Marketing Science Institute Working Paper Series 2013
Report No. 13-122

Pricing Online Content: Fee or Free?

Anja Lambrecht and Kanishka Misra

Report Summary

Plummeting advertising revenues have lead many online content providers to experiment with additional sources of revenues. Most often, firms aim to compensate for a loss in advertising revenues by charging consumers for access to online content. However, such a choice is not straightforward since subscription fees typically deter customers, further reducing advertising revenues. As of yet, academic research offers little guidance on whether firms indeed benefit from charging for content and how firms should optimally implement such a "fee" model.

Anja Lambrecht and Kanishka Misra examine *whether* and *how* firms should charge for access to online content. They build a unique data set from the sports website ESPN.com to empirically study this question. ESPN.com offers the majority of content for free but charges a membership fee for a subset of articles. The authors 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.

Using this data, Lambrecht and Misra 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. Their approach controls for a wide range of demand shifters and possible endogeneity of the number of articles the firm offers on any day. The results show that, on average, the firm should not adjust the amount of paid content. However, there are strong differences across sports' seasons: the marginal paid article increases revenue in the off season but decreases revenue in regular season. The authors suggest that this variation over time is largely due to a change in the number and type of unique visitors to the site.

An important implication of this research is that firms can increase revenue by flexibly adjusting the amount of content they offer against a fee instead of setting a static paywall as most often is the case. More broadly, the results suggest that when evaluating whether to charge for content firms should account for heterogeneity in consumer demand. Specifically, firms may benefit from using a dynamic strategy where they flexibly adjust the amount of paid content over time rather than always offering the same amount of content for free or against a fee. The insight that a dynamic policy may allow firms to significantly increase their revenues is especially interesting in light of the fact that a number of online content providers have recently started experimenting with fee-models but have rarely explored flexibly adjusting the amount of paid content.

Anja Lambrecht is Assistant Professor of Marketing at the London Business School, U.K. Kanishka Misra is Assistant Professor of Marketing at the Ross School of Business, University of Michigan.

Acknowledgments
We are grateful to Rajesh Chandy, S. Sriram, Catherine Tucker, seminar participants at Technical University Munich, Groningen University, MIT 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.

Introduction

The future of the media industry is widely believed to depend on the ability of media 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 content online (Edgecliffe-Johnson 2009). This is supported by research showing that consumers respond negatively to even small monetary fees (Shampanier et al. 2007; Ascarza et al. 2012), making it difficult for firms to charge even small amounts for digital content.

Yet, plummeting revenues across the media industry seem to leave no option but for 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 its continued existence (Hirschorn 2009). Most recently, on August 6, 2013 The Washington Post was sold to Jeff Bezos as "for much of the past decade, The Post has been unable to escape the financial turmoil that has engulfed newspapers" (Farhi 2013). Many other regional newspapers, such as the Miami Herald and the San Francisco Chronicle face financial trouble.[1]

Just how such new and additional sources of revenues should be opened, and paywalls potentially implemented, remains unclear. While charging for online content adds subscription revenue (Pauwels and Weiss 2007), it also deters consumers, leading to lower advertising revenues (Chiou and Tucker 2012). Additionally, paywalls may confer benefits on rivals in advertising markets (Athey et al. 2011). As of yet, academic research provides few definitive insights on whether firms should favor 'fee' or 'free' and what factors should guide their decision.

Acknowledging the trade-off between subscription and advertising revenues, firms have in recent years experimented with a wide range of revenue models that include giving away all content for free (e.g., washingtonpost.com), charging for all content (e.g., thetimes.co.uk) and giving away some content free of charge but charging for a subset of content (e.g. ESPN.com, faz.net, nyt.com)[2]. Some content providers have experimented with a variety of strategies: The

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

[2] ESPN.com charges for access to a daily subset of articles, faz.net charges for historic articles only, and nyt.com charges for any article that exceeds a monthly allowance of 10 free articles. For more details on the New York Times' decision to introduce a paywall, see Kumar et al. 2013.

NYT initially offered all content for free, switched to a paid model with 20 free articles per month and later reduced this to 10 free articles per month. The Wall Street Journal initially required a subscription, later changed to a largely free version but then reverted to a partly paid model.

Interestingly, while the industry norm is to follow a static rule on how much content is free or paid (e.g., all paid, 10 free per month), firms 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. Thus, a key determinant of a paid content strategy is whether to opt for a static policy where the amount of free content remains the same over time or, alternatively, dynamically vary the amount of free versus paid content. In sum, it is not clear *whether* and, if so, *how* charging for online content can improve media companies' revenues.

In this research, we empirically examine *whether* and *how* firms should charge for access to online content. We first ask '*fee or free*' i.e. whether a firm benefits from charging for online content. Here, we examine a firm's average trade-off between subscription and online advertising revenues. We then turn to the question '*static or dynamic?*' that is whether the firm may benefit from dynamically adjusting the amount of paid content instead of holding constant how much they offer against a fee.

We build a unique data set from the sports website ESPN.com to study this question. ESPN.com offers the majority of content for free but charges a membership fee for a subset of articles. The number of paid articles varies by day and by sport. Via a web crawler, we collect data on the number of free and paid articles per day and sport over a thirteen months period. We complement this data with the number of unique visitors, page views and time spent for each type of article per day and sport.

We estimate how the number of free and paid articles affects viewership of the site, controlling for a wide range of demand shifters and possible endogeneity of the number of articles the firm offers on any day. This allows us to empirically quantify the marginal impact of an additional paid article on the increase in subscribers, and the decrease in page views. We find, on average, that indeed paid articles increase the number of visitors to the paid section, and thus subscribers, while reducing overall advertising impressions from page views on the site. Using these estimates we evaluate the monetary impact of adding an additional paid article. On average, our estimates suggest that the marginal increase in subscription revenue is statistically indistinguishable from the marginal decrease in advertising revenue, suggesting that, on average, the firm should not adjust the amount of paid content. However, we find strong differences over time when accounting for factors that exogenously vary consumer demand. Specifically, we allow our results to vary by whether a sport is off season, in regular season or in post season. We find that the marginal paid article increases total revenue in off season, while in regular season it strongly decreases total revenue. We attribute these differences to a variation in consumer valuation for sports news across seasons that lead to a change in the number and type of unique visitors to the site.

Our results suggest that a static policy that does not account for exogenous variation of demand may not be optimal. Instead, firms should consider flexibly adjusting the amount of paid content they offer. We estimate that a dynamic policy, where ESPN.com adjusts the number of paid articles by day and sport can substantially increase firm revenues. A large amount of this increase can be captured by a simplified policy where the firm adjusts the number of paid articles by sport and season instead of by day. Our findings suggests that many online content providers that currently use a static policy may benefit from re-adjusting their pricing strategy.

More broadly, our results illustrate how firms need to adjust traditional price discrimination strategies to digital environments. Media firms traditionally sell a diverse bundle of content to a heterogeneous group of consumers. Newspaper subscription pricing responds to this heterogeneity through temporal price variation, that is by varying prices depending on the number and type (weekday/weekend) of days a subscription covers[3]. Our results suggest that by using the detailed data available online, firms may be able to implement a similar but potentially more granular form of price discrimination online. After all, media firms might be able to leverage 'digital' to their advantage.

Relationship to Previous Literature

Analytical modeling has long been the prevalent technique used to analyze a firm's choice of 'free' versus 'fee'. Shapiro and Varian (1998) and Bhargava and Choudhary (2001) show that offering both a paid and a free component can allow firms to implement quality differentiation,

3 Most newspapers offer daily, weekday, weekend and Sunday subscription options (e.g. see https://nytimesathome.com/hd/237)

versioning or second-degree price discrimination. Godes et al. (2009) relate the trade-off between greater advertising and subscription revenues to the degree of competition in the market. They find that greater competitive intensity may increase profits from charging for content and decrease profits from advertising. Athey et al. (2011) show that offering paid content can lead to both a loss in visitors and to a positional disadvantage in advertising markets since advertisers are willing to pay a premium to firms with a high expected share of loyal consumers. When analyzing the effectiveness of a firm's advertising or pay-per-view strategy, Prasad et al. (2003) assume that consumers are heterogeneous in their willingness to pay to avoid ads. They find that in most cases the firm should combine pay-per-view and advertising revenues, rather than exclusively relying on either revenue stream. In sum, there is no consensus on whether and when firms benefit from charging for access to content and few insights on how specifically such strategies should be implemented.

