Paywalls and the Demand for News

Lesley Chiou and Catherine Tucker*

April 17, 2013

Abstract

Given the preponderance of free content on the Internet, news media organizations face new challenges over how to manage access to and the pricing of their content. It is unclear whether content should be free or whether customers should pay via a \paywall." We use experimental variation from a media publisher's eld test of paywalls to examine demand for online news across several local media markets. We nd a 51 percent drop in visits after the introduction of a paywall and a far larger drop for younger readers.

 Economics Department, Occidental College, CA

*MIT Sloan School of Management, MIT, Cambridge, MA. and NBER

z

1 Introduction

Scholars have long argued that by providing information, news media promote community engagement and political participation (Putnam, 2000; Feddersen, 2004). Indeed, studies have shown that the di usion of news media is associated with positive spillovers on civic awareness and engagement (Mondak, 1995a,b; Lee and Wei, 2008; Oberholzer-Gee and Waldfogel, 2009; Gentzkow et al., 2009). Though the Internet facilitates the spread of information, it is not clear how or whether news organizations should charge for their content and how this

be interested in reaching different segments of the population through online and offline advertising, so differences in the characteristics of offline and online users may influence an advertiser's decision to multi-home and advertise on either side of the market (Athey et al., 2011).

We exploit a unique pricing experiment conducted by a media publisher. In July 2010, the Gannett Company introduced paywalls at the websites of three of its local newspapers—The Spectrum (Utah), The Greenville News (South Carolina), and The Tallahassee Democrat (Florida)—as part of an experiment. When readers navigated to the websites of these newspapers, they were prompted to sign up for a subscription to access online content as opposed to freely available content.[3] According to Gannett, if the experiment was a success, it would eventually introduce paywalls at its more than 100 news sites.

We use a rich dataset on consumer online behavior to empirically estimate how paywalls affect the demand for local news sites. In particular, our study investigates the number of visits and the demographic composition of visitors before and after the adoption of the paywalls, and we study users' subsequent behavior after navigating to a paywall. We use changes in traffic to other Gannett-owned news sites as a control for general trends on news consumption.

Overall visits to the websites after the introduction of a paywall sites fell by 51%. This decline is associated with a dramatic shift in demographics away from young readers: readership among 18-24 year olds falls by 99%. The introduction of paywalls disproportionately excludes the young, which undermines the creation of a comprehensive community (Putnam, 2000). We also find that visits for in-state and out-of-state readers fall by 50% and 56%, though the difference is not statistically distinguishable. Some evidence exists that visits

3

and 42%.

Our analysis serves as a case study, as today several papers have instituted some sort of paywall. Popular press articles suggest sizeable effects from the introduction of paywalls. The Times in UK lost almost 90% of its online readership since instituting its paywall Halliday

Our analysis focuses on paywalls erected at three local newspaper sites owned by Gannett Company: The Spectrum (Utah), The Greenville News (South Carolina), and The Tallahassee Democrat (Florida). Starting July 1, 2010, users could access content online by either purchasing a monthly subscription of $9.95 or a daily pass for $2. Subscribers of the print edition of the newspapers were o ered online access. Prior to the paywall, all content on the sites had been free of charge. Gannett indicated that these three sites would serve as an \experiment," and if deemed successful, Gannett would eventually introduce paywalls at its other news sites. Appendix A-1 provides an announcement issued by the publisher.

