DISCUSSION PAPER SERIES

DP13651

PAYWALLS AND THE DEMAND FOR ONLINE NEWS

Frode Steen

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

PAYWALLS AND THE DEMAND FOR ONLINE NEWS

Frode Steen

Discussion Paper DP13651
Published 04 April 2019
Submitted 22 March 2019

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

PAYWALLS AND THE DEMAND FOR ONLINE NEWS

Abstract

The digitisation of society has posed a challenge to news outlets. Seeking advertising revenues and facing competition for the attention of their readers, many news outlets entered the digital era with unrestricted access to their online content. More recently, news outlets have sought to restrict the amount of content available for free. We quantify the impact of introducing a paywall on the demand for news in Norway. The short-run average impact of a paywall is negative and between 3 and 4%, in the long run the effect increases to between 9 and 11%. We find heterogeneity in the response to paywalls. The largest news outlet within its market experiences larger effects than the other news outlets. After introducing a paywall, the largest news outlets face a long-run reduction in demand between 13 and 15%, as compared to the others who experience a decrease of between 8 and 11%. The timing of introducing a paywall does not seem to affect the demand response very much.

JEL Classification: L20, L82, D40

Keywords: Online news, paywalls, business models, two-sided markets

Frode Steen - frode.steen@nhh.no

Professor, Norwegian School of Economics and CEPR

Acknowledgements

Frode Skjeret, SNF Centre for Applied Research at NHH, E-mail: Frode.Skjeret@snf.no Frode Steen, Norwegian School of Economics and CEPR, E-mail: Frode.Steen@nhh.no Timothy G.A. Wyndham, Norwegian School of Economics E-mail: Timothy.Wyndham@nhh.no. We are grateful to all the Norwegian news outlets and media houses included in our study for providing access to data on paywall introductions, and, in particular, to TNS Gallup and NRK who provided us with data on online media consumption. We thank Hans Jarle Kind and seminar participants at NHH for comments and discussions. The usual caveats apply.

Paywalls and the demand for online news*

Frode Skjeret

SNF Centre for Applied Research at NHH

Frode Steen

NHH Norwegian School of Economics and CEPR

and

Timothy G.A. Wyndham

NHH Norwegian School of Economics

This version: March 2019

Abstract

The digitisation of society has posed a challenge to news outlets. Seeking advertising revenues and facing competition for the attention of their readers, many news outlets entered the digital era with unrestricted access to their online content. More recently, news outlets have sought to restrict the amount of content available for free. We quantify the impact of introducing a paywall on the demand for news in Norway. The short-run average impact of a paywall is negative and between 3 and 4%, in the long run the effect increases to between 9 and 11%. We find heterogeneity in the response to paywalls. The largest news outlet within its market experiences larger effects than the other news outlets. After introducing a paywall, the largest news outlets face a long-run reduction in demand between 13 and 15%, as compared to the others who experience a decrease of between 8 and 11%. The timing of introducing a paywall does not seem to affect the demand response very much.

JEL codes: L20, L82, D40

Keywords: Online news, paywalls, business models, two-sided markets

* Frode Skjeret, SNF Centre for Applied Research at NHH, E-mail: Frode.Skjeret@snf.no Frode Steen, Norwegian School of Economics and CEPR, E-mail: Frode.Steen@nhh.no Timothy G.A. Wyndham, Norwegian School of Economics E-mail: Timothy.Wyndham@nhh.no. We are grateful to all the Norwegian news outlets and media houses included in our study for providing access to data on paywall introductions, and, in particular, to TNS Gallup and NRK who provided us with data on online media consumption. We thank Hans Jarle Kind and seminar participants at NHH for comments and discussions. The usual caveats apply.

Introduction

The digital transformation of society has profoundly impacted news producing organisations. Digitisation has lowered the cost of getting content to readers, but has also increased the range of substitutes available to readers and advertisers. News outlets, as two-sided platforms, have always had the option to not charge their readers for news.[1] Most print newspapers have adopted a model where readers pay for news, the free newspapers often available in large cities being the exception.[2] Online news outlets, on the other hand, have frequently decided not to charge their readers.

In many countries the proportion of news going behind a paywall has increased. We adopt the paywall definition used in Chiou & Tucker (2013): a "digital mechanism that separates free content from paid content on a website". In 2011, according to Høst (2016), only five out of 194 Norwegian news outlets with an online presence had implemented any sort of paywall.[3] By 2015, he found that business models had evolved, and nearly two thirds of news outlets in Norway with an online presence had some sort of paywall.

News outlets have two sources of profits and revenues: readers and advertisers. Changing a business model by introducing a paywall is likely to have opposite effects on these sources. Whilst it will lead to new reader revenues, the number of readers and/or pages viewed is likely to fall, decreasing advertising revenues. If, as is most common, the advertising price is based on number of views, the news outlet will have fewer views and lower revenues.[4]

The business models on how to integrate and price online vs printed news are still not concluded. Though we have seen a trend towards introduction of paywalls across many markets, we also see reversals. In the US, some news outlets have reverted to a non-paywall model (Kim et al. 2018), thus suggesting that knowledge on the effects of paywalls is still developing and the optimal choice of business model is not clear-cut. This is something we also observe in the Norwegian market, where some news providers stick to a business model with no paywall. There is also an example of a paper introducing a paywall early, that later removed the wall.

[1] Casadesus-Masanell and Zhu (2010) provide a more general framework of business models when advertising sponsoring is possible.
[2] Two examples are the Metro and the Evening Standard in the United Kingdom who offer printed news to readers in urban areas for free (in the morning and the afternoon respectively).
[3] Most of our news providers are traditional newspapers that have also developed digital platforms. However, since some of the news providers are either fully digital web based outlets (e.g., Nettavisen), and others are television channels that also operate web based news pages (e.g. NRK), we refer to the group of news providers in our sample as 'news outlets'.
[4] Where advertising fees are per ad, as in printed newspapers, fewer expected readers would reduce the attractiveness of ads at that paper, and the ad demand curve would shift down.

The impact on reader demand is clearly a crucial determinant of the profitability of a paywall and therefore the choice of business model. We address two major questions. First, what has been the general quantitative impact of introducing paywalls for the average news outlet in the Norwegian market for online news? Second, is the impact different for news outlets that introduce paywalls after their rivals (timing heterogeneity), and does the effect differ with the outlet's local market position (news outlet heterogeneity)? More particularly, there are reasons to believe that the larger news outlets have a different readership base in their respective markets than the smaller outlets. This might lead to a heterogeneous response from readers faced with paywalls at different sized news outlets. We disentangle the response to the introduction of paywalls and estimate separate effects for the larger news outlets and for the others. We also estimate dynamic models, which allows us to quantify to which degree the effects differ in the short and the long run. To this end, we use longer, more frequent data and analyse a larger number of news outlets that are changing their business model than has previously been done. Distinguishing between short- and long run effects within the same analysis is also new to this literature.

To do this, we utilize weekly data from 122 news producers from January 2012 to December 2015 on the usage of electronic news outlets before and after the introduction of paywalls. Of these, 69 introduced paywalls during our sample period. Furthermore, we contrast these data to the consumption of news from a number of news providers that offered open access to their online content throughout the sample period, including the national public service broadcaster (NRK). NRK has produced online news from dedicated newsrooms tailor-made for all the regions throughout our data period.[5] Whereas regional news providers without payment walls are often smaller and without a strong presence online, NRK is typically among the largest regional news providers in its respective markets, and is as such a particularly well suited control group to the news providers that introduced paywalls over the period of study.[6] Given the nature of the Norwegian media topography, we are able to allocate all our news outlets to 13 well-defined markets. These comprise twelve regional markets and one national market for the outlets with a much wider spread.

Analysing a relatively large number of markets enables us to study heterogeneity both with regard to how differences in type of news outlet (relative size in their local markets) affect the introduction of paywalls, and how heterogeneity in the timing of the introduction affects the

[5] We differentiate between national and regional news providers, applying the regions used by NRK when defining the geographical scope of their district offices. See Figure 1 and Table A1 for more details.
[6] Note that also some of the news providers that do not impose paywalls are large online providers, examples are e.g., Nettavisen and TV2 that are ranked as number 4 and 5 in the national market (See Table A1 in the Appendix).

demand responses. We apply a difference-in-difference approach, where we look at how the introduction of paywalls affects the number of page hits, unique sessions and unique visitors.

Chiou & Tucker (2013) were the first to empirically investigate the quantitative impact of paywalls on digital news outlet readership. They used state level data from the USA for two periods; before and after the simultaneous introduction of paywalls at three local newspapers.[7] Employing a difference-in-difference strategy and a control group of 76 similar newspaper owned by the same newspaper group, they estimate that the introduction of the paywall led to a short-term decrease of readership of 51%. Pattabhiramaiah et al (2018) assess the impact of a paywall at the New York Times (NYT). They used five national newspapers to create a synthetic control group. In the thirteen months after the NYT introduced the paywall, the number of unique visitors compared to that of the synthetic control newspaper, fell by 16.8%. They also estimated a reduction in other engagement metrics such as number of visits, pages viewed per visit and duration of visit.

The closest paper to ours is perhaps Kim et al. (2018). They study the rollout of paywalls in 42 newspapers in the US between 2010 and 2015. They analyse the newspapers as a pooled group without being able to attribute them to different regional markets, but rather utilize a rich dataset on individual newspaper characteristics to control for marginal effects of paywall introductions. As opposed to us, they have no control groups that offer free online news. They find that most newspapers' paywalls have long-term negative effects (though they do not estimate short-run effects), but that the amount of the loss varies by reader demographics, newspaper characteristics and when the paywall is introduced. Of their 42 newspapers they find significant decreases in online demand for 36 newspapers. The estimates vary from -54% to -10% with an average decrease of 28.3%. For the remaining six newspapers estimates are positive, but non-significant (calculated from Kim et al. (2018), Table 4). Their dataset has more information on newspaper and readership characteristics than ours, allowing the estimation of a richer set of marginal effects in terms of heterogeneity in consumer responses to paywalls. Our analysis complements theirs in the sense that we can utilise the combination of a number of well-defined markets and the existence of a public, online, freely accessible news provider that has dedicated newsrooms for each regional market. This provides us with better controls throughout the data period and better information on the size and ranking of news outlets within their local markets.

A number of studies have analysed the optimal type of paywall. Lambrecht & Misra (2016) assess the dynamic question of what share of content to place behind the paywall. They find empirical evidence

[7] Each period contained visit data for four weeks.

showing that news outlets make more news available for free in periods of high demand. Aral & Dhillon (2017) use a series of natural experiments to evaluate the impact of changes to the amount and breadth of news behind a pay wall on cross channel demand and subscription rates.

Our paper is linked to previous qualitative work on the Norwegian roll-out of paywalls. Sjøvaag (2016) found that papers which provide some content free whilst keeping other content behind a paywall are more likely to place content of local relevance behind the paywall and leave more widely relevant news, such as 'wire copy' news and syndicated content open to all. This is in line with Kim et al. (2018) who quantify this and find that newspapers with more unique content tend to perform better after the roll-out than newspapers with more common content. Hognaland and Saebi (2015) investigate qualitative drivers of business model choices (full paywall, partial paywall, freemium etc.) in Norwegian newspapers. An interesting finding for this paper is that experimentation played a central role for news outlets' choice of business model. Hence, knowledge about the effects of introducing paywalls was most likely scarce at the time of introduction.

In our most basic average effect models, we find that the short-run average impact of a paywall is negative and reduces demand by between 3 and 4%, which is smaller than in previous studies. However, the effect is found to be much larger in the long run and when we control for news outlet heterogeneity. Our results suggest that readers' habits take some time to change. The average news outlet experiences between 9 and 11% long-run reduction in demand after a paywall introduction, suggesting that the longer the paywall exists, the stronger the impact from its reader-base will be.