An additional stream of analytical research has focused on free units as a sample of the paid product. Bawa and Shoemaker (2004) show that allowing consumers to sample content before a purchase can increase long-term sales. In addition, for digital goods, free samples, alongside high prices, can be used to signal superior quality (Boom 2010). But sampling enhances subscription demand only for intermediate levels of advertising effectiveness (Halbheer et al. 2013). Instead, under low advertising effectiveness, firms should offer only paid content and under high levels of advertising effectiveness it is optimal to offer all content for free.

As of yet, there is only sparse empirical evidence on whether a "fee" or a "free" strategy is most beneficial. Pauwels and Weiss (2007) show for an online content provider targeted towards marketing professionals that moving from free to fee can be profitable, despite loss of advertising revenue. Yet, Chiou and Tucker (2012) find that visits to an online news site fall significantly after the introduction of a paywall, particularly among younger consumers. But lacking detailed data on users' website activities, research to date has been unable to examine the trade-off between advertising and subscription revenues in detail.

To reduce the perceived costs of online content to consumers, firms have experimented with a variety of payment strategies, including micropayments where consumers pay only small amounts per article or visit to a site (Athey et al. 2011). However, even such strategies have not been successful. This can be linked to academic studies illustrating that consumers typically 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). Such preferences imply that consumers' utility loss from charging for content may be significantly higher than one would otherwise expect. Our research contributes to this area by empirically exploring the demand effect of charging a small positive versus a zero price for online content and identifying conditions under which consumers may be more willing to pay such prices.

In sum, while research so far has offered some broad guidelines on a content provider's choice of '*fee or free*', it does not provide conclusive evidence on whether and when a firm benefits from charging for online content. Specifically, findings as of yet provide little insight into the firm's choice of a '*static or dynamic*' strategy. Our research seeks to address these questions.

Data

Empirical setting

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

The ESPN website 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 empirical analysis, we focus on six different sports that typically offer both paid and free articles: College Basketball (CBA), College Football, Baseball (MLB), Basketball (NBA), Football (NFL) and Hockey (NHL). We abstract form sports such as NASCAR and tennis that did not offer paid articles during our observation period.

Website content and user activity

A typical challenge in analyzing the effectiveness of "free" versus "fee" strategies, is the difficulty to obtain data that discloses detailed usage information alongside pricing strategies (Pauwels and Weiss 2007) while also controlling for industry-wide demand. We circumvent this challenge by combining multiple data sets. Our data capture, for a period of thirteen 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. This means, that while we do not have user-level data, our data are disaggregate on the day-and-sport level. We next describe in more detail the different data sets we use.

Website content of ESPN.

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' homepages at ESPN from December 2010 to December 2011. As free articles, we collect all links with the url-format espn.go.com/sportname. As paid articles, we collect all links with the url-format insider.espn.go.com/sportname.[4] We then identify links that remain on a sport's homepage for a very 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 indicates, a sport's homepage displays 34 articles on average per day of which 25 are free and 9 paid.

We next explore the recency of articles. 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 have already been displayed the previous day. On any day, the average age of free

4 This metric abstracts away from content on other websites that the sport's homepage links to, such as Twitter, and blogs that come with a different url-format. These links are always free.

articles displayed is 11 days and the average age of paid articles is 7 days. This suggests that while the firm updates content over time updating happens gradually.

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 number of words) 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. While paid articles are on average longer, the standard deviation in article length is high and more so for free articles (Table 2). This is a result of a high number of very short free articles: 10% of free but no paid articles have less than 200 words. We compare all 274 paid articles to the top 274 free articles, by number of words, and find that in this subset free articles are on average longer. This suggests that both the paid and the free section feature detailed articles.

Lastly, we broadly look at the type of articles that are 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 news reports (e.g., interview with a coach). This makes sense since readers could easily substitute news articles by an article from a competing site whereas this is more difficult for editorial content or in-depth reporting.

User activity on ESPN

In our second data set, we obtain, for the same time period, daily data from Comscore on consumer activity by sport. This includes the number of unique visitors, the number of pages viewed and 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 that our data sometimes record zero visitors (mostly to the insider section or to one of the competing websites, see below) even though the true number for the US population is nonzero. Since these numbers are hard to interpret and since in our empirical estimation we take logs of the key variables, we exclude these 218 out of a total of 2,250 day-sport observations.

Table 1 reflects that significantly more individuals visit the free section 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. The time visitors spend per page is similar across paid and free articles.

User activity on competing sites

As a third data set, to control on a daily and per-sport level for industry-wide demand for sport news, we obtain from Comscore data per day and sport on website activities for the three competing sports websites sports.yahoo.com, cbssports.com and sportsillustrated.cnn.com (following Yahoo, CBS and SI)[5]. All three sites offer their content for free. The data include the number of unique visitors, the number of page views and total time spent per day and sport on all three websites. Table 3 documents that page views per visitor and time spent per visitor on competitive site are similar to those for free ESPN articles. It also illustrates that ESPN is the second most popular sports news site.

To further measure demand for ESPN news on a particular sport, we collect from 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.

Seasonalities

We next collect data on the seasons by sport and examine whether the demand for sport news varies by a sport's season. Each sport has three seasons. The off season is the period when no games are scheduled. Note that in the off season there are still sports news such as free agency signing and drafts, and scores for any pre-season games results of which 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 post season playoffs and a sport's final games are played

5 These are the top four sports website based on estimates by Google ad planner and Alexa.com. Google ad planner estimates the reach of ESPN.com is 12%, sports.yahoo.com is 18%, cbssports.com is 4% and cnnsi.com is 0.2%

(e.g., playoff in the professional sports MLB, NBA, NFL, NHL; the bowl season for college football and March madness in college basketball). Table 4 summarizes the key variables by season.

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 off season. Average demand is similar in regular and post season.

Lastly, we collect data on sport events as possible demand controls. This includes the number of games played in each sport on each individual day, the date of the final game within each sport, for professional sports, the dates of the draft and, for college sports, college signing day. We also collect the date of the NBA lockout in the 2011 season.

Subscription and advertising revenues

We next describe the subscription plans that ESPN offers. We then 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. Below we describe these calculations in detail. We will discuss the implications of this approach in Section 0.

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 $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. This suggests that 47% of customers choose the yearly plan, 35% choose the 2-year and 13% choose the monthly plan.[6] This gives us an average subscription revenue of $40.44 per year. Note that while 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.

We know that ESPN had 640,000 subscribers in 2011 (ESPN 2012) and, according to Comscore, a total of 55 million unique daily visitors[7] to the paid section. This means that each

6 Additionally, 4% signed up for a holiday offer in December 2010 and 1% for a trial in October 2011.

7 This number treats visits to each sport's homepage as separate unique daily visits.

subscriber returned 86.62 times a year to the paid section[8], or 7.22 per month; on average every 4.2 days. The effective price per day visited therefore amounts to $0.47.

ESPN features advertising on all webpages, including its homepage, the homepage for each sport and the page for each article. On each page it typically displays one ad, independently of whether an article is free or paid.[9] From Comscore we obtain estimates on ESPN's monthly advertising revenues as well as page views from December 2010 to December 2011.[10] We use this data to compute the monthly price per 1000 impressions. On average, ESPN's revenue per 1000 impressions is $11.51. Prices vary over time with a minimum of $8.34 (in April) and a maximum of $15.45 (in December 2011). We were able to verify the average, minimum and maximum advertising prices with ESPN.