The three newspapers that adopted paywalls had di erent circulation levels. The Spectrum is a community daily newspaper that is based in St. George, Utah and covers an area spanning 200 miles that includes communities in Arizona and Nevada. According to their website, they are the only daily newspaper between Provo, Utah and Las Vegas, Nevada. The Spectrum joined Gannett in 2000, and of the three Gannett newspapers, it has the smallest circulation (23,000 for its Sunday edition.)[6] The Tallahassee Democrat serves the Tallahassee, Florida region. As Tallahassee's only daily newspaper, it covers local news in Leon County and the surrounding counties in northwestern Florida and southern Georgia. The Tallahassee Democrat has an average circulation of 49,627 for its Sunday edition. The Greenville News is a daily newspaper published in Greenville, South Carolina and is the largest of the three Gannett sites we study. In 1995, Gannett purchased the newspaper as part of Multimedia, Inc. The Greenville News has a circulation of 103,298 for its Sunday edition; according to the Audit Bureau of Circulations, only one other newspaper (Columbia State) has a (slightly) larger circulation within South Carolina.

Table 1 reports the demographics of the online users for each of the three Gannett sites as reported by Experian Hitwise for April 2010. The readers of the newspapers are predominantly male and are older than the general US population. The Greenville News

---

[6]Audit Bureau of Circulations, eCirc.

Table 1: Demographic description of online users

Measure	The Spectrum	Greenville News	Tallahassee Democrat	Other Gannett
Male	59.95	57.06	62.09	51.90
Age 18-34	1.69	42.52	15.00	22.30
Age 35-44	9.39	7.76	13.05	15.83
Age 45-54	54.91	9.4	49.50	20.42
Age 55+	34.01	40.33	22.45	41.45
Income <30k	21.7	30.01	28.65	24.77
Income 30-60k	19.43	21.96	29.96	31.30
Income 60-100k	39.66	23.07	30.18	24.94
Income >100k	19.21	24.96	11.21	18.98

Source: Hitwise, April 2010

Notes: This table reports the fraction of online users of a particular website within each demographic category. Statistics are reported for online users of The Spectrum, the Greenville News, the Tallahassee Democrat, and all other Gannett websites.

captures a larger fraction of younger readers compared to the other two newspapers. The median income is similar across the three sites. The table also includes demographics for other sites owned by Gannett.[7]

According to the Newspaper Association of America, 48.5% of US newspaper readers are men during 2010. Approximately 30.0% are between the ages of 18-34 and 70.0% are older than 35 years of age. Forty-seven percent have an annual household income below $50,000; 31.9% have a household income between $50,000 and $100,000; 20.6% have an income above $100,000.

3   Data Description

Our primary data source is Experian Hitwise. Hitwise aggregates data from website logs created on Internet Service Providers and combines the information with data from opt-in

---

[7] The column \Other Gannett" includes all other sites owned by Gannett as well as one site that covers news in Guam. For our analysis of site visits in all the following sections, we focus on US sites. Please refer to Table A-1 in the Appendix for the full list of newspapers in our sample for site visits.

panels. The resulting dataset forms a geographically diverse sample with usage data from 25 million people worldwide. For further detail, Chiou and Tucker (2010) also use this data.

First, we identified sites of local newspapers owned by the Gannett Company. Hitwise reports the share of visits to a given website that originates from each state. For instance, we observe the fraction of visits that greenvilleonline.com receives from each of the 50 states. Hitwise defines a "visit" as a "series of one or more page requests by a visitor without 30 consecutive minutes of inactivity." The data are aggregated and reported over a period of four weeks, so we collect visit information for the two months before and the two months after the paywalls were implemented— four rolling weeks ending 5/29, 6/26, 7/24, and 8/28 in 2010. We observe state-level data for 79 sites. Table A-1 in the Appendix supplies the full list of newspapers in our sample.[8]

Since Hitwise reports traffic as a fraction of visits, we acquire additional data from Compete to estimate the number of visits to a site from each state. Compete collects data from a panel of 2 million consumers who have given permission to have their Internet clickstream behaviors observed and from opt-in survey responses. It estimates the total number of monthly visits for each site. To calculate the number of visits from each state, we multiply the fraction of visits that originate from each state (as reported in Hitwise) with the total number of visits (as reported in Compete).

We use the classification defined by Gannett Publishing on their official website to determine which states these "community newspapers" cover. The definition also us to categorize visits that originate from "in-state" and "out-of-state" readers.

