Pricing Online Content: Fee or Free?*

Anja Lambrecht[+]

Kanishka Misra[±]

May 7, 2014

Keywords: Paywall, Pricing, Online Media, Internet, Electronic Commerce, Free Content, Paid Content

* We are grateful to Rajesh Chandy, S. Sriram, Catherine Tucker, seminar participants at Groningen University, London Business School, MIT, Northwestern, Michigan, Technical University Munich, Washington University, and participants at the Theory and Practice in Marketing 2013 Conference, the Choice Symposium 2013 and at the 2013 Marketing Science conference for their comments and suggestions.
[+] Anja Lambrecht, London Business School, London, UK; alambrecht@london.edu.
[±] Kanishka Misra, Ross School of Business, University of Michigan; kanishka@umich.edu.

Pricing Online Content: Fee or Free?

Abstract

Many online content providers aim to compensate for a loss in advertising revenues by charging consumers for access to online content. However, such a choice is not straightforward, because subscription fees typically deter customers, further reducing advertising revenues.

In this research, we empirically examine and quantify a content provider's trade-off between advertising and subscription revenues. We build a unique data set from the sports' website ESPN.com, which offers the majority of content for free but charges a membership fee for a subset of articles. We collect data on the number of free and paid articles per day and sport, as well as demand for each type of article per day and sport over a 13-month period.

We estimate how the number of free and paid articles affects viewership of the site, and empirically quantify a firm's trade-off between advertising and subscription revenues, controlling for a wide range of possible demand shifters. We find that, on average, the firm should not adjust the amount of paid content. However, our results show strong differences across sports' seasons: the marginal paid article increases revenue in the off-season but decreases revenue in the regular season. This finding implies the firm can increase revenue by flexibly adjusting the amount of paid content it offers across sports' seasons. More generally, our results suggest that online content providers should carefully identify temporal variation in demand, and over time adjust the amount of paid content they offer rather than setting a static paywall.

1 Introduction

The future of the media industry is widely believed to depend on the ability of companies to monetize content online. However, for well over a decade, the prevalent view has been that "information wants to be free" and that consumers are unwilling to pay for online content (Edgecliffe-Johnson 2009). Research supports this view by showing that consumers respond negatively to even small monetary fees (Shampanier et al. 2007; Ascarza et al. 2012), which makes charging small amounts for digital content difficult.

Yet plummeting advertising revenues across the media industry (see Figure 1) force companies to identify new and additional sources of revenue: in December 2008, The Tribune, owner of the *Chicago Tribune* and *LA Times*, filed for bankruptcy protection. In 2009, the *New York Times*' credit crisis prompted a piece questioning the paper's continued existence (Hirschorn 2009). Then, in August 2013, *The Washington Post* was sold to Jeff Bezos, because "for much of the past decade, The Post has been unable to escape the financial turmoil that has engulfed newspapers" (Farhi 2013). Many regional newspapers, such as the *Miami Herald* and the *San Francisco Chronicle* face similar financial trouble.[1]

But how firms can open up new and additional sources of revenues remains unclear. Charging for online content adds subscription revenue but can significantly deter consumers, reducing advertising revenues. For example, Chiou and Tucker (2012) find that visits to websites of local newspapers fell by 73% after the introduction of a paywall.

Acknowledging the trade-off between subscription and advertising revenues, firms have in recent years experimented with a wide range of revenue models that

[1] http://www.realclearpolitics.com/lists/top_10_newspapers_in_trouble/miami_herald.html?state=play

include giving away all content for free (e.g., washingtonpost.com), charging for all content (e.g., thetimes.co.uk), and offering some content free of charge but charging for a subset of content (e.g., ESPN.com, faz.net, nyt.com).[2] Some firms experiment with different strategies: the *New York Times* initially offered all content for free but switched to a paid model with 20 and later 10 free articles per month. The *Wall Street Journal* required a subscription, later changed to a largely free version, but reverted to a partly paid model.

Importantly, although most firms decide to follow a static rule on how much content is free or paid (e.g., all paid, 10 free per month), they can more flexibly adjust the amount of paid content they offer. For example, at the *Wall Street Journal*, only subscribers can "unlock" a selection of articles, and this selection varies by day. Theoretically, such policies are promising if consumer demand for paid content varies over time. But as of yet, evidence is scarce regarding how the trade-off of advertising versus subscription revenues plays out, and whether firms benefit from varying the amount of paid content they offer.

In this research, we empirically examine and quantify a content provider's trade-off between advertising and subscription revenues. We evaluate whether a firm should follow a static policy or dynamically adjust the amount of paid content offered.

We build a unique data set from the sports website ESPN.com, which offers the majority of its content for free but charges a fee for a subset of articles. The number of paid articles varies by day and sport. Via a web crawler, we collect data on the number of free and paid articles per day and sport over a 13-month period. We complement these

[2] ESPN.com charges for a subset of articles, faz.net for historic articles, and nyt.com for any article that

data with the number of unique visitors, page views (the number of instances a user visited the firm's web pages), and time spent for each article type per day and sport.

We document how consumer demand for news at ESPN.com varies across sports' seasons, which are characterized by strong differences in consumer utility from sport news. We estimate how unique visitors to the firm's paid section, as a proxy for subscribers, and page views, as a proxy for advertising revenues, respond to paid content.

We find that, on average, paid articles increase unique visitors to the paid section, and thus subscribers, while reducing advertising revenues from page views on the site. Using these estimates, we quantify the impact of adding a paid article on the firm's revenues. We find that, on average, the increase in subscription revenue from a marginal paid article is statistically indistinguishable from the decrease in advertising revenue, suggesting the firm should not adjust the amount of paid content. However, we find strong differences across sports' seasons, a factor that appears to exogenously vary demand for sports news. Specifically, the marginal paid article increases revenues in the off-season but decreases revenues in the regular season. For the post-season, results further vary across sports. We confirm our results in a quantile regression and using the percentage of paid articles instead of their absolute number as independent variable.

Our findings suggest that adding paid content may sometimes—but not always—benefit firms. Indeed, based on our results, many online content providers that currently use a static policy may benefit from re-adjusting their pricing strategies to dynamically respond to consumer demand. These insights can also be useful for firms trying to learn about their competitors' likely decisions.

---

exceeds a monthly allowance of 10 free articles (on the *New York Times*, see also Kumar et al. 2013).

More broadly, our results suggest that to increase revenue from online content, firms need to examine in detail underlying consumer behavior and, ideally, identify in real-time periods of unusually high (or low) demand to then flexibly adjust the amount of paid content. To achieve this, firms would need to track consumer usage behavior. In some instances, firms may have access to individual-level data, such as for consumers who have signed up for the paid section. For nonsubscribers, individual-level data can often prove difficult to obtain because consumers use multiple devices, may not allow cookies, or cookies expire. Our empirical approach illustrates how firms can use aggregate data to align their paid-content offering with consumer demand.

2 Relationship to previous literature

Our research adds to three different streams of research. First, research using analytical modeling shows that offering a paid and a free component can allow firms to implement quality differentiation, versioning, or second-degree price discrimination (Shapiro and Varian 1998; Bhargava and Choudhary 2001). But paid content not reduces the number of page impressions and thus advertising revenues and can also lead to a disadvantage in advertising markets if advertisers are willing to pay a premium to firms with a high expected share of loyal consumers (Athey et al. 2011). Further, the trade-off between free and paid content varies with competition (Godes et al. 2009) and consumer heterogeneity. With heterogeneous consumers, firms should combine pay-per-view and advertising revenues but offer options to consumers (Prasad et al. 2003). Likewise, advertising effectiveness determines whether a firm should offer both paid and free content (Halbheer et al. 2013).[3] This research finds that under intermediate levels of advertising effectiveness firms should offer both. In our research we ask whether a firm that offers free and paid content would benefit from a marginal change to the current offering.

A second, emerging stream of research aims to provide empirical insights relevant to an online content provider's choice of revenue model (for an overview on online revenue models, see Lambrecht et al. 2013). Empirically, an online content provider targeted toward marketing professionals finds that moving from free to fee can be profitable (Pauwels and Weiss 2007). But other research shows that visits to online news sites can fall by as much as 73% following the introduction of a paywall (Chiou and Tucker 2012).[4] Yet, lacking detailed data on advertising revenues or users' website activities, research to date has been unable to examine whether an increase in subscription revenue would off-set such losses in advertising revenues from reduced page views. Additionally, empirical studies have been limited to a static setup in which either all or none of the content is free, and have been unable to explore the effect of demand variation on paid-content strategies.

More broadly, a third stream of literature that focuses on the state of the media industry motivates our work. This research is concerned with an increase in subscription prices alongside declining circulation of print newspapers (Seamans and Zhu 2013; Pattabhiramaiah et al. 2013) and the effect on their ability to price discriminate

[3] Broadly related, Bawa and Shoemaker 2004 show for physical goods that allowing consumers to sample content before a purchase can increase long-term sales.

[4] The strong decline in readership following the introduction of a monthly fee of $9.95 is perhaps not surprising given findings that consumers perceive the benefits associated with free products, compared to those of paid products, as higher than would be expected based on the price change alone (Shampanier et al. 2007; Ascarza et al. 2012).

(Angelucci et al. 2013). Additional research explores the product line pricing problem of a firm that offers the same content in print and digital formats (Kannan et al. 2009). In studying how media firms can monetize their online content, we aim to contribute to the discussion on how media firms can build sustainable revenue models, given that consumers' attention increasingly shifts online.

In sum, whereas research so far has offered some broad guidelines on a content provider's choice of "fee or free," it does not provide detailed insights into how a firms' trade-off between subscription and advertising revenues plays out. It also largely assumes that the firm follows a static policy, for example always offers only free or paid content. Our research seeks to fill this gap. We quantify an online-content provider's trade-off between subscription and advertising revenues, and examine whether the firm would benefit from dynamically adjusting the amount of paid content. We examine how such dynamic strategies can build on insights into heterogeneity in consumer willingness to pay across time and individuals.

3 Data

3.1 Empirical setting

Our empirical study is set in the context of the sports' website ESPN.com, which is the website of the US sports TV network ESPN and is owned by Disney. ESPN.com provides a wide range of coverage on sports and sports events, including news and background reports. In the following, we refer to ESPN.com simply as ESPN.

ESPN has a main homepage plus homepages for each sport. The homepages display only title and links to articles but no abstracts or full articles. Importantly, ESPN offers two types of articles: regular articles, available free of charge to all consumers (hereafter free articles), and “insider” articles (hereafter paid articles), available only to consumers who pay a membership fee. On each sport’s homepage, paid articles are easily recognizable through a small orange “in”-icon. The number of paid articles varies across days and sports.

In our analysis, we focus on six different sports that typically offer both paid and free articles: college basketball (CBA), college football (CFB), professional baseball (MLB), professional basketball (NBA), professional football (NFL), and professional H=hockey (NHL). We do not use data from other sports (such as soccer, NASCAR, golf, or tennis), because ESPN did not offer any paid articles during our observation period.

3.2 Website content and user activity

A typical challenge in analyzing the effectiveness of “free” versus “fee” strategies is the difficulty in obtaining data that disclose detailed usage information alongside pricing strategies while also controlling for industry-wide demand. We circumvent this challenge by combining multiple data sets. Our data capture, over 13 months, per day and sport the number of free and paid articles featured on the firm’s sport-specific homepage, the number of unique visitors to the paid and free sections on the firm’s website, and the number of page views in both sections. They also include, on a day and sport level, unique visitors and page views to competitive sites. Our data is thus disaggregated on the day and sport level. We next describe in detail the different data sets we use.

ESPN website content: First, we use a web scraper to collect on a daily basis the number of free and paid articles on each of the six sports’ home pages at ESPN from December 2010 to December 2011. For free articles, we collect all links with the url-format *espn.go.com/sportname*. For paid articles, we collect all links with the url-format *insider.espn.go.com/sportname*.[5] We then identify links that remain on a sport's homepage for a long time period (more than 100 days). These links typically do not represent content-based news articles but provide general information that often does not change over time (e.g., links to pages on the NBA draft for previous years, or games timetables). We count as articles all links that appear on the sport's home page for less than 100 days. As the first part of Table 1, Panel A indicates that a sport's home page displays 34 articles on average per day, of which 25 are free and nine paid.

We compare free and paid articles in more detail. For a sample period of seven days (November 9–15, 2011), we collect data on the length (measured as the word count) of all free and paid articles featured in the two most prominent sections of the sports' homepages (sections "Headlines" and "Top Stories") as well as in the "Insider" section that lists a selection of paid articles. Although paid articles are on average longer, the standard deviation in article length is high, and more so for free articles (Table 2). The large standard deviation is driven by a high number of very short free articles: 10% of free articles have less than 200 words. We compare all 274 paid articles with the top 274 free articles by word count and find that in this subset, free articles are on average longer. This finding suggests that both the paid and the free sections feature many detailed articles.

Lastly, we broadly look at the type of articles featured in both sections. We find that the free section includes both news and editorial content (e.g., comments on a team's performance), whereas the paid section focuses on editorial content and more in-depth

[5] This metric abstracts away from content on other websites to which the sport's home page links, such

reports (e.g., interview with a coach). This finding makes sense because readers could easily substitute news articles with an article from a competing site, whereas such substitution is more difficult for editorial content or in-depth reporting.

ESPN user activity: In our second data set, we obtain, for the same time period, daily data from Comscore on consumer activity by sport. These data include the number of unique visitors, the number of pages viewed, and the total time spent for both free and paid articles. We do not have access to consumer-level data. Consistent with our definition of free and paid articles, we use the url-formats espn.go.com/sportname and insider.espn.go.com/sportname to identify website activities.

Comscore collets its data based on an online panel of consumers whose web activities they follow. They then weigh the individual-level observations to obtain a data set that is representative of the US population. This approach means our data sometimes record zero visitors (mostly to the paid section) even though the true number for the US population is nonzero. Because interpreting these numbers is difficult, and because in our empirical estimation, we take logs of the dependent variables, we exclude these 218 out of a total of 2,250 day-sport observations.[6]

Panel A in Table 1 reflects that significantly more individuals visit the free than the paid section of the site. It also illustrates that each unique visitor to the free section visits on average 5.3 pages, and each visitor to the paid section visits 2.1 pages, in line with the fact that the site offers significantly more free than paid articles. Reflecting the difference between the number of free and paid articles, we find that whereas 4% of

as Twitter, and blogs that come with a different url-format. These links are always free.

unique visitors visit the paid section, only 2% of all page views come from the paid section. Lastly, the time visitors spend per page is similar across paid and free articles.

