Pricing and Consumption in Subscription Settings*

Pedro M. Gardete, Daniela Schmitt, Florian Stahl

December 2024

Abstract

This paper examines how subscription pricing impacts usage intensity, a critical factor for firms relying on subscription-based models. Our analysis focuses on data from an online news publisher, a context where promotional pricing is often seen as a way to attract new customers, but where the impact on usage is less clear. Traditional analyses suggest that lower prices should lead to lower per capita consumption levels, as lower-paying customers are presumed to place lower value on product consumption. In contrast, our analysis reveals that promotional subscribers may consume significantly more than those who pay regular prices, even after controlling for churn rates. This phenomenon may occur due to multiple reasons, including switching and/or multihoming costs. From the firm's (partial equilibrium) point of view, this pattern is akin to there existing a negative correlation between consumption value and willingness-to-pay. We develop and estimate a model that allows for a flexible relationship between subscription prices and consumption, and use it to recover the fundamental parameters and to investigate the impact of different pricing policies through their effect on revenues from subscriptions (via new customers) and from subsequent consumption (via advertising).

*Gardete and Schmitt: Nova School of Business and Economics; Stahl: University of Mannheim. Authors contributed equally and order is alphabetical. We thank André Trindade for helpful comments. This work was funded by Fundação para a Ciência e a Tecnologia (UIDB/00124/2020, UIDP/00124/2020 and Social Sciences DataLab - PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016).

1 Introduction

Ask any executive how pricing policies influence the demand for a product or service, and you'll get a confident, well-reasoned reply. Ask that same executive how pricing policies affect consumption—the extent to which customers use products or services that they've paid for—and you'll get a muted response at best.

– John T. Gourville and Dilip Soman, HBR 2002

This paper examines the phenomenon of subscription pricing and its relationship with subsequent usage intensity, a key relationship for many firms. For news publishers, streaming services, and software platforms, promotional pricing campaigns are more than a means to attract new customers: they are essential tools for promoting user engagement, building customer loyalty, and driving revenue. In subscription-based business models in particular, promoting engagement and resulting consumption are especially important, since firms often incur costs and/or derive value not only from subscriptions themselves but also from the intensity of product consumption.

The relationship between subscription pricing and consumers' usage is nuanced and can take place through several mechanisms. One such mechanism is that of demand composition or selection: The fact that changes to subscription prices attract different types of customers, who in turn have different preferences surrounding how much they would like to consume, conditional on subscribing. The traditional prediction from economic theory on this matter is relatively straightforward: Because individuals consider their consumption value before subscribing (i.e., they are forward looking), subscription prices are expected to be positively correlated with individuals' consumptions. When price decreases, the firm is able to attract consumers who would not be willing to subscribe at the regular price, and who, ceteris paribus, value consumption less than the current customers. This prediction is reinforced by the fact that the prices of most subscription offers are typically not significant enough to induce wealth effects (e.g., Vives, 1987 and Hayashi, 2008), suggesting that consumption is more likely to be determined by preferences rather than by income effects due to budget constraints and diminishing marginal utility of consumption of the outside good.[1]

Despite this, there also exist economic rationales that predict marginal subscribers will

[1]Moreover, the interpretation of classical budget constraints is often ambiguous when taken to the empirical context (see the varying definitions and discussions in Deaton and Muellbauer (1980), Kim, Allenby, and Rossi (2002), Chintagunta and Nair (2011)).

consume more than existing ones. Consumers who take advantage of promotions may belong to specific populations (e.g., retired and/or unemployed) who feature both a lower willingness to pay as well as a lower opportunity cost of time. In such a case, a company may find that its new subscribers, acquired through a price promotion, consume more per capita than their current customers. Similarly, across many contexts, price promotions may primarily attract only those consumers who do not yet subscribe to products in the category, such as streaming services and online news. Current subscribers of competitive products, on the other hand, may face switching costs or even costs of maintaining and monitoring multiple accounts, i.e., multihoming costs. As a result, the relationship between marketing activities devoted to acquiring new customers and their subsequent consumption levels may constitute primarily an empirical question.

We believe there already exists evidence in the Marketing field – albeit rarely made explicit – supporting the idea that changes in subscription prices may not always carry traditional implications to consumption patterns. For example, Danaher (2002) manipulates subscription prices of a telecommunications subscription service. One of his findings is the fact that individual usage on average decreased as a result of an increase in the subscription price of the service. Moreover, this results holds even after controlling for customer churn; that is, consumers who continued using their subscription despite higher prices exhibited lower consumption levels compared to those who remained active under regular pricing. The effect is significant, both in its magnitude as well as statistically. In the context of a music streaming service, Chou and Kumar (2024) find that segments are heterogeneous in terms of their willingness-to-pay and their usage. For example, women are reported to use the service less than men, despite exhibiting lower price elasticities. The authors propose that, despite their lower usage, women may face a higher valuation for leisure (or other competing activities). Basically, consumers hold preferences on one hand, but other factors also drive their willingness to pay. Relatedly, Runge, Levav, and Nair (2022) inquire whether price promotions may hurt consumer usage in the context of a game subscription. For example, it is possible that consumers associate lower prices as signals of lower quality. In fact, they find little if no evidence of future usage changes following a price promotion.

In this paper we investigate the consumption patterns of news of a leading European online news publisher. The publisher offers both free articles as well as premium articles, the latter being accessible via a monthly subscription. During the sample period, the publisher reduced its subscription price for new subscribers, which led to a marked increase in the number of new subscriptions. In addition, consumers who subscribed during the promotion period exhibited higher news consumption levels, i.e. read more news articles, than those consumers who subscribed just a few days before the promotion at the regular price.

Two striking results are also visible in our data. First, while the higher consumption by new subscribers is partially explained by their lower churn rate, the higher level persists even after conditioning on active users: Conditional on remaining active, promotion subscribers consume more news than their regular-price counterparts. Second, the difference in consumption levels does not decrease even after a year has elapsed. Rather than inducing temporary changes in consumer behavior, it appears the promotion effectively attracted new types of subscribers to the firm.

These data patterns are especially important in our setting since the online publisher accrues revenue from both subscriptions and advertising. In order to quantify these effects and consider the impact of different pricing policies, we develop and estimate a simple model of subscription and consumption behavior. This model takes into account the fact that the utility a customer derives from product consumption may not be enough to characterize consumption. We allow for consumers to hold a willingness to pay (w.t.p.), due to mechanisms like switching and multihoming costs as well as access to heterogeneous opportunity costs of time, as we describe later. Importantly, consumers' willingness to pay may be arbitrarily correlated with consumption value. Another advantage of this approach is that empirical work – such as this one – frequently has at its center a rich dataset from a single source and rarely contemplates all activities that may compete directly and indirectly to satisfy consumers' wants and needs, use of time and resources. By allowing for consumption utility to be correlated with w.t.p., we inform the relation between the act of subscribing and subsequent consumption directly from the data, regardless of the factor(s) that induced it. An alternative to our model considers the decision of subscribing and consuming as a joint problem, much in the spirit of discrete/continuous models (see for example Chintagunta (1993), Kim, Allenby, and Rossi (2002), Bhat (2005), Tuchman, Nair, and Gardete (2018)). We see two disadvantages with this approach. First, these models typically predict a positive correlation between subscription prices and consumption levels. Second, they ignore a unique aspect of subscription programs: that consumption takes place at a time when price has already been sunk, whereas discrete-continuous models typically deal with per-unit prices where the total price depends linearly on the quantity purchased. In contrast, because in our setting subscription prices are sunk at the time of consumption, we assume subscribers' types are given by the the number of articles they would like to consume conditional on subscribing, which may crucially correlate with their willingness to pay for the subscription.

After documenting the patterns in the data, we develop a model to characterize consumers' subscription decisions as well as the observed consumption. The working assumption is that the price promotion induced selection into consumption by customers who may behave differently from the company's current customers. The model features a statistical copula

that allows us to link consumers' observed consumption with their willingness to subscribe. We find that while the correlation parameter acts as a local, reduced-form characterization of demand faced by the focal firm due to potentially many mechanisms, it is nonetheless essential to produce even a reasonable fit with the data. We consider different promotional depths in the counterfactual analysis, and find that the firm priced very near to the optimal price point if only subscription revenue were taken into account, but underdiscounted once advertising revenue is taken into account. The firm would have been better off promoting aggressively in order to induce more consumption and higher advertising revenue in the future. This finding is robust even at lower advertising rates than the ones communicated to us by the management team of the firm. We then consider a counterfactual analysis in which the firm is assumed to be able to monitor actual consumption by subscribers; a scenario in which consumers cannot manipulate their true product consumption (e.g., implement a bot to generate fictitious consumption). In this case, firms may be able to introduce quantity discounts that effectively feature negative marginal prices. We find that providing discounts for high levels of (verifiable) consumption would be extremely profitable to the firm vis-a-vis the status quo for a reasonable range of advertising elasticities, owing primarily to the advertising subscription revenue generated by the program.