Table 2 reports the summary statistics for our sample. The number of observations reflects 4 months of data for 79 sites, i.e., 4 × 79 = 316. Each observation in the summary statistics represents a website-month combination. For instance, we observe the number of

---

[8]Some websites receive relatively low traffic and do not meet the minimum reporting standard. We focus our analysis on US newspapers and omit one newspaper with coverage in Guam.

Table 2: Summary statistics

	Mean	Std Dev	Min	Max
Number of visits	668480.5	982402.8	439.9	7419216
% visits in-state	0.23	0.11	0.047	0.66

Figure 1: Average number of visits to sites before and after paywalls implemented
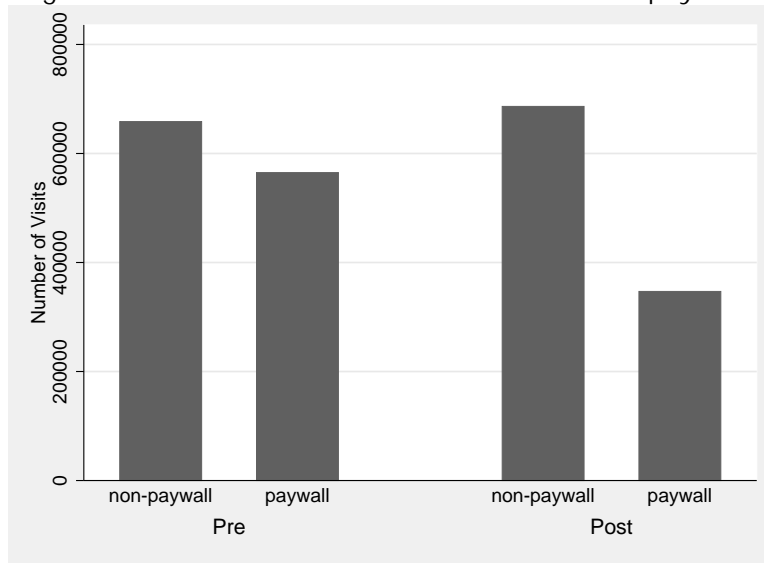
Notes: This figure shows the number of visits to sites before and after the implementation of paywalls.

$$+\ \beta_i + \gamma_t + \varepsilon_{it}$$

where *Paywall* is an indicator variable for whether a site is one of the three sites that adopted a paywall during our time period, and *Post* is an indicator variable equal to one after the paywalls were implemented. The coefficient represents website fixed effects, and the coefficient represents month fixed effects. We employ a semi-log regression to account for differences in scale across the different sites. Since some states report no visits to a given sites within a month, we add a small positive number before taking the logarithm of visits. We cluster our standard errors at the website level. The main effects for *Paywall* and *Post* are dropped due to collinearity with monthly dummies and website fixed effects.

   Given the semi-log specification, we interpret our estimated coefficients as the \ratio-of-ratios" (Mullahy, 1999). For instance, to determine the effect of paywall adoption on visits,

we compute the corresponding ratio-of-ratios:

$$\frac{\left\{\frac{E[visits|Paywall=1;Post=1]}{E[visits|Paywall=1;Post=0]}\right\}}{\left\{\frac{E[visits|Paywall=0;Post=1]}{E[visits|Paywall=0;Post=0]}\right\}} = exp(\beta_1): \tag{2}$$

For Equation (2) above, the fraction in the numerator (proportionately) compares the expected number of visits to a paywall site before and after the introduction of the paywall. The fraction in the denominator compares the expected number of visits to the control sites before and after the introduction of the paywall. The advantage of this interpretation is that we avoid the \retransformation bias" for estimating the number of visits from the semi-log regression. The expression offers a natural interpretation for the estimated coefficients directly (Mullahy, 1999). Consequently

Table 3: Visits before and after the implementation of paywalls

	(1) All	(2) In-state	(3) Out-of-state
Post ⨯ Paywall site	-0.723	-0.696	-0.819
	(0.156)	(0.186)	(0.180)
Website Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	316	316	316
R-Squared	0.965	0.963	0.961

Notes: Robust standard errors clustered at website level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The dependent variable is the logarithm of visits plus one. The main effects for *Paywall* and *Post* are dropped due to collinearity with monthly dummies and website fixed effects. The number of observations reflects 4 months of data for 79 sites, i.e., $4 \times 79 = 316$.