8 This estimate is across sports and days. Explicitly this means we have on average 86.62 day-sport visits for each subscriber. For example, visiting the NFL and NBA sport pages on the same day will count as two separate visits.

9 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. It is likely that revenue from sponsored links is negligible, so we do not include sponsored links. Comscore also does not provide estimates for revenues from sponsored links which further suggests that such revenues are negligible.

10 Comscore estimates are based on projected ad spend costs. This means the advertising revenue they report approximates the net advertising cost not the gross cost that is quoted on ratecards and often substantially higher. The data is predominantly inputted by agencies and so it reflects the actual payments to ESPN rather than gross pay-outs by advertisers that may include costs for agency services.

Results

Consumer response to paid articles

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, since our data is aggregate by day and sport, we observe variations in behavior both within and across time. Unlike previous research, we are therefore in the unique position that we can estimate the effect of paid content on both subscription revenues through the analysis of visits to the paid section of the site, and on advertising revenues through the analysis of page views on the site while controlling for variation in industry-wide demand.

Empirically, our objective is to generate insights into a firm's ability to extract revenues from paid content, the trade off between subscription and advertising revenues and, importantly, how this ability may vary over time. This means we aim to focus on the aggregate effect of paid articles on consumer behavior and ultimately revenues. Before laying out in more detail our empirical approach, we discuss the underlying conceptual model of consumers' response to paid articles, and the resultant implications for the firm's subscription and advertising revenues. Note that the lack of data that track consumers' individual-level visitation and viewing behavior across subscribers and non-subscribers to the free and the paid section means that we cannot identify the full behavioral model. Instead, we use our conceptual model as the basis for our reduced-form estimation approach and acknowledge that a more detailed modeling of consumer behavior alongside individual-level data could provide further insights into consumer response to paid content.

Our work rests on the primitive that consumers derive utility from articles the firm posts on its website.[11] We assume that consumers are heterogeneous in their valuation of online sports content. This heterogeneity has a horizontal 'taste' component (what information the consumer values) and a vertical 'value' component (how much the consumer values the information).

---

[11] We assume that consumers have an outside option that they can trade off against reading articles

Every day, consumers make two decisions for every available sport: (1) Should they visit the free section and the paid section of the site, where the latter implies the decision to subscribe? (2) Which, and therefore how many, articles should they view in each section? Both decisions depend on the number of free and paid articles the firm posts on each sport's website. These decisions have a direct relationship with firm revenues. Subscription revenue comes from consumers visiting the paid section of the site (decision 1) and advertising revenue comes from the total articles viewed (depends on both decision 1 and 2). We structure our subsequent discussion around these two economically relevant outcome variables for a firm and discuss the decision to visit the site as it directly relates to these two outcomes.

Impact on total unique visitors to the paid section: A consumer's decision to subscribe is based on the utility they derive from paid articles. Since consumers pay a lump-sum to visit the paid section, 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 pro-rated subscription fee.[12]

As the firm adds a paid article, the consumer's expected utility from subscribing to the paid section marginally increases. As a result, the consumer becomes more likely to subscribe. It is this marginal impact of a paid article on unique visitors to a sport on a day that we focus on in our estimation.

There are two behavioral mechanisms by which a consumer's utility from subscribing may increase and which would lead to a positive relationship between the number of paid articles and unique visitors to the paid section.[13] 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. 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. The expected utility from subscription increases as the probability of finding a paid article that fits increases, which in turn increases with the

---

[12] The consumer signs a contract over an extended time period (mostly one year). Therefore, the consumer's total discounted utility for the contract is the discounted sum of utilities per day minus the fee for the contract. As such, a subscription decision relies on the expected number of paid articles the firm offers for the period of the contract. This expectation may be based on past experience or foresight into the amount of articles offered in the future. We do not model this expectation in greater detail.

[13] There are several ways by which consumers can learn about the availability of paid articles for a sport. The headline of a paid article may be featured on the firm's homepage or on the sport's homepage. Also, link to an article reported by search engines such as Google indicates whether an article is paid.

number of paid articles offered. In our empirical analysis, we focus on how the number of paid articles impacts the number of unique paying visitors and so jointly capture the effect of both mechanisms on subscription revenues.

Impact on total page views across both sections: Advertising revenues for the firm generally increase with page views as each page view allows the firm to display more ads. There are two potential ways by which an increase in paid articles may affect total page views. First, paid articles affect the number of unique visitors to the site. While subscribers are more likely to visit, the utility from visiting for non-subscribers may decrease when only less appealing content is offered for free. As a result, they may visit the site less often. Since 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.

Second, paid articles may reduce page views per visiting non-subscriber as the firm offers little appealing content for free. Alternatively, they may be upset about the sheer number (or share) of paid articles and as a result of reactance view less pages. Put differently, non-subscribers may have a disutility from paid articles that increases in the number of paid articles the firm offers.

Since we do not have individual-level data that allows us to disentangle in detail the underlying behavioral mechanism we are unable to quantify how paid articles affect the number of subscribers versus non-subscribers that visit the free section of the site or the number of page views per unique visitor for each customer type. Instead, we focus on the total effect on page views as the relevant outcome variable for the firm.[14] In Section 0, however, we will use an aggregate analysis of total unique visitors to the site to better understand how a change in unique visitors contributes to the total decline in page views, relative to a change in page views per visitor.

Next, we estimate the effect of paid articles and quantify the trade off between the positive effect on visitors to the paid section and the negative effect on page views.

---

14 We acknowledge that there may be other long-term effects of paid articles 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. In our analysis we abstract from such possibilities.

Average effect of paid articles

Effect on visitors to paid articles: We pool data across sports and days. We take logs of the dependent variable and estimate a linear regression. Our first key dependent variable is the number of unique visitors to the paid section of sport *i* on day *t*, $UniqVisPaid_{it}$:

$$\begin{aligned}\ln\left(UniqVisPaid_{it}\right) = {} & \beta_1 PaidArt_{it} + \beta_2 FreeArt_{it} + \beta_3 PercGame_{it} + \beta_4 GameDay_{it} + \\ & \beta_5 Draft_{it} + \beta_6 Lockout_{it} + \beta_7 FinalGame_{it} + \beta_8 NonworkingDay_t + \beta_9 Google_{it} + \\ & \beta_{10j} \ln\left(UniqVisComp_{ijt}\right) + \beta_{11} RegSeason_{it} + \beta_{12} PostSeason_{it} + \beta_{13} Sport_i + \varepsilon_{it}\end{aligned} \tag{1}$$

where $PaidArt_{it}$ represents the number of paid and $FreeArt_{it}$ the number of free articles for sport *i* on day *t*.

We include a number of control variables. $PercGame_{it}$ represents the percentage of all games in sport *i* that is being played on day *t*. $GameDay_{it}$ captures whether any games are played in sport *i* on day *t*. $Draft_{it}$ controls for whether, during off season, there is a draft of players in a sport on a particular day (for college sports, it captures the national sign-up day) and $Lockout_{it}$ controls for the NBA lockout during the 2011/12 season. The variable $FinalGame_{it}$ controls for whether on a day there was the final game in a sport. $NonworkingDay_t$ captures whether day *t* is a weekend day or a public holiday. We further control for demand shocks that may not be captured by our controls so far but similarly affect all firms in the market using as controls the number of visitors to a competing site *j*, including Yahoo, CBS and SI, for sport *i* on day *t*. We additionally use data captured from Google Trends to control for demand shocks that may be unique to the focal firm, ESPN. $Google_{it}$ measures the number of Google searches for 'ESPN + sport' scaled between 0 and 100. To further account for a possible shift in demand by season, we include dummies for whether a sport is in regular season or post season. Lastly, we include fixed effects by sport.

Column (1) in Table 5 displays the results. It confirms that when the firm offers more paid articles, consumers are more likely to visit the paid section of the site. Specifically, after controlling for demand, increasing the number of paid articles by one, increases the number of unique paying visitors by 5%. The estimates suggest that, at the mean of 28,309 unique visitors to the paid section, an additional paid article increases viewership in the paid section by 1,552 customers, implying greater subscription revenues. Offering more free articles also increases the number of unique visitors to paid content.