The ability to characterize the relationship between pricing and subsequent consumption is not only relevant to understand profit tradeoffs for firms, but also essential to assess the broader societal implications of subscription pricing. Our findings contribute to understanding the extent to which news consumption can be affected by pricing strategies — a critical issue in a world where misinformation is widespread and online news publishers face significant financial challenges. By shedding light on how pricing policies influence readership, this paper offers insights that are relevant for policymakers, media regulators, and publishers aiming to promote access to reliable and accurate information.

In the next section we discuss the potential mechanisms in play in the relation between subscriptions and consumption. Section 3 describes the dataset patterns and Section 4 presents the model, identification and estimation. Section 5 and 6 present the estimation results and counterfactual analyses, respectively, and Section 7 concludes.

2 Subscriptions and Consumption

While the goal of this paper is not to separate out specific mechanisms, it is useful to clarify a few avenues through which subscription decisions can relate to subsequent consumption. We first describe classical mechanisms that likely predict a positive correlation between subscription prices and consumption. We then explain that there exist a number of reasonable

mechanisms that may predict the opposite direction. In the appendix we include a toy model exemplifying how a positive correlation between prices and consumption may arise in the market, and describe a few mechanisms that are less likely to impact the relationship in most analyses, including the case of wealth/income effects.

2.1 Positive Correlation Between Subscription Prices and Consumption

Consumer Preferences. We start by explaining that simple consumer preferences can only explain a positive correlation between subscription prices and subsequent consumption. The reasoning is straightforward: When consumers make decisions solely based on preferences, a price drop leads to extramarginal consumers (i.e. those whose valuations fall below the regular price) to select into buying, leading to an overall drop in average consumption. Proposition 1 formalizes this argument (proof in the appendix):

Proposition 1 ***Simple Consumer Preferences.*** *Let $u(v_i)$ be a utility function and v_i be the optimal consumption of consumer i conditional on a purchase at price p so that the consumer subscribes iff $u(v_i) \geq p$. Then, a price decrease leads simultaneously to higher demand and lower average consumption.*

This is a relatively straightforward result that, nonetheless, may be overturned by some of the factors we describe in the next section.

Mental Accounts. The idea that consumers may hold separate mental accounts has received significant attention (see for example Thaler (1985) for discussions). It is likely that consumers who appreciate certain activities more are also more likely to keep larger mental budgets for those activities. Similarly, consumers with higher wealth levels are also more likely to hold larger mental accounts. As a result, it is likely that mental accounts mean that higher prices induce more consumption due to willingness-to-pay selection of consumers. While this is an intuitive result, in the next section we explain, however, that the presence of mental accounts may also induce negative correlation between subscription prices and consumption in some cases.

Sunk Costs. The theory of sunk costs predicts that costs incurred today affect future behaviors despite their payoff irrelevance. The relationship with usage is also relatively straightforward. For example, Arkes and Blumer (1985) find that consumers who obtain

a price promotion for Opera season tickets decrease their attendance in the future, due to the lower sunk costs incurred. The same theory predicts that a price increase will lead consumers to value their past investment more in the future, leading them to take advantage of more consumption opportunities. Overall, the sunk cost fallacy predicts that subscribers are expected to distort their subsequent consumption decisions in the direction of the change in subscription price.

2.2 Negative Correlation Between Subscription Prices and Consumption

Switching Costs, Multihoming Costs, and Opportunity Cost of Time. The phenomenon of switching and multihoming costs, which make it costly for consumers to switch between or maintain multiple offers, has been well established in the literature (e.g., Klemperer (1987), Hartmann and Viard (2008), Villas-Boas (2015)). Moreover, consumers also face smaller but relevant costs during consumption with significant implications to competing firms (see Esteves-Sorenson and Perretti (2012)). These factors can induce persistence and cause high-value consumers to refrain from taking advantage of price promotions. For example, a price promotion may be effective in attracting low-type consumers who did not subscribe to competing services before, but fail to attract high-type consumers who already subscribe to a competitor offering a premium alternative. For example, a price promotion of a streaming service may attract cheapskates as well as 'movie buffs', but it may also fail to attract moderate consumers who find it challenging and wasteful to manage several streaming accounts simultaneously.

The presence of switching and multihoming costs may be bolstered by other factors, such as mental accounts and heterogeneous time constraints. For example, consumers with lower incomes may assign lower amounts to specific categories, such as leisure and entertainment. This does not mean they would not derive consumption value from the category; rather, other activities may take priority due to financial reasons, for example. A price promotion may therefore provide license to such consumers to opt into a subscription offer without having to exceed their mental account for the category. Relatedly, a price promotion may motivate consumers to change the way they allocate their time. Even consumers with a high opportunity cost of time may react by reducing the time spent in competing activities and consume more from the promoted offer.

3 Data Patterns

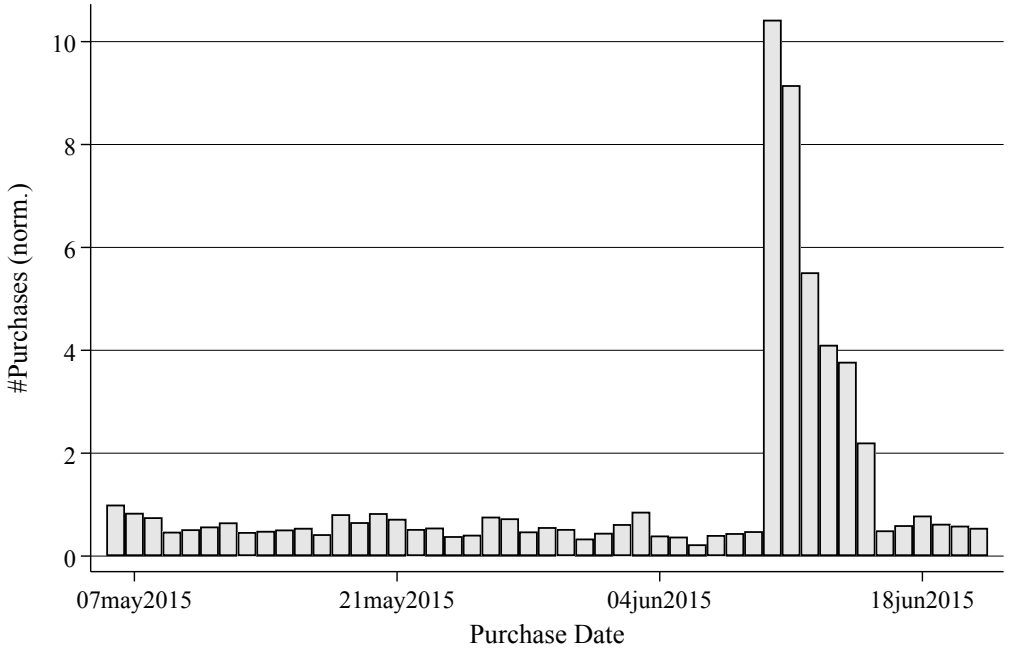
The dataset originates from a leading European digital news publisher that implements a freemium business model, i.e. the publisher provides access to a combination of free and premium content on its website. While visitors can read free articles at no cost, premium content is accessible only to paying subscribers. The subscription is priced at €4.99 per month, providing unlimited access to all content. The publisher also generates revenue by displaying advertisements on both free and premium articles, with ad revenue directly tied to the number of articles viewed. Based on meetings with the management team, the marginal cost for the publisher to accommodate additional readership is considered irrelevant.

Our dataset focuses on consumers who subscribed between May 6^{th}, 2015 and June 15^{th}, 2015, providing detailed individual-level data on the browsing and subscription behavior of over 10,000 new subscribers. For each subscriber, the dataset captures the timing and frequency of website visits, along with subscription initiation and cancellation dates. Our analysis examines behavior during the one-year period following each user's original subscription, offering a comprehensive view of their engagement patterns over time.

In mid-June 2015, the publisher ran a previously-unannounced promotional campaign. The six-day promotion, running from June 10^{th} to June 15^{th}, temporarily reduced the monthly subscription price from €4.99 to €2.00 and was available exclusively to first-time subscribers. Customers who subscribed during this promotional period continued to pay the reduced price of €2.00 per month until they canceled their subscription.

Figure 1 presents the number of new subscriptions during the sample period (for completeness, it also includes six days subsequent to the promotion, which are ignored in the main analysis). The data reveal a stark increase in new subscriptions during the six-day promotional period. Prior to the promotion, subscription activity remained relatively stable, but the launch of the temporary price reduction from €4.99 to €2.00 triggered a pronounced spike in subscriptions with no detectable anticipation effects. This momentum diminished gradually over the promotional period and is undetectable after the promotion ended, reflecting the well-defined impact of its influence on subscription rates.

Figure 1: New Subscriptions Per Day



Note: Number of daily subscriptions normalized to 1 based on the subscriptions on May 6^{th}, 2015.

The promotional effects appear proportional to the magnitude of the promotion: The 40% discount generated more than ten times the number of new subscribers the publisher had attracted in any day before the promotion. The effects of the promotion are also clearly diminishing over time. This effect is in line with the notion of Conlisk, Gerstner, and Sobel (1984) that some types of consumers may subscribe only during promotional periods. As a result, the firm may benefit from running promotions long enough to attract them but not too long so as not to allow other types to take advantage of it.