Table 4: Falsification check of visits before the implementation of the paywalls

	(1) All	(2) In-state	(3) Out-of-state
June ⨯ Paywall site	-0.126	-0.199	0.139
	(0.110)	(0.122)	(0.0827)
Website Fixed Effects	Yes	Yes	Yes
Month Fixed Effects	Yes	Yes	Yes
Observations	158	158	158
R-Squared	0.992	0.993	0.988

to users located inside of the state. [11] However, the coe cients are not precise enough to statistically distinguish the di erence.

Next we examine three additional measures of demographics: age, income, and gender. We focus on these three categories, since Hitwise reports visits by age, income, and gender.[12] For each website, Hitwise reports the share of monthly visitors that fall into each demographic category. We run regressions similar to Equation (1) for each demographic category. Tables 5-7 report the results of the regressions. For instance, Column (1) of Table 5 reports the results of the regression for visits by users in the 18-24 age group during a given month; an observation reports the number of visits by users in the 18-24 age group to a particular site during a given month.

In Table 5, a general pattern emerges where visits fall to a much greater extent for younger visitors after the implementation of paywalls. For instance, visits by users of 18-24 years of age fall by 99 percent while visits by users over the age of 55 fall by 46 percent.[13] In Table 6, the four columns report visits by each of the four income groups: <$30,000, $30,000-$59,999, $60,000-$99,999, and >$100,000.[14] For the lowest income users in Group 1, the coe cient on $Post \quad Paywall$ is signi cant at the 10.1% signi cance level. Given the magnitude of the coe cients, visits by users in the lowest income group fall by a larger amount than visits by higher income groups. For instance, visits by users with incomes below $30,000 fall by 91 percent while visits by users with incomes above $100,000 decline by 51 percent. [15] Finally, in Table 7, visits by male users fall by more than visits by female users

---

[11]To calculate the ratio-of-ratios for in-state users, we nd that visits are 50 percent of their previous levels ($exp($ 0:696$) = 0$:50); in other words, visits fall by 50 percent ($= 1$ 0:50 = 0:50). For out-of-state users, visits are 44 percent of previous levels ($exp($ 0:819$) = 0$:44); visits fall by 56 percent ($= 1$ 0:44 = 0:56).

[12] We also perform falsi cation checks for each of the demographic categories, and we verify that no pre-existing trend exists in the dataset.

[13]To calculate the ratio-of-ratios, we nd that visits by ages 18-24 are 1 percent of their previous levels ($exp($ 4:

Table 5: Visits by age group (1=lowest, 5=highest)

		(1) Group 1	(2) Group 2	(3) Group 3	(4) Group 4	(5) Group 5
Post	Paywall	-4.895	-1.832	-1.300	-1.025	-0.608

Table 7: Visits by gender

	(1) Male	(2) Female
Post Paywall	-0.858	-0.538
	(0.159)	(0.203)
Website Fixed Effects	Yes	Yes
Month Fixed Effects	Yes	Yes
Observations	316	316
R-Squared	0.963	0.958

Notes: Robust standard errors clustered at website level. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. The dependent variable is the logarithm of visits plus one. The main effects for *Paywall* and *Post* are dropped due to collinearity with monthly dummies and website fixed effects. The number of observations reflects 4 months of data for 79 sites, i.e., $4 \times 79 = 316$.