A key strength of our OLS specification is that it controls for a wide range of demand shifters. For example, increased demand on days that new players are drafted into the teams would be captured by the variable $Draft_{it}$. Or, if people tend to watch more TV and consume less online sport news on weekends, then this should be captured by the dummy $NonworkingDay_{t}$. Alternatively, if there is a piece of unexpected sport news that generally affects demand, such an effect would be captured by the activity on competing sites whereas shocks that would affect the attractiveness of the ESPN site only should be captured by our variable $Google_{it}$ derived from Google Trends.

However, even after controlling for a wide range of observable demand shocks there could possibly be demand shocks that ESPN observes but not the researcher. ESPN could then use this information when deciding on the number of free or paid articles on that day. An example could be a breaking news story that is unique to ESPN (e.g. ESPN signing a new Monday Night Football deal with the NFL on September 8, 2011). Note that anecdotal evidence suggests that rather than knowing the revenue-optimizing paywall, firms experiment with respect to their paid content strategy.[15] Nonetheless, we turn to an instrumental variable estimation to control for such possible endogeneity. We use as an instrument the number of free and paid articles that ESPN displayed the previous day that is the day before such news were known.

Our 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. On average across days and sports, 75% of all paid and 61% of all free articles were retained from the previous day. Displaying an article for more than a single day makes sense as long as potential readers do not visit 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

[15] 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 as of yet are not necessarily aware of the optimal paid content strategy.

to display an existing article is attractive for the firm since it incurs zero marginal cost of production on the second day.

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

Using as instrument for the number of paid and free articles on any day the number of paid and free articles displayed the previous day 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 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.[16] Note that our instrumenting strategy also assumes that the firm is myopic, meaning that the decision whether or not to publish an article may be affected by the expected demand on the same day but not by expected demand on the following day.

We estimate our model using the two-step efficient GMM estimator. The results in Column (2) indicate that the number of free and paid articles is endogenous in this regression. An F-test of the significance of excluded instruments strongly rejects zero. The Kleibergen-Paap rk statistic suggests that we can reject the hypothesis that the first stage is underidentified. Our instruments are therefore good predictors of the number of free and paid articles.

The effect of paid articles on visitors to the paid section is slightly smaller in size with a larger standard error than in Column (1) but our main effect still holds. Specifically, the estimate now suggests that, on average, an additional paid article increases viewership in the paid section

[16] 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. In an alternative specification we include as additional control unique visitors to the paid section the previous day, respectively page views the previous day. Our results hold, though the standard deviation of the regular season estimate in Table 10 increases.

by 1,194 unique visitors. In a further unreported instrumental variable specification we find that our results are robust to including as additional instrument the age of the previous day's free and paid articles. This likewise applies to all other results represented in Sections 0 and 0.

We test the robustness of our results allowing for a correlated error structure across both the estimation of unique visitors and page views using 3SLS. Column (3) presents the set of results pertaining to the effect of articles on unique visitors to the paid section. Again, our results hold.

Effect on total page views: We similarly evaluate the effect of paid articles on all page views within the firm's website:

$$\begin{aligned}\ln\left(PageViewsOwn_{it}\right) = \beta_1 PaidArt_{it} + \beta_2 FreeArt_{it} + \beta_3 PercGame_{it} + \beta_4 GameDay_{it} + \\ \beta_5 Draft_{it} + \beta_6 Lockout_{it} + \beta_7 FinalGame_{it} + \beta_8 NonworkingDay_t + \beta_9 Google_{it} + \\ \beta_{10j} \ln\left(PageViewsComp_{ijt}\right) + \beta_{11} RegSeason_{it} + \beta_{12} PostSeason_{it} + \beta_{13} Sport_i + \varepsilon_{it}\end{aligned} \tag{2}$$

Here, $PageViewsOwn_{it}$ represents the total number of page views on the firm's website and $PageViewsComp_{ijt}$ represents the number of page views on each of the three competitor's websites. We include similar controls as in the previous regression.

Table 6 summarizes the results. Column (1) illustrates that while increasing paid articles by one, *reduces* the overall number of page views by 0.6%, this effect is not statistically different from zero.

We turn to an instrumental variable estimation using the same set of instruments as in the previous section that is the number of paid and free articles on the previous day. Similarly to before, the identifying assumption is that after controlling for a wide range of demand shifters, the number of yesterday's free and paid articles affect today's page views only through the number of free and paid articles the firm displays today.

The endogeneity test in Column (2) rejects the null hypothesis that the specified endogenous regressors can be treated as exogenous. Similarly to before, the F-test of the significance of excluded instruments strongly rejects zero and the Kleibergen-Paap rk statistic again suggests that we can reject the hypothesis that the first stage is underidentified. The IV estimates indicate a marginally significant negative effect of paid articles on page views. Specifically, they show that increasing paid articles by one, reduces the overall number of page views by 1.0%. Additionally, Column (3) includes the corresponding set of results from a 3SLS estimation.

Again, our results hold. As before, our results are robust to non Gaussian errors with a quantile regression specification. In sum, our results confirm that paid articles decrease advertising revenues.

Effect on the firm's revenue: We evaluate whether offering an additional article is profitable for the firm. We use several data points in addition to our parameter estimates. First, we use the fact that alongside each article the firm displays on average one ad, meaning that each page view leads to one ad impression. Second, as laid out in Section 0 we compute the weighted average of subscription fees across plans as $3.37 per month. Third, as illustrated in Section 0, we rely on the fact that $0.47 revenue can be attributed, on average, to a visit to the paid section.

Note that this analysis translates unique visitors into subscribers holding constant unique monthly visits per subscriber and so the implicit price per visit. We acknowledge that 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 0 will examine this possibility.

We compute the minimum advertising revenue per thousand impressions that will compensate foregone revenues for the firm from not offering an additional insider article. We draw from the distribution of the two main coefficients, meaning we draw the effect of paid articles on unique visitors to the paid section and the effect of paid articles on page views. Multiplying the increase in visitors with the revenue per visit, that is $0.47, gives us the additional revenue the firm could earn from offering one more insider article. In our discussion, we focus on the results of our instrumental variable specification.

On average, an additional paid article increases visitors to the paid section by 4%. This means, it increases subscription revenues by $557. The decrease in page views from offering an additional insider article gives us the number of ad impressions that the firm looses by offering an additional paid article (since each page displays on average one ad). At the average effect of an article this is a loss of 36,346 page views and ad impressions. This suggests that adding an additional paid article would be profitable if the firm would earn less than $15.3 per 1,000 impressions. Only then, the added subscription revenue would exceed the potential advertising revenue. We find a similar value, if instead of evaluating the effect at the mean, we evaluate the effect at randomly drawn day-sport observations.

As discussed in the data-section, the firm currently charges an average price of $11.51 per 1000 impressions which is close to the break-even value we computed. Therefore, it is not clear whether the firm should, on average, add an additional paid article. This would only be profitable if subscription revenues for paid articles (on average $15.3) would not compensate for the loss in advertising revenue ($11.51).[17]

To understand the uncertainty in the estimate of $15.3, we bootstrap from the estimated asymptotic distribution of the impact of insider articles on insider visitors (Table 5) and page views (Table 6). We then compute the minimum advertising revenue per thousand impressions that compensates foregone revenues for the firm from not offering an additional insider article for each draw. The bootstrapping results show that 69% of observations lie above $11.51, the value at which it is advisable to add another paid article. Figure 2 displays the distribution of the dollar values we obtain from bootstrapping. The red dashed line represents the average revenue per 1000 impressions of $11.51. We similarly evaluate whether the firm should add paid articles at the price of $8.34 and of $15.45 per 1000 impressions which we identified as the minimum and the maximum monthly ad prices during our observation period. Table 7 again shows that the large majority of observations lie very close to the respective cut-off value, suggesting that the firm would not benefit from adding or removing paid articles. In Figure 2 these further cut-off values are represented by red dotted lines. The results are robust to results obtained from alternative specifications where in our regression the percentage of paid articles, instead of the number of paid and free articles, enters as endogenous variable.