Turning to the relative ranking and size of the news outlets, we find some evidence of heterogeneity in the responses to paywalls. The largest news outlet experiences larger demand effects than the other news outlets within their regional (or national) market. The largest news outlets face a long-run reduction in demand of between 13 and 15% after paywall introductions, as compared to between 8 and 11% decrease in demand for the other outlets.

When estimating the effect of the introduction of a paywall, we control for the share of hits behind other competing news outlets' paywalls. As competing news outlets install paywalls, the share of freely available online news is reduced. If this share is increased by 10% (implying a reduction in freely available news), our models predict a general and significant increase in online consumption for the remaining free news outlets of between 3.6 to 4.4%.

We also analyse to what extent the timing of introducing a paywall affects the demand response. By allowing the effect to differ according to the share of the market that is behind the competitors' paywalls, we find that timing is not very important. However, for the largest news outlets, we find some indication of an increased negative demand effect as more and more competing news outlets introduce payment walls.

We show that our results are robust both when focusing only on changes around the paywall introductions (event-study) and when we allow for alternative market compositions.

In the next section, we describe our data. We then present our descriptive analysis before describing our empirical strategy. We present our results and robustness analysis before we discuss our results and conclude.

Data and market definitions

We combine data on the usage of electronic news outlets before and after the introduction of paywalls with data on the consumption of news from the regional and national pages of the public service broadcaster, NRK.

Our data on the usage of electronic news outlets is from Kantar TNS. We have removed sites that cannot be regarded as news-media sites, e.g. weekly magazines, special interest group sites and various news aggregators. Our dataset contains only news outlets that actually produce news in-house. We also have data on the usage of NRK's internet sites, both for the nationwide site, and its regional news outlets.[8] Although our data on private news outlets dates back to 2009, we restrict our data to span from week 1, 2012 to week 52, 2015, in order to maintain consistency with the NRK data. In addition, many of the smaller private news outlets were included in the data from 2013 and onwards.

The dataset from Kantar TNS consists of three measures of weekly internet media consumption:

- Hits: this is the total number of hits (all articles and front-page) from all visitors.
- Unique sessions: Unique sessions are measured as the number of sequences of hits by all unique visitors from first visit to site until leaving the site.
- Unique visitors: All visitors to a news outlet site are uniquely identified and counted.

The three measures are highly correlated; we typically find a correlation coefficient ranging from 0.98-0.99, all measuring online news outlet activity.[9]

---

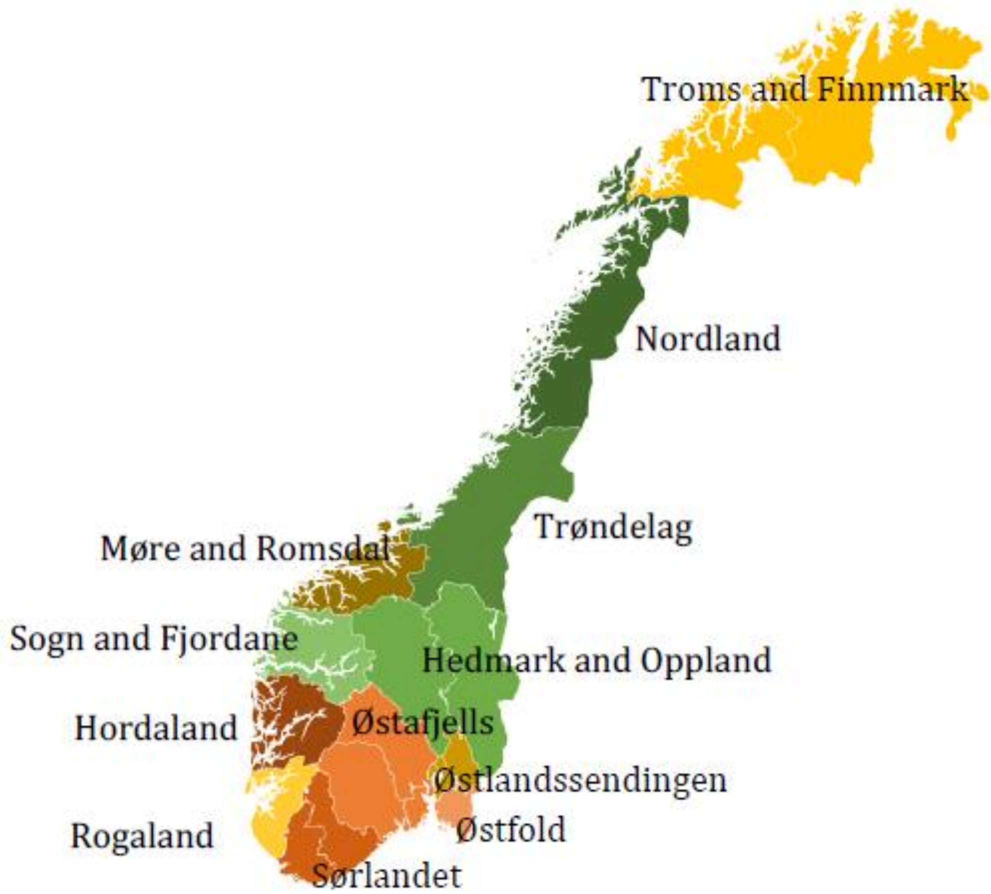
[8] We only have aggregate NRK data for Buskerud, Telemark and Vestfold before week 42, 2013, thus we aggregate these three counties from week 43, 2013 by summing them. This way we have a consistent market for this area (Østafjells) throughout our sample period. We have also checked that there is no apparent shift in the long-run trend for other outlets in the three areas before and after the aggregation takes place. From week 10, 2015 we have separate NRK data for Finnmark and Troms, we thus aggregate these two counties for the last period (week 11, 2015 - week 52, 2015) to obtain a consistent regional market control group.

[9] The correlation between Hits and Visitors is 0.978, between Hits and Sessions 0.987, and, finally, between Sessions and Visitors 0.994. All are significant at the 1% level.

Our final dataset includes 122 news-producing media-sites. We extend this news outlet level data by adding the dates of introduction of paywalls in the Norwegian market by 69 news outlets. We obtained this data by contacting the individual news outlets, or their owners.

Data on the geographical coverage of the news outlets in the sample is provided by Medietilsynet (the Norwegian Media Authority). We use the NRK district offices as our definition of geographical markets. When the location of a news outlet in our data from Medietilsynet is within the regional boundaries of an NRK district office, the news outlet is assumed to belong to that regional market. Thus, we use the following 12 regional markets: *Hedmark and Oppland, Østafjells, Østlandssendingen, Østfold, Sørlandet, Rogaland, Hordaland, Sogn and Fjordane, Møre and Romsdal, Trøndelag, Nordland*, and *Troms and Finnmark*.

Figure 1: The regional markets used for our analysis



In addition, we define a national market where all larger, national news outlets are included. Since news providers that are attributed to the national market are larger by orders of magnitude than the regional market providers, allocating these national players to regional markets might lead to biases in our models. However, as a robustness check, we estimate models in which we leave out the national market, i.e. only estimating models for the 12 regional markets, and our predictions are stable across these models. In Figure 1 we show a map illustrating our 12 regional markets.

Descriptive analysis

While the previous chapter introduced the data used, this chapter describes the dataset in greater detail. Table 1 provides an overview of our markets, number of news outlets and paywalls introduced, the average weekly hits and visitors and Herfindahl-Hirchman concentration Index (HHI) based on hits across markets. A more detailed list providing the names and relevant figures of all outlets is found in Table A1 in Appendix A.

Table 1: Markets, news providers, and usage of data (Week 1, 2012- Week 52, 2015).

Market	Number of news providers	Paywalls introduced	Average weekly hits	Average weekly visitors	HHI
National Market	12	6	32 092 240	4 663 477	2 007
Hedmark and Oppland	6	3	492 595	163 736	3 563
Østfold	7	6	693 226	173 507	2 711
Østlandssendingen	11	8	718 570	233 670	2 584
Østafjells	22	20	1 106 149	344 066	1 614
Sørlandet	7	3	673 722	180 121	3 306
Rogaland	3	2	938 422	238 308	4 173
Hordaland	11	5	1 760 959	419 341	3 722
Sogn and Fjordane	5	4	839 009	144 623	7 798
Møre and Romsdal	10	3	725 251	184 844	2 754
Trøndelag	12	3	1 442 963	337 078	4 195
Nordland	8	5	782 982	242 228	2 769
Troms and Finnmark	8	1	847 941	266 123	3 389
Total	122	69			
Average local markets			918 482	243 970	3 548

Note: Hit and visitor data in this table are aggregates of the weekly average figures in Table A1 in the Appendix. At the market level, the visitor data will be an overestimate since some readers will visit more than one site.

We see that the different markets display different sizes and compositions. News outlets in the national market have substantially more hits and visitors than news outlets in the regional markets. News outlets in the regional markets are reasonably homogenous in terms of weekly hits, but less so in terms of the number of news outlets within each region. We also observe that the number and proportion of paywall introductions differ quite a bit across regional markets. For half of the markets the percentage of news outlets having introduced paywalls is between 43 and 67%, whereas for some markets nearly all news outlets have done so. For Troms and Finnmark, only one of the news outlets has introduced a paywall. Market concentration is high; measured by HHI using hits per news outlet, the average across the regional markets is as high as 3 548, which is well above competition authorities' 'worrying' threshold. Even the national market has a high concentration index.

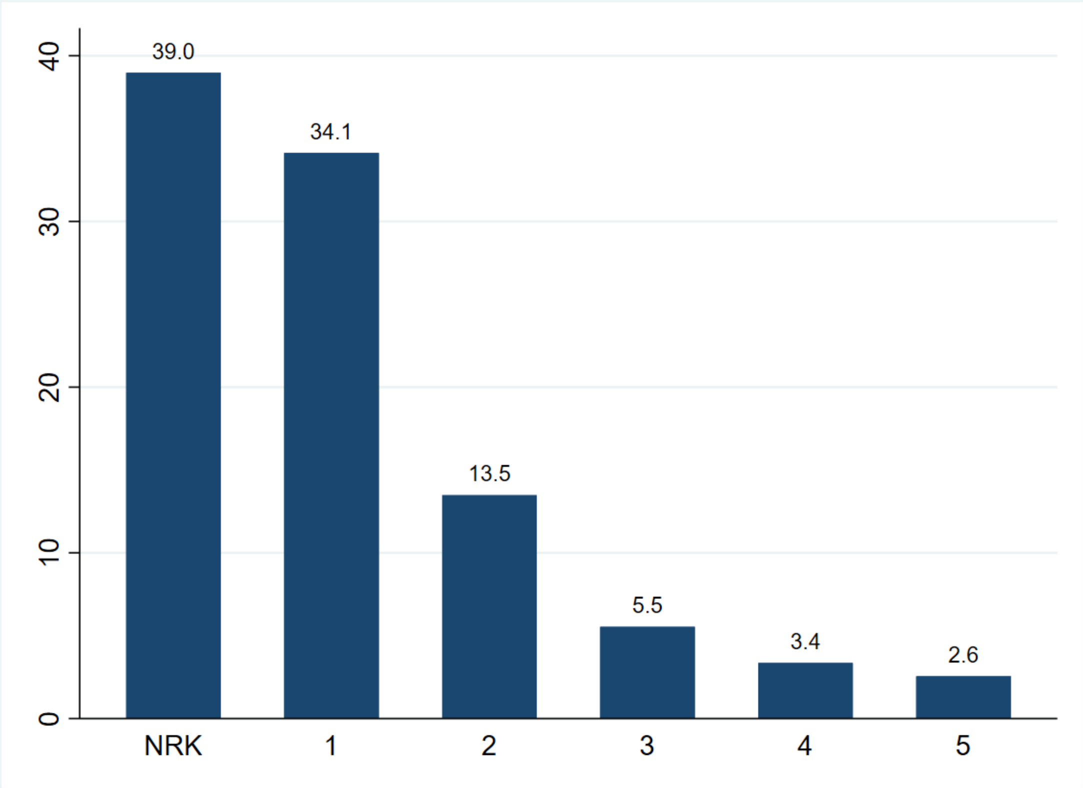
Looking at the data for individual outlets' figures from Table A1 in the Appendix, we observe that in half of the regions, the NRK district page has the highest number of hits. The remaining half consists of the national market, and the regional markets Hordaland, Trøndelag, Rogaland, Sørlandet, Østfold and Østlandssendingen. Using 'unique visitors', it should be noted that only in the national market is a private news outlet (VG.no) larger than NRK. Indeed, the NRK district pages account for 25% of visitors but only 10%of hits. When excluding the national market from the analysis, the NRK regional pages comprise 56 % of visitors on average, but 34 % of hits. In terms of hits per visitor, the NRK district pages receive 2.2 hits per visitor on average, as compared to 3.5 hits per visitor for the regional news outlets. The national market and the NRK national page show a higher ratio, perhaps reflecting a greater breadth of coverage, e.g. all regional news and national common interest topics. In this market, we observe 7.0 hits per visitor per week, on average.