Competitor user activity: Third, to control on a day and sport level for industry-wide demand for sport news, we obtain data on website activities from Comscore for the main competing sports website, sports.yahoo.com. Yahoo offers its content for free. The data include the number of unique visitors, the number of page views, and the total time spent per day and sport. Table 1, Panel B documents that page views per visitor and time spent per visitor at Yahoo is comparable to those for free ESPN articles (5.5 vs 5.2 for page views per visitor; 6.2 vs 5.2 for time spent per visitor).

ESPN demand control: Fourth, we collect Google Trends data on the number of searches for “ESPN + sport” for every day in our data. We scale the data to numbers between 0 and 100. This scaling will later allow us to control for demand for news on a particular sport specific to ESPN.

Further demand controls: Fifth, we collect a comprehensive set of additional demand controls. Most importantly, this set includes the start and end of sports’ seasons. Each sport has three seasons. The off-season is the period when no games are scheduled. Note that in the off-season, sports news is still available, such as free-agency signing and drafts, and scores for any pre-season games, results that are not considered in the teams’ final performance. The regular season is the period when scheduled games are played. Participation in these games is based on the planned schedule and so is independent of performance. During the post-season, playoffs and a sport’s final games are played (e.g.,

[6] We separately estimate linear models that do not require us to take logs of the dependent variables to check whether our results are robust to excluding the observations. We find that the coefficients in a specification that includes them are not significantly different from one that excludes them.

playoff in the professional sports MLB, NBA, NFL, NHL; the bowl season for college football and March madness in college basketball).

Table 3 illustrates that the number of free and paid articles displayed varies more strongly within than across seasons. As would be expected, we observe a large variation in demand for articles across seasons. All our measures indicate that demand for news is lowest in the off-season. Average demand is similar in regular and post-seasons.

Additionally, we collect data on sport events. This includes whether or not a game was played in that sport on a day and the date of the final game within each sport. For professional sports, the data include the dates of the draft, and for college sports, they include college-signing day. We also collect the dates of the NBA lockout in the 2011 season and record whether a day was a non-working day, because consumption of online news may be different on weekends and public holidays than on working days.

3.3 Subscription and advertising revenues

We next describe the subscription plans ESPN offers. We then explain in detail how we estimate an implicit average price per visit to the paid section using the weighted average subscription price and information on visitors to the paid section. In Section 6, we discuss the robustness of the underlying assumption that monthly visit frequency remains constant over time.

Customers can sign up for one of three membership plans to access paid articles. A two-year membership costs $2.50 per month, a yearly membership plan charges $3.33 per month, and a monthly membership charges $6.95 per month. We obtain data from Comscore on the number of customers that sign up for each of the membership plans for December 2010 to December 2011. They suggest that 47% of customers choose the yearly plan, 35% choose the two-year plan, and 13% choose the monthly plan,[7] giving us an average subscription price of $40.44 per year.

ESPN.com had 640,000 paid subscribers in 2011 (ESPN 2012) and a total of 55 million unique daily sport-visits to the paid section (Comscore data). On average, each subscriber had 86.62 day-sports visits per year (or 7.22 day-sports visits per month) to the paid section for each sport they visited. Note that this estimate is across sports and days; for example, visiting the NFL and NBA sport pages on the same day will count as two separate visits. The effective price per visit on a day to a sport amounts to $0.47.

ESPN features advertising on all webpages, including its home page, the home page for each sport, and the page for each article. Each page typically displays one ad, independently of whether an article is free or paid.[8] From Comscore, we obtain additional data that lists on a monthly level ESPN's advertising revenues and page views from December 2010 to December 2011.[9] We use these data to compute the monthly revenue per 1,000 impressions. On average, ESPN's revenue per 1,000 impressions is $11.51. This value varies over time with a minimum of $8.34 (in April) and a maximum of $15.45 (in December 2011). Importantly, we were able to informally verify the average,

[7] Additionally, 4% signed up for a holiday offer in December 2010 and 1% for a trial in October 2011. Although our data give us reliable information about the average attractiveness of the plans, the number of individuals signing up for any plan in any month is low, so we are unable to report representative data on total monthly new subscribers at ESPN.

[8] We counted the number of display ads per article for MLB and NBA on a single day (June 28, 2011). On average, these articles display one ad. We do not include sponsored links, because revenue from them is likely negligible. Comscore also does not provide estimates for revenues from sponsored links, which further suggests such revenues are negligible.

[9] Comscore estimates are based on projected ad spend costs and so the advertising revenue they report approximates the net advertising cost, not the gross cost that is quoted on rate cards and is often substantially higher. The data are predominantly inputted by agencies and so reflect the actual payments to ESPN rather than gross pay-outs by advertisers that may include costs for agency services.

minimum, and maximum advertising prices with ESPN. In our later analysis, we look at the robustness of our results to these different advertising prices.

4 Model-free Evidence

The key strength of our empirical setting is that it allows us to combine three types of data: first, detailed data on consumers' usage of *both* free and paid content; second, data that capture variation in pricing—here, through the amount of paid content offered per day and sport, third, variables that measure, and so allow us to empirically control for, industry-wide demand. Importantly, because our data are on the day and sport levels, we observe variation in behavior both within and across time. Unlike previous research, we are in the unique position of being able to estimate the effect of paid content on both subscription revenues through the analysis of visits to the paid section, and on advertising revenues through the analysis of page views on the site, while controlling for variation in industry-wide demand. We next provide model-free evidence that examines current firm policy. We then turn to the relationship of paid content to consumer behavior that translates into an online content provider's revenue streams.

4.1 Firm policy

We first analyze the distribution of revenue for the firm for different levels of paid articles the firm may offer. We use our data on unique visitors to the paid section and on page views, the price per visit to the paid section of $0.47, and the price per 1,000 impressions of $11.51. On a day and sport level, we calculate the daily revenue as *(unique visitors to paid section)*$0.47 + (page views)*$11.51/1000*. The density of log revenues displayed in Figure 3 illustrates a large degree of variation in estimated revenues. Interestingly, the data appear to be bimodal with one mode at 10.3 and one at 11.3. In Figure 4, we split the data by season and find this bimodality is a result of variation in revenue across sports' seasons. In the off-season, the mode of daily revenue by sport lies at 10.0, and in regular and post-seasons it lies at 11.2 and 11.0, respectively. This finding suggests that ESPN faces different revenue levels across seasons.

Second, we try to identify empirical regularities that shed light on how the firm allocates paid content to its website. We estimate a set of regressions with the number of paid articles by day and sport as dependent variables. We use three sets of independent variables. First, we include only controls for sports; second, controls for sport and seasons; and third, controls for sports, season, and the full set of demand controls we collected. The results in Table 4 indicate that the number of articles mainly varies across sports. Adding controls for seasons or additional demand shifters hardly increases the explanatory power. This finding is consistent with the observation that ESPN has Insider columnists by sport.

Table 4 further summarizes a second set of regressions with revenue per day and sport as dependent variables. Here we find that both sports and seasons explain a large amount of the variation of observed revenue. Together, the results in Table 4 indicate that temporal variation in demand across sports' seasons appears to significantly affect revenue but not the firm's decision on how many paid articles to display. Instead, firm policy mainly varies across sports.

4.2 Paid content and consumer behavior

Unique visitors to the paid section: We build on the premise that consumers derive utility from reading articles. Thus a consumer's net utility of a visit to the paid section on any day is the sum of the utilities from all paid articles minus the time prorated subscription fee. As the firm adds a paid article, the expected utility from subscribing to the paid section marginally increases. As a result, the consumer becomes more likely to subscribe, which means the number of visitors to the paid section should increase in the number of paid articles.

Two behavioral mechanisms may explain why consumers' utility from subscribing may increase in the number of paid articles. First, the expected utility from visiting the paid section on any day increases in the number of paid articles offered, because more content is available to view, and consumers pay a lump sum to visit the paid section. A second and complementary view is that consumers have heterogeneous preferences for articles. As the number of available articles increases, the likelihood that a consumer finds an article that fits their preferences increases and the expected utility from subscription increases. In our empirical analysis, we focus on how the number of paid articles affects the number of unique paying visitors, and so jointly capture the effect of both mechanisms.

We examine in our data the relationship between the number of paid articles on a day and sport and unique visitors to the paid section for that sport that day. Figure 5, Panel A plots for all day-sport observations the normalized unique visitors to the paid section. We obtain this measure by computing the difference between the number of unique visitors on any day in a sport and the corresponding number of unique visitors on the Yahoo site. We further normalize this number by sport by subtracting the difference between the average number of unique visitors to the paid section and the average number of visitors to Yahoo. As expected, we find that unique visitors increase in the number of paid articles, with a slope of the trend line of 0.09 ($p < 0.01$).

Total page views: We analyze how page views change with the number of paid articles. Because each page view allows the firm to display ads, advertising revenues increase with page views. Several behavioral mechanisms suggest that an increase in paid articles could reduce page views.

First, “cannibalization” could occur due to limited web real estate on a sport’s home page or to resource constraints of the firm that can either produce paid or free articles. If little content is available to view for free, non-subscribers may simply view fewer articles, reducing overall page views. In our data, we find an economically low and statistically insignificant correlation between free and paid articles (0.030, p=0.172). However, consistent with the findings in the retail assortment literature (see Chernev (2012) for a review), as paid articles increase consumers might perceive fewer free articles and focus instead on paid articles.

Second, non-subscribing customers may have disutility from paid articles and therefore reduce page views. For example, if an article of great interest to a visitor is paid, that visitor might choose to visit another website for all news on that sport on that day.

Third, paid articles may affect the number of unique visitors to the site. Although subscribers are more likely to visit, the utility from visiting for non-subscribers may decrease when only less appealing content is offered for free.[10] As a result, visitors may

[10] Consumers have several ways in which they can learn about the availability of paid articles for a sport. The headline of a paid article features on the firm’s overall home page or on the sport’s home page. Also, links to an article reported by search engines indicate whether an article is paid.

decide not to visit the site. Because in our sample, despite the high number of subscribers in absolute terms, 96% of visitors are non-subscribing and 98% of page views are in the free section, we expect the negative impact on non-subscribing consumers to dwarf a positive effect from subscribers on page views. [11]

We examine in our data whether, as theory predicts, page views fall with the number of paid articles. Figure 5, Panel B plots the corresponding relationship for paid articles and the median of total page views on the focal site for every level of paid articles, where page views are normalized by page views on Yahoo for that sport and by the average number of page views in that sport on our focal site and on Yahoo. The results confirm the hypothesized relationship between paid articles and page views: page views appear to decline for higher levels of paid articles (slope of the trend line is -0.02, $p < 0.01$), potentially resulting in lower advertising revenues.

Whereas the results in Figure 5 present correlational, not causal, evidence, they support the basic trade-off we expect online content providers to face. As firms increase the number of paid articles, unique visitors to the paid section, and thus subscription revenues, appear to increase, but page views appear to decline, potentially resulting in lower advertising revenues.

Analysis by sports' seasons: We examine whether the relationship between paid articles and unique visitors to that section, and total page views, holds across sports' seasons. Figure 6 plots the relationship between paid articles and unique visitors, and

[11] We acknowledge that paid articles may have other long-term effects on consumer behavior. For example, consumers may learn about the firm's policy over time, and their expectations about the future number of paid articles may affect their behavior. Here, we abstract from such possibilities.

page views, normalized as before, for each article level by whether a sport is in its off-season, regular season, or post-season.

We find that the relationship between the number of paid articles and unique visitors to the paid section, and page views, varies across seasons. The increase of unique visitors in paid articles is weakest in the off-season (slope 0.03, $p<0.01$) and regular season (slope 0.05, $p<0.01$) and most pronounced in the post-season (slope 0.10, $p<0.01$). Although the slope-coefficients for the off-season and regular season do not differ significantly, the coefficient for the post-season differs significantly ($p<0.01$) from that in both other seasons.

Page views mirror this effect. In the off-season, page views do not change as the number of articles increases (slope of 0.00, $p=0.53$). In regular and post-season, a negative relationship is present between paid articles and page views (slope of -0.02, $p<0.01$ and -0.02, $p<0.10$, respectively). Due to the large variation in the post-season, the slope coefficient is greater than in the regular season with a p-value of only 0.07.

We propose that the pattern we observe across seasons is a result of the variation in consumer demand and valuation for sport news over time. In the off-season, demand is generally lowest. Individuals with a particularly high valuation for sport news ("sport junkies") are probably the most likely to visit the site. The lack of any negative effect on page views is consistent with the high utility of news, which leads to a greater likelihood of signing up for the paid section, dwarfing the disutility from an increase in paid articles for these consumers. It could additionally relate to a lack of outside options providing sport news, such as TV.

Although a large number of consumers join the site in the regular season (Table 3), Figure 6 documents that this increase in viewers does not translate into an increase in visitors to the paid section, possibly because these consumers have, on average, a lower valuation of sport news or because during the regular season, many additional ways to consume sport news, such as TV, are available. In line with both interpretations, page views respond more negatively to the number of paid articles the firm posts on its site in the regular season than in the post-season.

In the post-season, paid articles appear to become more effective in attracting consumers relative to the off-season or regular season, perhaps because for some consumers, the valuation of sport news strongly increases when the most important games are played, for example, for those consumers whose teams are in the finals. The negative relationship between paid articles and page views in the post-season, however, suggests that a large group of consumers may still respond negatively to paid content.

Our interpretation is based on correlational evidence that does not control for many factors that could affect a firm's decision on how many articles to offer and the number of consumers who visit the site. For example, if the firm posts more (or fewer) paid articles on a game day and more consumers visit on such days, a correlation between paid articles and unique visitors would be endogenous and a causal interpretation inappropriate. The next section presents a comprehensive empirical analysis that controls for a wide range of covariates to fully capture the variation in the data, quantifies the effect of paid articles on subscription and advertising revenues, and confirms the results in a quantile regression. We then go on to evaluate whether a firm would benefit from varying paid content across seasons.