In sum, these results demonstrate that, on average, the firm should not change the number of paid articles displayed. Table 7 shows that the OLS results are higher in absolute dollar value and in the percentage of observations above the $11.51 threshold. By contrast, the results obtained with our 3SLS specification are close to those obtained in the IV estimation.

17 We assume that the firm charges the same price for a page view in the free and the paid section. We cannot conclusively rule out that ESPN charges a higher price for advertising in the paid section. But since ESPN already offers access to a highly targeted audience it is unlikely they would charge a significant premium for access to subscribers. Note also that an advertiser's willingness to pay is often lower for smaller audiences (Athey et al. 2011), suggesting further that the firm would not be able to charge a premium to advertisers in the paid section.

Exogenous variation of demand

We next ask whether there are conditions under which the aggregate effect of adding an additional paid article is positive. We propose that the value of news to a consumer varies exogenously with a sports' season. In the context of sports news, the importance of events, specifically games played, affects the value of news to consumers. In off season, no games are played. The most important games take place in post season and lead to higher value news than regular season games.

We suggested earlier that consumers' valuation of news affects their response to paid articles. Such valuations are likely to vary across seasons. Different valuations of news are likely to both affect a consumer's willingness to pay for sport news and the degree of negative response to paid content on a web page. We next empirically tease apart the effect of paid articles on unique visitors to the paid section and on page views by season. We then evaluate their impact on the firm's revenues.

Effect on visitors to paid articles: We estimate a similar model as above but allow the effect of free and paid articles to vary by season. Table 8 presents the results. Column (1) suggests that indeed the positive effect of paid articles is greatest during post season.

In Column (2) we again display an IV specification, using as instrument the number of free and paid articles by season on the previous day. Again, the null hypothesis of exogeneity of the endogenous regressors is rejected. The remaining statistics are similar to those previously reported. The IV estimates confirm that the effect of paid articles on unique visitors is greatest during post season. Column (3) reports the results of a 3SLS estimation pertaining to the effect of articles on unique visitors by season. Again, they confirm the robustness of our estimates.

Effect on total page views: We similarly evaluate the effect on page views, using both OLS and an IV specification. Table 9 displays the results, including similar statistics as previously reported. Importantly, the IV estimates indicate that the negative effect of paid articles on page views is lowest during off season, suggesting that consumers' response to paid articles varies significantly across seasons. Column (3) reports the results of a 3SLS estimation pertaining to the effect of articles on unique visitors by season. Again, they confirm the robustness of our estimates.

Effect on the firm's revenue: We compute the effect on the firm's revenue separately for off season, regular season and post season. Table 10 summarizes the results. For completion we report both the OLS and the IV estimates but focus in our discussion on the IV estimates.

Both the estimated dollar values and the percentage of observations above the cut-off values indicate that in off season the firm should increase the number of paid articles. This is a result of a positive effect of paid articles on unique visitors during off season (Table 8) alongside the lack of a significant negative effect of paid articles on page views at that time (Table 9). By contrast, the marginal effect of a paid article is negative during regular season, as indicated by the significant amount of observations below the cut-off values. This means that the firm should decrease the number of paid articles during regular season. The first part of Figure 3 displaying the density of the estimated break-even dollar value graphically supports this finding as the large majority of observations lie to the left of the red line.

Interestingly, this negative effect disappears when we turn to the post season. Figure 3 further illustrates that in post season a large proportion of observations indeed lie to the right of the break-even values. We conclude that during the post season the firm should not change the number of paid articles from its current value. As before, we find similar results for a 3SLS specification. Likewise, the results are robust to results obtained from alternative specifications where in our regression the percentage of paid articles enters as endogenous variable.

Discussion

Conceptualization: At first blush, our result that the firm should increase the number of paid articles during off season, decrease during regular season and keep at its current level during post season seems surprising. To provide conceptual support for these findings we turn to our raw data.

Figure 4 illustrates the number of unique visitors to the free and paid sections across seasons. We find that in going from off season to regular season, the number of unique visitors to the free section increases while unique visitors to the paid section remain approximately at the same level. This result suggests an influx of additional visitors in regular season that are, however, not signing up for the paid section of the site. But when going from regular to post season, visitors to the free section remain approximately constant while the number of visitors to the paid section increases: free visitors are now converting to paying visitors.

The descriptive results suggest that, on average, off-season customers are different from customers that only join during regular season. We propose that customers visiting in off season have a relatively high valuation of sport news – they are interested in sport news even when comparatively little happens. They are relatively likely to subscribe to paid content and so we observe little negative effect of paid content. During regular season, there is an influx of low-valuation customers who are not willing to pay for content, or at least not for the type of paid content the firm offers, and who respond negatively to paid articles. This means that during a period of high demand the firm should offer more content for free. Such a policy is consistent with countercyclical pricing observed for consumer packaged goods (Chevalier et al. 2003; Nevo and Hatzitaskos 2006; Guler et al. 2013).

However, when going from regular to post season, free visitors are being converted to paying visitors. This observation is in line with the intuition that news during post season, when the season's winning team will be determined, are particularly highly valued. It suggests that the valuation of previously low-value customers increases and that in going from regular to post season the firm should adopt a regular, cyclical pricing policy.

We next use our data to better understand what causes the drop in page views. This links back to our discussion of the behavioral mechanisms in Section 0. We estimate the effect of paid articles on page views separately for the paid and the free section using the same estimation framework as before. As is to be expected, the negative effect on page views is entirely due to a reduction of page views of free articles while adding paid articles increases page views in the paid section.

We then aim to disentangle the two factors that may cause the drop in page views: the reduction in unique visitors or in page views per visitor. In an instrumental variables specification similar to before, we estimate the impact of paid articles on total unique visitors for each season, using as controls the same variables as in our main specifications. We find that during regular season, 57.8% of the negative effect of paid articles on page views can be attributed to a reduction in unique visitors and 42.2% to a reduction in page views per visitor. This relative effect reverses during post season where only 27.4% of the reduction can be attributed to a change in unique visitors. The finding that the relative effect on unique visitors is greater during regular than during post season is consistent with our analysis of the raw data: regular-season visitors have, on average, a lower willingness to pay than post-season visitors and are therefore more likely to respond negatively to paid articles than post-season visitors.

Further robustness check: Throughout our empirical analysis we translate unique visitors into subscribers holding constant the number of unique monthly visits. It is, however, possible that the increase in unique visitors we observe as a result of an increase in paid articles is not 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 focus on unique monthly visitors. The data presented in Figure 4 suggest that the number of unique visitors on any day is greatest during post season. If this effect is indeed due to an increase in subscribers, then the result should hold when using more aggregate monthly data. We use data on unique monthly visitors to the paid section by sport. Here, an individual is counted once per month and sport, independently of how often they visit the site during that month. In a linear regression, we regress season dummies on monthly unique visitors by sport controlling for sports. We find a positive and significant effect of post season (p=0.002) on unique monthly visitors and insignificant effects for off season and regular season. This confirms that variation in the number of unique daily visitors that we observe across seasons translates into a variation of monthly visitors and so most likely can be traced back to an increase in subscribers.

Second, we use data on a consumer's average number of visits per sport and month. We estimate a linear regression with the average visit frequency as dependent and unique monthly visitors per sport as independent variable, controlling for sports. Our results suggest that visit frequency does not vary with unique monthly visitors (coefficient -0.000019, p=0.889). This result further supports that the variation in unique daily visitors we observe indeed comes from an increase in subscribers.

