year fixed effects. Column (5) adds state-year fixed effects. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 7. Effect of paywalls on political knowledge - First principal component

Dependent Variable:	1st Principal Component of Correct Answers (Standardized)				
	(1)	(2)	(3)	(4)	(5)
Paywall Exposure > Median	-0.250*** (0.078)	-0.172*** (0.053)	-0.152** (0.060)	-0.151** (0.062)	-0.143** (0.069)
County FE	✓	✓	✓	✓	✓
DMA-Year FE	✓	✓	✓	✓	✓
State-Year FE					✓
<b>Controls x Year FE:</b>					
Individual Characteristics		✓	✓	✓	✓
County Demographics			✓	✓	✓
County Partisanship				✓	✓
Observations	419,222	418,670	418,670	418,623	418,623
R <sup>2</sup>	0.092	0.291	0.271	0.289	0.301

*Notes:* Coefficients for static (pre-post) version of Equation (4): Regressions of first principal component of correctly answered political knowledge questions by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects and DMA-year fixed effects. Column (2) adds survey respondent characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Column (3) adds county demographics, and Column (4) adds NaNDA partisanship index, both at 2010-levels, interacted with year fixed effects. Column (5) adds state-year fixed effects. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 8. Effect of paywalls on political knowledge - Exclude counties by population

Dependent Variable: Exclude counties by population:	Share Correct Answers					
	Exclude smallest			Exclude largest		
	10%	20%	30%	10%	20%	30%
	(1)	(2)	(3)	(4)	(5)	(6)
Paywall Exposure > Median	-0.042*** (0.010)	-0.030*** (0.009)	-0.023** (0.010)	-0.029*** (0.010)	-0.028** (0.011)	-0.033** (0.015)
County FE	✓	✓	✓	✓	✓	✓
DMA-Year FE	✓	✓	✓	✓	✓	✓
Indiv. Charac. x Year FE	✓	✓	✓	✓	✓	✓
Population Cutoff	42,730	92,527	159,943	2,321,014	1,293,825	854,848
Observations	435,621	389,714	342,727	432,501	382,052	330,716
R <sup>2</sup>	0.290	0.202	0.200	0.318	0.309	0.191

*Notes:* Coefficients for static (pre-post) version of Equation (4), excluding counties by deciles of population size. Regressions of share correctly answered political knowledge questions by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects, DMA-year fixed effects, and survey respondent characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 9. Effect of paywalls on political knowledge - Index decomposition

Dependent Variable:		Indicator Correct Answer							
Question:		Knows Party				Knows Majority			
		Governor (1)	Representative (2)	Senator 1 (3)	Senator 2 (4)	US House (5)	US Senate (6)	St. House (7)	St. Senate (8)
Paywall Exposure > Median		-0.049** (0.019)	-0.035*** (0.012)	-0.051*** (0.009)	-0.035*** (0.013)	-0.008 (0.017)	0.009 (0.015)	-0.031*** (0.011)	0.000 (0.015)
Dep. Var. Mean		0.825	0.735	0.781	0.781	0.807	0.784	0.528	0.535
County FE		✓	✓	✓	✓	✓	✓	✓	✓
DMA-Year FE		✓	✓	✓	✓	✓	✓	✓	✓
Indiv. Charac. x Year FE		✓	✓	✓	✓	✓	✓	✓	✓
Observations		478,598	475,661	478,047	478,020	474,271	471,144	426,682	426,946
R <sup>2</sup>		0.216	0.189	0.224	0.232	0.215	0.192	0.156	0.179

*Notes:* Coefficients for static (pre-post) version of Equation (4), using as the outcome each separate component of the political knowledge index. Regressions of indicator for correct answer by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure, controlling for county fixed effects, DMA-year fixed effects, and survey respondent characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 10. Effect of paywalls on political knowledge - Heterogeneity by individual characteristics

Dependent Variable:		Share Correct Answers					
Sample Split by:		Yearly Family Income			Highest Attained Degree		
		< 50k (1)	50k - 80k (2)	> 80k (3)	High School (4)	Some College (5)	College (6)
Paywall Exposure > Median		-0.037** (0.015)	0.005 (0.025)	-0.017 (0.023)	-0.055*** (0.010)	0.016 (0.015)	-0.014 (0.010)
County FE		✓	✓	✓	✓	✓	✓
DMA-Year FE		✓	✓	✓	✓	✓	✓
Indiv. Charac. x Year FE		✓	✓	✓	✓	✓	✓
Observations		196,769	109,582	122,696	146,361	118,722	216,175
R <sup>2</sup>		0.291	0.338	0.309	0.317	0.331	0.287

*Notes:* Coefficients for static (pre-post) version of Equation (4), estimated separately for different subgroups defined by yearly family income (Columns 1–3) and highest attained educational degree (Columns 4–6). Regressions of share correctly answered political knowledge questions by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects and DMA-year fixed effects. All columns include survey respondent characteristics interacted with year fixed effects. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 11. Effect of paywalls on political knowledge - Heterogeneity by county characteristics