A more enduring consequence of the promotion, however, is its impact on consumption patterns. To explore these patterns further, we classify subscribers into two groups: regular subscribers who joined at the regular price and promotional subscribers who subscribed during the promotion period. Table 1 provides moments of the consumption behaviors during first year following each consumer's subscription decision. On average, promotional subscribers consumed almost three times as many articles as their regular counterparts. This pattern persisted throughout the first subscription year, as shown in Figure 2. When aggregating all articles consumed by subscribers over a maximum subscription period of 12 months, promotional subscribers demonstrated significantly higher consumption levels compared to regular subscribers (10.68 thousand vs. 3.69 thousand articles, $p<0.001$).

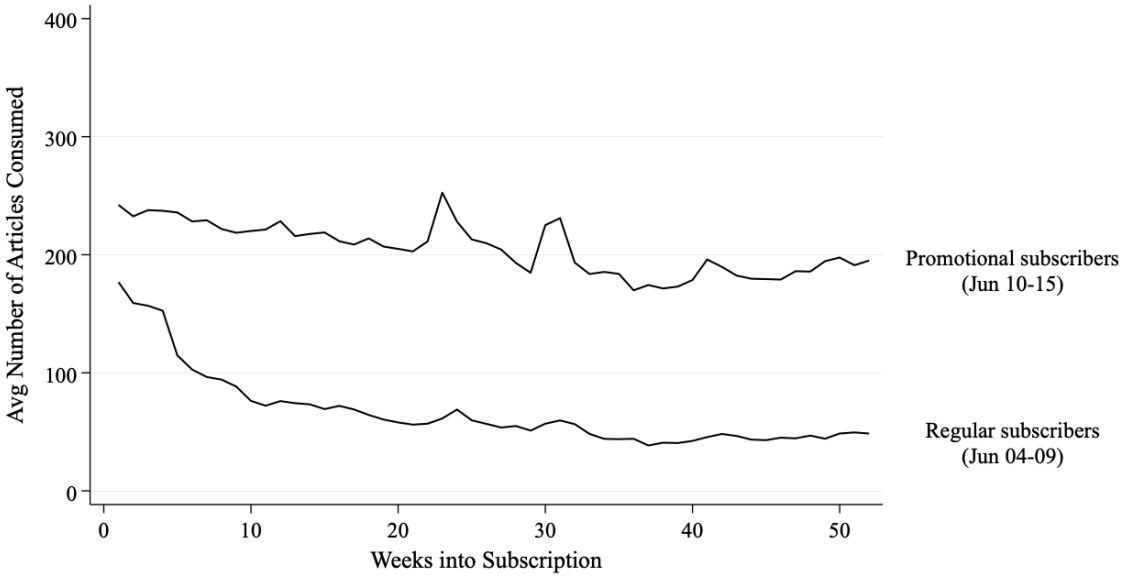
Table 1: Summary Statistics of News Articles Consumed Per Week

	Min	Mean	Median	Max	SD	N
#Articles consumed/week:						
Subscribed at regular price	.02	71.10	17.82	6,183.9	172.81	3,620
Subscribed at promotion price	.02	205.33	134.23	3,865.8	253.04	6,615
Total	.02	157.85	79.21	6,183.9	236.77	10,235

Note: Above, weekly summary statistics during the sample period, May 6^{th}, 2015 to June 15^{th}, 2015.

Figure 2 presents data only from regular consumers who subscribed a week before the onset of the price promotion, ensuring greater compatibility with promotional subscribers. Across the 52-week period depicted in Figure 2, promotional subscribers demonstrate consistently higher article consumption compared to regular subscribers. Their consumption rates began with a large gap and this difference stabilized, with the groups maintaining parallel trends over time. This persistent disparity implies that promotional subscribers exhibit distinct engagement behaviors that are not solely attributable to the reduced subscription price.

Figure 2: Number of Articles Consumed up to One Year since Subscription



Note: Weekly news readership over a 12-month period after subscribing.

Table 2 highlights the differences in contract duration between promotional and regular subscriber groups. Promotional subscribers remained active on the platform for nearly twice as long as regular subscribers.

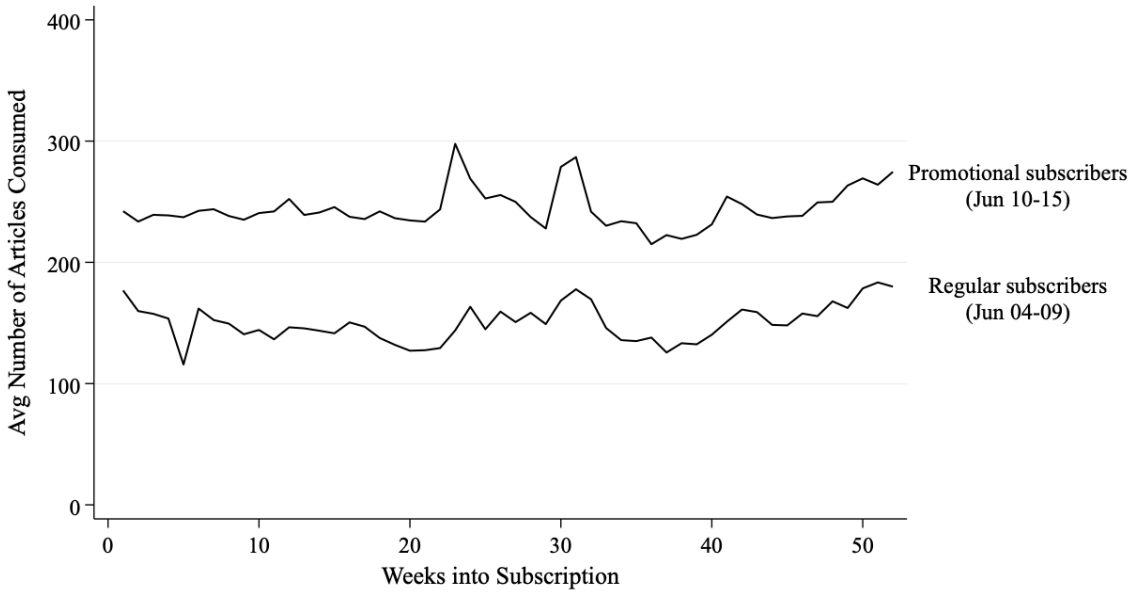
Table 2: Summary Statistics of Subscription Length

	Min	Mean	Median	Max	SD	N
Subscription length in months:						
Subscribed at regular price	1	4.89	1	12	4.83	3,620
Subscribed at promotion price	1	10.19	12	12	3.76	6,615
Total	1	8.32	12	12	4.88	10,235

Note: Length of users' subscription, measured in months, over the 12-month period following their initial subscription.

This difference in retention needs to be considered when analyzing consumption patterns, as it could be argued that the higher consumption observed among promotional subscribers is simply a consequence of their lower churn rates. For Figure 3, again only includes regular consumers who subscribed a week prior to the price promotion. Notably, as illustrated in Figure 3, even after controlling for contract duration—i.e., conditional on subscribers not churning—promotional subscribers still displayed substantially higher engagement levels, consuming approximately 50% more news articles than regular subscribers. The combination of extended retention and increased consumption likely contributed to higher advertising revenue opportunities, as the heightened activity of promotional subscribers resulted in more frequent exposure to advertisements. These findings underscore the potential revenue advantages of using strategic promotions to target high-consumption users.

Figure 3: Number of Articles Consumed by Active Subscribers up to One Year since Subscription



Note: Weekly news readership over a 12-month period after subscribing, conditional on subscribers remaining active (i.e., not having churned).

The data suggest that promotions can attract a distinct consumer segment characterized by higher engagement and retention. We develop a model that incorporates these demand patterns. This model also allows for counterfactual analyses, enabling us to explore the potential outcomes of alternative promotional strategies, including variations in pricing levels and plan menus.

4 Model

We consider a simple model that features heterogeneous consumers along two dimensions. The first dimension, noted as v_i, is the quantity that subscriber i would like to consume once subscription costs are sunk. In other words, if consumer i decides to subscribe, her unconditional consumed quality is v_i. Note that only the v_i levels of subscribers are observed by researchers, much in the spirit of Heckman (1979). Consumption level v_i is relevant for the consumer as it informs whether it is worth subscribing to the firm's plan, via a classical utility function.

In addition to the indirect utility function, consumers also hold a willingness-to-pay constraint that depends on forces other than direct consumption utility. The w.t.p. condition is different from the effect of price and income in the indirect utility (notice that the latter

already incorporates wealth/income and price, by definition). Our assumption is that there exist unobservable equilibrium factors (such as unobserved competition or opportunity costs of time) that induce heterogeneity in the willingness to pay of consumers. Consumer i's willingness to pay for a subscription is given by w_i, and essentially. A central novelty of the model is that we allow the consumption preference v_i and the w.t.p. w_i to be arbitrarily correlated. This is made possible, as we discuss later, from the fact that the subscription price is sunk once consumption takes place.