5 Estimation results

5.1 Average effect of paid articles

SUR: We estimate the effect of paid articles on unique visitors to the paid section and on page views on the site. In our estimation, we analyze separately (1) the decision to visit the firm's site and (2) the decision to visit the paid section. We also estimate the effect of articles on page views (3). We estimate a seemingly unrelated regression (SUR) to allow for a correlated error structure across all three equations. The estimation is on the day-sport level:

$$
(1)\quad
\begin{aligned}
&share(Firm)_{it} = \beta_1^F PaidArt_{it} + \beta_2^F FreeArt_{it} + \beta_{3-7}^F Controls_{it} + \beta_8^F \ln\left(UniqVisYa_{it}\right) + \\
&\quad \beta_9^F Google_{it} + \beta_{10}^F RegSeason_{it} + \beta_{11}^F PostSeason_{it} + \beta_{12}^F Sport_i + \varepsilon_{it}^F \\
&share(PaidS)_{it} = \beta_1^P PaidArt_{it} + \beta_2^P FreeArt_{it} + \beta_{3-7}^P Controls_{it} + \beta_8^P \ln\left(UniqVisYa_{it}\right) + \\
&\quad \beta_9^P Google_{it} + \beta_{10}^P RegSeason_{it} + \beta_{11}^P PostSeason_{it} + \beta_{12}^P Sport_i + \varepsilon_{it}^P \\
&PageViews_{it} = \beta_1^V PaidArt_{it} + \beta_2^V FreeArt_{it} + \beta_{3-7}^V Controls_{it} + \beta_8^V \ln\left(PageViewsYa_{it}\right) + \\
&\quad \beta_9^{PV} Google_{it} + \beta_{10}^V RegSeason_{it} + \beta_{11}^V PostSeason_{it} + \beta_{12}^V Sport_i + \varepsilon_{it}^V .
\end{aligned}
$$

The first part of Equation (1) captures the effect of paid articles on visitors to the firm's site (covariates with superscript *F*). The dependent variable $share(Firm)_{it}$ is the log-transformation of the ratio of unique visitors to ESPN for a sport *i* on day *t* to the number of all US citizens that search for sports news in 2011, a total of 162,027,797. (We obtain this number by multiplying two figures, the US population in 2011 of 311,591,917, and the share of the population that look for sports news online of 52%; see http://edition.cnn.com/2010/TECH/03/01/social.network.news/.) It thus represents the share of the total relevant population ESPN attracts that day for that sport.

The second part of Equation (1) analyzes visits to the paid section of the site. Here, the dependent variable, $share(PaidS)_{it}$, represents the log-transformed share of all ESPN visitors to a sport *i* on a day *t* that visited the paid section of that sport that day (covariates with superscript *P*). It thus captures unique visitors to the paid section, conditional on the number of individuals that visit the free section of the site.

The third part of Equation (1) has $PageViews_{it}$ as a dependent variable. It captures the effect of paid articles on page views within a sport *i* on day *t* on the firm's website (covariates with superscript *V),* including those in the free and the paid section.

Across all three equations, $PaidArt_{it}$ represents the number of paid, and $FreeArt_{it}$ represents the number of free articles for sport *i* on day *t* (we omit superscripts for ease of discussion). Of course, the firm might strategically determine the number of free or paid articles based on the expected demand on a day. For example, the firm may decide to offer more (or less) paid content on game days. We therefore include a vector of demand controls, $Controls_{it}$. This vector includes (1) a dummy that captures whether any games are played in sport *i* on day *t*, (2) whether a draft of players in a sport occurs during the off-season on a particular day (for college sports, it captures the national sign-up day), (3) a control for the NBA lockout during the 2011/12 season, (4) whether the final game in a sport *i* occurs on a day *t*, and (5) a variable that captures whether day *t* is a non-working day, that is, a weekend day or a public holiday. Such a control would, for example, capture if consumers tend to watch more TV and consume less online sports news on days they are not at work.

We acknowledge that even after controlling for such events, further demand shocks may occur that affect both the firm's policy and consumer demand. To account for shocks that would shift demand in the entire market, such as a piece of unexpected sport news, we use as additional controls the number of visitors to Yahoo as the major competing site for sport *i* on day *t*, $UniqVisYa_{it}$, and the number of page views on Yahoo for sport *i* on day *t*, $PageViewsYa_{it}$.

Additional demand shocks may uniquely affect the attractiveness of ESPN and may not be reflected in behavior on the Yahoo site. For example, a piece of news reported on ESPN TV may likewise be discussed on the website. We therefore use as an additional control the variable $Google_{it}$, which measures the number of Google searches for "ESPN + sport" scaled between 0 and 100. Lastly, we include dummies for whether a sport is in its regular or post-season and fixed effects by sport.

We initially estimate Equation (1) excluding demand controls. Table 5, Column (1) illustrates that the number of paid articles increases the share of paying visitors. However, we also obtain the somewhat surprising result that page views increase in paid articles, suggesting that some of the observed correlation may be due to omitted variable bias. Because Table 4 documented that paid articles and firm revenues vary by sport, we introduce controls for sports in Column (2) of Table 5. We find a large increase in the R^2 of all three equations. Consistent with theory, the estimate for the effect of paid articles on page views now becomes negative. This result suggests that not controlling for sports biases this estimate upward, which is consistent with the observation that the sport with the highest number of page views, the NFL, also has the highest number of paid articles.

Adding controls for season in Column (3) and additional demand controls in Column (4) further increases the R^2 for all three equations. Whereas the estimates for the effect of paid articles on the share of visitors to the paid section remains stable with the addition of further controls, the estimates for page views and visitors to ESPN increase in Columns (3) and (4), suggesting that not accounting for demand controls biases our results downward. To see why, consider the case of the NBA lockout: during this period, paid articles decreased and page views increased.

The results in Column (4), Table 5, which include the full set of demand controls, show that the number of paid articles does not, on average, attract consumers to the firm's website. However, in line with our model-free analysis, paid articles are indeed effective in converting users who visit the free section into visitors to the paid section and so ultimately into subscribers. Additionally, we estimate that paid articles decrease page views. Overall, these results are consistent with a behavioral mechanism whereby paid articles increase visits to the paid section of the site but reduce page views.

We use our estimates to quantify the effect of paid articles on visitors to the paid section and page views. Based on the logit share formulation, we determine the percentage change in unique visitors to the paid section from adding a paid article as $PercChangeVis = \beta_1^F \left(1 - share(Firm)_{median}\right) + \beta_1^P \left(1 - share(PaidS)_{median}\right)$. After controlling for demand, we find that an additional paid article increases the number of unique paying visitors by 5%. At the median of 19,269 unique visitors to the paid section, an additional paid article increases viewership in that section by 991 customers.

We obtain that an additional paid article decreases page views by 0.8% and so, at the median of 2,190,945 page views, leads to a loss of 18,108 page views. This result illustrates the trade-off apparent from our earlier descriptive analysis that paid content attracts subscription revenues but decreases page views and thus advertising revenues.

Our results likewise suggest that free articles fulfill three roles: First, they attract consumers to the firm's website. Second, the results indicate that free articles may convert consumers into subscribers, potentially because sampling high-quality free content increases the probability that a consumer will sign up for the paid section. Third, they increase the number of page views on the firm's site.

In Column (5), we present a specification that includes the percentage of paid articles instead of the number of free and paid articles, while controlling for the total number of paid articles. Our results are robust to this alternative specification.

Effect on the firm's revenue: The key outcome variable of interest to the firm is the aggregate revenue effect (broadly similar to Pauwels et al. 2011). We therefore evaluate whether offering an additional article is profitable for the firm. We make three assumptions. First, we translate unique visitors into subscribers, holding constant unique monthly visits per subscriber and so the implicit price per visit. In theory, the increase in unique visitors we observe as a result of an increase in paid articles might be due to a change in visit frequency by the same set of subscribers rather than to an increase in subscribers. Section 6 will examine this possibility. Second, we assume that article quality remains constant for a marginal paid article. Third, we assume that small changes in page views do not affect advertising prices.

To compute the financial impact for the firm of adding one marginal paid article, we bootstrap from the distributions of the three key coefficients: the effect of paid articles on $share(Firm)_{it}$, on $share(PaidS)_{it}$, and on $PageViews_{it}$. For each day-sport observation, we take 10,000 draws from these distributions.

We estimate the percent change of visitors to the paid section as $PercChangeVis = \beta_1^F\left(1 - share(Firm)_{it}\right) + \beta_1^P\left(1 - share(PaidS)_{it}\right)$ for each draw of all day-sport observations. For each draw, we multiply the percent change in visitors with market size to obtain the estimated marginal increase in paid visitors for that day-sport. Multiplying the increase in visitors with the average revenue for a unique day-sport visit of $0.47 (see Section 3.3) gives us the additional subscription revenue for the firm from offering one more paid article.

To understand the impact on advertising revenues, we multiply for each draw the change in page views from adding a paid article and the price per 1,000 page views.[12] We initially focus on the average price of $11.51 but later check the robustness of our results to the prices of $8.34 and of $15.45 per 1,000 impressions, the minimum and the maximum monthly ad prices during our observation period.

Table 6, Column (1) presents the median change in total revenue that accounts for both subscription and advertising revenue across day-sport observations. We also include the percentage of day-sport observations with a statistically significant negative or positive revenue impact. Figure 7, Panel A plots the distribution of the marginal revenue impact of a paid article at the average price per 1,000 impressions of $11.51.

Together, these results illustrate that at an average advertising price of $11.51 per 1,000 impressions, adding a paid article would increase revenues by $489, but large variation exists in the data with many observations being significantly greater or lower

[12] We assume the firm charges the same price for a page view in the free and the paid sections. We cannot conclusively rule out that ESPN charges a higher price for advertising in the paid section and acknowledge this as a limitation. But because ESPN already offers access to a highly targeted audience, they are unlikely to charge a significant premium for access to subscribers.

than zero. This finding holds when we check the robustness of the results to advertising prices per 1,000 impressions of $8.34 and $15.45 (Table 6).

This calculation assumes the firm adds or removes a paid article, holding constant the number of free articles. The assumption that the firm does not have limited "web real estate" but can easily add an article appears reasonable given the great variation in the total number of articles we observe. However, an alternative policy for the firm would be to add a paid article as it removes a free article and vice versa. Column (2) illustrates that the results broadly hold under this assumption, though the average benefit of adding a paid article is less pronounced. In Column (3), we then estimate the value of adding a paid article based on the specification that accounts for the percentage of free and paid articles. Figure 7, Panel A plots the corresponding distribution. Again, our results hold.

Quantile regression: We check the robustness of our results in an alternative specification that has as a dependent variable the revenue by sport and day, computed based on the advertising price of $11.51 per 1,000 impressions and the price per visit to the paid section of $0.47. We use as controls the full set of regressors from Equation (1), including both unique visitors and page views on Yahoo for sport i on day t.

First, we use OLS to estimate a homogenous effect of paid articles on revenue. In Figure 10, the solid red line indicates the OLS estimate for the effect of paid articles on revenue. The dashed lines illustrate that the 95% interval does not differ significantly from zero. Because our estimation does not account for potential heterogeneity across revenue quantiles, this effect is by definition consistent across the entire range of revenue quantiles indicated on the x-axis.

We then check whether this effect of paid articles holds across different revenue quantiles. We estimate a regression where the quantiles, τ, of the potential outcome distributions conditional on covariates are denoted by $Q_{\log(\mathrm{Re}venue_{it})|PaidArt_{it},X_t(\tau)}$ $with\ \tau(0,1)$ (Koenker 2005) and the effect of the treatment, here paid articles $PaidArt_{it}$, on different points of the marginal distribution of the potential outcome is

(2) $$QTE_{\tau} = \frac{\partial Q_{\log(\mathrm{Re}venue_{it})|PaidArt_{it},X_t(\tau)}}{\partial PaidArt_{it}}.$$

The quantile treatment effect, QTE_{τ}, measures the impact of paid articles on the total revenue to the firm after controlling for the observed heterogeneity in the explanatory variables. Specifically, we estimate a heterogeneous treatment effect of paid articles for 19 revenue quantiles ranging from the 5% to the 95% revenue quantile, conditional on controlling for the full set of covariates. In Figure 10, the dashed black line plots the quantile treatment effect for paid articles on revenue and the grey area indicates the 95% confidence interval.

The quantile regression suggests the insignificant effect in the OLS results and the large variation in the data presented in Table 6 and Figure 7 may be due to a heterogeneous treatment effect of paid articles across revenue quantiles. Our earlier descriptive analyses suggested that consumer response to sports news as well as firm revenues vary across sports' seasons. This insight motivates us to next empirically tease apart the effect of paid articles across seasons.

5.2 Exogenous variation of demand

SUR: We estimate the effect of paid articles on the number of visitors to the firm, to the paid section, and on total page views as a SUR, but now allow the effect of paid articles to vary by season.

Column (1) of Table 7 presents the estimates excluding demand controls. Surprisingly, the effect of paid articles on visitors to the paid section varies strongly across seasons, and paid articles increase page views throughout. After controlling for demand conditions in Columns (2), (3), and (4), the coefficients show the expected sign. The inclusion of demand controls also changes the relative size of the estimated impact of paid articles in all three equations. Consider the share of visitors to the paid section. In Column (1), which does not include demand controls, the estimated impact is largest in the off-season and smallest in the post-season. But when controlling for demand shifters in Column (4), we find the effect is greatest in the post-season, whereas the off-season estimate hardly changes. This pattern is similar for page views as a dependent variable. It indicates that the omitted variables bias is greatest in the post-season and least in the off-season. Adding demand controls further significantly increases the R^2 in all equations.

Based on the estimates in Column (4), we compute the aggregate effect of paid articles on unique visitors to the site by season using the logit share formulation similar to Section 5.1. We find that paid articles increase the share of visitors to the paid section by 5% in the off-season, 4% in the regular season, and 9% in the post-season. At the median number of unique visitors per season, our estimate translates into an increase in 991 visitors in the off-season, 782 in the regular season, and 2,184 in the post season.