Dependent Variable: Sample Split by:	Share Correct Answers					
	County Partisanship			County Population Density		
	Democratic (1)	Moderate (2)	Republican (3)	High (4)	Medium (5)	Low (6)
Paywall Exposure > Median	-0.083*** (0.016)	-0.006 (0.022)	-0.022 (0.027)	-0.063*** (0.018)	-0.064** (0.027)	-0.005 (0.048)
County FE	✓	✓	✓	✓	✓	✓
DMA-Year FE	✓	✓	✓	✓	✓	✓
Indiv. Charac. x Year FE	✓	✓	✓	✓	✓	✓
Observations	167,423	151,401	162,434	166,158	163,563	151,537
R <sup>2</sup>	0.039	0.325	0.347	0.284	0.315	-0.023

Notes: Coefficients for static (pre-post) version of Equation (4), estimated separately for different county types defined by terciles of NaNDA partisanship index (Columns 1–3) and terciles of population density (Columns 4–6). Regressions of share correctly answered political knowledge questions by CES survey respondent on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects and DMA-year fixed effects. All columns include survey respondent characteristics interacted with year fixed effects: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 12. Effect of paywalls on electoral participation

Dependent Variable: Election:	Voted in Election			
	Congressional Primary		Congressional	
	(1)	(2)	(3)	(4)
Paywall Exposure > Median	-0.027 (0.019)	-0.024 (0.017)	-0.011 (0.017)	-0.015 (0.017)
Dep. Var. Mean	0.423	0.423	0.673	0.673
P-value	0.168	0.167	0.502	0.382
County FE	✓	✓	✓	✓
DMA-Year FE	✓	✓	✓	✓
Individual Characteristics		✓		✓
Observations	342,535	342,002	287,125	286,691
R <sup>2</sup>	0.114	0.207	0.165	0.257

Notes: Coefficients for static (pre-post) version of Equation (4). Outcomes are indicators for whether survey respondent voted in the respective election. Columns (1) and (2) denote voting in any congressional primary election. Columns (3) and (4) refer to voting for either US Senator or US House Representative in the corresponding congressional elections. Regressions on indicator denoting whether respondent's county exceeded threshold of paywall exposure of 0.72, controlling for county fixed effects and DMA-year fixed effects. Columns (2) and (4) additionally include survey respondent characteristics: sex, age, ethnicity, education, income. Estimated using Sun-Abraham estimator. Survey weights included. Standard errors in parentheses are clustered by state. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## References

- Angelucci, Charles and Julia Cagé**, “Newspapers in times of low advertising revenues,” *American Economic Journal: Microeconomics*, 2019, 11 (3), 319–64.
- Ansolabehere, Stephen and Brian Schaffner**, “CES Common Content, 2021,” 2022.
- Aral, Sinan and Paramveer S Dhillon**, “Digital paywall design: Implications for content demand and subscriptions,” *Management Science*, 2021, 67 (4), 2381–2402.
- Baker, Andrew C, David F Larcker, and Charles CY Wang**, “How much should we trust staggered difference-in-differences estimates?,” *Journal of Financial Economics*, 2022, 144 (2), 370–395.
- Bartik, Timothy J**, “Who benefits from state and local economic development policies?,” 1991.
- Besley, Timothy and Robin Burgess**, “The political economy of government responsiveness: Theory and evidence from India,” *The quarterly journal of economics*, 2002, 117 (4), 1415–1451.
- Bhuller, Manudeep, Tarjei Havnes, Jeremy McCauley, and Magne Mogstad**, “How the internet changed the market for print media,” *American Economic Journal: Applied Economics*, 2024, 16 (2), 318–358.
- Blanchard, Olivier Jean, Lawrence F Katz, Robert E Hall, and Barry Eichengreen**, “Regional evolutions,” *Brookings papers on economic activity*, 1992, 1992 (1), 1–75.
- Cagé, Julia**, “Media competition, information provision and political participation: Evidence from French local newspapers and elections, 1944–2014,” *Journal of Public Economics*, 2020, 185, 104077.
- Callaway, Brantly and Pedro HC Sant'Anna**, “Difference-in-differences with multiple time periods,” *Journal of econometrics*, 2021, 225 (2), 200–230.
- Campante, Filipe, Ruben Durante, and Francesco Sobrino**, “Politics 2.0: The multifaceted effect of broadband internet on political participation,” *Journal of the European Economic Association*, 2018, 16 (4), 1094–1136.
- Chenoweth, Megan, Mao Li, Iris N Gomez-Lopez, and Ken Kollman**, “National Neighborhood Data Archive (NaNDA): Voter Registration, Turnout, and Partisanship by County, United States, 2004-2018,” *Ann Arbor, MI: Inter-university Consortium for Political and Social Research[distributor]*, 2020, pp. 11–04.
- Chiou, Lesley and Catherine Tucker**, “Paywalls and the demand for news,” *Information Economics and Policy*, 2013, 25 (2), 61–69.
- Chung, Doug J, Ho Kim, and Reo Song**, *The comprehensive effects of a digital paywall sales strategy*, Harvard Business School, 2019.
- Cook, Jonathan E and Shahzeen Z Attari**, “Paying for what was free: Lessons from the New York Times paywall,” *Cyberpsychology, behavior, and social networking*, 2012, 15 (12), 682–687.
- DellaVigna, Stefano and Ethan Kaplan**, “The Fox News effect: Media bias and voting,” *The Quarterly Journal of Economics*, 2007, 122 (3), 1187–1234.
- and Matthew Gentzkow, “Persuasion: empirical evidence,” *Annu. Rev. Econ.*, 2010, 2 (1), 643–669.
- Djourelova, Milena, Ruben Durante, and Gregory J Martin**, “The impact of online competition on local newspapers: Evidence from the introduction of Craigslist,” *Review of Economic Studies*, 2024, p. rdae049.
- Drago, Francesco, Tommaso Nannicini, and Francesco Sobrino**, “Meet the press: How voters and politicians respond to newspaper entry and exit,” *American Economic Journal: Applied Economics*, 2014, 6 (3), 159–88.
- Durante, Ruben, Paolo Pinotti, and Andrea Tesei**, “The political legacy of entertainment TV,” *American Economic Review*, 2019, 109 (7), 2497–2530.

- Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya**, “Media and political persuasion: Evidence from Russia,” *American Economic Review*, 2011, 101 (7), 3253–85.
- Falck, Oliver, Robert Gold, and Stephan Heblisch**, “E-lections: Voting Behavior and the Internet,” *American Economic Review*, 2014, 104 (7), 2238–65.
- Ferraz, Claudio and Frederico Finan**, “Exposing corrupt politicians: the effects of Brazil’s publicly released audits on electoral outcomes,” *The Quarterly journal of economics*, 2008, 123 (2), 703–745.
- Gao, Pengjie, Chang Lee, and Dermot Murphy**, “Financing dies in darkness? The impact of newspaper closures on public finance,” *Journal of Financial Economics*, 2020, 135 (2), 445–467.
- Gavazza, Alessandro, Mattia Nardotto, and Tommaso Valletti**, “Internet and politics: Evidence from UK local elections and local government policies,” *The Review of Economic Studies*, 2019, 86 (5), 2092–2135.
- Gentzkow, Matthew**, “Television and voter turnout,” *The Quarterly Journal of Economics*, 2006, 121 (3), 931–972.
- and Jesse M Shapiro, “Introduction of Television to the United States Media Market, 1946–1960,” 2008.
- , —, and Michael Sinkinson, “The effect of newspaper entry and exit on electoral politics,” *American Economic Review*, 2011, 101 (7), 2980–3018.
- , —, and —, “Competition and ideological diversity: Historical evidence from us newspapers,” *American Economic Review*, 2014, 104 (10), 3073–3114.
- George, Lisa M and Joel Waldfogel**, “The New York Times and the market for local newspapers,” *American Economic Review*, 2006, 96 (1), 435–447.
- Gerber, Alan S, Dean Karlan, and Daniel Bergan**, “Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions,” *American Economic Journal: Applied Economics*, 2009, 1 (2), 35–52.
- Jr, James M Snyder and David Strömberg**, “Press coverage and political accountability,” *Journal of political Economy*, 2010, 118 (2), 355–408.
- Kim, Ho, Reo Song, and Youngsoo Kim**, “Newspapers’ Content Policy and the Effect of Paywalls on Pageviews,” *Journal of Interactive Marketing*, 2020, 49, 54–69.
- Pattabhiramaiah, Adithya, S Sriram, and Puneet Manchanda**, “Paywalls: Monetizing online content,” *Journal of marketing*, 2019, 83 (2), 19–36.
- Pew Research Center**, “State of the News Media 2023,” 2023. Accessed on January 23, 2025.
- Sant’Anna, Pedro HC and Jun Zhao**, “Doubly robust difference-in-differences estimators,” *Journal of econometrics*, 2020, 219 (1), 101–122.
- Seamans, Robert and Feng Zhu**, “Responses to entry in multi-sided markets: The impact of Craigslist on local newspapers,” *Management Science*, 2014, 60 (2), 476–493.
- Strömberg, David**, “Radio’s impact on public spending,” *The Quarterly Journal of Economics*, 2004, 119 (1), 189–221.
- Sun, Liyang and Sarah Abraham**, “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects,” *Journal of Econometrics*, 2021, 225 (2), 175–199.
- U.S. Census Bureau**, “American Community Survey 5-Year Estimates: Comparison Profiles 5-Year,” 2010. Accessed on April 22, 2024.