As is evident from Tables 1 and A1, there is great variation in our dataset. First, in the market Østafjells, there are 22 news outlets, while in the market Rogaland, there are only three. The markets also differ when it comes to the relative size of the news outlets. While the number of hits for the smallest news outlet in Rogaland (Haugesund Avis) is slightly less than one third relative to that of the largest outlet (Aftenbladet), the situation is different in the market Østafjells. While NRK Østafjells – the largest outlet – has more than 340 000 hits per week, the 13 smallest outlets each has less than 20 000 hits per week. We also see that the dates of introduction of paywalls vary greatly. While more than half of the news outlets introduced a paywall in 2015, many of the major ones introduced paywalls already in 2012 and 2013 (Fædrelandsvennen, VG, Dagbladet, Aftenbladet, Bergens Tidende, Aftenposten and Agderposten). Hence, while many news outlets introduced a paywall late, a high fraction of the reader-base in our sample had experienced that one or more of their main news outlets introducing a paywall early.

Figure 2 illustrates the average size distribution of media providers. We see that NRK is the most important news source, and their news outlets are still freely accessible. We also see a clear pattern where the largest regional news outlets are substantially larger and have a market share that is two to three times as high as the number two outlets.[10]

[10] The numbers are much the same if we exclude the national market: 40.2% (NRK), 34.4% (no.1), 13.2% (no.2), 5.2% (no.3), 3.0% (no.4) and 2.2% (no.5). The major difference in the national market is that here there is somewhat less dispersion between the news outlets. The largest news outlet is a newspaper (*Verdens Gang*)

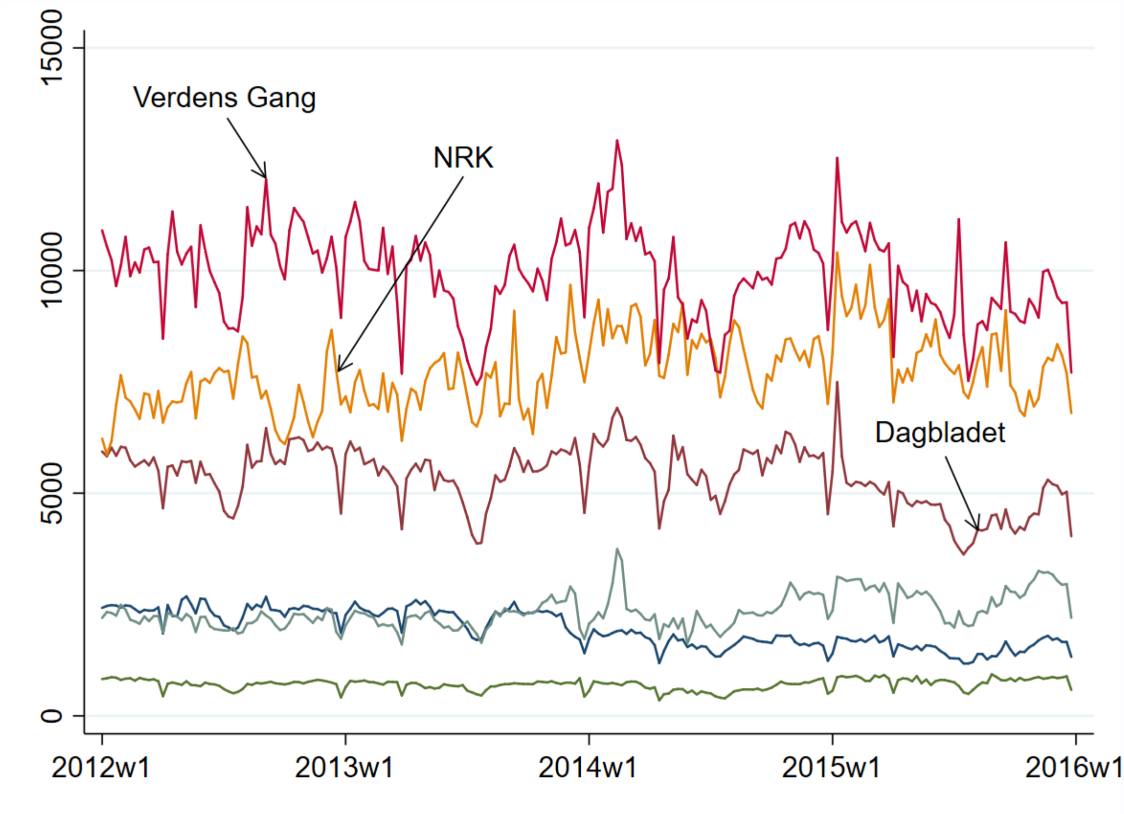
Figure 2: Market share of NRK and the five largest news outlets, average of market shares in 12 regional markets and the national market (based on number of hits).



This picture suggests that the number one news outlets have a different market position in most of the markets. Later we discuss their content profile as compared to that of the other news providers and analyse whether their consumers respond differently to the introduction of paywalls as compared to those of the others. There is also great variation in the weekly development of hits. In addition, the weekly patterns of hits for news outlets in similar markets display similar movements, see Figure 3 below.

that has 30.9% of the market, whereas NRK is number 2 with a market share of 24.2%. Still the number two newspaper that comes in as number 3 has half the market share of the largest newspaper (16.7%).

Figure 3: Development for selected papers of weekly hits in the national market from 2012 to the end of 2015, in thousands.

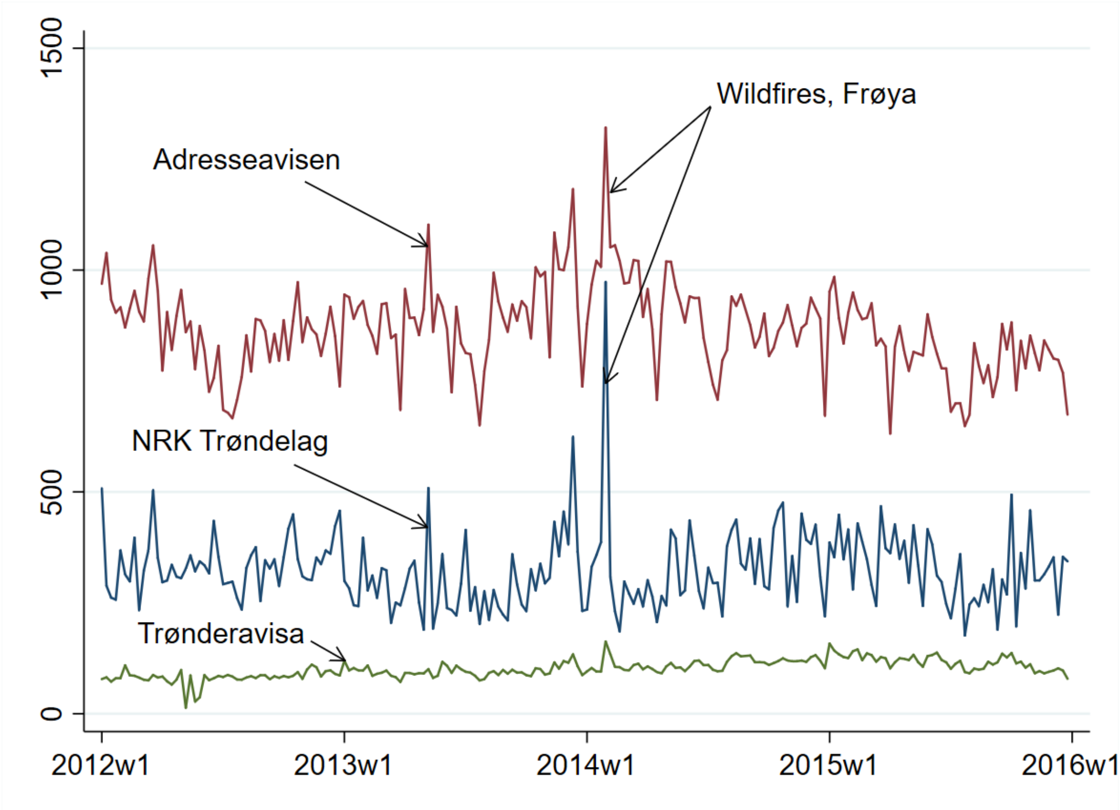


Note: The news outlets in the figure are Verdens Gang, NRK, Dagbladet, TV 2, Aftenposten, and Dagens Næringsliv.

There is no apparent trend in the data, although there does appear to be some seasonality. From Figure 3, we can clearly identify a reduction in demand for news during week 52 (Christmas) and during the Easter holidays. There is also a significant drop in demand for news during the summer holidays, in particular during the month of July. In addition to the seasonal variation, there are instances where the national news outlets are strongly correlated, while there are times when this strong correlation breaks down. Hence, visual inspection of Figure 3 indicates that the national news outlets belong to a common market. In particular, we see that the largest news outlets in the national market – Verdens Gang, NRK and Dagbladet – experience similar increases and decreases of demand for news during the period we analyse. However, there is also news-outlet specific variation in the time-series. This is expected, since the news-producing media will create outlet-specific demand. Finally, there does not seem to be a long-term trend when it comes to hits in our dataset, neither increase nor decrease in the variables *hits*, *sessions* and *visitors*.

Turning to the regional markets, we observe the same pattern as in the national market. The seasonal cycles are clearly present, and there are periods when regional markets correlate strongly, particularly when there are major events taking place in the regions. We give two examples of this below (Figures 4 and 5). We also observe the difference in magnitudes in terms of viewings. Whereas the national news outlet *Verdens Gang* (Figure 3) has around 10 million weekly hits, the largest news outlet in Trøndelag, *Adresseavisen* (Figure 4) has 1 million hits. (See also Table 1A).