This variation across seasons reflects our earlier analysis in Figure 6, Panel A, which s that the increase of unique visitors to the paid section is strongest during the post-season.

Conversely, the effect of paid articles on page views turns more negative as we move from the off-season to regular and post-season. Specifically, an additional paid article decreases page views by 2.9% in the regular and 4.0% in the post-season. At the median per season, our estimates suggest that a paid article reduces page views by 145,052 in the regular and 154,019 in the post-season and translates into a loss of advertising revenues of $1,670 and $1,773, respectively, when assuming an advertising price of $11.51 per 1,000 impressions.

Importantly, these results are consistent with Figure 6, Panel B, which demonstrated an increasingly negative effect of paid articles on page views when moving from the regular to post-season. Although we cannot conclusively rule out that article quality changes across seasons, this possibility seems unlikely given that the firm does not appear to be strategic in managing article quantity across seasons (Table 3).

We likewise examine the role of free articles. The first part of Table 7 that focuses on $share(Firm)_{it}$ illustrates that free articles are effective in attracting visitors to the site. They are most effective in the off-season, possibly because visits during the regular and post-seasons are less driven by articles and more by recent events. Interestingly, only during the regular season are paid articles effective in converting consumers into paying visitors. In Column (5), we use the share of paid articles by season as the independent variable, controlling for the number of paid articles by season. Again, the estimation results are broadly similar in direction and significance to those in Column (4).

Our demand shifters cover a wide range of possible events that can affect firm policy and consumer demand including major sports events, whether a day was a working day, or unobserved demand shifters captured through behavior on the Yahoo site or Google searches. As a result, in our main specifications, paid articles seem unlikely to be endogeneous. Nonetheless, we discuss in a Web Appendix a specification in which we instrument for free and paid articles using as instruments the number of free and paid articles on the previous day. The results are consistent with the findings presented here.

Effect on the firm's revenue: We compute the effect on the firm's revenue separately for the off-season, regular season, and post-season. We follow the same approach as in Section 5.1. For any day-sport observation, Table 8 summarizes the median financial impact of an additional paid article across all draws for the day-sport observations. It also reports the percentage of day-sport observations that are significantly different from zero—either positive or negative—at a 5% significance level.

Column (1) displays the results if the firm adds or subtracts a paid article, holding constant the number of free articles that day. Column (2) assumes the firm instead holds constant the total number of articles, substituting free for paid articles as the number of paid articles changes. Column (3) is based on the estimates that rely on the percent instead of the number of paid articles. Columns (2) and (3) both assume the firm substitutes the paid article with an average-quality free article. The revenue effect of transforming the high-quality paid article into a high-quality free article may be higher.

The results demonstrate that during the off-season, the firm would benefit from adding a paid article. For an additional paid article, revenue would increase between \$373 and \$630 per day and sport. Additionally, the vast majority of observations show a significant positive effect of paid articles. Figure 9, Panel A displays the density of the median revenue effect across all day-sport observations based on the data used in Column (1), Table 8. It confirms that the large majority of observations are positive. During the off-season, the firm should apparently increase the number of paid articles.

By contrast, we find a largely negative revenue effect of an additional paid article in the regular season of between \$628 and \$1591 on average. Figure 9, Panel B confirms the negative revenue effect for most day-sport observations. This result suggests the firm would benefit from decreasing the number of paid articles during the regular season.

In the post-season, the pattern is less clear. The total revenue effect of a marginal paid article varies between a loss of \$931 and a gain of \$129. Additionally, less than 50% of the observations differ significantly from zero. Figure 9, Panel C confirms that the firm appears to benefit little from changing the number of paid articles in the post-season.

Since 46% of our observations lie in the off-season and 44% in the regular season, the firm should increase the number of articles on roughly half of the days and decrease on roughly half of the days. Thus, accounting for variation across time is important in optimizing a firm's revenue streams from offering content online.

Quantile regression: We again check the robustness of our results in a quantile regression. Similar to Section 5.1 we estimate the impact of paid articles by season on the log of revenue initially as OLS and then in a quantile regression.

Figure 10 displays the effect of paid articles on revenue, controlling for all covariates that we use in our main regressions. We find a significant positive effect of the OLS estimation in the off season. This result is consistent with the corresponding quantile regression that shows a largely positive effect with little variation across revenue quantiles. Turning to the regular season, the OLS estimates illustrate a significant negative effect of paid articles on revenue, a result that holds in a quantile regression for the large majority of revenue quantiles. Lastly, for the post season, the OLS estimate is not significantly different from zero. There is likewise a great variation across revenue quantiles, though the effect is mostly not significantly different from zero.

6 Robustness checks

This section presents three robustness checks. First, we explore whether the effectiveness of paid articles varies by sport. Second, we examine our assumption that we can translate unique visitors to the paid section into subscribers. Third, we evaluate the total effect on revenue if the firm follows our policy recommendation.

Variation of revenue effects by sport: Our analyses indicate a great variation in the effect of a paid article on revenue for the post season (see Table 8 and Figure 9). We ask whether this variation might be due to variation across sports, and plot the data from Figure 9 by season and sport. We omit displaying the results for the off-season, because they consistently reflect the pattern in Figure 9 with close to 100% positive observations.

Figure 11 illustrates that across all sports, the majority of observations are positive in the regular season, meaning the firm should decrease the number of paid articles throughout. Interestingly, Figure 12 documents that in the post-season, the effect of paid articles on revenue varies across sports. Specifically, the results indicate the firm would benefit from decreasing the number of paid articles in the post-season for all sports except the NFL and college football.

Translating unique visitors to the paid section into subscribers: Throughout our empirical analysis, we translate unique visitors into subscribers, holding constant the number of unique monthly visits. The increase in unique visitors we observe as a result of an increase in paid articles, however, might not be due to an increase in subscribers but instead to a change in the visit frequency by the same set of subscribers. Using additional data provided by Comscore, we conduct two analyses that support our assumption that an increase in daily unique visitors indeed translates into a greater number of subscribers.

First, we provide support for the assumption that the monthly visit frequency is constant. We check whether consumers' monthly visit frequency varies with the number of unique monthly visitors. Using data on the average number of visits per sport and month, we estimate a linear regression with the average visit frequency as the dependent variable and unique monthly visitors per sport as the independent variable, controlling for sports. Our results suggest that visit frequency does not vary with unique monthly visitors (coefficient -0.000019, p=0.889) which supports that the variation in unique daily visitors we observe indeed comes from an increase in subscribers.

Second, to provide support for the assumption that the conversion rate from unique monthly visitors to new sign-ups is constant across seasons, we check whether the conversion rate varies across seasons. Using additional Comscore data, we compute the monthly conversion rate that is the sum of unique daily visitors to the paid section divided by total new signups that month. We then compute the percentage of visitors per month who went to a sport that was in off-season, regular season or post-season and estimate whether the conversion rate varies with the share of visitors that visited a sport in each season. We find no significant differences of conversion rates by seasons.

Evaluation of firm policy and recommendation: We estimate the total revenue change for the firm if it were to marginally increase paid content in the off-season and decrease paid content in the regular season. We find that with such a flexible policy, the firm's overall revenue would increase by 2.1% with a standard deviation of 0.3%. Although our estimates focus on the marginal effect of a paid article, a greater change to the number of paid articles might affect revenue to an even larger extent. We conclude that if the firm wishes to increase revenue from providing online content, it should account for the temporal heterogeneity in consumer response to paid content and flexibly adjust the amount of content by sports seasons.

7 Conclusion

The last decade has seen media companies, such as newspapers, struggle, and the general belief is that their futures hinge on their ability to implement a sustainable revenue model online. However, solving the trade-off that lies in gaining subscription revenues by offering paid content at the cost of lower advertising revenues is not obvious.

This research empirically examines and quantifies a content provider's trade-off between advertising and subscription revenues and evaluates whether a firm should follow a static policy or flexibly adjust the amount of paid content it offers.

We build a unique data set from the sports website ESPN.com in which we combine data on the number of free and paid articles offered per sport over time, with metrics of consumer demand such as unique visitors and page views, for both free and paid articles. We also track usage at the major competitive website. We estimate how the number of free and paid articles affects viewership of the site. We empirically quantify the impact of the number of paid articles on the increase in subscribers and the decrease in page views, and evaluate whether the company would benefit from adding paid content, controlling for a wide range of demand shifters.

Our results suggest that on average, the increase in subscription revenue from a marginal paid article is statistically indistinguishable from the decrease in advertising revenue. However, we find strong differences when we allow our results to vary by an indicator of exogenous demand variation. Specifically, the marginal paid article increases revenues in the off-season but decreases revenue in the regular season. For the post-season, we further find that the effectiveness of paid content varies across sports.

In sum, our findings indicate that adding paid content may sometimes—but not always—be good for firms. We conclude that online content providers that have recently experimented with fee models but typically use a static policy may benefit from re-evaluating their pricing strategies to flexibly respond to consumer demand. Whereas in our setting, variation in demand over time arises from sports' seasonalities, demand shocks can tilt the balance between advertising and subscription revenues in many other instances. Traffic on a general news site might greatly increase on election days because individuals who typically do not consume political news—and would not sign up for paid content—visit. In fact, both the *New York Times* and *Wall Street Journal* lifted their paywalls during the 2012 election. Similarly, following the 2013 bombings in Boston, the *Boston Globe* temporarily lifted its paywall. Our results suggest that for these sites, additional advertising revenues from the sudden influx of low-valuation customers (who would not pay a subscription) may have outweighed the loss of subscription revenues.

One implication of our findings is that online content providers can greatly benefit from investing in data analytics to identify in real time periods of unusually high (or low) demand and to respond by adjusting the amount of paid content. Partly, such analytics may rely on consumer-level data; for example, firms might track behavior of subscribers across their sites. However, often firms cannot track individual consumers: consistently identifying nonsubscribers through cookies may be difficult if consumers do not allow cookies, cookies expire, or consumers use multiple devices. Our approach illustrates how firms can use aggregate data to align their paid content offering with demand.

Of course, our work has limitations. Our study focuses on the immediate, short-term effects from offering paid content. Additional, long-term effects may exist for which we are not able to account. We also focus on one dimension along which firms can optimize their online content revenue model – whether firms benefit from a static or dynamic policy. There are many other important questions that we leave to future research and managerial practice to explore such as whether content providers should bundle their online and offline offering, whether a paywall should be 'leaky' (allowing free access through social media or search engines), what optimal price to set, and whether such prices should vary with article type or quality. While our analysis focuses on data that is aggregate across consumers, questions that focus on understanding consumer-level trade-offs would require access to consumer-level data. Lastly, our study is set in an industry in which many firms (still) offer all content for free. In settings where all or most competitors charge for access to content, a different kind of subscription model may be optimal. We leave such questions for future research.

Bibliography

Angelucci, Charles, Julia Cage, and Romain De Nijs (2013), "Price Discrimination in a Two-Sided Market: Theory and Evidence from the Newspaper Industry," Working paper.

Ascarza, Eva, Anja Lambrecht, and N.J. Vilcassim (2012), "When Talk is "Free": An Analysis of Subscriber Behavior under Two- and Three-Part Tariffs," Journal of Marketing Research, 49 (6), 882-99.

Athey, Susan, Emilio Calvano, and Joshua S. Gans (2011), "The Impact of the Internet on Advertising Markets for News Media," Working paper.

Bawa, Kapil and Robert Shoemaker (2004), "The Effects of Free Sample Promotions on Incremental Brand Sales," Marketing Science, 23 (3), 345–63.

Bhargava, Hemant K. and Vidyanand Choudhary (2001), "Information Goods and Vertical Differentiation," Journal of Management Information Systems, 18 (2), 89-106.

Chernev, Alexander (2012), “Product Assortment and Consumer Choice: An Interdisciplinary Review,”Foundations and Trends in Marketing, 6 (1).

Chiou, Lesley and Catherine Tucker (2012), "Paywalls and the Demand for News," Information Economics and Policy, forthcoming.

ESPN (2012), "ESPN Digital Media Breaks Ground in 2011," Press release (http://espnmediazone.com/us/press-releases/2012/01/espn-digital-media-breaks-ground-in-2011/).

Godes, David, Elie Ofek, and Miklos Sarvary (2009), "Content vs. Advertising: The Impact of Competition on Media Firm Strategy," Marketing Science, 28 (1), 20-35.

Halbheer, Daniel, Florian Stahl, Oded Koenigsberg, and Donald R. Lehmann (2013), "Digital Content Strategies," Working paper.

Hirschorn, Michael (2009), "End Times," The Atlantic, January/February 2009.

Kannan, P.K., Barbara Kline Pope, Sanjay Jain (2009), "Pricing Digital Content Product Lines: A Model and Application for the National Academies Press," Marketing Science, 28 (4), 620-636.

Koenker, Roger (2005), "Quantile Regression," Cambridge Books, Cambridge University Press.

Kumar, Vineet, Bharat Anand, Sunil Gupta, and Felix Overholzer-Gee (2013), "The New York Times Paywall," Harvard Business School Case.

Lambrecht, Anja, Avi Goldfarb, Alessandro Bonatti, Anindya Ghose, Dan Goldstein, Randall Lewis, Anita Rao, Navdeep Sahni, and Song Yao (2014), "How Do Firms Make Money Selling Digital Goods Online?," Working paper.

Pattabhiramaiah, Adithya, S. Sriram, and Shrihari Sridhar (2013), "Rising Prices under Declining Preferences: The case of the U.S. Print Newspaper Industry," Working paper.

Pauwels, Koen and Allen Weiss (2007), "Moving From Free to Fee: How Marketing Can Stimulate Gains and Stem Losses for an Online Content Provider," Journal of Marketing, 72 (3).

Pauwels, Koen, Peter S.H. Leeflang, Marije L. Teerling and K.R. Eelko Huizingh, "Does Online Information Drive Offline Revenues? Only for Specific Products and Consumer Segments!" Journal of Retailing, 87, 1-17.

Prasad, Ashutosh, Vijay Mahajan, and Bart Bronnenberg (2003), "Advertising versus pay-per-view in electronic media," International Journal of Research in Marketing, 20, 13-30.