We assume each consumer has a single chance to subscribe during the sample period, at the posted subscription price p_i. Those consumers who do subscribe can consume as much as they would like with no associated costs. Consumers take both their consumption utilities and w.t.p. into account when deciding whether to subscribe.

Consumer i subscribes if and only if both of the following conditions are met:

$$\alpha v_i + \varepsilon_i^1 \geq p_i + \varepsilon_i^0 \qquad \text{(indirect utility)} \tag{1}$$

$$w_i \geq p_i \qquad \text{(w.t.p. constraint)} \tag{2}$$

where inequality (1) is the classic indirect utility function and inequality (2) is the w.t.p. constraint. Parameter α is a preference parameter that converts news consumption to monetary units, v_i is consumer i's optimal news consumption in case she subscribes. Shocks ε_i^1 and ε_i^0 are extreme-value type 1 distributed preference shocks, p_i is the price observed by consumer i and w_i is consumer i's willingness to pay.

We provide a visual representation of this simple model in Figure 4. Panel 1) depicts the space of consumption v_i and willingness to pay w_i that justifies subscribing. The top-right area, outlined by an orange line, reveals that only consumers who are both interested in consuming enough news ($v_i \geq p/\alpha$) and are willing to pay enough for that service for non-utility reasons ($w_i \geq p$) will subscribe (notice we abstract away from the preference shocks here for simplicity). The second panel illustrates the effect of introducing a lower price, $p' < p$. The subscription region grows down and to the left, a combination of the preference effect and the willingness to pay effect. Panel 3 introduces the support of a distribution of individuals' $\{w_i, v_i\}$ pairs, with correlation parameter $\rho > 0$. In this case, it is clear that the price reduction increases the volume of subscribers, given by the intersection of the distribution's volume over its support with the new region. Consumers in this region had not subscribed at the regular price, either because they exhibit relatively weak preferences for consumption or because they hold a low willingness to pay related with other factors. Finally, panel 4) depicts the analogous case of a price decrease when the correlation of consumption and will-

ingness to pay is negative. In this case, the price reduction captures consumers that tend to take great interest in reading the news but other factors – such as the appeal of unobservable competitive offers – discourages them from subscribing at the regular price. Consider the mass points A and B in Panel 3. It is clear that, following a price reduction, the new subscribers will pull average consumption down. This contrasts with Panel 4 of the figure, where if only mass points of consumers A and C existed, the price reduction would lead to an increase of the average consumption due to the new subscriber mass C. The formulation described above is relatively simple and readily interpretable. It also helps illustrate that the central reason we are able to identify the correlation parameter ρ is the fact that price variation affects consumption levels only through selection into the subscription, since subscription costs are sunk during consumption. As a result, we are able to depict consumer v_i as consumers' types, in contrast with typical discrete-continuous models of demand that tend to depend exclusively on structure to link the extensive and intensive margins.

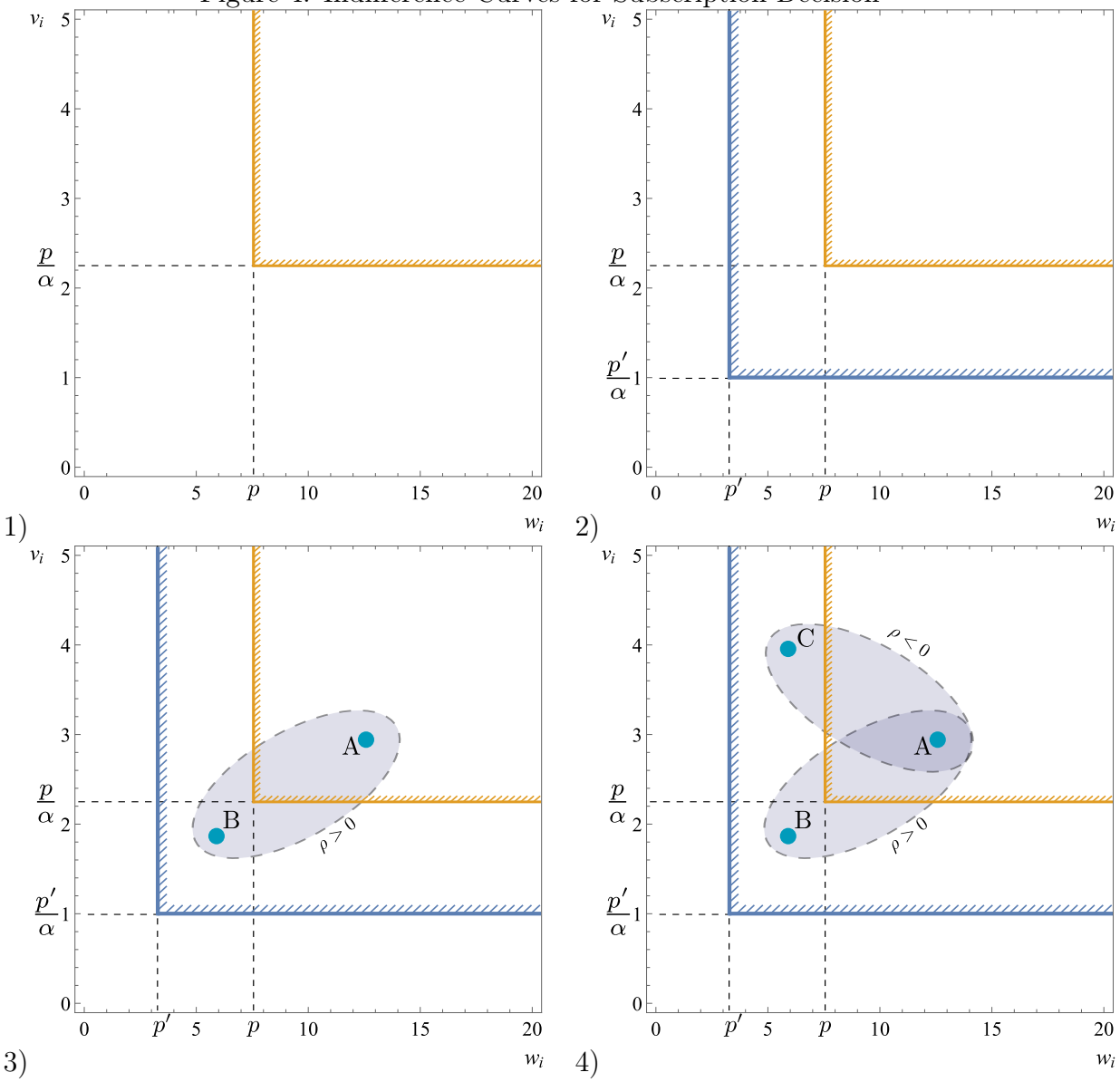
Distribution of v and w. In addition to the parameters presented in equations (1) and (2), we also consider the parameterization of the joint distribution of v and w. We start by noting that the distribution of w is unobservable to the researchers. Indeed, being able to account for this unobserved latent variable is part of the contribution of this paper. As for the levels of consumption v_i, they are indeed observable from the data, but only for subscribers. In other words, we observe the distribution

$$F_{v_i|subscribe_i} = F_{v_i|\alpha v_i + \varepsilon_i^1 \geq p_i + \varepsilon_i^0 \wedge w_i \geq p_i} \tag{3}$$

where p_i may belong to one of two price levels. Figure 5 presents histograms of annual readership levels for consumers who subscribed at the regular (price: €4.99) and at the promotional (price: €2) prices. To be precise, we define v_i as the consumption level by consumer i over a 12-month period after the initial subscription.[2]

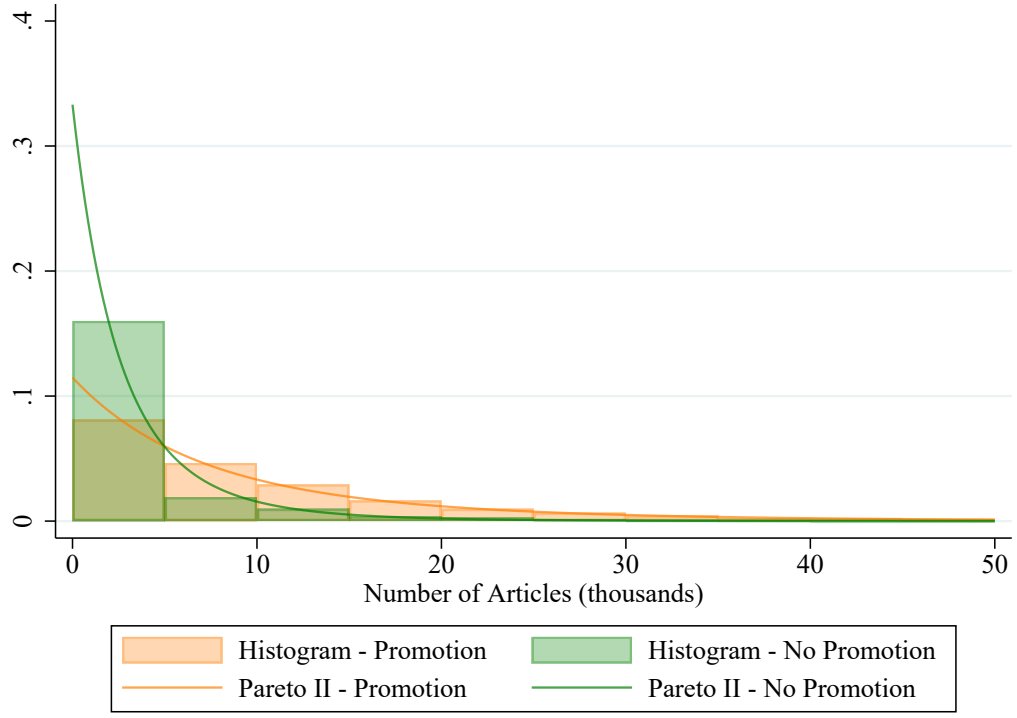
[2]We focus on the sample of consumers who, within each price group, falls within three standard deviations from the mean in terms of news readership. This removed a few very high values, that may (or not) be associated with non-human activities such as bots, scraping tools, etc.