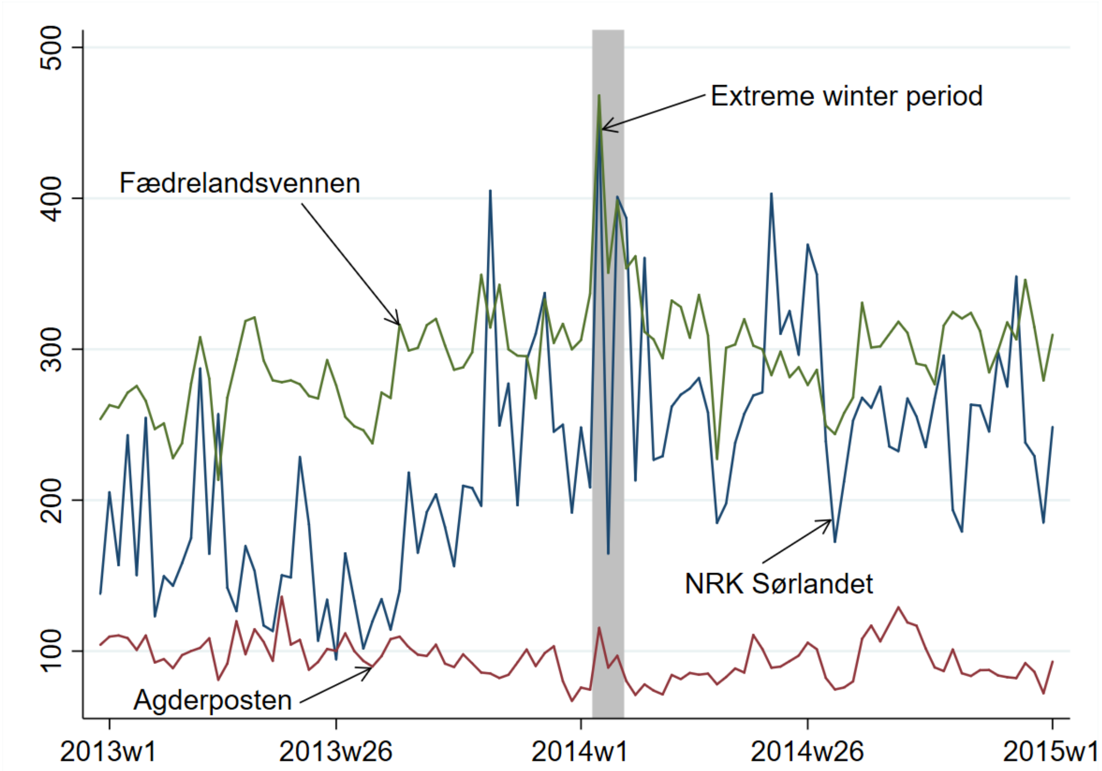
Figure 4: Weekly hits for selected papers in the regional market *Trøndelag* from 2012 to the end of 2015, hits in thousands.



We now focus on the regional market Trøndelag in Figure 4. In early 2014, there were a number of wildfires in the western parts of Norway, notably the fires in Lærdal, Flatanger and Frøya. The latter of these, the fire on the island of Frøya, started at midday on January 29th, and due to extremely dry weather and windy conditions, the fire quickly spread across large areas of the island. There was widespread evacuation of people living in areas affected by the fire. Thus, the fire created a strong demand for news, and, as observed in Figure 4, hits at the news outlets in the region of Trøndelag increased strongly. From Figure 4, we also see that there are other periods where the demand for news from NRK Trøndelag and Adresseavisen correlates in the same manner. This is an indication that these news outlets belong to the same market. There are also periods where the smaller news outlets in the region are clearly strongly correlated; however, this is due to differences in scale and therefore not easily observed in Figure 4. The correlation coefficient between NRK Trøndelag and Adresseavisen is about 0.36.

A similar situation occurred in the same month in the region of Sørlandet. In January 2018, the southern parts of Norway experienced a massive snowfall during a short period of time, and the region of Sørlandet was heavily hit. For several days, roads were kept closed due to lack of snow-removal personnel. In addition to the snowfall, the region was hit by a storm, making it difficult to move around. The extreme weather period increased the demand for news about weather forecasts, traffic-operations and general news about the region. This is also observed by visual inspection of Figure 5. Most of the news outlets in the region witnessed all-time high observations of number of weekly hits and number of unique visitors. This was particularly so for *Fædrelandsvennen* – the largest news outlet of the region - and for NRK Sørlandet.

Figure 5: Weekly hits in the Sørlandet-region for selected papers, in thousands. Data period from the start of 2013 to end of 2015.



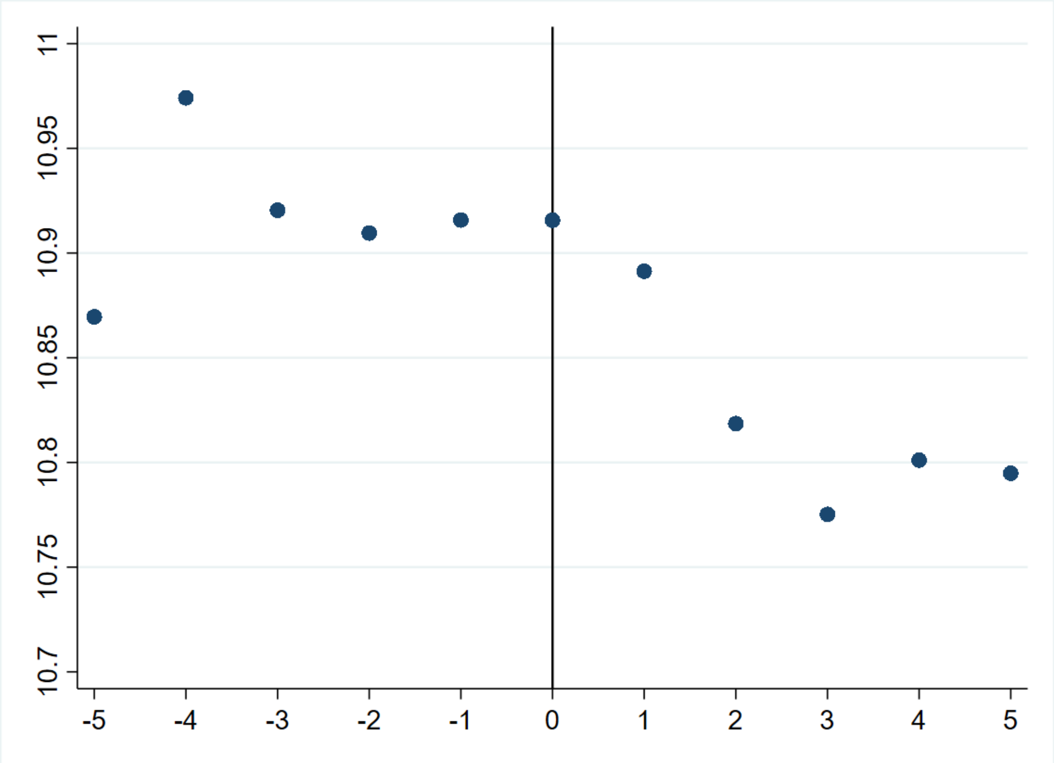
As seen from the discussion above, there are clear links between news providers' weekly hits both in the national market and in the regional markets. The other important observation is that the online activity across news outlets and markets correlates well with that of the NRK sites. The NRK sites have always been freely available, and have a significant readership and, as such, NRK is a well suited control group when analysing private news outlets introducing paywalls.

This correlation in online activity within markets, together with the fact that there are no apparent differences in trends across markets, supports a common trend assumption in our difference-in-difference modelling. Since we later define the treatment group as all news outlets introducing paywalls sequentially over time we cannot perform traditional common trend tests here. In the empirical models, we also include a full set of weekly time dummies to account for seasonality due to holidays etc.

To get a first impression of how the introduction of paywalls affects consumption across all our introductions we scrutinize an event-window of ten weeks around all introductions. In Figure 6 we pool all the news outlets in our sample that introduced a paywall in the period of study. We normalize the week of the introduction to zero and look at the average hits before and after the introduction.

Given that introducing paywalls makes it harder to access news from a site, we expect that news consumption will decrease as a result. Figure 6 does indeed suggest that we see fewer hits after the introduction of a paywall.

Figure 6: Average weekly hits 5 weeks before and after the introduction of the paywall, hits measured in logs.



In the next section we present our empirical strategy and models that will be used to analyse our research questions and hypotheses.

Econometric framework and models estimated

Our main research question asks whether introducing a paywall affects the pattern of consumption of news at news-producing media sites. To test for this we apply a difference-in-difference framework.

Main models - average effects across all news outlets

In order to empirically test the relationship between paywalls and news consumption, we estimate the following generic fixed effect model for the logarithm of our time-series of data on hits, unique sessions and visitors, $ln(x_{i,j,t})$:

$$(1)\quad ln(x_{i,j,t}) = \alpha_0 + \alpha_1 ln(x_{i,j,t-1}) + \alpha_2 ln(x_{i,j,t-2}) + \beta_1 Post_{i,j,t} + \beta_2 Share_{-i,j,t} + \gamma_t + \delta_i + \epsilon_{i,j,t},$$

where subscript *i* refers to online news outlet, j to regional (or national) market, and *t* to time period (week). To account for serial correlation in our time series we specify an autoregressive process of second order AR(2), allowing two lags of the dependent variable on the right hand side. The parameter δ_i is the news outlet-specific fixed effect parameter. A fixed effect for all 52 weeks in our dataset, γ_t, accounts for common demand shocks across all news outlets such as holidays. The $\epsilon_{i,t}$, is the standard error-term, anticipated to have the standard properties and being *iid*.

All models will be estimated with robust standard errors where we allow the error term to be clustered for each news provider.

We account for the general effect of the number of news providers in a market that are behind a paywall. The number of news outlets implementing paywalls increases over time. To take account of this, we explicitly consider the paywall share of our regional news markets. The variable '$Share_{-i,j,t}$' measures the share of readers in a market behind competing papers' paywalls. That is, when paper i introduces a paywall, the 'Share'-variable increases with the market share of paper i for all *other* papers. When paper i introduces a paywall, this variable is unaffected for paper i.

The variable $Post_{i,j,t}$ is an indicator-variable taking the value 1 in all periods t after paper i introduces a paywall, and is equal to 0 otherwise. We expect the parameter measuring this effect, β_1, to be negative, since the paywall introduces restrictions on the news readership. This parameter also serves as the difference-in-difference parameter in our model. Since we apply a fixed effect regression, the treated term used in ordinary regression difference-in-difference models is omitted from the model. All firm-specific differences will be picked up by the δ_i-parameters, and will be omitted due to collinearity if included in a fixed effect regression. In our dataset, treatments do not take place simultaneously. Rather, the introduction of paywalls takes place throughout our sample period. In this respect, we have two types of control groups. First, papers that introduce paywalls late in the sample period, will act as control group for papers that introduce paywalls early. The second type of control group outlets includes papers that never introduce a paywall, and the NRK internet sites (national and regional). NRK is a significant player both in the regional markets as well as in the national market, with dedicated newsrooms with tailor-made news. NRK is financed via a mandatory licence fee for all households that own a television, and is not allowed to charge for their online news services. These sites will act as control group for all news outlets that introduce a paywall.

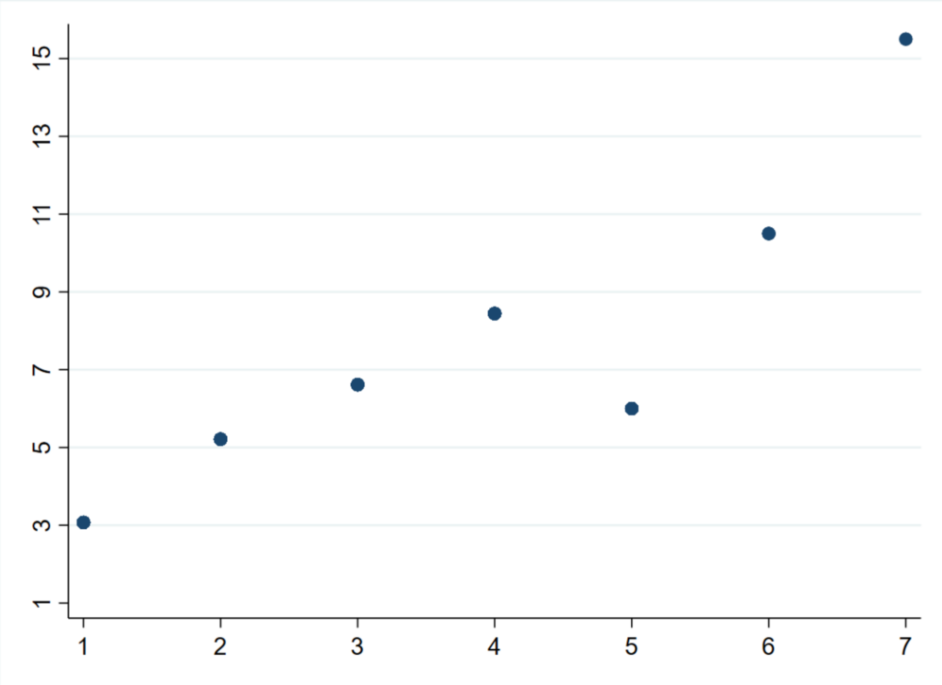
In line with the results found by Kim et al. (2018) we expect that online demand response will differ according to several factors, such as reader demographics, news outlet characteristics and the time of introduction of the paywall. Thus, we will extend our model to account for heterogeneity on both timing and news outlet size.