Lee and Wei (2008) find that a decrease in newspaper readership among 17- to 24- year olds is associated with a decline in political participation.

Our result is related to prior work that finds the Internet more generally has had the effect of drawing younger and more educated readers out of the market for traditional news (George, 2008). We find that younger readers are more price sensitive.

The decline in young readers also has potential implications for media organizations. In the short-run, high income readers remain and can subscribe to paywalls, generating revenues and preventing local newspapers from failing. However, in the long run, the ability to attract new readers and maintain circulation may be diminished. Losing a youthful audience on one side of the market could also affect pricing on the other side of the market. While advertising rates from Gannett were held fixed during their paywall experiment, some evidence exists that advertisers are willing to pay higher advertising rates for young audiences in other media such as television (Goettler, 1999).

Ultimately, the welfare implications and political externalities of paywalls will depend on what people do instead of visiting the paywall sites. It is particularly important that the trend to implement paywalls centers on local newspapers. Such local media presumably has fewer substitutes, so it may be likely that paywalls will lead to less local news consumption

overall. Other studies have found that lowering the costs of consumption can shift attention towards local news (Athey and Mobius, 2012).

6   Conclusion

Granting access and charging for online content has been a controversial issue. The debate over how to provide and whether to charge for information is particularly heated, since the dissemination of information by the Internet has also coincided with declining print circulation and advertising revenues. To our knowledge, our paper is the first to empirically study whether charging for content shifts the quality and composition of readers to media sites. On the one hand, some argue that because technological advances— such as the Internet— have created the plethora of alternative sources of information and lowered consumers' costs of searching for news, consumers will not be willing to pay for content. On the other hand, others argue that information is highly differentiated and can be targeted to a specific geographic market or audience that readers will be willing to pay for such differentiated products.

We study a unique pricing experiment by a publisher that implemented paywalls at three of its local media sites. We find that imposing paywalls leads to a large decline in readership, particularly among young readers.

Our study has implications for the future of media. The introduction of paywalls disproportionately excludes young readers, which undermines policymakers' attempts to create a comprehensive community. Scholars have emphasized that newspaper readership as imperative to the promotion of democracy and civic engagement (Putnam, 2000; Feddersen, 2004; Oberholzer-Gee and Waldfogel, 2009; Gentzkow et al., 2009).

References

Athey, S., E. Calvano, and J. Gans (2011). The Impact of the Internet on Advertising Markets for News Media. working paper.

Athey, S. and M. Mobius (2012). The Impact of News Aggregators on Internet News Consumption: The Case of Localization. working paper.

Chandra, A. (2009, March). Targeted Advertising: The Role of Subscriber Characteristics in Media Markets. *The Journal of Industrial Economics 57*(1), 58{84.

Chiou, L. and C. Tucker (2010). Copyright, Digitization, and Aggregation. working paper.

Feddersen, T. (2004). Rational Choice Theory and the Paradox of Not Voting. *Journal of Economic Perspectives 18*, 99{112.

Gentzkow, M., J. Shapiro, and M. Sinkinson (2009). The Eect of Newspaper Entry and Exit on Electoral Politics. *American Economic Review Forthcoming*.

George, L. (2008). The Internet and the Market for Daily Newspapers . *The B.E. Journal of Economics Analysis and Policy (Advances) 8*(1), Article 26.

Goettler, R. (1999). Advertising Rates, Audience Composition, and Competition in the Network Television Industry. working paper.

Grant, D. (2009, December 11). Newsday.com Sees Pay Wall-Induced Drop in Trac. *Fishbowl NY*.

Halliday, J. (2010, July 20). Times loses almost 90 percent of online readership . *The Guardian*.

Lee, T. and L. Wei (2008). How Newspaper Readership Aects Political Participation. *Newspaper Research Journal 29*, 8{23.