Figure 4: Indifference Curves for Subscription Decision



Note: Parameter values used for the graphs above: $\alpha = 1, \sigma = 0.75$, $p_1 = 5$, and $p_2 = 10$.

Figure 5: Histograms of the Number of Articles Consumed up to One Year since Subscription



Note: Histograms of news readership over a 12-month period after subscribing. The lines plotted above are fitted Pareto II densities with shape parameter $\alpha = 5$. Readership capped at the 50,000 articles for readability.

Both histograms follow exponentially decaying densities, with consumers in the promotion condition consuming more news than their counterparts, in line with the results of the preliminary data analysis. Table 3 presents summary statistics of the sample used for model estimation.

Table 3: Summary Statistics of News Articles Consumed

	Average	Standard Deviation	Min	Max	N:
Promotional Price	9.635	9.672	0.001	51.104	6,480
Regular Price	3.066	5.116	0.001	30.73	3,569

Note: Consumption values for thousands of news articles consumed over the period of 12 months after subscribing.

One of the striking features of the figures in Table 3 is the fact that the standard deviation exceeds the average, given that these are distributions bounded below at zero. This is a sign of extremely heavy-tailed distributions. To see this, notice that the exponential distribution

– often used to model cases of long tails – does not admit the standard deviation to fall above its mean. Indeed, its mean and standard deviation are equal to $\frac{1}{\lambda}$, and in cases where v is truncated from below at $a > 0$, the standard deviation falls strictly below the mean ($\mathrm{E}(v|\, v > a) = a + \frac{1}{\lambda}$; $\mathrm{Std.Dev.}(v|\, v > a) = \frac{1}{\lambda}$). It follows that the moments in Table 3 are incompatible with the exponential distribution. We address the heavy tail of news readership by assuming that v is distributed as a Pareto type II distribution, also known as the Lomax distribution. This distribution has various uses across literatures, but it is mainly characterized by its heavy positive tail. As can be seen in Figure 5, it can fit the observed news consumption data quite well. We opt for the following specification for the density of v:

$$f_v(v_i) = \frac{\alpha_v}{\lambda_v}\left(1 + \frac{v_i}{\lambda_v}\right)^{-(1+\alpha_v)}, v_i \geq 0, \alpha_v > 0, \lambda_v > 0 \tag{4}$$

where α_v is the shape parameter and λ_v is the scale parameter.

Willingness to pay is unobservable by the researchers. We assume it follows a normal distribution with a mass point at zero, that is,

$$w_i = \begin{cases} Normal(\mu_w, \sigma_w), & w.p.\ \gamma \\ 0, & w.p.\ 1-\gamma \end{cases}, \gamma \in [0,1]$$

This interpretation of this specification is that there exists a mass of consumers of size $1-\gamma$ that is willing to pay nothing to access premium news.[3] This assumption is inspired on the relatively low share of subscribers to the online news publisher in comparison with the very large number of free users; it is likely that most consumers do not even consider the possibility of subscribing in the first place. The use of a mass point is also consistent with the "pain-of-paying" literature, by which even the smallest positive price will deter a large portion of consumers from buying. Finally, the normal specification is focal among distributions, and can be justified by a central limit motivation, owing to the fact that willingness to pay can be thought of as aggregating over many articles, time units and/or multiple benefits of subscribing.

The fundamental parameter in our analysis is the correlation between consumption and willingness to pay. Because we do not observe the joint distribution of v and w, we employ a statistical copula to characterize their relationship. This allows us to assume a flexible

[3]Note that we could have set the mass point at any negative value with identical results, since consumers always face positive prices in counterfactual analyses.

bivariate density of v and w, approximated by a gaussian copula with density

$$c_{v,w}(u_1,u_2)=\frac{1}{\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2}\begin{pmatrix}\Phi^{-1}(u_1)\\ \Phi^{-1}(u_2)\end{pmatrix}'\cdot\left(\begin{bmatrix}1&\rho\\ \rho&1\end{bmatrix}^{-1}-I\right)\cdot\begin{pmatrix}\Phi^{-1}(u_1)\\ \Phi^{-1}(u_2)\end{pmatrix}\right) \quad (5)$$

where

$$u_1=F_w(w)$$
$$u_2=F_v(v)$$

Above, $\Phi(\cdot)$ is the standard normal distribution and I is a 2×2 identity matrix. This specification allows us to have a clear definition of correlation parameter $\rho\in(-1,1)$, which captures the correlation between the percentiles of v and w.

Estimation. For estimating the model, we employ a method of moments estimator that matches the moments in the data with the simulated moments of the model:

$$m_1(\theta)=(share|\,p^r)-\left(\widehat{share}(\theta)\Big|\,p^r\right) \quad (6)$$
$$m_2(\theta)=\left(share|\,p^d\right)-\left(\widehat{share}(\theta)\Big|\,p^d\right) \quad (7)$$
$$m_3(\theta)=(\overline{v}|\,p^r)-E(\widehat{v(\theta)}|\,p^r) \quad (8)$$
$$m_4(\theta)=\left(\overline{v}|\,p^d\right)-E(\widehat{v(\theta)}|\,p^d) \quad (9)$$
$$m_5(\theta)=std.dev(v|\,p^r)-\sigma(\widehat{v(\theta)}|\,p^d) \quad (10)$$
$$m_6(\theta)=std.dev(v|\,p^r)-\sigma(\widehat{v(\theta)}|\,p^d) \quad (11)$$

where θ is the set of parameters to be estimated. Moments m_1 and m_2 match the market shares observed in the data with the ones predicted by the model at the regular (p^r) and discount $\left(p^d\right)$ prices. Moments m_3 and m_4 match the average news consumption of subscribers at the different price levels, and moments m_5 and m_6 match the standard deviations. The estimator above is very convenient to compute, but each moment contemplates different numbers of observations. As such, the GMM approach is not immediately applicable. We employ a method of moments estimator with scaling (Greene, 2000, see Section 5.5 and p. 479):

$$obj(\theta)=\sum_{i=1}^{6}m_i'(\theta)^2 \quad (12)$$

where $m_i'=\frac{1}{\phi_i}m_i$ is a scaled version of moment m_i. We scale each moment equation by

dividing it by its square. For example,

$$m_1^{'}(\theta)^2 = \left(\frac{(share|\, p^r) - \left(\widehat{share}(\theta) \Big|\, p^r \right)}{(share|\, p^r)} \right)^2 = \frac{1}{(share|\, p^r)^2} m_1(\theta) \tag{13}$$

This weighting scheme ensures that moment deviations are equally penalized per percentage unit of the moment in the data. For example, a market share prediction of 4.4% vs. an actual market share of 4% is penalized as much as an average consumption of 9,900 predicted at 9,000. This weighting method is especially useful given the very different scales of the moments to be matched in the data. Finally, standard errors of the estimated parameters are calculated via 50 bootstrap replications.

Simulation and estimation algorithm. In each iteration of the set of parameters θ, the model is used to simulate consumption and willingness-to-pay levels, and subscription decisions. We follow the following procedure:

1. Consider some guess of parameters θ.
2. Take K draws of v_k by inverting of the candidate cumulative distribution function F_v.
3. Take K draws of $w_k|\, v_k$ by use of the candidate gaussian copula.
4. For each price level p^t, $t \in \{r, d\}$, calculate whether each simulated consumer satisfies the utility and willingness to pay conditions, i.e., . $I^v_{t,k} = \alpha v_k + \varepsilon^1_k \geq p_t + \varepsilon^0_k$ and $I^w_{t,k} = w_r \geq p_t$. For example, $I^v_{r,k} = 1 \wedge I^w_{r,k} = 1$ means that consumer k is willing to subscribe at the regular price, whereas $I^v_{r,k} = 0 \wedge I^w_{r,k} = 1$ means that the consumer is unwilling to subscribe at the regular price due to insufficient consumption value.
5. Select the appropriate simulations based on the indices above to construct the moments. For example, the predicted average readership from consumers who subscribed at the regular price is obtained by averaging the set $\left\{v : I^v_{r,k} = 1 \wedge I^w_{r,k} = 1\right\}$. The market share at the regular price is obtained by dividing the number of simulations that satisfy conditions $I^v_{r,k} = 1 \wedge I^w_{r,k} = 1$ by the number of simulations, multiplied by the market share parameter γ .
6. Given the predicted moments, the objective function (12) is computed and the optimizer either generates a new guess for parameters θ or stops at the candidate minimum.