Heterogeneity due to timing of introduction of paywall

The average effect Model (1) accounts for the general effect on consumption of an increasing number of news providers implementing paywalls. However, we do not allow for any changes in the difference-in-difference effect due to this development. We might expect that the first news outlet that introduces a paywall in a particular market will potentially experience a different negative impact than a news outlet that introduces a paywall after a large fraction of news outlets have already done so. The pool of freely available substitutes is decreasing in the share of the market behind a wall. Studies suggest that consumers will be less price sensitive when there is a smaller number of substitutes (*e.g.*, Gumus, Kaminsky, Mathur 2016). Another argument put forward by Kim et al. (2018) relates to the reference price research, which seems to suggest that the more accustomed consumers are to paying, the less price sensitive they become. Or, as in the case of paywalls, where the product was originally provided for free, with paywalls slowly evolving into a new 'normal', where most providers charge for access. This development will potentially affect consumers' product choice (Mazumdar, Raj, and Sinha 2005). News providers that adopt a paywall late are likely to experience a smaller reduction in demand. Both because late movers have a smaller number of substitutes to compete against and because consumers' references have changed. On the other hand, when a news outlet enters late, consumers may already have purchased access to news outlets that have introduced paywalls early, and are thus potentially less likely to purchase additional subscriptions. This effect will work in the opposite direction of those put forward by Kim et al. (2018).

When we look at the data, we see heterogeneity in paywall introductions. For instance, first-moving papers typically have a relatively strong position within their regional market. The largest news outlets (in our dataset) were typically the first outlets to introduce a paywall in eight of twelve regional markets.[11] We see a clear pattern where the largest news outlets introduce paywalls early. This is illustrated in Figure 7. The horizontal axis measures the order of paywall introduction in a market, (1 means first mover), whereas the vertical axis is the average size (in rank form) of that order. Thus, the average rank of the first mover was 3, and for the second mover it was 5. More generally, the graph shows a clear pattern where the larger news outlets move earlier than the smaller ones.

Figure 7: Average rank of regional news outlets vs their timing of introducing a paywall.



Note: Horizontal axis: The order of papers according to when they introduced a paywall. Vertical axis: The average size (measured by rank) of the ordered introductions.

In sum, there might be a selection regarding the timing of introducing paywalls. We thus introduce an additional model where we allow for an interaction between the '*Share*' and '*Post*' variables. This variable ($Post_{i,j,t} \cdot Share_{-i,j,t}$) measures whether there is a marginal difference in the paywall effect depending on whether the outlet is an early or a late mover. Thus, we expand (1) and also estimate:

$$(2)\quad ln(x_{i,j,t}) = \alpha_0 + \alpha_1 ln(x_{i,j,t-1}) + \alpha_2 ln(x_{i,j,t-2}) + \beta_1 Post_{i,j,t} + \beta_2 Share_{-i,j,t} + \beta_3 Post_{i,j,t} \cdot Share_{-i,j,t} + \gamma_t + \delta_i + \epsilon_{i,j,t},$$

[11] The eight regional markets and the news providers are as follows: Hedmark and Oppland (Oppland Arbeiderblad), Hordaland (Bergens Tidende), Møre and Romsdal (Sunnmørsposten), Sogn and Fjordane (Firda and one other news outlet in the same week), Sørlandet (Fædrelandsvennen), Trøndelag (Adresseavisen) and Østafjells (Drammenstidende) (see Table A1 for details on paywall introductions and news outlets' size).

where β_3 measures the additional effect of news outlet i introducing a paywall early or late: Early, the share of competing firms already behind a paywall will be low, later it will be higher. For instance, the average effect of introducing a paywall will be given by β_1 in Model (1) and by (β_1+$\beta_3 \cdot \overline{Share_{-\imath,\jmath,t}}$) in Model (2), where $\overline{Share_{-\imath,\jmath,t}}$ refers to the mean of $Share_{-i,j,t}$. Since the $Share_{-i,j,t}$ variable differs according to local and national markets, this formulation is more precise in measuring potential time heterogeneity in local demand responses, than simply introducing a time trend across all wall-introductions and markets (which is e.g., what is done in Kim et al. (2018)).

Heterogeneity within news outlets due to readership and news outlet size

The ability to analyse a set of markets also allows us to analyse the effect of news providers' rank and relative size within their relevant markets. In particular, the definition of markets allows us to precisely define which news provider is number one in its relevant market. As we saw above, (see Figure 2) the largest news outlet stands out by being more than double the size of its number two competitor across most markets. This might make the largest news outlets different in terms of what effects to anticipate after the introduction of a paywall.

There seems to be a positive correlation between news providers' size and breadth of content coverage. It has long been common to see news providers differentiate news content to their readers' preferences (Litman and Bridges 1986). Larger news providers thus aim towards covering broader and more general news in order to attract as many consumers as possible. These outlets typically have less unique content and are attractive for their breadth rather than for special news from local areas or local interest groups. This is the case for most of the larger regional news outlets, as well as for the major national news outlets covering the national market. In this respect, one could argue that the larger news outlets have a different readership group than more specialized news outlets.

A number of (often smaller) news providers specialize towards (smaller) more narrowly defined reader groups, either through political slant (*e.g.*, towards more extreme political groups) or through adapting their content to certain groups (*e.g.*, religious news outlets), or simply by having much more local coverage for a smaller population group/area. For instance, Gentzkow and Shapiro (2010) have shown that news outlets respond to their readership's political opinions and tailor their political orientation to differentiate themselves from other papers. Some readers will have a particular interest in local news from e.g., their home municipality, where dedicated and narrower news outlets cover the local market for more in-depth news than regional or national papers are able to. This connection between local coverage and circulation is shown by Lacy and Sohn (1990). In a more recent study, Mitchell, Holcomb, and Page (2015) find evidence that a large majority of readers follow local news very, or somewhat, closely. All this suggests that consumers are less price-sensitive when a brand, in our case a news outlet, has a more unique positioning, which is also shown by Nagle, Hogan, and Zale (2016).

In general, we therefore anticipate that the degree of content uniqueness might influence the effect of introducing a paywall. This is also in line with Kim et al. (2018) who quantify whether the response to introducing paywalls differs according to content uniqueness. Indeed, they find empirical evidence supporting the idea that more unique content providers tend to perform better after paywall rollout than more general content providers. Previous qualitative work on the Norwegian rollout of paywalls also seems to support this finding. Sjøvaag (2016) found that papers which provide some content for free whilst keeping other content behind a paywall, are more likely to place content of local relevance behind the paywall and leave more widely relevant news, such as syndicated content open to all.

Since the largest news providers within our regional markets and the national market typically have a broader scope for their news coverage, we would anticipate that they observe a larger reduction in hits following the introduction of a paywall than the news providers with a larger proportion of unique content.

We thus expand Model (1) to include an estimate for separate effects for the largest and the other news outlets:

$$(3) \quad ln(x_{i,j,t}) = \alpha_0 + \alpha_1 ln(x_{i,j,t-1}) + \alpha_2 ln(x_{i,j,t-2}) + \beta_1^L Post_{i,j,t}^L + \beta_1^O Post_{i,j,t}^O$$

$$+\beta_3 Share_{-i,j,t} + \gamma_t + \delta_i + \epsilon_{i,j,t},$$

where superscript '*L*' refers to the largest news outlet in market j introducing a paywall, and superscript '*O*' refers to all other news outlets introducing paywalls in market j, hence parameters β_1^L and β_1^O quantify the effect of introducing a paywall for the largest news outlets (*L*) and for the other outlets (*O*) respectively.

Empirical results

In this section, we present our empirical results. We start out by estimating and discussing our two main models. Then we extend these to allow for potential heterogeneity between the largest and the other news outlets.

Main models - average effects across all news outlets

Model (1) is presented in Table 2. The model perform well, we explain between 34 and 48% of the variation in the data and all parameters come in significant. The autoregressive components, $x_{i,t-1}$ and $x_{i,t-2}$ are highly significant.[12]

Table 2: Difference-in-difference results, logarithm of hits, unique sessions and unique visitors, Model (1), clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.501***	0.404***	0.372***
	(0.026)	(0.025)	(0.025)
$ln(x_{i,j,t-2})$	0.198***	0.213***	0.207***
	(0.012)	(0.016)	(0.016)
$Post_{i,j,t}$	-0.035***	-0.040***	-0.040***
	(0.010)	(0.010)	(0.010)
$Share_{-i,j,t}$	0.036*	0.039**	0.044**
	(0.019)	(0.020)	(0.020)
Long-term effect	-0.117***	-0.104***	-0.094***
	(0.030)	(0.022)	(0.021)
Constant	3.402***	3.895***	4.092***
	(0.312)	(0.327)	(0.332)
r2	0.475	0.383	0.343
N	18 451	18 451	18 451
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

Across all volume measures we find an adjustment speed of around 30 and 42%, suggesting that when

[12] To determine the lag length we used STATA's varsoc routine for all news providers independent time series on the right side variable (note that this was only doable for 89 news outlets due to sample lengths) and calculated optimal number of lags for the autoregressive process based on the distribution of the AKAIKE information criterion (Akaike, 1974). For the majority of the news outlets the criterion suggested at most two lags for the autoregressive process. See histogram in Figure A1 in Appendix A.

we see changes in volumes in the short run it takes some time before we return to the long run trend.[13] As expected, the difference-in-difference '*Post*'-variable is highly significant and negative, thus, introducing a paywall entails a reduction in hits for the papers as well as in the number of sessions and unique visitors falls. The average short-term effect of introducing a paywall is relatively small, reducing volumes by 3.5 to 4.0%. This result is considerably lower than previous studies have found, but it is a short-run effect and clearly in line with the expectation that volumes are reduced when a news outlet introduces a paywall.

In addition, the $Share_{-i,j,t}$ variable is positive and significant, suggesting that as more competing news outlets move behind paywalls, volumes generally increase for those that do not introduce paywalls. One implication is thus, that there is a significant increase in consumption of news on free internet media sites after a competing paper introduces a paywall.

We use the log-log framework in our model, hence the parameter estimates are elasticities, and changes could be interpreted as percentage changes, suggesting that a 10% increase in the Share variable increases volumes by between 3.6 to 4.4%.[14]

In row 5 of Table 2 we have calculated the long-run elasticity for introducing a paywall.[15] In line with the low adjustment speed the estimated long-term effects are substantially larger than the short-term effects. All long-run effects are significant, and suggest a reduction in traffic between 9 to 12%. The difference between short- and long-run effects is highly significant for all measures.[16] The highest figures are found for number of hits. These figures are more in line with previous findings but remain smaller. Chiou & Tucker (2013) found an effect as high as 51% and Pattabhiramaiah et al. (2018) had an estimate of 17%. However, these two studies include very few news outlets in the treatment group (one and three, respectively) and they have either fewer periods before and after the introduction of paywalls, or less relevant control groups. We look at 69 paywall introductions, and have a larger control group operating in the same markets, and use high frequency data for a longer period. Comparing

[13] The adjustment speed can be written as (1-α_1-α_2), which means that for e.g., hits we find the adjustment speed to be 1-0.501-0.198=0.301.

[14] Note that since $Share_{-i,j,t}$is a percentage variable, a one percentage change is not the same as a one percentage-point increase, e.g., a change in the Share from 0.50 to 0.51 represents a 2% increase when we refer to the elasticity.