Mataconis, D. (2010, July 19). London Times Web Tra c Falls 66 percent after Paywall Goes Up . *Outside the Beltway*.

Mondak, J. (1995a). Media Exposure and Political Discussion in U.S. Elections. *The Journal of Politics 57*, 62{85.

Mondak, J. (1995b). Newspapers and Political Awareness. *American Journal of Political Science 39*, 513{27.

Mullahy, J. (1999, November). Interaction e ects and di erence-in-di erence estimation in loglinear models. NBER Technical Working Papers 0245, National Bureau of Economic Research, Inc.

Oberholzer-Gee, F. and J. Waldfogel (2009). Media Markets and Localism: Does Local News en Espanol Boost Hispanic Voter Turnout? *The American Economic Review 99*(5), 2120{2128.

Putnam, R. (2000). Bowling Alone: The Collapse and Revival of American Community.

Rysman, M. (2009). The Economics of Two-Sided Markets. *Journal of Economic Perspectives 23*, 125{144.

Seamans, R. and F. Zhu (2011). Technology Shocks in Multi-sided Markets: The Impact of Craigslist on Local Newspapers. Working Paper.

Table A-1: Gannett media sites in sample

Newspaper	Location
Argus Leader, Sioux Falls	Sioux Falls, SD
Asbury Park Press	Asbury, NJ
Asheville Citizen-Times	Asheville, NC
Battle Creek Enquirer	Battle Creek, MI
Chillicothe Gazette	Chillicothe, OH
Coshocton Tribune	Coshocton, OH
Courier-Post, Cherry Hill	Cherry Hill, NJ
Daily Press & Argus, Livingston County	Howell, MI
Daily Record, Parsippany	Parsippany, NJ
Daily World, Opelousas	Opelousas, LA
Detroit Free Press	Detroit, MI
Florida Today, Brevard County	Brevard, FL
Fort Collins Coloradoan	Fort Collins, CO
Great Falls Tribune	Great Falls, MT
Green Bay Press-Gazette	Green Bay, WI
Hattiesburg American	Hattiesburg, MS
Herald Times Reporter, Manitowoc	Manitowoc, WI
Indianapolis Star	Indianapolis, IN
Iowan City Press Citizen	Iowa City, IA
Journal and Courier, Lafayette	Lafayette, IN
Lancaster Eagle-Gazette	Lancaster, OH
Lansing State Journal	Lansing, MI
Marsh eld News-Herald	Marsh eld, WI
News Herald, Port Clinton	Port Clinton, OH
News Journal, Mans eld	Mans eld, OH
Observer & Eccentric, Newspapers, Livonia	Livonia, MI
Oshkosh Northwestern	Oshkosh, WI
Palladium-Item, Richmond	Richmond, IN
Pensacola News Journal	Pensacola, FL
Poughkeepsie Journal	Poughkeepsie, NY
Press & Sun-Bulletin, Binghamton	Binghamton, NY
Reno Gazette-Journal	Reno, NV
Rochester Democrat and Chronicle	Rochester, NY
Spring eld News-Leader	Spring eld, MO
St. Cloud Times	Saint Cloud, MN
Star-Gazette, Elmira	Elmira, NY
Statesman Journal, Salem	Salem, OR
Stevens Point Journal	Stevens Point, WI
Tallahassee Democrat	Tallahassee, FL
Telegraph-Forum, Bucyrus	Bucyrus, OH
The Advocate, Newark	Newark, OH
The Arizona Republic, Phoenix	Phoenix, AZ
The Baxter Bulletin	Mountain Home, AR
The Burlington Free Press	Burlington, VT
The Cincinnati Enquirer	Cincinnati, OH
The Clarion-Ledger, JacVu4Mountain Home, AR	

Journal and33(y)32(A(y)97()-1en-Times)-1l,   832(y)97()Journal and33(y)32(AIthaca3484(Lans)1(i)-14986AIthaca-6.974 Td