Implications for firm policy: We assess the implications of our findings for pricing policies. Given the nature of our instrumental variable estimation, we cannot easily estimate the profit-maximizing number of articles. Instead we include a back-of-the-envelope calculation to roughly assess the revenue impact for the firm if, instead of using its current pricing policy, it adjusted the number of paid articles by day and sport.[18]

We find that a dynamic solution where the firm chooses the optimal number of articles by sport and day can increase firm revenues by 7% compared to current revenues. Note, however, that it may be operationally difficult for a firm to optimize the number of paid articles by sport and day. We therefore consider an alternate policy where the firm optimizes the number of paid articles by sport and season, instead of on a daily level. We find that this would achieve 93% of the revenue benefit.

By-season versus by-sport variation: We further check our finding that the firm should adjust its policy to vary the number of paid articles across seasons, by regressing the optimal number of daily paid articles against only the estimated season fixed effects. We find a R^2 of 0.71. When alternatively regressing the optimal number of paid articles against only sport fixed effects, the R^2 is 0.12. This result suggests that the firm would indeed benefit from varying the number of paid articles across seasons rather than across sports.

The model specification that we used to obtain the estimates that entered above interacts the number of articles with seasons but not with sports. To check that our results do not merely reflect this specification, we reestimate our regressions interacting the number of paid and free articles with sports instead of seasons. We then use these estimates to compute the optimal number of articles and next regress the optimal number of articles on either season or sport fixed effects. Again, we find that season fixed effects explain a greater share of the variation in the optimal number of articles than sport fixed effects (R^2 0.33 vs 0.13).

We turn to the data to evaluate whether firm policy during our observation period reflects a variation across seasons or sports. We regress the actual number of articles against season dummies and then against sport dummies. We find that season fixed effects explain 0.35% of the variation in the number of actual articles whereas sport fixed effects explain 50%.

Lastly, we optimize revenues using the estimation results obtained in the model including by-sport (instead of by-season) interactions. Here, the projected revenue increase is 2% -

[18] We use the second-stage estimates of the log-linear IV model to simulate the optimal number of paid articles by day and sport, keeping the number of free articles as currently in our data. We account for the uncertainty of our estimates by bootstrapping from the asymptotic distributions of our parameter estimates. We allow the number of paid articles per sport and day to vary between 5 and 13 that is the mean that we observe in our data +/- 1 standard deviation.

significantly smaller than the 7% revenue increase we obtained in the optimization using sport fixed effects. These results further suggest that the firm would benefit from changing from a policy where paid content tends to vary across sports to a policy where paid content varies across seasons.

Conclusion

The last decade has seen many media companies, such as newspapers, struggle and it is generally believed that their future hinges on their ability to implement a sustainable revenue model online. However, solving the basic trade-off, offering free content and maximizing advertising revenues versus offering paid content at the cost of advertising revenues, is not obvious.

In this research, we ask *whether* and, if so, *how* firms should charge for online content and specifically whether firms benefit from dynamically varying the amount of paid content over time.

We build a unique data set from the sports website of ESPN where we combine data on the number of free and paid articles offered per sport over time with different metrics of consumer demand including both unique visitors and page views, for both types of articles. We also control for industry-wide demand by tracking usage at the major competitive websites. 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 the number of subscribers, and the decrease in total page views and evaluate whether the company would benefit from adding an additional article.

Our results suggest that on average the marginal paid article decreases firm revenues. However, we find strong differences when we allow our results to vary by an indicator of exogenous demand variation. Specifically, we evaluate the revenue effect of paid content by whether a particular sport is off season, in regular season or in post season. We find that the marginal paid article increases revenues in the off season but decreases revenue in regular season. We propose that this is due to the variation in the value of sport news across seasons. We suggest that a simple conceptual model can rationalize our empirical results: customers visiting in off season have a relatively high valuation of sport news and are willing to pay for content. In regular season, there is an influx of customers with a lower valuation. As a result, paid content is ineffective and the firm should turn to a policy akin to countercyclical pricing. By contrast, post season sees increasing valuation by even customers with a low valuation. Hence, the firm benefits from offering more paid content relative to the regular season.

Our results have important managerial implications. They suggest that in evaluating whether to charge for content firms should account for heterogeneity in consumer demand. This means that firms benefit from using a dynamic strategy where they flexibly adjust the amount of paid content over time rather than always offering the same amount of content for free or against a fee. One recent example illustrates how firms can flexibly respond to changes in demand. Immediately following the 2013 bombings in Boston the Boston Globe temporarily lifted its paywall. Our results suggest that the site may have concluded that additional advertising revenues from the sudden influx of low-valuation customers (who would not pay a subscription) may outweigh loss of subscription revenues. Our result that a dynamic policy may allow firms to significantly increase their revenues are especially interesting in light of the fact that a number of firms have recently started experimenting with fee-models but have, unlike the Boston Globe, rarely explored flexibly adjusting the amount of paid content.

More generally, our work has implications for firms using a 'freemium' pricing policy with two sources of revenue, one that is related to activities by all users (or viewers, e.g. advertising), and one that comes from the subset of consumers who purchase (or subscribe). We suggest that in identifying the optimal point at which a 'free' offering becomes 'premium', firms should carefully consider customer heterogeneity in demand, identify dimensions along which such heterogeneity varies (e.g., time) and then use these insights to design price discrimination strategies.

Of course there are limitations to our work. Our study focuses on the immediate, short-term effects from offering paid content. There may be additional, long-term effects that we are not able to account for. Further, our study is set in an industry where many firms (still) offer all content for free. It is possible that in settings were all or most competitors charge for access to their content a subscription model may more generally appear to be optimal. This would then raise a new set of questions, such as how consumers trade off between fee-paying online sites and other media, e.g. cable or satellite television. We leave such questions for future research.

Bibliography

Ascarza, Eva, Anja Lambrecht, and Naufel 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.

Boom, Anette (2010), ""Download for Free" - When Do Providers of Digital Goods Offer Free Samples?," Working paper.

Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi (2003), "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review* 93 (1).

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

Edgecliffe-Johnson, Andrew (2009), "Media wants to break free," *Financial Times*.

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/).

Farhi, Paul (2013), "Washington Post to be sold to Jeff Bezos, the founder of Amazon," *Washington Post*.

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.

Gourville, John T. (1998), "Pennies-a-Day: the Effect of Temporal Reframing on Transaction Evaluation," *Journal of Consumer Research* 24 (March), 395-408.

Guler, Umut, Kanishka Misra, and Naufel Vilcassim (2013), "Guler, Misra and Vilcassim (2013) Countercyclical Prices: A Consumer Heterogeneity Explanation," Working paper.

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.

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

Nevo, Aviv and Konstantinos Hatzitaskos (2006), "Why Does the Average Price of Tuna Fall During High Demand Periods?," NBER Working Paper No. W11572.

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).

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.

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

	Mean	Std. Dev.	Min	Max
Articles				
All	33.8	17.8	8.0	121.0
Free	24.8	17.3	4.0	113.0
Paid	9.0	3.8	1.0	21.0
Unique Visitors				
All pages	649,082	505,135	27,429	3,786,197
Paid pages	28,307	42,865	139	1,015,253
Page views				
All pages	3,666,901	3,876,976	79,314	43,372,082
Free pages	3,598,434	3,855,771	75,318	43,295,967
Paid pages	68,467	159,403	162	3,909,487
Page views per unique visitor				
All pages	5.2	2.4	1.6	30.2
Paid pages	2.1	1.7	1.0	35.8
Time spent (min)				
All pages	3,586,429	3,994,947	65,785	49,637,376
Free pages	3,531,348	3,977,205	58,476	49,587,368
Paid pages	55,081	137,296	18	3,923,605
Time spent per page (min)				
All pages	0.9	0.3	0.3	2.9
Free pages	0.9	0.3	0.3	3.2
Paid pages	0.8	0.5	0.0	8.1

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: Activity on Competitive Sports News Sites