Seamans, Robert and Feng Zhu (2013), "Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers," Working paper.

Shampanier, Kristina, Nina Mazar, and Dan Ariely (2007), "Zero as a Special Price: The True Value of Free Products," Marketing Science, 26 (6), 745-57.

Shapiro, Carl and Hal H. Varian (1998), "Versioning: The Smart Way to Sell Information," Harvard Business Review, 76(Nov/Dec), 106-14.

Table 1: Articles and Activity on ESPN and Yahoo

			Percentiles		
Panel A: ESPN	Mean	Std. Dev.	5th	Median	95th
Articles					
All	33.8	17.8	15	30	70
Free	24.8	17.3	10	19	62
Paid	9.0	3.8	4	8	17
Unique Visitors					
All pages	648,713	504,733	107,745	490,503	1,564,764
Paid pages	28,309	42,819	1,504	19,269	78,292
Page views					
All pages	3,665,127	3,873,991	391,988	2,190,945	10,663,158
Free pages	3,596,655	3,852,824	361,481	2,126,869	10,602,879
Paid pages	68,472	159,223	1,871	34,319	225,185
Page views per unique visitor					
All pages	5.2	2.4	2.6	4.5	9.7
Paid pages	2.1	1.7	1.1	1.6	4.4
Time spent (min)					
All pages	3,583,409	3,991,603	289,428	2,032,662	10,862,223
Free pages	3,528,311	3,973,873	272,430	1,985,387	10,772,211
Paid pages	55,098	137,155	772	26,303	183,661
Time spent per page (min)					
All pages	0.9	0.3	0.6	0.9	1.4
Free pages	0.9	0.3	0.6	0.9	1.4
Paid pages	0.8	0.5	0.2	0.8	1.8
Time spent per unique visitor					
Free pages	5.2	3.4	2.3	4.3	10.8
Paid pages	1.7	1.4	0.3	1.4	3.7
Panel B: Yahoo					
Unique Visitors	879,752	989,628	54,791	562,754	2,826,432
Page views	4,451,955	5,244,159	285,452	2,231,225	14,160,282
Page views per unique visitor	5.5	3.3	2.2	4.6	12.2
Time spent (min)	5,073,084	6,404,012	209,335	2,429,279	16,166,617
Time spent per unique visitor	6.2	5.3	2.1	4.9	14.1

N=2032

Table 2: Length of Free and Paid Articles

	Mean (word count)	Std.dev. (word count)	N
Articles overall			
Free	965	837	824
Paid	1332	654	274
Top 274 per type (by length)			
Free	1832	921	274
Paid	1332	654	274
Category: Top Stories			
Free	1392	980	402
Paid	1241	538	139
Category: Headlines			
Free	615	404	481
Paid	1561	1047	46
Category: Insider			
Paid	1404	587	148

Note: Word counts for 11/9 - 11/15; sometimes articles are listed in more than one category.

Table 3: Data by Sports Seasons

	Off season			Regular season			Post season		
	Median	Mean	Std. dev.	Median	Mean	Std. dev.	Median	Mean	Std. dev.
Number of articles									
All	24	25.6	9.5	34	40.9	19.6	35	40.1	22.2
Free	16	16.6	8.0	25	31.8	19.5	25	31.6	21.2
Paid	8	8.9	4.3	9	9.1	3.4	9	8.5	3.1
Unique visitors									
ESPN all pages	299,375	362,540	282,980	884,217	899,314	542,883	851,424	856,498	434,384
ESPN paid pages	18,975	30,060	56,203	18,817	24,588	22,822	23,792	36,363	37,386
Yahoo	259,349	534,452	752,980	910,998	1,163,744	1,102,731	963,030	1,208,856	912,929
Page views									
ESPN all pages	1,134,415	1,522,005	1,663,506	5,084,165	5,765,398	4,511,573	3,858,686	4,268,230	2,687,423
ESPN free pages	1,063,301	1,437,158	1,550,101	4,986,436	5,719,069	4,497,247	3,702,846	4,178,381	2,667,945
ESPN paid pages	36,446	84,847	220,913	30,697	46,329	52,391	45,585	89,849	116,629
Yahoo	1,048,170	1,933,525	2,667,499	5,531,379	6,835,527	6,278,702	4,911,359	5,520,920	3,754,829
N	932			891			209		

Table 4: Explanatory Power of Demand Shifters

	R_2 when controlling for ...		
Dependent variable	Sports	Sports, Seasons	Sports, Seasons, additional demand controls*
Number of paid articles	44.8%	44.9%	46.9%
log(Revenue)	30.36%	61.98%	79.21%

* Additional controls are nonworking day, gameday, scaled Google searches, indicators for draft, NBA lockout and whether a final game was played that day.

Table 5: Effect of Paid Articles on Visitors and Page Views

	(1) No controls, DV is number of articles		(2) Sports controls, DV is number of articles		(3) Sports and season controls, DV is number of articles		(4) All controls, DV is number of articles		(5) All controls, DV is percent of articles	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Share ESPN										
Free Articles	0.016	0.001 ***	0.019	0.001 ***	0.006	0.001 ***	0.003	0.001 ***		
Paid Articles	0.062	0.004 ***	-0.018	0.005 ***	-0.010	0.003 ***	0.003	0.003		
Percent Paid Articles									-0.165	0.090 *
Regular Season					0.981	0.024 ***	0.453	0.025 ***	0.448	0.025 ***
Post Season					0.975	0.036 ***	0.396	0.031 ***	0.390	0.031 ***
Nonworkingday							-0.240	0.017 ***	-0.240	0.017 ***
Gameday							0.072	0.025 ***	0.071	0.025 ***
Googlescaled							0.023	0.001 ***	0.023	0.001 ***
ln(Yahoo Unique Visitors)							0.165	0.009 ***	0.163	0.009 ***
Draft							-0.075	0.100	-0.070	0.100
Lockout							-0.007	0.039	-0.010	0.039
Final Game							-0.165	0.148	-0.170	0.148
Constant	-6.798	0.049 ***								
Share Paid Section										
Free Articles	-0.010	0.001 ***	-0.008	0.001 ***	0.004	0.001 **	0.003	0.002 **		
Paid Articles	0.019	0.007 ***	0.063	0.008 ***	0.057	0.008 ***	0.050	0.008 ***		
Percent Paid Articles									1.429	0.268 ***
Regular Season					-0.973	0.053 ***	-0.817	0.074 ***	-0.776	0.075 ***
Post Season					-0.621	0.080 ***	-0.510	0.094 ***	-0.461	0.094 ***
Nonworkingday							-0.079	0.050	-0.081	0.050
Gameday							-0.115	0.075	-0.104	0.075
Googlescaled							-0.006	0.003 **	-0.006	0.003 **
ln(Yahoo Unique Visitors)							-0.046	0.029	-0.048	0.029
Draft							1.251	0.299 ***	1.270	0.300 ***
Lockout							-0.637	0.117 ***	-0.639	0.117 ***
Final Game							0.079	0.442	0.019	0.442
Constant	-3.284	0.074 ***								
ln(Page Views)										
Free Articles	0.021	0.001 ***	0.024	0.001 ***	0.006	0.001 ***	0.002	0.001 **		
Paid Articles	0.060	0.006 ***	-0.034	0.006 ***	-0.023	0.005 ***	-0.008	0.003 **		
Percent Paid Articles									-0.350	0.118 ***
Regular Season					1.354	0.031 ***	0.588	0.033 ***	0.577	0.033 ***
Post Season					1.117	0.048 ***	0.349	0.041 ***	0.337	0.041 ***
Nonworkingday							-0.268	0.022 ***	-0.267	0.022 ***
Gameday							0.172	0.033 ***	0.170	0.033 ***
Googlescaled							0.033	0.001 ***	0.033	0.001 ***
ln(Yahoo Page Views)							0.186	0.013 ***	0.185	0.013 ***
Draft							0.052	0.132	0.050	0.132
Lockout							0.123	0.051 **	0.122	0.051 **
Final Game							-0.408	0.194 **	-0.397	0.194 **
Constant	13.555	0.061 ***								
Sport controls	no		yes		yes		yes		yes	
Controls for Total Articles, by Season	no		no		no		no		yes	
N	2032		2032		2032		2032		2032	
R-2 Share ESPN	0.1923		0.4864		0.7305		0.853		0.854	
R-2 Share Paid Section	0.0234		0.1598		0.2802		0.300		0.298	
R-2 ln(Page Views)	0.1737		0.3973		0.6887		0.832		0.832	

*** p<0.01, **p<0.05, *p<0.1

Table 6: Financial Impact of Adding a Paid Article

Price per 1000 impressions	Statistic	(1) Number of articles	(2) Number of articles, limited space	(3) SUR: Percent of articles, limited space
8.34	Median $ per day	$789	$435	$151
	% days sig (5%) negative	28%	36%	44%
	% days sig (5%) positive	72%	64%	56%
11.51	Median $ per day	$489	$115	($146)
	% days sig (5%) negative	36%	45%	53%
	% days sig (5%) positive	64%	55%	47%
15.45	Median $ per day	$147	($180)	($495)
	% days sig (5%) negative	44%	53%	62%
	% days sig (5%) positive	56%	47%	38%

Table 7: Effect of Paid Articles on Visitors and Page Views by Season

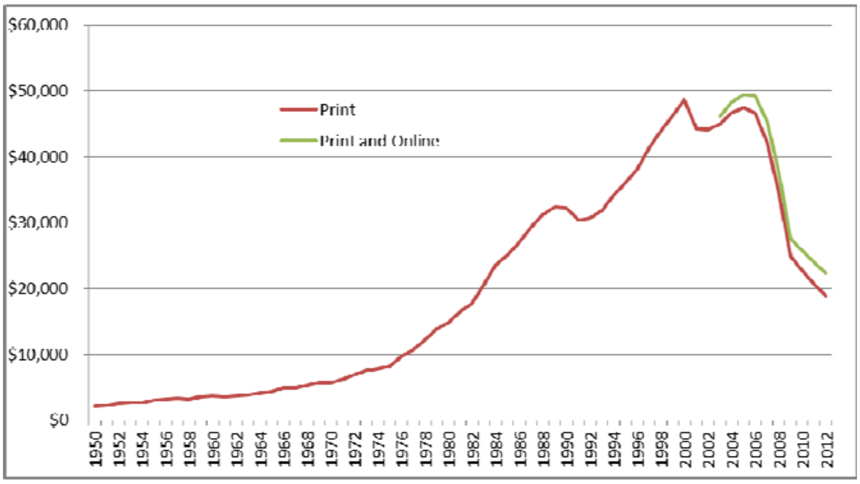
	(1) No controls, DV is number of articles		(2) Sports controls, DV is number of articles		(3) Sports and season controls, DV is number of articles		(4) All controls, DV is number of articles		(5) All controls, DV is percent of articles	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Share ESPN										
Free Articles -Off Season	0.007	0.003 ***	0.000	0.002	0.015	0.002 ***	0.010	0.001 ***		
Free Articles -Regular Season	0.007	0.001 ***	0.010	0.001 ***	0.004	0.001 ***	0.002	0.001 ***		
Free Articles -Post Season	0.010	0.002 ***	0.010	0.002 ***	0.006	0.002 ***	0.004	0.001 ***		
Paid Articles -Off Season	0.031	0.005 ***	-0.040	0.004 ***	-0.002	0.004	0.018	0.003 ***		
Paid Articles -Regular Season	0.107	0.005 ***	0.026	0.004 ***	-0.022	0.005 ***	-0.016	0.004 ***		
Paid Articles -Post Season	0.113	0.010 ***	0.027	0.008 ***	-0.024	0.011 **	-0.025	0.008 ***		
Percent Paid Articles - Off Season									0.267	0.114 **
Percent Paid Articles - Regular Season									-0.710	0.125 ***
Percent Paid Articles - Post Season									-0.778	0.368 **
Regular Season					1.386	0.067 ***	0.917	0.052 ***	1.139	0.077 ***
Post Season					1.298	0.105 ***	0.849	0.078 ***	1.017	0.142 ***
Nonworkingday							-0.242	0.016 ***	-0.240	0.016 ***
Gameday							0.046	0.025 *	0.046	0.025 *
Googlescaled							0.025	0.001 ***	0.025	0.001 ***
ln(Yahoo Unique Visitors)							0.158	0.008 ***	0.158	0.008 ***
Draft							-0.067	0.098	-0.074	0.098
Lockout							-0.029	0.039	-0.024	0.039
Final Game							-0.217	0.145	-0.188	0.145
Constant	-6.673	0.047 ***								
Share Paid Section										
Free Articles -Off Season	0.021	0.004 ***	0.013	0.004 ***	-0.002	0.004	0.000	0.004		
Free Articles -Regular Season	-0.004	0.002 **	0.000	0.002	0.006	0.002 ***	0.004	0.002 **		
Free Articles -Post Season	0.006	0.003	-0.004	0.003	0.000	0.003	0.003	0.003		
Paid Articles -Off Season	0.036	0.007 ***	0.082	0.009 ***	0.046	0.009 ***	0.037	0.010 ***		
Paid Articles -Regular Season	-0.014	0.008 *	0.020	0.009 **	0.065	0.011 ***	0.060	0.011 ***		
Paid Articles -Post Season	-0.004	0.016	0.074	0.017 ***	0.133	0.024 ***	0.121	0.024 ***		
Percent Paid Articles - Off Season									0.918	0.348 ***
Percent Paid Articles - Regular Season									1.769	0.382 ***
Percent Paid Articles - Post Season									4.457	1.126 ***
Regular Season					-1.311	0.149 ***	-1.106	0.160 ***	-1.160	0.235 ***
Post Season					-1.361	0.234 ***	-1.268	0.238 ***	-1.835	0.433 ***
Nonworkingday							-0.071	0.050	-0.079	0.050
Gameday							-0.091	0.076	-0.084	0.076
Googlescaled							-0.008	0.003 **	-0.007	0.003 **
ln(Yahoo Unique Visitors)							-0.038	0.029	-0.044	0.029
Draft							1.266	0.299 ***	1.291	0.299 ***
Lockout							-0.623	0.119 ***	-0.644	0.119 ***
Final Game							0.207	0.444	0.146	0.445
Constant	-3.561	0.074 ***								
ln(Page Views)										
Free Articles -Off Season	0.003	0.003	-0.002	0.002	0.017	0.002 ***	0.009	0.002 ***		
Free Articles -Regular Season	0.010	0.001 ***	0.012	0.001 ***	0.004	0.001 ***	0.000	0.001		
Free Articles -Post Season	0.007	0.003 **	0.010	0.002 ***	0.006	0.002 ***	0.000	0.001		
Paid Articles -Off Season	0.018	0.006 ***	-0.064	0.006 ***	-0.019	0.006 ***	0.009	0.004 **		
Paid Articles -Regular Season	0.123	0.006 ***	0.029	0.006 ***	-0.032	0.006 ***	-0.029	0.005 ***		
Paid Articles -Post Season	0.121	0.012 ***	0.016	0.011	-0.035	0.014 **	-0.040	0.011 ***		
Percent Paid Articles - Off Season									0.156	0.151
Percent Paid Articles - Regular Season									-0.923	0.165 ***
Percent Paid Articles - Post Season									-1.177	0.488 **
Regular Season					1.715	0.089 ***	1.093	0.069 ***	1.306	0.101 ***
Post Season					1.452	0.139 ***	0.923	0.102 ***	1.148	0.187 ***
Nonworkingday							-0.271	0.021 ***	-0.268	0.021 ***
Gameday							0.147	0.033 ***	0.145	0.033 ***
Googlescaled							0.034	0.001 ***	0.034	0.001 ***
ln(Yahoo Page Views)							0.178	0.013 ***	0.179	0.013 ***
Draft							0.055	0.130	0.038	0.130
Lockout							0.107	0.052 **	0.117	0.052 **
Final Game							-0.451	0.192 **	-0.412	0.193 **
Constant	13.766	0.056 ***								
Controls for Sports		no		yes		yes		yes		yes
Controls for Total Articles, by Season		no		no		no		no		yes
N		2032		2032		2032		2032		2032
R-2 Share ESPN		0.362		0.6755		0.7373		0.861		0.861
R-2 Share Paid Section		0.159		0.2551		0.2862		0.305		0.302
R-2 ln(Page Views)		0.412		0.6326		0.693		0.839		0.838