[15] In the estimated models the long-run solution is the steady-state solution, implied from equalizing all periods and thus treating variables as equal regardless of their time lag. In our case this implies that we assume that $x_{i,j,t} = x_{i,j,t-1} = x_{i,j,t-2} = x_{i,j}$. Hence, the long-run effect in Model (1) is given as $\beta_1^* = \beta_1/(1-\alpha_1-\alpha_2)$. Since the long-run parameters are non-linear combinations of the short-run parameters we use the "delta method" to approximate standard errors.

[16] The differences between short- and long run are all significant; 0.082 (0.021), 0.064 (0.014) and 0.055 (0.012) for hits, session and visitors respectively. Delta calculated standard errors in parentheses.

these numbers to the findings of Kim et al. (2018), we are in the lower end of their distribution of long run individual news outlet estimates in the range of -54 to -10%.

We now scrutinize the timing heterogeneity by allowing the difference-in-difference effects to differ with the $Share_{-i,j,t}$ variable.

Heterogeneity due to timing of entry of paywall

Model (2) allows for timing heterogeneity in the difference-in-difference effects. The results are tabulated in Table 3. Our results parallel the results discussed above. Most of the joint parameters and explanatory power stay the same. The '*Post*' variable is still significant and in the same range as in Model (1). The interactions between '*Post'* and '*Share'* are not significant however.

Table 3: Difference-in-difference results, logarithm of hits, unique sessions and unique visitors, Model (2), clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.501***	0.404***	0.372***
	(0.026)	(0.025)	(0.025)
$ln(x_{i,j,t-2})$	0.198***	0.213***	0.207***
	(0.012)	(0.016)	(0.016)
$Post_{i,j,t}$	-0.037***	-0.039***	-0.041***
	(0.013)	(0.010)	(0.012)
$Post_{i,j,t} \cdot Share_{-i,j,t}$	0.003	-0.003	0.004
	(0.035)	(0.037)	(0.040)
$Share_{-i,j,t}$	0.035*	0.040*	0.043**
	(0.019)	(0.020)	(0.021)
Short-term effect	-0.036***	-0.039***	-0.040***
	(0.009)	(0.007)	(0.008)
Long-term effect	-0.119***	-0.103***	-0.096***
	(0.028)	(0.018)	(0.018)
Constant	3.403***	3.895***	4.093***
	(0.311)	(0.326)	(0.331)
r2	0.475	0.383	0.343
N	18451	18451	18451
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

To measure the effect of introducing the paywall we still need to consider both the Post and the interaction parameter between *Post* and *Share*.[17] Since the interaction variable is very small, the

[17] We estimate the effect (elasticity) of introducing the paywall for the average *Share*, i.e.,

elasticities for introducing a paywall for the competitors' average '*Share*' behind the wall are much the same as what we found for Model (1), ranging from 3.6 to 4%, and the difference is not significant.

In row 7 of Table 3, the long-run effects for the difference-in-difference results for Model (2) are tabulated. They are of similar size to those found in Model (1), suggesting only marginally higher long-run elasticities in the order of 10 to 12%.

Hence, when allowing also the difference-in-difference effects to vary with the number of competing news providers behind a paywall, we still find a relatively low short-term negative effect of 4% from introducing a paywall for the average news outlet in our sample. The long-term effect increases somewhat, but is still lower than what has been found in earlier studies.

The elasticities from Model (2) are calculated using the average *Share* of competitors behind a paywall across the sample (=0.210). Since the interaction parameter is very small, and insignificant, the difference between adopting a payment early or late is negligible, and not significant.

We have also estimated a model where we introduce a time trend (*'Trend')* and an interaction between '*Trend'* and our difference-in-difference variable *'Post'*. This specification is closer to the model estimated by Kim et al. (2018). The interaction ($Post_{i,j,t} \cdot Trend_t$) measures whether there is a potential overall linear trend in the effects from introducing paywalls, but now across all paywall introductions across all markets.[18] The results are presented in the Appendix in Table A2. In this trend model we find similar but lower and less precise elasticities for the average news outlet, in both the short- and the long run, suggesting a short-term decrease in demand between 2 to 3% and a long run decrease between 4 to 10%. However, now the whole difference-in-difference effect is picked up by the interaction parameter, which is only significant (at a 5%-level) for the two models. The positive effect from the share of competitors behind the wall found in Model (2) seems to be picked up by the trend-parameter (these variables have a correlation coefficient of 0.596). Since the significant interaction terms come out negative, this suggest that, if anything, the negative effect of introducing paywalls increases linearly over time. Note however, that in our setting with 13 different markets the trend model is biased due to the fact that the sequence of entering payment walls in local markets is

---

$\mu_{paywall} = \beta_1 + \beta_3 \cdot \overline{Share_{-\iota,J,t}}$

[18] We expand Model (1) to include the $Trend_t$and the $Post_{i,j,t} \cdot Trend_t$ interaction:

$$ln(x_{i,j,t}) = \alpha_0 + \alpha_1 ln(x_{i,j,t-1}) + \alpha_2 ln(x_{i,j,t-2}) + \beta_1 Post_{i,j,t} + \beta_2 Trend_t + \beta_3 Share_{-i,j,t} + \beta_4 Post_{i,j,t} \cdot Trend_t + \gamma_t + \delta_i + \epsilon_{i,j,t}.$$

The $Trend_t$ takes the value 1 in the first week of observation for all news outlets. The variable increases by (1/52) for each week, taking the value 2 after one year of observations, 3 after three years etc. The parameter β_4 measures the interaction effect; how does the difference-in-difference effect change over time across markets.

not captured through a 'national' trend variable. The trend formulation picks up a more general trend across the whole country, rather than the development in each regional market. Thus, we are reluctant to place too much weight on this formulation of our models.

In sum, timing heterogeneity in the sense of when to introduce a paywall does not seem to be very important in the Norwegian market. As we argue above, Model (2), where we allow for timing heterogeneity and measure the effect through the local market based share of competing news outlets behind a wall, measures timing heterogeneity more precise than a general linear trend across all markets. For instance, the trend model assumes that the introduction of paywalls in the Sørlandet-region in the South of Norway is as important for the local market in the North as if regional news outlets in the North (more than 3000 km away) had introduced paywalls. Thus, we need to be careful when interpreting the trend-model, and conclude that in terms of timing heterogeneity, we believe this effect to be very modest in the Norwegian data.

Heterogeneity within news outlets due to readership and news outlet size

Generally, Kim et al. (2018) show that the paywall effect differs according to news outlet demographics and readership. We are able to expand on their analysis, using more detailed information on the relative market position of the news outlets in terms of size distribution in their relevant markets. In this subsection, we thus open up for size heterogeneity among the news outlets.

In Models (1) and (2) we have looked at average effects for any news outlet introducing a paywall, but have allowed the difference-in-difference parameter to change as more competing news outlets go behind a paywall. As we discussed above, aside from NRK, the largest news outlets are significantly larger than their competitors within their respective markets. In order to serve a large fraction of consumers within the market, the largest news outlets typically have a more general news content, suggesting heterogeneous demand responses to paywalls for the largest providers as compared to the others. In Model (3), we will allow for this heterogeneity by estimating separate effects for the largest news outlets and the others. The results are shown in Table 4.

In essence, the results are similar to those of Model (1). This comes as no surprise since the only difference from Model (1) is the way in which we divide the *Post*-variable into two separate variables. However, we still find significant difference-in-difference effects for the introduction of the paywall across all models and parameters, but we now observe some signs of heterogeneous responses across groups. For the largest news outlets, the short run effect is a decrease in demand of around 5%, whereas it is around 3% for the others, suggesting that the largest news outlets take a bigger relative hit when introducing walls. The differences between groups are between 1 and 2 percentage points across the models, but they are not significant.[19]

Table 4: Difference-in-difference results for largest and other news providers, Model (3), logarithm of hits, unique sessions and unique visitors, clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.500*** (0.026)	0.404*** (0.025)	0.372*** (0.025)
$ln(x_{i,j,t-2})$	0.198*** (0.012)	0.213*** (0.016)	0.207*** (0.016)
$Post^{L}_{i,j,t}$	-0.045*** (0.013)	-0.052*** (0.009)	-0.055*** (0.011)
$Post^{O}_{i,j,t}$	-0.032*** (0.012)	-0.036*** (0.012)	-0.035*** (0.012)
$Share_{-i,j,t}$	0.034* (0.019)	0.037* (0.020)	0.041** (0.020)
Long-term: Largest	-0.150*** (0.042)	-0.137*** (0.021)	-0.131*** (0.026)
Long-term: Other	-0.107*** (0.037)	-0.094*** (0.028)	-0.083*** (0.026)
Constant	3.405*** (0.311)	3.899*** (0.326)	4.097*** (0.331)
r2	0.475	0.383	0.343
N	18 451	18 451	18 451
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

Turning now to the long-run effects from Model (3), these are tabulated in rows 6 and 7 of Table 4. The long-run negative effects for the largest news outlets are now larger, suggesting a reduction in online activity between 13 and 15%. For the other news providers the figures vary between 8-11%. The differences are thus much larger and vary between 4 to 5%, but are still not statistically significant.[20] Hence, when accounting for heterogeneity in size and market rank, we generally see larger negative responses from the demand side following introductions of paywalls.

[19] For hits, sessions and visitors, the differences are -0.013 (0.016), -0.017 (0.012) and -0.020 (0.014) respectively, delta calculated standard errors in parentheses.
[20] For hits, sessions and visitors, the differences are -0.044 (0.054), -0.043 (0.033) and -0.048 (0.035) respectively, delta calculated standard errors in parentheses.

To explore both types of heterogeneity simultaneously we also estimate a fourth model where we introduce interaction effects between the difference-in-difference and the share variables for the two groups of news outlets. Thus, this model allows us to also differentiate the timing (share) effect for the largest news outlets and the others:[21]

The results are tabulated together with long-run elasticities in Table A3 in the Appendix. Generally, we find much the same when it comes to the "*other"* group, and somewhat larger long-run effects for the largest news outlets' elasticities, that now range from 15 to 17%, as compared to the results for Model 3, which range from 13 to 15%. As in Model (2), the interaction terms are mostly insignificant. We find significance in two cases, for larger news outlets for sessions and visitors, suggesting an increased negative effect as more competing news outlets are behind a wall. As we discussed above, the largest news outlets (in our dataset) in eight of twelve regional markets were typically the first outlets to introduce a paywall (see also footnote 11). Hence, we ascribe this larger and significant effect from the interaction term to be driven by a few news outlets, and we are therefore careful in the interpretation of the results in this combined extended model. Generally, as argued, we typically find only weak evidence for heterogeneity in timing. To the extent that we do find some heterogeneity, it seems that the largest news outlets benefit from being early adopters of payment walls.

Summing up, we find only weak evidence of timing heterogeneity. When looking at heterogeneity within news outlets we find stronger evidence. From Model (3) we find that larger news outlets that typically have a broader news content and a broader readership, will take a bigger hit from introducing a paywall, this is true even when controlling for timing heterogeneity. Our long-run results vary between 8 and 17 percentage points, which is more in line with average estimates found earlier in the literature.