	Mean	Std. Dev.	Min	Max
Unique visitors				
Yahoo	879,119	989,635	12,274	10,276,782
CBS	81,697	104,332	766	1,421,150
SI	78,299	70,473	210	670,433
Page views				
Yahoo	4,453,604	5,248,149	33,312	49,977,896
CBS	467,218	943,677	702	14,878,230
SI	302,938	601,566	277	15,921,859
Page views per visitor				
Yahoo	5.5	3.3	1.5	32
CBS	4.4	3.8	0.9	62
SI	3.5	5.5	0.8	191
Time spent				
Yahoo	5,073,535	6,408,316	20,166	57,668,366
CBS	535,442	1,286,355	53	18,219,379
SI	265,243	482,941	125	12,496,384
Time spent per page				
Yahoo	6.2	5.3	0.4	7.6
CBS	4.6	5.5	0.1	4.9
SI	3.0	5.2	0.1	4.3

N=2027

Table 4: Data by Seasonality

	Off season		Regular season		Post season	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Number of articles						
All	25.6	9.5	40.9	19.6	40.1	22.2
Free	16.6	8.0	31.8	19.5	31.6	21.2
Paid	8.9	4.3	9.1	3.4	8.5	3.1
Unique visitors						
ESPN all pages	362,745	283,342	899,457	543,171	856,649	435,426
ESPN paid pages	30,061	56,284	24,596	22,833	36,350	37,476
Yahoo	534,209	752,875	1,163,516	1,103,330	1,202,714	910,793
CB	41,996	44,848	109,353	97,901	140,674	206,154
SI	52,854	58,553	100,184	74,023	98,294	65,880
Page views						
ESPN all pages	1,523,344	1,666,011	5,765,791	4,514,094	4,259,936	2,691,224
ESPN free pages	1,438,461	1,552,429	5,719,445	4,499,762	4,170,139	2,671,714
ESPN paid pages	84,883	221,250	46,347	52,417	89,797	116,907
Yahoo	1,935,053	2,669,757	6,835,588	6,282,232	5,510,175	3,760,665
CB	152,585	323,375	728,851	1,108,320	752,990	1,466,642
SI	183,766	583,253	406,233	383,600	393,220	1,125,635
N	932		891		209	

Table 5: Effect on Unique Visitors to the Paid Section

	(1) OLS+			(2) IV+			(3) 3SLS		
	Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
Free Articles	0.007	0.002	***	0.014	0.003	***	0.014	0.002	***
Paid Articles	0.053	0.009	***	0.041	0.012	***	0.040	0.009	***
PercGames	-0.039	0.039		-0.009	0.040		-0.007	0.044	
Gameday	-0.030	0.081		-0.110	0.087		-0.117	0.081	
Draft	0.988	0.312	***	1.048	0.358	***	1.047	0.304	***
Lockout	-0.579	0.157	***	-0.521	0.167	***	-0.515	0.117	***
Final game	-0.124	0.226		-0.221	0.230		-0.228	0.426	
Nonworkingday	-0.286	0.050	***	-0.302	0.051	***	-0.307	0.049	***
Googlescaled	0.016	0.004		0.015	0.004		0.015	0.003	***
ln(Ya Uniqvisitors)	0.113	0.034	***	0.099	0.035	***	0.097	0.031	***
ln(CB Uniqvisitors)	0.003	0.039		0.005	0.039		0.000	0.035	
ln(SI uniqvisitors)	0.100	0.031	***	0.107	0.031	***	0.102	0.029	***
Reg Season	-0.366	0.079	***	-0.405	0.083	***	-0.406	0.074	***
Post Season	-0.149	0.109		-0.160	0.110		-0.159	0.095	*
R-2		0.1597			0.1512			0.4157	
N		2032			2007			2007	
Endogeneity test ($\chi 2$)					11.022	***			
Kleibergen-Paap rk LM statistic ($\chi 2$)					76.267	***			
Significance of first-stage regressions				significant at 0.001					
F-test of significance of excluded instruments				significant at 0.001					
First stage R-2					0.62 - 0.72				

Fixed effects by sport included but not displayed for readability.

+Standard errors corrected for heteroskedasticity and autocorrelation using Newey-West.

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

Table 6: Effect on Page Views

	(1) OLS+			(2) IV+			(3) 3SLS		
	Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
Free Articles	0.001	0.001		0.003	0.001	***	0.002	0.001	**
Paid Articles	-0.006	0.004		-0.010	0.005	*	-0.010	0.004	**
PercGames	-0.072	0.016	***	-0.064	0.016	***	-0.065	0.018	***
Gameday	0.111	0.037	***	0.105	0.037	***	0.110	0.034	***
Draft	-0.032	0.140		0.000	0.152		0.001	0.128	
Lockout	0.233	0.062	***	0.248	0.065	***	0.244	0.049	***
Final game	-0.397	0.192	**	-0.429	0.182	**	-0.426	0.180	**
Nonworkingday	-0.209	0.023	***	-0.217	0.023	***	-0.216	0.021	***
Googlescaled	0.027	0.002		0.027	0.002		0.027	0.001	***
ln(Ya Pageviews)	0.172	0.019	***	0.160	0.020	***	0.160	0.016	***
ln(CB Pageviews)	0.081	0.014	***	0.077	0.014	***	0.077	0.012	***
ln(SI Pageviews)	0.149	0.013	***	0.152	0.013	***	0.150	0.010	***
Reg Season	0.449	0.039	***	0.438	0.039	***	0.441	0.031	***
Post Season	0.184	0.049	***	0.179	0.049	***	0.183	0.040	***
R-2		0.8107			0.8102			0.8585	
N		2032			2007			2007	
Endogeneity test ($\chi 2$)					10.459	***			
Kleibergen-Paap rk LM statistic ($\chi 2$)					71.399	***			
Significance of first-stage regressions				significant at 0.001					
F-test of significance of excluded instruments				significant at 0.001					
First stage R-2					0.62 - 0.72				

Fixed effects by sport included but not displayed for readability.

+Standard errors corrected for heteroskedasticity and autocorrelation using Newey-West.

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

Table 7: Monetary Effect

	OLS	IV	3SLS
Dollar value	30.6	15.2	15.5
Above 8.34	**100%**	87%	93%
Above 11.51	98%	69%	75%
Above 15.45	91%	49%	50%

Table 8: Effect on Unique Visitors to the Paid Section by Season

	(1) OLS+		(2) IV+		(3) 3SLS	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Free Articles						
- Off season	0.009	0.004 **	0.012	0.005 **	0.012	0.004 ***
- Regular season	0.006	0.002 ***	0.015	0.004 ***	0.017	0.003 ***
- Post season	0.007	0.003 **	0.010	0.003 ***	0.010	0.004 ***
Paid Articles						
- Off season	0.051	0.012 ***	0.044	0.014 ***	0.045	0.011 ***
- Regular season	0.047	0.012 ***	0.035	0.015 **	0.034	0.013 ***
- Post season	0.101	0.024 ***	0.081	0.028 ***	0.080	0.029 ***
PercGames	-0.033	0.039	-0.006	0.039	-0.005	0.044
Gameday	-0.032	0.082	-0.127	0.091	-0.138	0.084 *
Draft	1.016	0.315 ***	1.036	0.358 ***	1.028	0.306 ***
Lockout	-0.585	0.163 ***	-0.467	0.181 ***	-0.449	0.126 ***
Final game	-0.038	0.213	-0.105	0.228	-0.105	0.432
Nonworkingday	-0.280	0.050 ***	-0.306	0.053 ***	-0.313	0.050 ***
Googlescaled	0.015	0.004 ***	0.015	0.004 ***	0.016	0.004
ln(Ya Uniqvisitors)	0.111	0.034 ***	0.097	0.036 ***	0.097	0.031 ***
ln(CB Uniqvisitors)	0.010	0.039	0.012	0.040	0.007	0.036
ln(SI uniqvisitors)	0.103	0.031 ***	0.108	0.032 ***	0.103	0.030 ***
Reg Season	-0.264	0.178	-0.400	0.209 *	-0.422	0.186 **
Post Season	-0.543	0.261 **	-0.377	0.269	-0.346	0.262
R-2		0.1622		0.1493		0.4131
N		2032		2007		2007
Endogeneity test ($\chi 2$)				12.794 **		
Kleibergen-Paap rk LM statistic ($\chi 2$)				105.174 ***		
Significance of first-stage regressions			significant at 0.001			
F-test of significance of excluded instruments			significant at 0.001			
First stage R-2				0.76 - 0.98		

Fixed effects by sport included but not displayed for readability.