The heavy tail of the distribution led us to set $K = 10^8$ simulations of v_i and w_i, at which point different seeds for random number generation had a negligible effect on the estimates.

Given the discrete nature of simulations, we employ the Nelder-Mead method (Nelder and Mead, 1965), and obtain the standard errors for the parameters via 50 bootstrap sample draws. In each bootstrap sample, we draw (with replacement) from subscribers and non-subscribers. We were provided with the number of unique visitors to the firm's website during the period at hand as well as the number of current subscribers in the beginning of the period. We use the difference of the two numbers (withheld for confidentiality) as the potential market size in the model.

Simulation of v and w pairs is a central part of estimation, and takes advantage of the copula correlation structure. We start by taking K independent standard normal draws of vectors Z_1 and Z_2 (we use matrix form below for simplicity), such that

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim N(0, I) \tag{14}$$

At each set of candidate parameters θ, we multiply the draws by the (lower triangular) cholesky decomposition matrix of the gaussian copula correlation matrix (matrix L, defined below) to obtain correlated normal draws, i.e.,

$$\begin{pmatrix} Z_w \\ Z_v \end{pmatrix} = L. \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} . \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \tag{15}$$

Finally, we obtain draws of v and w by applying the standard normal c.d.f. to Z_w and Z_v, and then applying their respective inverse marginal distributions, i.e.,

$$\begin{pmatrix} W \\ V \end{pmatrix} = \begin{pmatrix} F_w^{-1}(\Phi(Z_w)) \\ F_v^{-1}(\Phi(Z_v)) \end{pmatrix} \tag{16}$$

where F_W and F_v are the marginal distributions of w and v, respectively, at the candidate parameters θ.

Identification. We start by normalizing two parameters of the marginal distributions of v and w. For the distribution of v we normalize the shape parameter $\alpha_w = 5$. The reason for normalizing this parameter is that it disciplines the existence of moments of the Pareto type II distribution. Specifically, moment $E\left(v^k\right)$ (and lower) exists if and only if $\alpha > k$. By setting $\alpha = 5$, we focus on Pareto type II distributions that have well-defined first four moments, that is, mean, variance, skewness and kurtosis. While α_w is identified from the higher-order moments of v in the data, strictly speaking (via functional form), those moments only exist if the true value of α_w is indeed high enough. Rather than running the risk of employing

an invalid estimator, we prefer to normalize the value of α_w and utilize only the first two moments of v in the objective equation(12). This issue of lack of existence of moments is common when dealing with heavy-tailed distributions. Changes to the normalization of α_w yielded small effects on the predicted moments of the model.

We also normalize the value of the mean parameter of w_i, μ_w, to zero. We have found that model fit is not especially affected by the mean of w, whose function of capturing the number of non-purchasing consumers is already captured by parameter γ. We prefer to normalize these parameters because 1) the parameters we do include already have focal roles in matching moments in the data, and 2) while we could include higher order moments of v in the objective function to ensure that the number of estimation moments was at least as high as the number of parameters, experimentation revealed very little effect of these parameters in terms of their ability to improve model fit. The parameters of interest are given in Table 4.

Table 4: Model Parameters

Parameter	Description
α	Utility parameter converting readership level v_i to monetary units
λ_v	Scale parameter of v
σ_w	Standard deviation of the willingness to pay
ρ	Correlation parameter of readership and w.t.p. percentiles
γ	Proportion of consumers with w.t.p. different from zero

Identification. We now explain that the moments presented above are identified given the assumptions of our model. Market shares at the regular and discount prices are primarily matched by the utility parameter α and the share of consumers who are willing to pay a strictly positive amount to subscribe, γ. Parameter γ affects the overall share across price conditions, whereas parameter α primarily affects relative market shares across price conditions. Figure 4 is useful to see the effect of α, since this parameter directly impacts how low the values of v consistent with subscribing across price levels can be (abstracting away from the preference shocks). The correlation parameter ρ is identified by the difference in average readership across price levels. The reason is that the extent to which the average readership increases as price decreases is determined by the correlation between v and w, as is also apparent by comparing panels 3) and 4) of Figure 4. Parameters λ_v and σ_v primarily match the mean and standard deviations of readership at the different price levels. A higher value of λ_v translates to a higher mean and standard deviation of v across price levels, a result that follows trivially from the moments of the Pareto II distribution. However, it is necessary to not only match the mean and standard deviation of v overall, but also across

price levels. Parameter σ_w plays the role of capturing the 'decay' of these moments of v as price increases, following the assumed curvature of the normal distribution. Referring back to Figure 4, it is clear that if w is very spread out, the moments of v vary less across price levels than if w is very concentrated near zero (i.e., a low value of σ_w), ceteris paribus.

Final Assumptions. The model above abstracts away from consumers' subscription renewal decisions over the course of a year, focusing only on the first subscription decision and on the overall consumption during the same period. In reality, consumers can decide to stop their subscription during that time frame, a decision that is linked with the other ones. While it is straightforward to incorporate the option value of renewals, we believe it brings unwarranted complexity into the model, without producing an obvious benefit. In section 6.2 we introduce a predictive hazard model of renewals based on subscription time and overall consumption to consider the implications of specific counterfactual scenarios.

Another assumption we make is the fact that v_i accounts for the consumption of all articles by consumers, rather than just the premium ones. This is useful for counterfactual analyses where advertising revenue is relevant. Since different types of articles accrue the same revenue *per reader*, accounting for the full consumption of each consumer takes into account the overall effects of the pricing policies of interest. For example, if we considered only the consumption of premium articles, then a pricing policy could appear profitable by increasing consumption of paid articles (and subsequently advertising revenue), but hide a substitution of fewer free articles consumed.

5 Results

Table 5 summarizes the maximum likelihood estimates of the model:

Table 5: Model Estimates

Parameter	Estimate
α	152.133*
	(15.943)
λ_v	122.457**
	(5.988)
σ_w	2.455**
	(0.016)
ρ	-0.672**
	(0.015)
γ	0.01**
	(0.000)
N: 10,049	
Objective Function:	0.028

Note: Standard errors in parentheses. Significance levels: † p≤0.10, * p≤0.05, ** p≤0.01. During estimation, parameters λ_v and σ_w were applied an exponential transformation to impose a positive support. Parameter ρ was kept between -1 and 1 through transformation $2/(1+\exp(-\rho^{'}))$-1. Standard errors obtained via 50 bootstrap samples. Figures above rounded to three decimal places with trailing zeros omitted.

All parameter signs are in line with expectation. Consumers draw positive value from news consumption ($\alpha > 0$), and exhibit a standard deviation in w.t.p. of 2.46 euros. The scale parameter of consumption is positive as expected, and about 1% of consumers exhibit a non-zero willingnesses to pay for online news. While this number appears small, when multiplied by the multiple millions of potential consumers in the market, 1% of consumer, it still results in hundreds of thousands of potential customers.

It is useful to understand the extent to which subscriptions are driven by consumption value alone. The value of parameter α translates the number of articles consumed to utility units. We interpret the magnitude of this parameter by analyzing its effect on demand. We consider the case in which consumers subscribe only due to direct preferences rather than WTP effects. In this case, the ratio of demand between the discounted price and the promotional price would be given by the expected ratio of logit expressions:

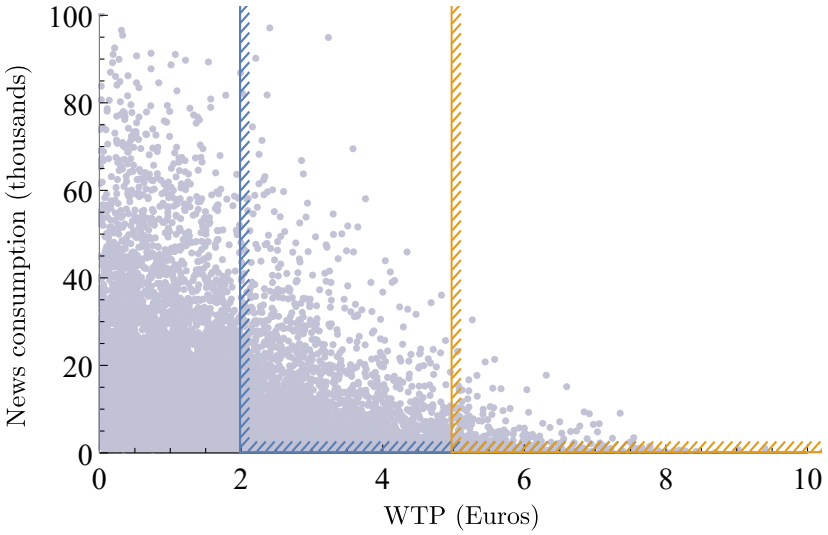
$$\int_0^\infty \frac{\frac{1}{1+\exp\left(-\left(\widehat{\alpha}v-p^d\right)\right)}}{\frac{1}{1+\exp(-(\widehat{\alpha}v-p^r))}} f_v\left(v;\widehat{\lambda_v}\right)dv = \int_0^\infty \frac{1+\exp\left(-\left(\widehat{\alpha}v-p^r\right)\right)}{1+\exp\left(-\left(\widehat{\alpha}v-p^d\right)\right)} f_v\left(v;\widehat{\lambda_v}\right)dv \tag{17}$$

$$= 1.0108 \tag{18}$$

In other words, a rough interpretation for the estimate of α is that a price reduction from €4.99 to €2 translates to a short-term increase in market share – as determined by consumption value – of approximately 1.1%.