Our dynamic autoregressive distributed lag models will account for likely serial correlation in the left hand side variables. However, we have also performed the Box-Pierce test on the residuals in Model (3) to test for potential autocorrelation in the error terms. Since we have a panel, we undertake tests for our individual news outlets. We can keep the null-hypothesis of no autocorrelation for 87% of the

---

[21] The combined model accounting for both types of heterogeneity is thus:

$$ln(x_{i,j,t}) = \alpha_0 + \alpha_1 ln(x_{i,j,t-1}) + \alpha_2 ln(x_{i,j,t-2}) + \beta_1^L Post_{i,j,t}^L + \beta_1^O Post_{i,j,t}^O$$

$$+\beta_2 Share_{-i,j,t} + \beta_3^L Post_{i,j,t}^L \cdot Share_{-i,j,t} + \beta_3^O Post_{i,j,t}^O \cdot Share_{-i,j,t} + \gamma_t + \delta_i + \epsilon_{i,j,t}$$

The first part of this combined model matches Model (3). However, the last to parameters, β_3^L and β_3^L measure the interaction effect outlined in Model (2), but where we now differentiate between large and small news outlets. For instance, the average effect of introducing a paywall for a large news outlet is β_1^L in Model (3) and by (β_1^L+$\beta_3^L \cdot \overline{Share_{-\iota,J,t}}$) and in this combined model.

news outlets on a 99% significance level.[22] In the next Section, we will use Model (3) when exploring the robustness of our results.

Robustness analysis

Here we focus on two issues. First, we look at the time dimension in the sense that we would like to know whether the longevity of the sample is driving our results. That is, one might think that there is a strong underlying time trend that influences our difference-in-difference results since we consider relatively long before and after periods for quite a few of the news outlets, and the variation in these periods could be influenced by underlying long-run trends. Obviously, the inclusion of an autoregressive process as well as a full set of week dummies will account for such effects. However, to make sure that a significant change in demand takes place around the introduction of the paywall, we also undertake an event study looking at only a limited period before and after each news outlets' paywall entry. Second, we look more closely at our market definition. We have so far included the national market as a thirteenth market, thereby treating the national market similarly to the regional markets. Of our 18 451 week-news-provider observations, the national market amounts to 2 049 observations (11%). To investigate the robustness of this assumption, we estimate Model (3) where we exclude the national market. We will also estimate a model where we reduce our control group and take out all NRK sites.

Event analysis

In this part, we define the week in which a news outlet introduced its paywall to be week 0. Then we include all five-week-periods prior to introducing a paywall (the estimation window), and all five-week-periods after the introduction (the event window). We use the average effect Model (1), and apply an identical difference-in-difference structure, but we now drop the competitors' share behind the wall variable and the weekly-indicator variable, and do not include a lag structure.

The results are shown in Table 5. We replicate our main findings from Model (1) in Table 2 even when we look at the ten-week window. We find a significant reduction in demand ('*Post*') across all three online demand measures. The difference-in-difference effect is now measuring what happens within a five-week span, and, as such, is neither directly comparable to our short-run or our long-run estimates. The estimates in the event-model range between 8 and 12%, which is somewhere between our short- and long-run estimates. Hence, even with a five-week event window we confirm our findings in the much longer main sample.[23]

---

[22] For 105 news outlets we keep the null of no autocorrelation of first order on a 99%-level. For 91 we keep the null on a 95% level.

[23] This is the case if we also include the competitors' share behind a payment wall (not reported).

Table 5: Event analysis five weeks before/after introduction of paywall, Difference-in-difference results, logarithm of hits, unique sessions and unique visitors, reduced Model (1) without competitors' share behind payment walls, clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$Post_{i,j,t}$	-0.119***	-0.088***	-0.083***
	(0.029)	(0.018)	(0.019)
Constant	10.929***	9.581***	9.207***
	(0.016)	(0.010)	(0.010)
r2	0.078	0.066	0.053
N	748	748	748

* p<0.10, ** p<0.05, *** p<0.01

An alternative market definition

We re-estimate Model (3) where we include only our 12 regional markets, excluding the potentially different national market. This reduces our estimation sample by 11%. The results are presented in Table 6.

Generally, we find similar results to those shown for the whole sample in Table 4. We lose only around 1 percentage-point of explanatory power, and we mostly keep the significance structure on our parameters. The AR(2) variables come in significant with very similar values as for the whole sample. The same is true for the difference-in-difference variables and elasticities. Only the long-run elasticities for the 'other' group are between one and two percentage points lower in magnitude. The difference-in-difference elasticities lose some significance, but are still significant on a 5%-level. The share variable, though similar in numbers, is no longer significant. Apart from this, we fully confirm the results from Table 4 where we estimate across the whole sample, also including the national market as one of our 13 defined markets.

Table 6: Difference-in-difference results Model (3), logarithm of hits, unique sessions and unique visitors, robust standard errors, all regional markets, national market excluded.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.496***	0.400***	0.370***
	(0.027)	(0.026)	(0.026)
$ln(x_{i,j,t-2})$	0.195***	0.210***	0.205***
	(0.013)	(0.017)	(0.016)
$Post^{L}_{i,j,t}$	-0.048***	-0.055***	-0.057***
	(0.014)	(0.009)	(0.012)
$Post^{O}_{i,j,t}$	-0.025**	-0.030**	-0.031**
	(0.012)	(0.012)	(0.013)
$Share_{-i,j,t}$	0.033	0.034	0.036
	(0.021)	(0.022)	(0.022)
Long-term: Largest	-0.155***	-0.140***	-0.135***
	(0.044)	(0.022)	(0.027)
Long-term: Other	-0.080**	-0.076**	-0.073**
	(0.037)	(0.030)	(0.029)
Constant	3.384***	3.846***	4.007***
	(0.319)	(0.330)	(0.333)
r2	0.462	0.368	0.331
N	16 402	16 402	16 402
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

As an additional robustness test of our market definition, we estimate our Model (3) where we leave out NRK regional sites from the control group. Since NRK is publicly funded and might have a different regional and national role, they might differ from the other commercial regional news outlets. Thus, in this model, the control group consists of all other news outlets in our sample. The results are tabulated in Table 7. We find very similar results as we have in Table 4, the explanation power increases for all three models, and all our major predictions come through.

Table 7: Difference-in-difference results Model (3), logarithm of hits, unique sessions and unique visitors, robust standard errors, all regional markets, NRK-sites excluded.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.545***	0.436***	0.393***
	(0.025)	(0.032)	(0.033)
$ln(x_{i,j,t-2})$	0.197***	0.231***	0.223***
	(0.014)	(0.016)	(0.016)
$Post^{L}_{i,j,t}$	-0.038***	-0.043***	-0.047***
	(0.011)	(0.007)	(0.010)
$Post^{O}_{i,j,t}$	-0.028***	-0.025***	-0.025**
	(0.010)	(0.009)	(0.010)
$Share_{-i,j,t}$	0.030	0.013	0.017
	(0.019)	(0.018)	(0.020)
Long-term: Largest	-0.148***	-0.128***	-0.122***
	(0.044)	(0.022)	(0.026)
Long-term: Other	-0.110***	-0.076***	-0.064***
	(0.037)	(0.026)	(0.024)
Constant	2.879***	3.287***	3.615***
	(0.221)	(0.275)	(0.343)
r2	0.544	0.459	0.402
N	15 991	15 991	15 991
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

Summing up our robustness section, our results do not change much, neither when we focus on an event-window around the paywall introductions, nor when we scrutinize the market definition. The national market does not appear to be very different from the set of regional markets, and changing the control group by excluding NRK does not change our conclusions.

Discussion and conclusions

The current paper analyses the impact of introducing paywalls on the demand for online news. We estimate models for number of hits, unique sessions and unique visitors. Applying a difference-in-difference framework, we utilize very detailed weekly data across 122 news outlets from January 2012 to December 2015 on the usage of electronic news outlets before and after the introduction of paywalls. Of these, 69 introduced paywalls during our sample period. We contrast this to the consumption of news from a number of news providers that offered open access to their online content throughout the sample period. The national public service broadcaster (NRK) was providing freely attainable news throughout the whole period. Given the Norwegian media topography, we are able to allocate all our news outlets into 13 well defined local (and a national) market(s). NRK was providing online news tailor made for these regions throughout our data period.

Using a dynamic autoregressive distributed lag framework we are also able to provide both short and long run results on the effects of introducing payment walls.

Analysing a relatively large number of markets enables us to study heterogeneity with regard to both how differences in type of news outlet (relative size in their local markets) affect the paywall introductions, and how heterogeneity in the timing of paywall introductions affects demand responses.

We find that the short run average impact of a paywall introduction on the number of hits is negative and, between 3 and 4%, which is smaller than in previous studies. However, the effect is found to be much higher in the long run and when we control for news outlet heterogeneity. After a paywall introduction, the long-run reduction in demand for the average news outlet is between 9 and 11%. This difference between short- and long run impact may seem surprising at first glance. We find that the longer the paywall exists, the stronger the impact from its reader base will be. Readers' habits take some time to change. In our model, the short run is defined as only a couple of weeks, and though readers will spend less time visiting the news outlet, it will take some time before a new long-run equilibrium is reached. First, knowledge about suitable alternatives will increase as time passes from the introduction of the paywall. While some readers will emigrate to alternative news outlets immediately, others are less knowledgeable about alternatives, and will emigrate later. Second, in some instances, the paywall was extended over time: That is, the first weeks after a news outlet introduced a paywall, a fairly small fraction of articles published was located behind the paywall. As time went by, this fraction increased, and readers without a subscription were forced to reduce their consumption of news on this news outlet. Thus, both demand (learning) and supply (paywall-introduction structure) can explain the higher long run responses in demand. This is also supported by the findings in our event study, where the effects across a five-week window are higher than our estimated short run effects, but lower than our estimated long run effects.

Turning to the relative ranking and size of the news outlets, we also find some evidence for heterogeneity in the paywall responses. The largest news outlet within its regional (national) market experiences larger effects than the other news outlets. The largest news outlets face a long-run reduction in demand between 13 and 15% after paywall introductions, as compared to the others that experience between 8 and 11% decrease in demand.

As competing news outlets install paywalls, the share of freely available online news is reduced. Thus, we control for the competitors' share of hits, unique sessions and unique visitors when estimating the effect of paywall introductions. As the percentage of news content behind the competitors' paywall increases by 10%, we find a general and significant increase in online consumption between 3.6 to 4.4%.

We also analyse the extent to which the timing of introducing a paywall affects demand. By allowing the effect to differ according to the amount of the market that is behind the competitors' paywalls, we generally find that the effect does not change significantly. However, for the largest news outlets, we find some indication of an increased negative demand effect as more and more competing news outlets introduce payment walls.

We show that our results are robust to autocorrelations, and we also replicate our results in an event study where we only focus on the five weeks before and after the introduction of paywalls.

Our results seem to suggest that paywalls do indeed reduce demand, but to a lesser extent than what is found in other studies. Compared to these, our study has more detailed data over a longer period of time for many more news outlets and paywall introductions. One weakness in our data is that heterogeneity beyond relative size and timing is not controlled for beyond fixed effects. A future study should aim to also include additional information on readership and reader demography.

References

Aral, S., and Dhillon, P. (2017). Digital Paywall Design: Effects on Subscription Rates & Cross-Channel Demand. Working paper.

Casadesus-Masanell, R., and Zhu, F. (2010). Strategies to fight ad-sponsored rivals. *Management Science*, 56(9), 1484-1499.

Chiou, L., and Tucker, C. (2013). Paywalls and the demand for news. *Information Economics and Policy*, 25(2), 61-69.

Gentzkow, M., and Shapiro J.M. (2010) .What drives media slant? Evidence from U.S. daily newspapers. *Econometrica,* 78(1):35-71.

Gumus, M., Kaminsky P., and Mathur S. (2016). The impact of product substitution and retail capacity on the timing and depth of price promotions: Theory and evidence. *International Journal of Production Research,* 54(7), 2108-2135.

Hognaland, I., and Saebi, T. (2015). Business Model Selection and Innovation: An Empirical Investigation of the Norwegian Newspaper Industry. Beta. *Scandinavian Journal of Business Research*, 29(01), 50-73.

Høst, S. (2016). Papiravisene lever, og de satser digital. Status ved utgangen av 2015. Høgskulen i Volda (Volda University College).

Kim, H., Song, R., and Kim, Y. (2018). Free-to-Paid Transition of Online Newspapers: The Impact of Paywalls on Website Pageviews. Working Paper, October 2018

Lambrecht, A., and Misra, K. (2016). Fee or Free: When Should Firms Charge for Online Content?. *Management Science*, 63(4), 1150-1165.