*** p<0.01, **p<0.05, *p<0.1

Table 8: Financial Impact of Adding a Paid Article by Season

Price per 1000 impressions	Statistic	(1) Number of articles Season Off	(1) Number of articles Season Regular	(1) Number of articles Season Post	(2) Number of articles, limited space Season Off	(2) Number of articles, limited space Season Regular	(2) Number of articles, limited space Season Post	(3) Percent of articles, limited space Season Off	(3) Percent of articles, limited space Season Regular	(3) Percent of articles, limited space Season Post
8.34	Median $ per day	$550	($628)	($137)	$377	($706)	($204)	$466	($649)	$129
	% days sig (5%) negative	0%	73%	27%	0%	77%	30%	0%	79%	6%
	% days sig (5%) positive	100%	1%	13%	86%	0%	8%	92%	0%	13%
11.51	Median $ per day	$590	($1,050)	($444)	$375	($1,143)	($502)	$496	($1,016)	($256)
	% days sig (5%) negative	0%	84%	38%	0%	87%	40%	0%	89%	11%
	% days sig (5%) positive	100%	0%	5%	81%	0%	3%	90%	0%	5%
15.45	Median $ per day	$630	($1,591)	($775)	$373	($1,686)	($931)	$523	($1,481)	($656)
	% days sig (5%) negative	0%	92%	49%	0%	93%	52%	0%	94%	16%
	% days sig (5%) positive	100%	0%	2%	74%	0%	0%	85%	0%	2%

Figure 1: US Newspaper Advertising Revenue, 1950–2012



Source: Newspaper Association of America.

Figure 2: Screenshot of ESPN website Displaying Insider-icon

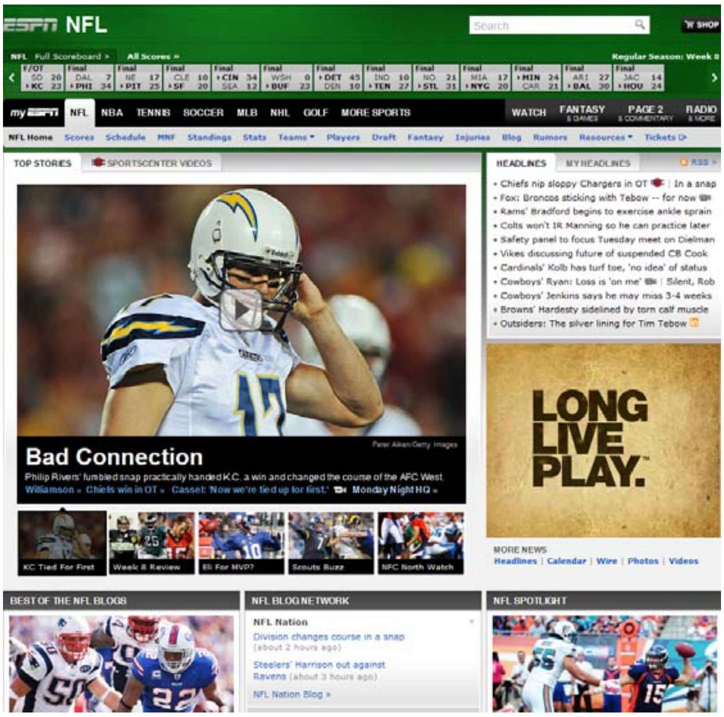


Figure 3: Density of ESPN Revenues across Day-Sports Observations

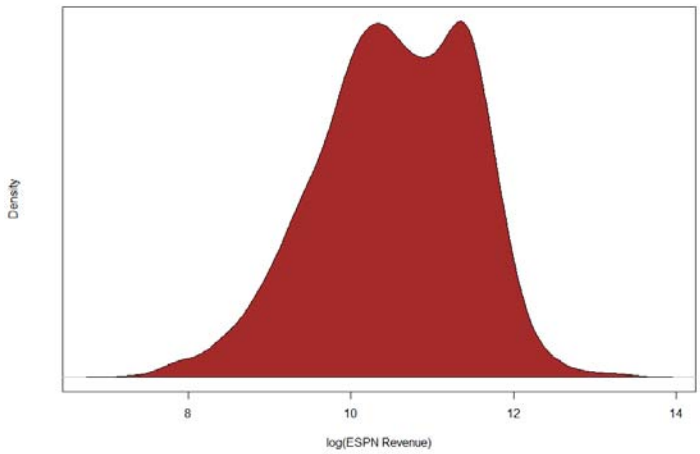


Figure 4: Density of ESPN Revenues across Day-Sport Observations, by Season

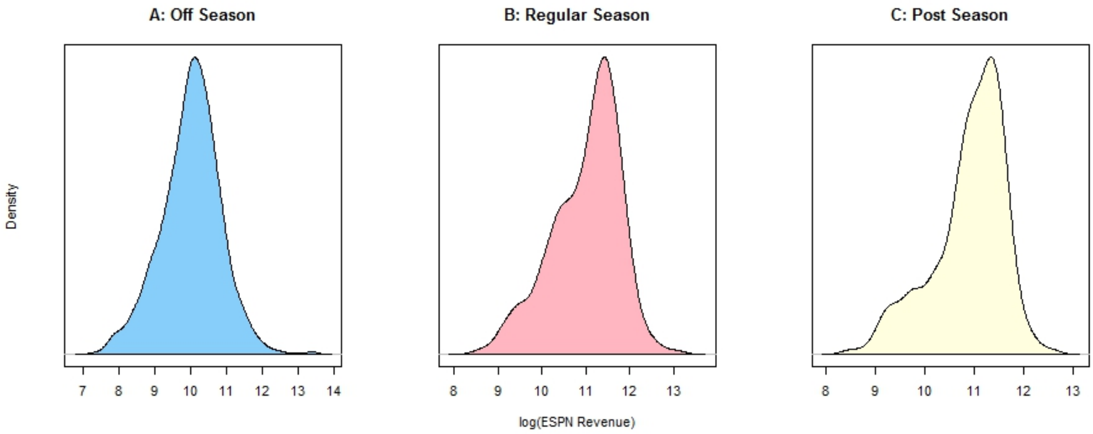


Figure 5: Relationship between Paid Articles and Unique Paying Visitors, Respectively Total Page Views

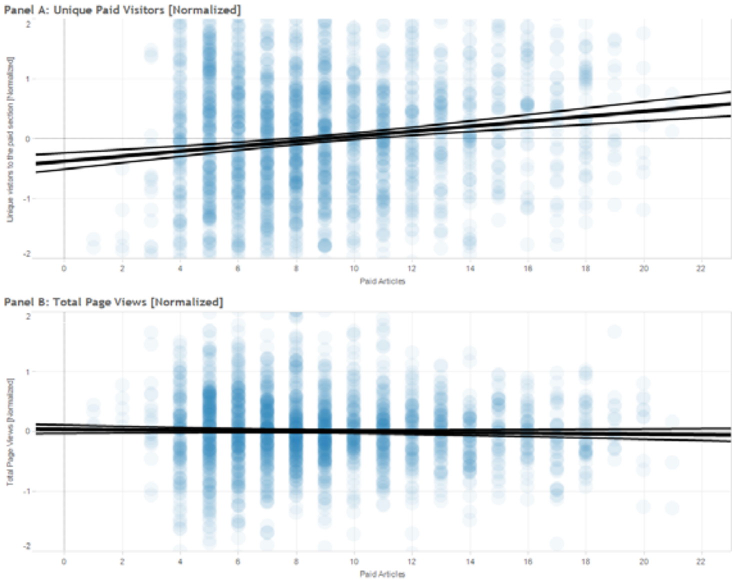


Figure 6: Relationship of Paid Articles to Unique Visitors to the Paid Section, Respectively Total Page Views, by Season