The main parameter of the model, ρ, should be interpreted as a relatively strong negative correlation between news consumption and willingness to pay. Figure 6 plots simulations of $\{v, w\}$ pairs in at the estimated parameter values.

Figure 6: Scatter plot of consumption and willingness-to-pay



Note: Scatter plot obtained from simulating 10 million joint draws of v_i and w_i.

The figure reveals a clear negative correlation between news consumption and willingness-to-pay, in line with the negative estimate of parameter ρ. Although the negative value of ρ is necessary for a price discount to lead to an increase in the average news consumption level, it is not sufficient. To validate whether the model predicts such an effect, we turn to comparing the moments of the data with the ones predicted by the model.

Comparison of Model and Data Moments. We now compare the data moments with the ones predicted by the model.

Overall, the model fits the data moments reasonably well. The relatively low market shares – owing to the large potential market – are closely matched by the model for both segments. As for articles read, all moments are well approximated, with the largest deviation being an underestimation of the average number of articles read by the promotional segment and an overestimation of the standard deviation. Importantly, the model is able to replicate two most fundamental patterns in the data: First, both standard deviations of readership

Table 6: Comparison of Data and Predicted Moments

	Data Moments	Model Prediction
Subscription at regular price (N: 3,569)		
Subscription rate:	0.02%	0.02%
Articles read (Mean):	3.028	3.132
Articles read (Std. Deviation):	5.031	4.693
Subscription at promotional price (N: 6,480)		
Subscription rate:	0.21%	0.208%
Articles read (Mean):	9.46	8.476
Articles read (Std. Deviation):	9.468	10.5

Note: Figures above rounded to three decimal places with trailing zeros omitted.

exceed the segment means and, second, the average readership is higher for the promotional segment than the regular one. Matching these two patterns lends credibility to the model in terms of capturing the economic forces at play. We now turn to the counterfactual analyses.

6 Counterfactuals

6.1 Effect of correlation parameter ρ

Marketing scientists often perform demand analyses ignoring potential correlations between willingness to pay and optimal consumption levels. In this section we document the effects of the correlation parameter ρ on the model estimates and its implications.

Table 7: Model Estimates

Parameter	Original Estimate	Estimate with $\rho = 0$
α	141.5*	174.779**
	(63.124)	(5.481)
λ_v	124.04**	16.128**
	(6.484)	(0.524)
σ_w	2.444**	2.386**
	(0.017)	(0.02)
ρ	-0.67**	–
	(0.016)	–
γ	0.01**	0.011**
	(0.000)	(0.000)
N: 10,049		
Objective Function:	0.028	0.68

Note: Standard errors in parentheses. Significance levels: † p≤0.10, * p≤0.05, ** p≤0.01. During estimation, parameters λ_v and σ_w were applied an exponential transformation to impose a positive support. Parameter ρ was kept between -1 and 1 through transformation $2/(1 + \exp(-\rho^{'}))$-1. Standard errors obtained via 50 bootstrap samples. Figures above rounded to three decimal places with trailing zeros omitted.

Table 7 compares side-by-side estimates of the model parameters with and without a normalization of the correlation parameter ρ to zero. When the correlation between preferences and w.t.p. is assumed away, we obtain a larger estimate of parameter α, associated with the extensive margin (e.g., the act of subscribing), and a much lower estimate of parameter λ_v, associated with the intensive margin of consuming news. When parameter ρ is normalized to zero, the model attempts to match the moments of the data by severely underpredicting the demand for news consumption. To understand this effect, it is useful to compare the moments of the constrained and unconstrained models. The new predicted moments are presented in Table 8, together with those from the data and the original model. While the market share moments remain well matched, we observe large differences for the news consumption predictions. These differences are relatively modest for consumers subscribing at the regular price, but the model now underpredicts the mean and the standard deviation of articles consumed by subscribers at the promotional price by a factor of two. In trying to match readership levels across segments, the model finds a middle ground in terms of mischaracterizing readership patterns for both segments: It underpredicts mean readership of the promotional segment and overpredicts readership of the regular one. Straddling both

Table 8: Comparison of Data and Predicted Moments

	Data Moments	Model	Model, $\rho = 0$
Regular Subscribers (N: 3,569)			
Subscription rate:	0.02%	0.02%	0.019%
Articles read (Mean):	3.028	3.132	4.062
Articles read (Std. Deviation):	5.031	4.693	5.188
Promotional Subscribers (N: 6,480)			
Subscription rate:	0.21%	0.208%	0.214%
Articles read (Mean):	9.46	8.476	4.043
Articles read (Std. Deviation):	9.468	10.5	5.222

Note: Figures above rounded to three decimal places with trailing zeros omitted.

readership levels is now impossible, due to the normalization of ρ. In effect, the model can no longer predict that subscription levels increase with the price reduction. Given this, it finds a middle ground between fitting the first readership moment of the segments, failing to match either.[4]

6.2 Effect of Promotional Price Level on Profits

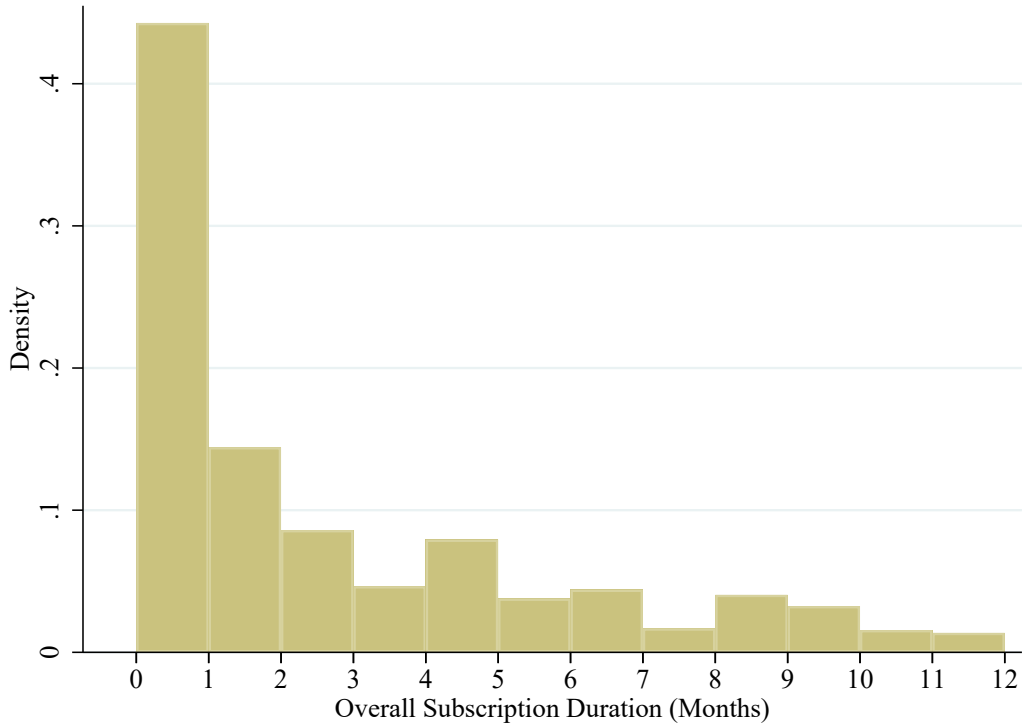
We now simulate firm profitability at different price promotion levels. In line with the data, we consider the consumers acquired during the estimation period of the data (May 6^{th} until June 15^{th} 2015), with the regular price of €4.99 being practiced until June 9^{th}, followed by a promotional price – which we vary in the counterfactual analyses – starting June 10^{th} until June 15^{th}. We first consider the effect of different promotional prices over the course of a year. Given that the firm's cost structure is relatively constant with the number of premium subscribers, we think of the revenues below as very good proxies for firm profits.