Standard errors corrected for heteroskedasticity and autocorrelation using Newey-West.

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

Table 9: Effect on Page Views by Season

	(1) OLS+			(2) IV (inst: articles)+			(3) 3SLS		
	Coef.	Std. Err.		Coef.	Std. Err.		Coef.	Std. Err.	
Free Articles									
- Off season	0.007	0.002	***	0.007	0.002	***	0.007	0.002	***
- Regular season	0.000	0.001		0.002	0.001		0.002	0.001	
- Post season	0.000	0.001		0.001	0.001		0.001	0.002	
Paid Articles									
- Off season	0.001	0.005		-0.002	0.006		-0.001	0.005	
- Regular season	-0.017	0.005	***	-0.019	0.006	***	-0.019	0.005	***
- Post season	-0.014	0.011		-0.030	0.012	**	-0.030	0.012	**
PercGames	-0.074	0.016	***	-0.068	0.017	***	-0.070	0.018	***
Gameday	0.111	0.037	***	0.105	0.039	***	0.112	0.035	***
Draft	-0.019	0.138		-0.002	0.151		0.003	0.128	
Lockout	0.215	0.064	***	0.240	0.069	***	0.230	0.052	***
Final game	-0.398	0.200	**	-0.443	0.187	**	-0.445	0.181	**
Nonworkingday	-0.211	0.023	***	-0.220	0.023	***	-0.218	0.021	***
Googlescaled	0.028	0.002	***	0.028	0.002	***	0.028	0.002	
ln(Ya Pageviews)	0.172	0.019	***	0.161	0.019	***	0.162	0.016	***
ln(CB Pageviews)	0.076	0.014	***	0.071	0.014	***	0.072	0.012	***
ln(SI Pageviews)	0.140	0.013	***	0.141	0.013	***	0.139	0.010	***
Reg Season	0.751	0.083	***	0.691	0.094	***	0.713	0.078	***
Post Season	0.457	0.134	***	0.563	0.140	***	0.563	0.110	***
R-2		0.814			0.813			0.8607	
N		2032			2007			2007	
Endogeneity test (χ2)					15.052	**			
Kleibergen-Paap rk LM statistic (χ2)					97.884	***			
Significance of first-stage regressions				significant at 0.001					
F-test of significance of excluded instruments				significant at 0.001					
First stage R-2					0.76 - 0.98				

Fixed effects by sport included but not displayed for readability.

Standard errors corrected for heteroskedasticity and autocorrelation using Newey-West.

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

Table 10: Monetary Effect by Season

		OLS	IV	3SLS
Off season				
	Dollar value	>100	>100	>100
	Above 8.34	**100%**	100%	100%
	Above 11.51	100%	100%	100%
	Above 15.45	100%	**99%**	100%
Regular season				
	Dollar value	5.9	3.7	3.6
	Above 8.34	19%	7%	5%
	Above 11.51	6%	2%	1%
	Above 15.45	2%	1%	1%
Post season				
	Dollar value	30.6	11.5	11.2
	Above 8.34	**99%**	**73%**	71%
	Above 11.51	**95%**	**50%**	48%
	Above 15.45	86%	30%	28%

Figure 1: Screenshot of ESPN website Displaying Insider-icon

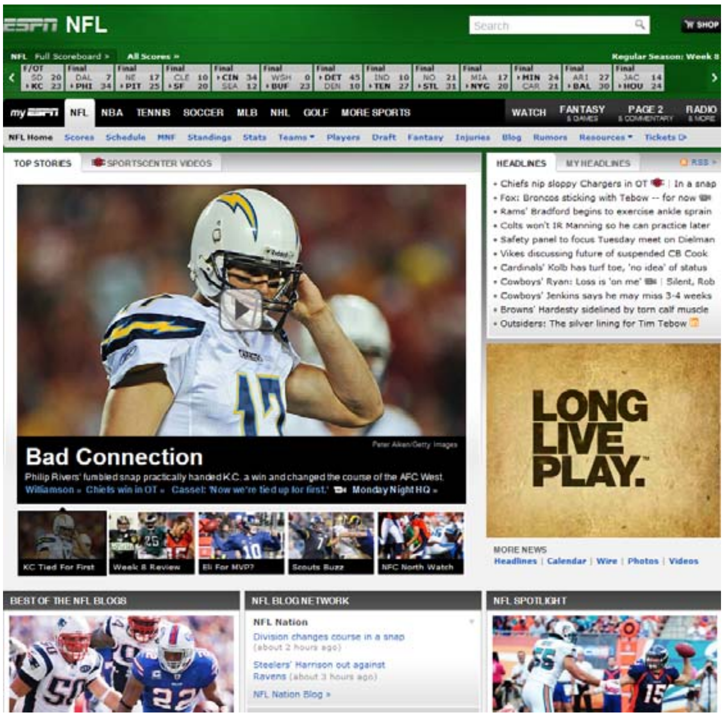

Figure 2: Density of Break-Even Dollars per Article

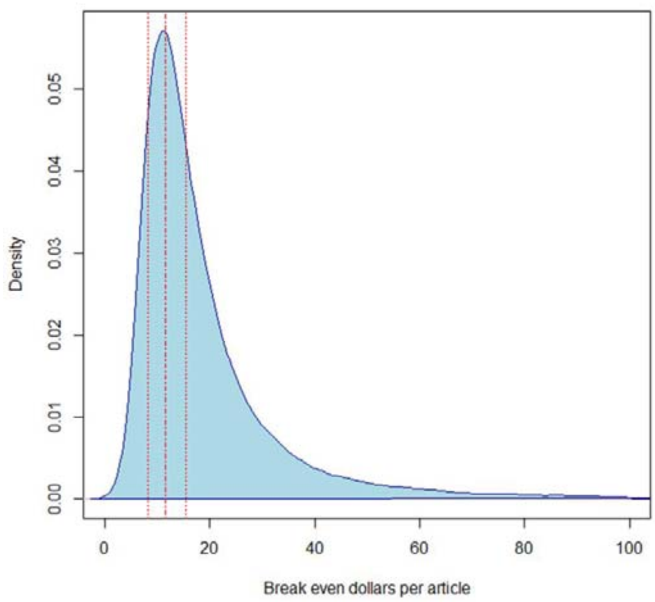

Figure 3: Estimated Break-Even Dollars per Article by Season

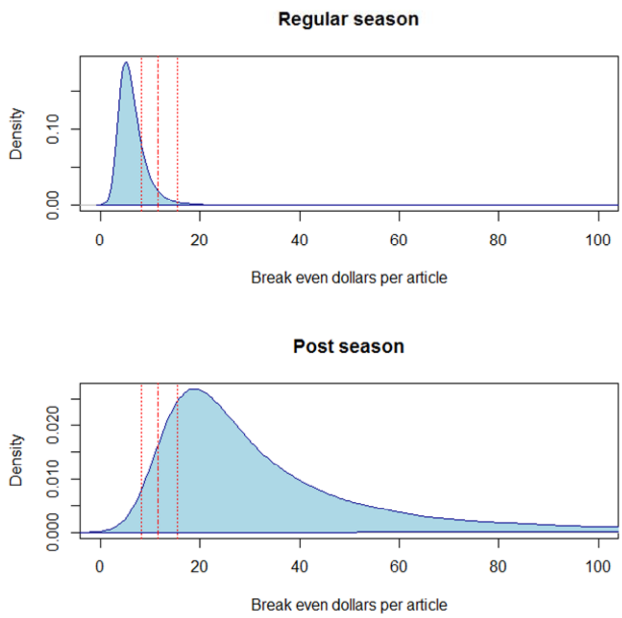

Figure 4: Unique Visitors to Free and Paid Sections Across Seasons