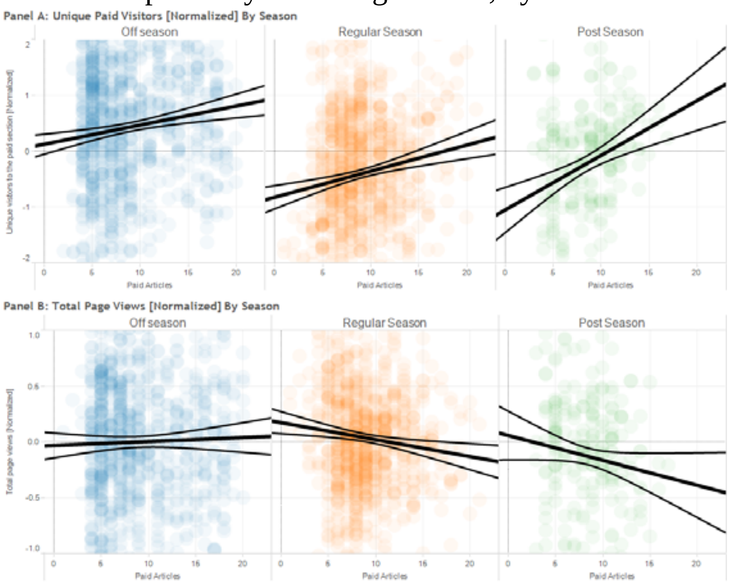


Figure 7: Financial Impact of Adding a Paid Article

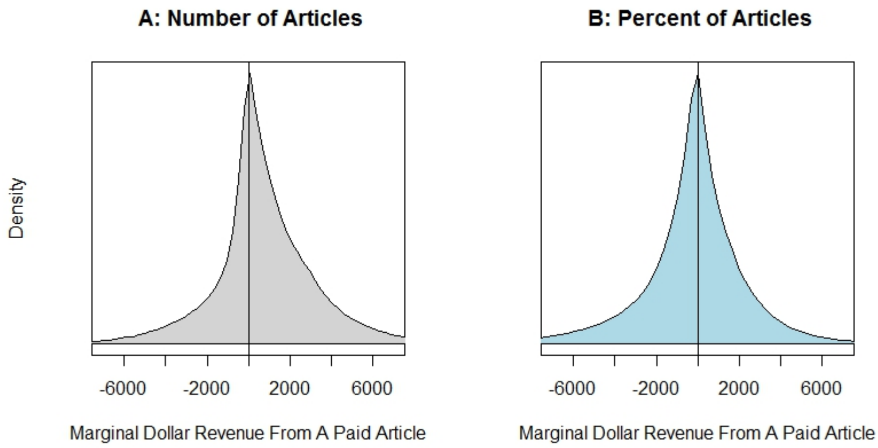


Figure 8: Impact of Marginal Paid Article on Revenue, OLS and QR

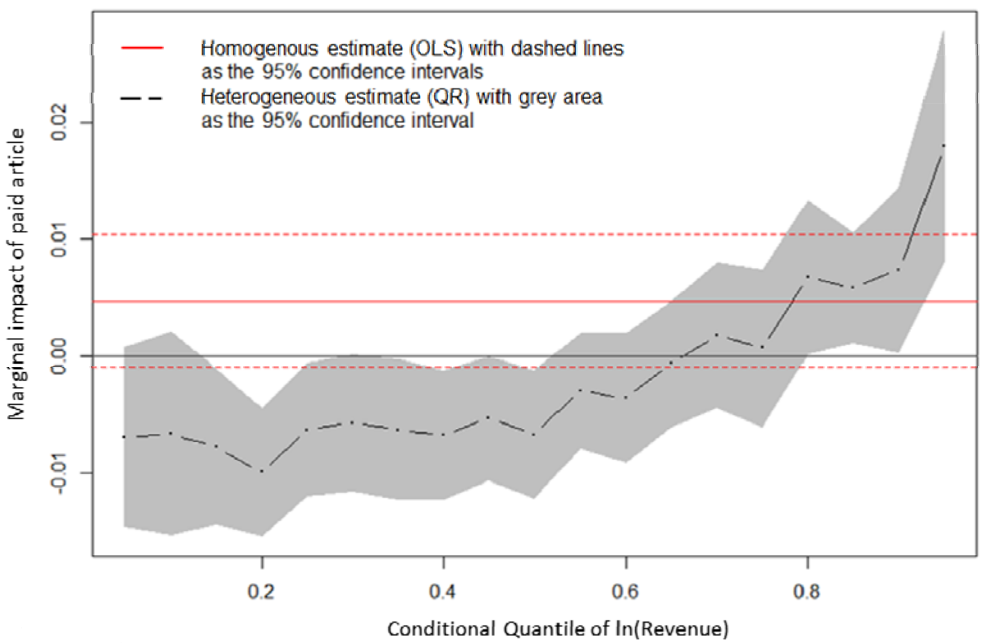


Figure 9: Financial Impact of Adding a Paid Article by Season

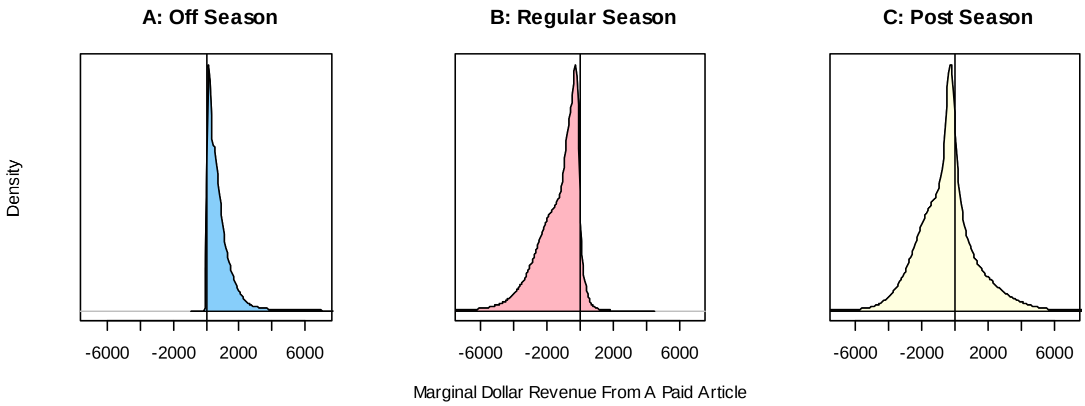


Figure 10: Impact of Marginal Paid Article on Revenue, OLS, and QR, by Season

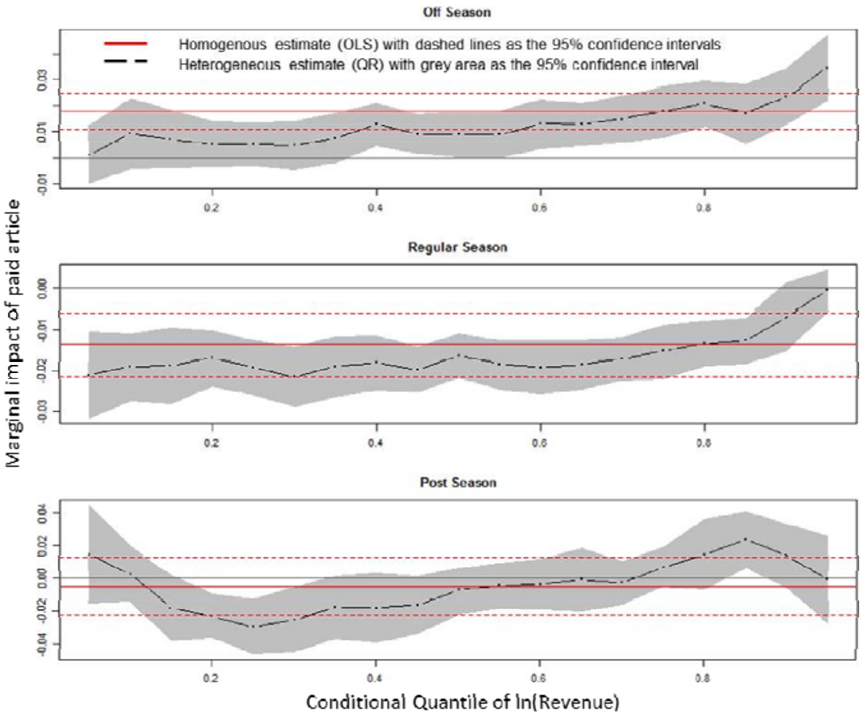


Figure 11: Financial Impact of Adding a Paid Article by Sport, Regular Season

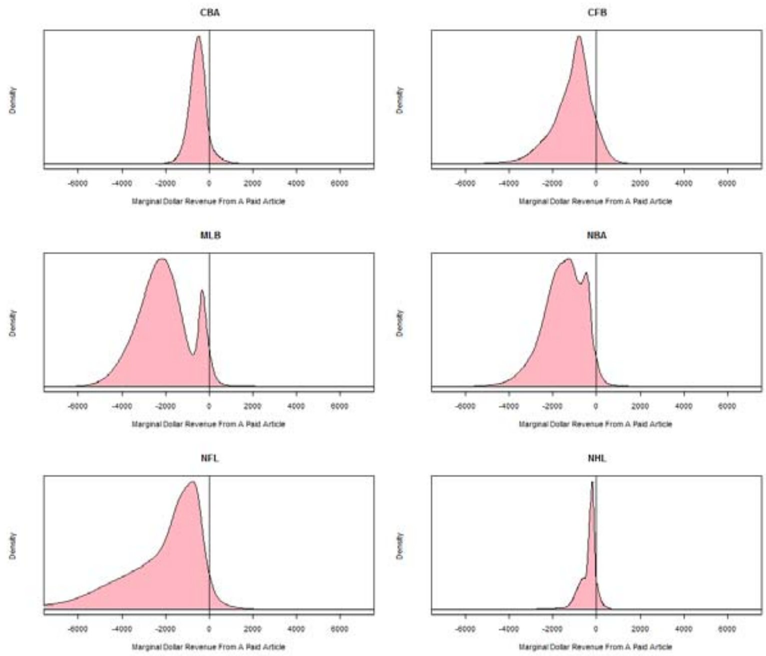
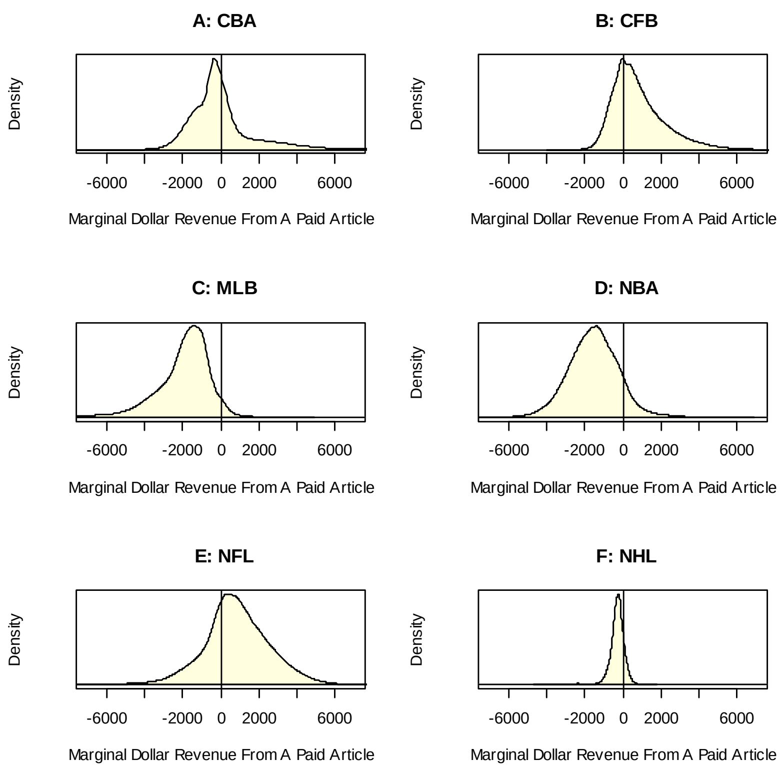


Figure 12: Dollar Effect of an Additional Paid Article by Sport for Post-Season



Web Appendix

This Appendix summarizes additional results of a three-stage least squares estimation that instruments for the number of articles.

A.1 Overall

The specification presented in the main paper controls for a wide range of demand shifters including those specific to overall demand in a sport ($UniqVisYa_{it}$, $PageViewsYa_{it}$), those specific to a sport at ESPN ($Google_{it}$), and a range of variables that capture additional demand shocks related to specific events.

These demand shifters cover a wide range of possible events that can affect both firm policy and consumer demand. Moreover, anecdotal evidence suggests that rather than knowing the revenue-optimizing paywall, firms experiment with respect to their paid content strategy.[13] As a result, we do not believe that in our main specifications, the number of paid articles the firm displays is likely to be endogeneous. Yet we cannot conclusively rule out the possibility that even after controlling for a wide range of observable demand shocks, ESPN might observe demand shocks that the researcher does not, and that ESPN might use this information when deciding on the number of free or paid articles. An example could be a news story that is unique to ESPN (e.g., ESPN signing a new Monday Night Football deal with the NFL on September 8, 2011) if such an event would not be fully captured by the variable $Google_{it}$ that measures "ESPN + sport" specific searches on Google.

13 The design of the *New York Times* paywall seems to be based more on trial and error than robust optimization (http://www.poynter.org/latest-news/mediawire/167147/changes-to-new-york-times-paywall/). The examples provided in the introduction further suggest that firms are not yet necessarily aware of the optimal paid-content strategy.

To check whether our main results are robust, we turn to an estimation that controls for such possible endogeneity. We use as an instrument the number of free and paid articles ESPN displayed the previous day, that is, the day before such news was known.

This instrumenting strategy builds on the insight that on any given day, the firm does not update the full set of articles it displays for any sport, but instead retains a subset of articles that were displayed the previous day. Specifically, on average across all days and sports, 39% of free and 25% of paid articles displayed every day are new content, whereas 61% of free and 75% of paid articles were displayed the previous day. On any day, the average age of free articles displayed is 11 days and the average age of paid articles is seven days. This finding suggests that whereas the firm updates content over time, such updating happens gradually.

Displaying an article for more than a single day makes sense as long as potential readers do not visit the site every day. Indeed, our data indicate that customers visit the firm's website on average every 4.2 days, meaning that an article initially displayed the previous day will still be of interest to many customers visiting today. Continuing to display an existing article is attractive for the firm because it incurs zero marginal cost of production on the second day.

In the case of the ESPN-NFL deal on September 8, 2011, we would use the number of free and paid articles displayed on September 7, 2011, to instrument for the number of free and paid articles on September 8, 2011. Indeed, although ESPN featured an article on the deal on September 8, it displayed no such report on the day before the deal was announced.

Using the number of paid and free articles displayed the previous day in a sport as an instrument for the number of paid and free articles on a day in a sport means we assume that, after controlling for the extensive set of our demand shifters, yesterday's free and paid articles affect unique visitors and page views today only through the number of articles today and not in any other way. This assumption would be problematic if the firm would not immediately publish a new article but rather would delay publishing until the next day, possibly because of anticipated demand that day. For example, it would be a problem if, hypothetically, ESPN would expect greater traffic to its site on September 9 and hold back reporting the news for a day. However, the market for online news is highly competitive and, by its nature, competes on real-time information. As a result, delaying news does not seem a likely strategy.[14] Note that our instrumenting strategy also assumes the firm is myopic, meaning the expected demand on the same day but not the following day may affect the decision of whether to publish.

We re-estimate the specification in our main paper, accounting for the possible endogeneity of the number of paid and free articles, using as instruments the lagged number of paid and free articles in that sport on that day. Table A-1 displays the results of the 3SLS estimation with endogenous explanatory variables. The results are similar to those in Table 5 of the main paper. Specifically, the estimates now suggest that, calculated at the median value, an additional paid article increases viewership in the paid section by 766 unique visitors. Similarly to before, paid articles significantly reduce page

[14] Articles forwarded through email or social media are likely consumed by recipients on the same day. However, we cannot rule out that consumption may occur the next day. Then our identifying assumption that yesterday's articles affect unique visitors and page views today only through the number of articles today would not be valid. We checked whether this assumption affects the results. We rerun the by-season specification we discuss in Section 5.2 and include as additional control

views on the site overall, here by 24,394. In a further unreported specification we find that our results are robust to including as additional instrument the age of the previous day's free and paid articles.

We evaluate the effect of adding a paid article on firm revenue similar to Section 5.1 in the main paper. Table A-2 displays the results. Column (1) reports the results if the firm changes the number of paid articles, holding constant the number of free articles while Column (2) assumes that the firm substitutes between paid and free articles. Together, the results indicate that the median effect on revenue from a marginal paid article lies between a loss of $1,343 and a gain of $162. Similar to the results displayed in Table 6 of the main paper, the majority of day-sport observations is significantly different from zero, either positive or negative. Figure A-1 confirms that the revenue effect across day-sport observations varies greatly.

A.2 By Season

Building on the specification in Section 5.2 in the main paper, we estimate a 3SLS specification in which we let the number of articles vary by season, using as instruments the lagged number of free and paid articles by season. Table A-3 displays the results. Across all three equations, the estimated coefficients are similar in sign, significance, and size to those summarized in Column (4) of Table 7 in the main paper.

Again, we evaluate the effect of a marginal paid article on the firm's revenues. Together, the results in Columns (1) and (2) (for paid and free articles, respectively) that assume the firm holds constant the number of paid articles confirm the key result of the

unique visitors to the paid section and page views the previous day. The coefficients are broadly similar in sign, significance, and size to those we present in Table 7.

main paper: during the off-season, the firm clearly benefits from increasing the number of paid articles. During the regular season, the firm should decrease the number of paid articles. Again similar to the main paper, the results are mixed during the post-season. Figure A-2 displays the density of day-sport observations, which again confirms these findings.

Lastly, we confirm the robustness of our finding from Section 6 that the variation in revenue effects we observe in the post-season comes from differences across sports. Figure A-3 displays the density for revenue across day-sport observations and again illustrate that the result this holds across sports. Consistent with our results in Section 6, Figure A-4 confirms the firm would benefit from decreasing paid content in all sports except college football and the NFL.