Lacy, S. and Sohn A.B. (1990). Correlations of newspaper content with circulation in the suburbs: A case study. *Journalism & Mass Communication Quarterly,* 67(4): 785-793.

Litman, B.R. and Bridges J. (1986). An economic analysis of daily newspaper performance. *Newspaper Research Journal,* 7, 9-26.

Mazumdar, T., Raj S.P., and Sinha I. (2005). Reference price research: Review and propositions. *Journal of Marketing* 69(4), 84-102.

Mitchell, A., Holcomb J., and Page D. (2015). Local news in a digital age. Pew Research Center.

Nagle, T.T., Hogan J.E., and Zale J. (2016). *The Strategy and Tactics of Pricing: A Guide to Growing More Profitably*, 5th ed. (Routledge, New York).

Pattabhiramaiah, A., Sriram, S., and Manchanda, P. (2018). Paywalls: Monetizing Online Content. Georgia Tech Scheller College of Business Research Paper No. 17-8.

Sjøvaag, H. (2016). Introducing the Paywall: A case study of content changes in three online newspapers. *Journalism Practice*, *10*(3), 304-322.

Appendix

Table A1: News outlets and their regional markets, averages 2011-2015.

Paper	Introduction		Hits	Visitors	Market Share
	Year	Week			
National Market					
Verdens Gang	2012	24	9 927 362	1 338 830	30.93
NRK			7 742 327	996 398	24.13
Dagbladet	2013	11	5 368 873	807 286	16.73
Nettavisen			2 907 840	504 571	9.06
TV 2			2 345 140	442 629	7.31
Aftenposten	2013	47	1 950 088	357 144	6.08
Hegnar Online			1 048 468	70 514	3.27
Dagens Næringsliv	2014	11	710 115	123 562	2.21
Nationen	2014	13	49 040	10 930	0.15
Vårt Land			42 041	11 206	0.13
Morgenbladet			900	390	0.00
Klassekampen	2015	38	46	17	0.00
Hedmark and Oppland					
NRK Hedmark og Oppland			270 921	121 387	55.00
Oppland Arbeiderblad	2015	5	76 340	14 123	15.50
Østlendingen	2015	16	62 111	13 746	12.61
Glåmdalen	2015	17	54 005	8 730	10.96
Avisen Hadeland			17 200	2 859	3.49
Avisa Valdres			12 018	2 891	2.44
Østfold					
NRK Østfold			210 621	106 363	30.38
Fredrikstad Blad	2015	17	270 934	31 858	39.08
Moss Avis	2015	16	80 999	14 179	11.68
Sarpsborg Arbeiderblad	2015	17	62 419	9 489	9.00
Halden Arbeiderblad	2015	16	37 766	5 828	5.45
Smaalenenes Avis	2015	16	24 548	4 814	3.54
Rakkestad Avis	2015	24	5 939	976	0.86
Østlandssendingen					
NRK Østlandssendingen			228 214	128 633	31.76
Romerikes Blad	2015	16	250 956	43 662	34.92
Budstikka.no			120 486	27 808	16.77
Østlandets Blad	2015	19	47 621	10 032	6.63
Dagsavisen	2014	46	36 074	15 457	5.02
Akershus Amtstidende	2015	19	10 976	2 279	1.53
Indre Akershus Blad	2015	16	7 058	1 852	0.98
Eidvoll Ullensaker Blad			6 658	2 120	0.93
Vestby Avis	2015	25	5 866	968	0.82
Enebakk Avis	2015	25	2 487	398	0.35
Ås Avis	2015	25	2 174	461	0.30
Østajells					
NRK Østafjells			346 089	197 086	31.29
Drammens Tidende	2014	38	186 134	35 859	16.83
Tønsbergs Blad	2014	39	141 894	29 856	12.83
Telemarksavisa	2015	35	101 697	19 739	9.19
Østlandsposten	2015	6	61 821	9 991	5.59

Varden	2014	46	52 603	10 590	4.76
Sandefjords Blad	2014	38	45 022	7 295	4.07
Ringerikes Blad	2015	5	40 497	6 687	3.66
Lågendalsposten	2014	38	32 226	6 084	2.91
Porsgrunns Dagblad	2015	24	19 286	3 551	1.74
Hallingdølen			16 197	3 458	1.46
Gjengangeren	2015	6	16 115	3 308	1.46
Bygdeposten	2015	6	12 377	2 498	1.12
Telen	2015	6	10 413	2 239	0.94
Kragerø Blad Vestmar	2015	4	7 694	1 895	0.70
Røyken og Hurums Avis	2015	24	5 585	1 460	0.50
Jarlsberg Avis	2015	6	3 574	748	0.32
Svelviksposten	2015	25	1 804	456	0.16
Llierposten	2015	25	1 766	484	0.16
Øyene	2015	25	1 666	279	0.15
Sande Avis	2015	25	1 289	363	0.12
Eiker Avis	2015	24	400	140	0.04
Sørlandet					
NRK Sørlandet			219 366	100 365	32.56
Fædrelandsvennen	2012	11	304 193	50 285	45.15
Agderposten	2013	50	92 429	17 262	13.72
Lindesnes Avis			18 072	3 962	2.68
Llister24			18 035	4 658	2.68
Tvedestrandsposten	2015	24	11 122	2 020	1.65
Aust-Agder Blad			10 505	1 569	1.56
Rogaland					
NRK Rogaland			237 651	119 350	25.32
Aftenbladet	2013	21	531 353	90 499	56.62
Haugesunds Avis	2014	38	169 418	28 459	18.05
Hordaland					
NRK Hordaland			293 648	159 309	16.68
Bergens Tidende	2013	42	882 458	157 262	50.11
Bergensavisen	2015	49	537 260	90 110	30.51
Dagen			16 043	5 029	0.91
Kvinnheringen	2015	24	10 568	2 240	0.60
Hardanger Folkeblad	2015	24	9 580	2 140	0.54
Avisa Nordhordland	2015	24	7 757	1 814	0.44
Strilen			1 050	380	0.06
Bygdanytt			975	449	0.06
Vestnytt			904	363	0.05
Askøyværingen			716	245	0.04
Sogn and Fjordane					
NRK Sogn og Fjordane			737 691	126 390	87.92
Firda	2014	39	64 668	11 549	7.71
Firdaposten	2014	39	21 385	3 164	2.55
Fjordens Tidende	2015	19	8 494	2 036	1.01
Fjordingen	2015	19	6 771	1 484	0.81
Møre and Romsdal					
NRK Møre og Romsdal			282 274	107 948	38.92
Sunnmørsposten	2014	45	207 099	36 422	28.56
Romsdals Budstikke			135 233	20 189	18.65
Tidens Krav	2015	4	60 307	10 664	8.32
Vikebladet-Vestposten			12 235	2 625	1.69
Aura Avis	2015	24	7 608	1 478	1.05

Driva			7 112	1 932	0.98
Åndalsnes Avis			5 855	1 424	0.81
Mørenytt			5 770	1 564	0.80
Sunnmøringen			1 758	598	0.24
Trøndelag					
NRK Trøndelag			323 871	149 487	22.44
Adresseavisen	2014	25	868 379	136 861	60.18
Trønderavisa	2015	2	101 495	23 529	7.03
Namdalsavisa	2015	3	49 189	7 954	3.41
Stjørdalens Blad			30 947	4 553	2.14
Fosnafolket			16 952	3 456	1.17
Avisa Sørtrønderlag			14 224	3 279	0.99
Hitra Frøya			13 986	2 670	0.97
Trønderbladet			8 896	2 220	0.62
Opdalingen			8 643	1 686	0.60
Malvikbladet			4 242	840	0.29
Innherred			2 139	543	0.15
Nordland					
NRK Nordland			358 855	177 717	45.83
Avisa Nordland	2015	8	170 177	26 814	21.73
Lofotposten	2014	45	58 779	6 888	7.51
Rana Blad	2015	6	57 050	8 194	7.29
Fremover	2015	3	45 676	7 390	5.83
Vesterålen online	2014	50	44 678	6 519	5.71
Helgeland Arbeiderblad			30 520	5 680	3.90
Brønnøysunds Avis			17 247	3 026	2.20
Troms and Finnmark					
NRK Troms og Finnmark			388 414	174 204	45.81
Nordlys			294 425	57 590	34.72
itromsø			60 230	12 358	7.10
Harstad Tidende	2015	10	32 538	6 305	3.84
Altaposten			23 237	5 436	2.74
iFinnmark			19 977	3 794	2.36
Troms Folkeblad			18 769	3 966	2.21
Framtid i Nord			10 351	2 470	1.22

Note: Market shares calculated from number of hits

na

Table A2: Difference-in-difference results (Model 2), logarithm of hits, unique sessions and unique visitors, clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.500***	0.400***	0.367***
	(0.026)	(0.026)	(0.026)
$ln(x_{i,j,t-2})$	0.198***	0.209***	0.202***
	(0.012)	(0.017)	(0.016)
$Post_{i,j,t}$	-0.014	0.016	0.035
	(0.034)	(0.029)	(0.029)
$Trend_t$	0.010**	0.019***	0.022***
	(0.005)	(0.007)	(0.007)
$Post_{i,j,t} \cdot Trend_t$	-0.007	-0.016**	-0.021**
	(0.008)	(0.008)	(0.008)
$Share_{-i,j,t}$	0.010	-0.008	-0.010
	(0.022)	(0.023)	(0.024)
Constant	3.385***	3.926***	4.132***
	(0.313)	(0.340)	(0.346)
Short-run effect	-0.030*	-0.025*	-0.018
	(0.016)	(0.013)	(0.012)
Long-run effect	-0.101*	-0.063*	-0.041
	(0.052)	(0.033)	(0.029)
r2	0.475	0.386	0.347
N	18 451	18 451	18 451
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

Table A3: Difference-in-difference results Model (3), logarithm of hits, unique sessions and unique visitors, clustered standard errors, all markets.

	Log of hits b/se	Log of sessions b/se	Log of visitors b/se
$ln(x_{i,j,t-1})$	0.500*** (0.026)	0.404*** (0.025)	0.372*** (0.025)
$ln(x_{i,j,t-2})$	0.198*** (0.012)	0.212*** (0.016)	0.206*** (0.016)
$Post_{i,j,t}^{L}$	-0.027** (0.012)	-0.034*** (0.008)	-0.034*** (0.012)
$Post_{i,j,t}^{O}$	-0.037 (0.022)	-0.032* (0.019)	-0.035* (0.020)
$Post_{i,j,t}^{L} \cdot Share_{-i,j,t}$	-0.115 (0.071)	-0.121*** (0.039)	-0.136*** (0.047)
$Post_{i,j,t}^{O} \cdot Share_{-i,j,t}$	0.008 (0.050)	-0.011 (0.052)	-0.001 (0.054)
$Share_{-i,j,t}$	0.036* (0.020)	0.041** (0.020)	0.045** (0.021)
Short-term; Largest	-0.051*** (0.015)	-0.059*** (0.010)	-0.063*** (0.012)
Short-term; Other	-0.035** (0.014)	-0.035*** (0.011)	-0.036*** (0.011)
Long-term; Largest	-0.170*** (0.049)	-0.154*** (0.024)	-0.149*** (0.027)
Long-term; Other	-0.116** (0.046)	-0.090*** (0.030)	-0.085*** (0.028)
Constant	3.411*** (0.310)	3.905*** (0.326)	4.104*** (0.330)
r2	0.475	0.383	0.343
N	18 451	18 451	18 451
Fixed effect; news outlet	Yes	Yes	Yes
Fixed effect; time	Yes	Yes	Yes

* p<0.10, ** p<0.05, *** p<0.01

Figure A1: Histogram of VARSOC lag suggestions for the autoregressive process based on Akaike information criterion