Figure A-1: Density of Marginal Revenue from a Paid Article, 3SLS

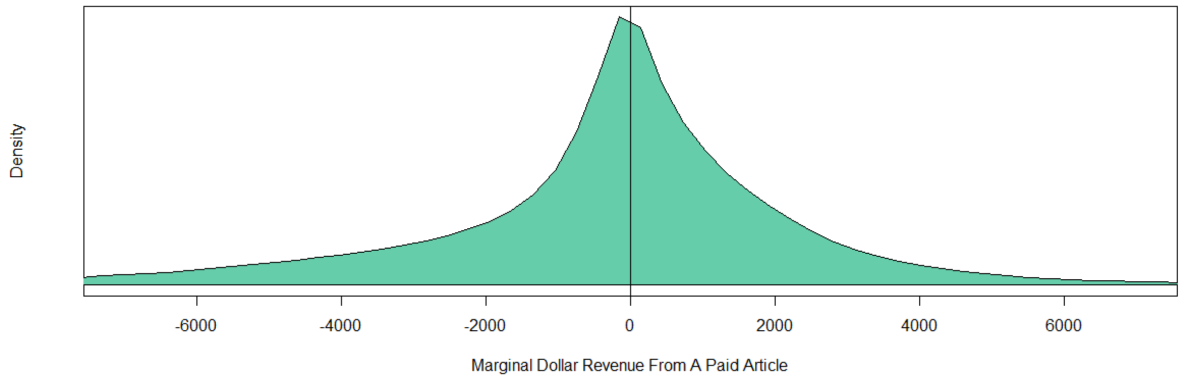


Figure A-2: Density of Marginal Revenue from a Paid Article, 3SLS

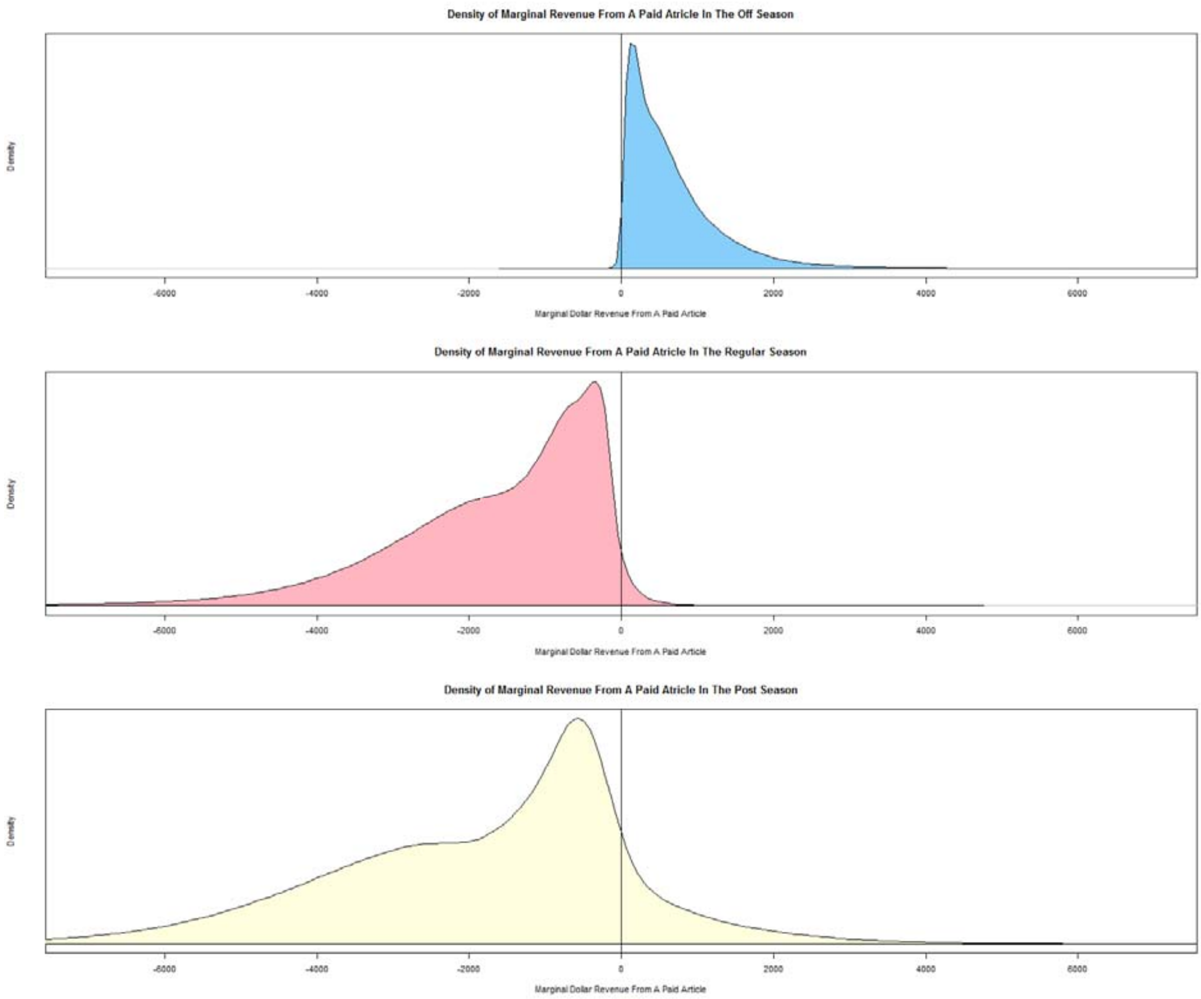


Figure A-3: Density of Marginal Revenue from a Paid Article, Regular Season, 3SLS

Figure A-4: Density of Marginal Revenue from a Paid Article, Post-Season, 3SLS

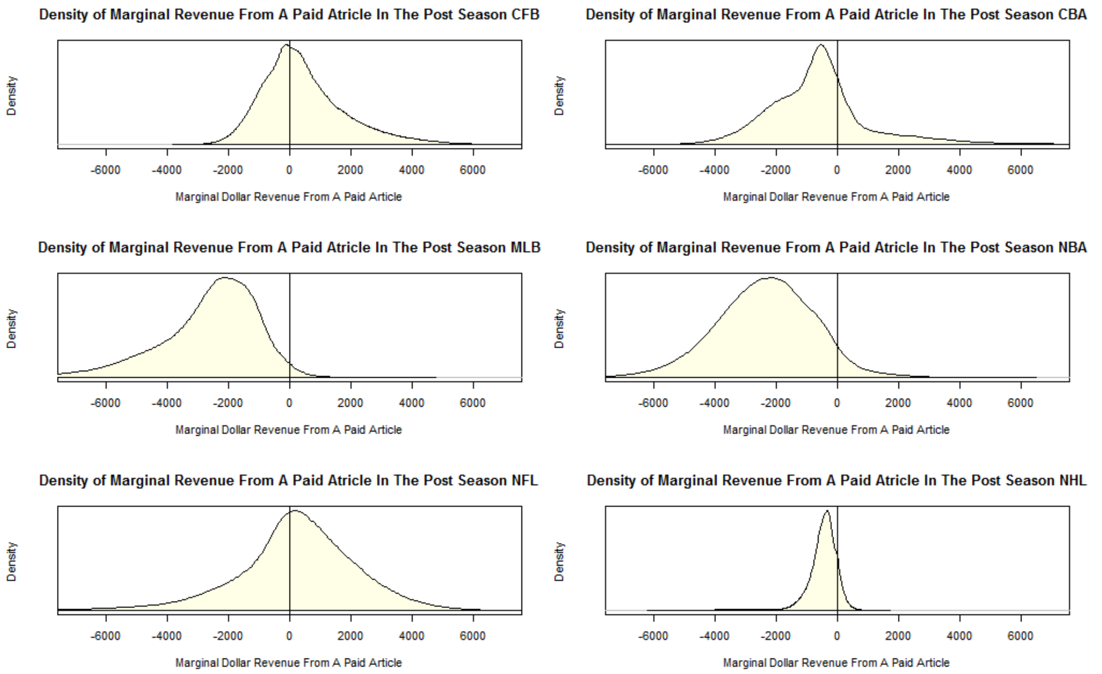


Table A-1: Effect of Paid Articles on Visitors and Page Views, 3SLS

	3SLS		
	Coef.	Std. Err.	
Share ESPN			
Free Articles	0.006	0.001	***
Paid Articles	0.002	0.003	
Regular Season	0.425	0.025	***
Post Season	0.373	0.032	***
Nonworkingday	-0.248	0.017	***
Gameday	0.046	0.026	*
Googlescaled	0.024	0.001	***
ln(Yahoo Unique Visi	0.161	0.009	***
Draft	-0.090	0.106	
Lockout	0.034	0.040	
Final Game	-0.199	0.149	
Constant	-8.662	0.119	***
Share Paid Section			
Free Articles	0.008	0.002	***
Paid Articles	0.040	0.009	***
Regular Season	-0.834	0.076	***
Post Season	-0.513	0.095	***
Nonworkingday	-0.085	0.050	*
Gameday	-0.158	0.078	**
Googlescaled	-0.006	0.003	**
ln(Yahoo Unique Visi	-0.054	0.030	*
Draft	1.319	0.315	***
Lockout	-0.602	0.119	***
Final Game	0.005	0.443	
Constant	-2.752	0.399	***
ln(Page Views)			
Free Articles	0.004	0.001	***
Paid Articles	-0.011	0.004	***
Regular Season	0.566	0.033	***
Post Season	0.333	0.041	***
Nonworkingday	-0.270	0.022	***
Gameday	0.163	0.034	***
Googlescaled	0.033	0.001	***
ln(Yahoo Unique Visi	0.187	0.014	***
Draft	0.091	0.138	
Lockout	0.149	0.052	***
Final Game	-0.438	0.194	**
Constant	11.184	0.202	***
N		2007	
R-2 Share ESPN		0.851	
R-2 Share Paid Section		0.297	
R-2 ln(Page Views)		0.832	
Significance of first-sta	significant at 0.001		
First stage R-2	0.65 - 0.84		

Fixed effects by sport included but not displayed for readability.
*** p<0.01, **p<0.05, *p<0.1

Table A-2: Financial Impact of Adding a Paid Article, 3SLS

Price per 1000 impressions	Statistic	(1) 3SLS	(2) 3SLS, limited space
8.34	Median $ per day	$162	($392)
	% days sig (5%) negative	43%	65%
	% days sig (5%) positive	57%	34%
11.51	Median $ per day	($130)	($797)
	% days sig (5%) negative	52%	74%
	% days sig (5%) positive	48%	26%
15.45	Median $ per day	($457)	($1,343)
	% days sig (5%) negative	44%	53%
	% days sig (5%) positive	39%	19%

Table A-3: Effect of Paid Articles on Visitors and Page Views, by Season, 3SLS

	3SLS		
	Coef.	Std. Err.	
Share ESPN			
Free Articles -Off Season	0.010	0.001	***
Free Articles -Regular Season	0.006	0.001	***
Free Articles -Post Season	0.005	0.001	***
Paid Articles -Off Season	0.018	0.004	***
Paid Articles -Regular Season	-0.019	0.004	***
Paid Articles -Post Season	-0.034	0.010	***
Regular Season	0.849	0.062	***
Post Season	0.940	0.087	***
Nonworkingday	-0.254	0.017	***
Gameday	0.015	0.026	
Googlescaled	0.026	0.001	***
ln(Yahoo Unique Visitors)	0.154	0.009	***
Draft	-0.102	0.104	
Lockout	0.033	0.042	
Final Game	-0.247	0.147	*
Constant	-8.799	0.118	***
Share Paid Section			
Free Articles -Off Season	0.002	0.005	
Free Articles -Regular Season	0.011	0.003	***
Free Articles -Post Season	0.005	0.004	
Paid Articles -Off Season	0.031	0.011	***
Paid Articles -Regular Season	0.050	0.013	***
Paid Articles -Post Season	0.107	0.029	***
Regular Season	-1.205	0.188	***
Post Season	-1.164	0.264	***
Nonworkingday	-0.087	0.051	*
Gameday	-0.152	0.081	*
Googlescaled	-0.007	0.003	**
ln(Yahoo Unique Visitors)	-0.048	0.030	
Draft	1.306	0.316	***
Lockout	-0.537	0.128	***
Final Game	0.157	0.448	
Constant	-2.701	0.400	***
ln(Page Views)			
Free Articles -Off Season	0.009	0.002	***
Free Articles -Regular Season	0.003	0.002	*
Free Articles -Post Season	0.002	0.002	
Paid Articles -Off Season	0.006	0.005	
Paid Articles -Regular Season	-0.031	0.006	***
Paid Articles -Post Season	-0.060	0.013	***
Regular Season	1.022	0.080	***
Post Season	1.051	0.113	***
Nonworkingday	-0.276	0.022	***
Gameday	0.135	0.035	***
Googlescaled	0.035	0.001	***
ln(Yahoo Page Views)	0.179	0.014	***
Draft	0.077	0.136	
Lockout	0.141	0.055	**
Final Game	-0.507	0.193	***
Constant	11.036	0.202	***
N		2007	
R-2 Share ESPN		0.858	
R-2 Share Paid Section		0.299	
R-2 ln(Page Views)		0.838	
Significance of first-stage regres:	significant at 0.001		
First stage R-2	0.77 - 0.98		

Fixed effects by sport included but not displayed for readability. *** p<0.01, **p<0.05, *p<0.1

Table A-4: Financial Impact of Adding a Paid Article, by Season, 3SLS

Price per 1000 impressions	Statistic	(1) 3SLS Season Off	Regular	Post	(2) 3SLS, limited space Season Off	Regular	Post
8.34	Median $ per day	$475	($857)	($754)	$273	($1,228)	($963)
	% days sig (5%) negative	0%	82%	52%	0%	93%	57%
	% days sig (5%) positive	97%	0%	0%	69%	0%	0%
11.51	Median $ per day	$499	($1,338)	($1,421)	$264	($1,772)	($1,577)
	% days sig (5%) negative	0%	90%	63%	0%	98%	67%
	% days sig (5%) positive	95%	0%	0%	57%	0%	0%
15.45	Median $ per day	$533	($1,952)	($2,137)	$248	($2,444)	($2,394)
	% days sig (5%) negative	0%	95%	72%	0%	99%	76%
	% days sig (5%) positive	93%	0%	0%	45%	0%	0%