We consider the effects of the prices practiced during the window above on the profitability of customers up to one year after their subscription dates. So far, we have opted to model subscription decisions and consumption rates. In order to forecast subscription revenues over the course of a year, it is also useful to incorporate consumer churn into the analysis. By construction, consumption levels are a good predictor of consumer churn, and so we use them in the forecasting model. Indeed, the correlation between duration until churn and number of articles read is 0.42, which is to be expected since consumption decisions are intimately linked with churn decisions. The working assumption for the counterfactual analyses is that the relationship between subscription durations and consumption during that period is

[4]The second moment is actually matched better when $\rho = 0$ for the regular segment, but at a higher mismatch for the same second moment in the case of the promotional segment.

stable.[5] We plot the histogram of churn dates in Figure 7, in order to understand whether there are additional factors that may be taken into account in this prediction model.

Figure 7: Histogram of Churn Durations since Consumers' Subscription Dates



Note: Histogram above represents only consumers who churned within a 1-year period following the subscription decision, i.e., approximately 41% of subscribers.

The histogram above reveals a monotonously decreasing churn risk over time since the initial subscription. The monotonic decreasing trend is amenable to being modeled by a Cox survival model to predict how long consumers stay with the firm over the course of a year, and ultimately, how that translates to yearly revenue (consumers who churn earlier produce less subscription revenue).

We regress a Cox survival model on the churn data, including as predictors the total number of articles read by a consumer (i.e., the consumer type) as well as dummy variables to capture potentially higher churn risk near exact month thresholds. Table 9 reports

[5]Note that in the data there exists heterogeneity in terms of how fast consumption takes place over time across individuals. We abstract from this effect to keep the counterfactual analyses parsimonious. Because we are interested in overall profitability rather than its distribution across consumers, we believe that an average analysis, conditional on the regressors, provides a good first-order approximation to the total profitability effects. Alternatively, one would simulate multiple churn dates from the Cox model conditional on predictors to take heterogeneity into account, and average the revenues across simulations to predict the revenue of each individual.

Table 9: Cox Estimated Hazard Rates

Regressor	**Parameter Estimate**	**Standard Error**	
Number of Articles (×1000)	-0.249**	0.000	
Log-likelihood	-35,046.73		
Conditional Baseline Survival/Renewal Probabilities			
1^{st} month	0.609	7^{th} month	0.266
2^{nd} month	0.487	8^{th} month	0.255
3^{rd} month	0.417	9^{th} month	0.230
4^{th} month	0.381	10^{th} month	0.211
5^{th} month	0.322	11^{th} month	0.202
6^{th} month	0.296	12^{th} month	0.197

Note: Standard errors in parentheses. Significance levels: † p≤0.10, * p≤0.05, ** p≤0.01. Figures above rounded to three decimal places with trailing zeros omitted.

the results of the Cox regression model as well as the baseline survival probabilities. As expected, the parameter estimate associated with the number of articles read is negative, which means that the risk of churning decreases with consumption. The conditional baseline survival probabilities fall over time, such that the baseline probability (before adjusting for consumption) of renewing equals approximately 61% after one month and 19.7% after 12 months. We calculate consumer i's expected number of subscription payments as for the twelve renewal decisions plus the initial subscription decision as:

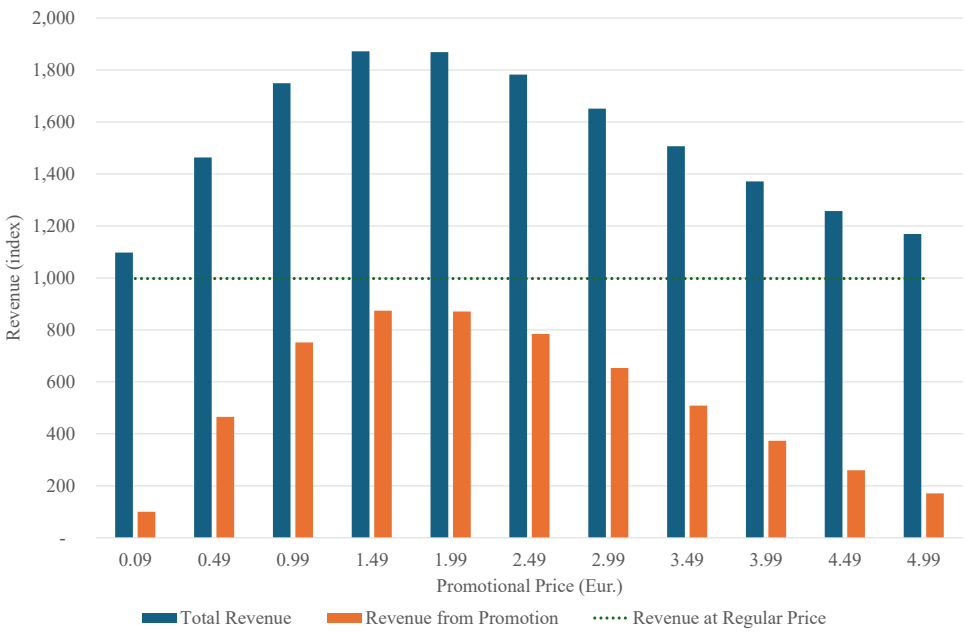
$$n_i = floor\left(\sum_{t=1}^{12} S\left(t\right)^{\exp(\beta v_i)}\right) + 2 \tag{19}$$

where each term $S\left(t\right)^{\exp(\beta v_i)}$ is the conditional baseline survival probability of period t, adjusted by the number of articles read by consumer i weighted by its parameter estimate, following the Cox model specification. We employ the floor function to the sum and add two units to count fractional month occurrences as a single payment and to take the initial subscription into account as well. Parameter n_i is used during counterfactual analyses to construct an estimate of yearly subscription revenue from each consumer.

Figure 8 presents counterfactual firm revenues over the course of a year, at different levels of the promotional price.[6]

[6]Note that in the original promotion consumers lock in the promotional price. We keep this feature in the counterfactual analysis: Consumers may renew their subscription at price they originally subscribed at.

Figure 8: Subscription Revenue as a Function of Promotional Price



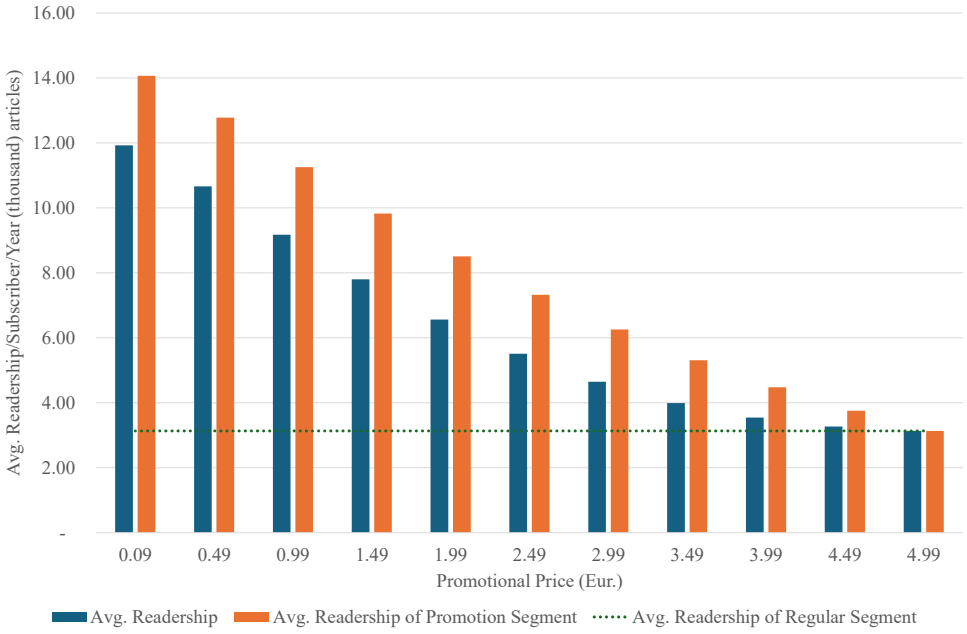
Note: Above, revenue of promotional segment at the €0.09 price level normalized to 100. The results above pertain to the sample period.

Revenue is concave on price, surprisingly, achieving its maximum at the actual promotional price observed in the data. We remind the reader that there is not structural or modeling reason that imposes this result. It seems that the firm's promotional price was set very close to the revenue-maximizing level. Above, we plot a dashed line as a reference for the regular-price period. In comparison, it is clear that the revenue from the promotional segment is quite significant, especially when the promotional price is relatively similar to €1.99. Finally, when no promotion is introduced (rightmost bars), the difference between promotion revenue and the dashed lines are explained by the difference in regular and promotion periods alone (35 vs. 6 days).

The revenue from the promotion is asymmetric around the €1.99 level. At values near €1.99, the revenue curve is shallower to the left, meaning that it is worth overdiscounting by a little than underdiscounting. However, when larger deviations from €1.99 are considered, we notice that revenue decreases more slowly with underdiscounting than overdiscounting. So, from a managerial perspective, it is best to make small overdiscounting errors than large overdiscounting ones. Another way to think about this result is that, when managers are experienced or have high-quality information about discount effects, and so their pricing mistakes tend to be small, they are better off undercutting the optimal price. In contrast, when

they are less experienced or their information is poor, they are better off underpromoting.

Figure 9: Average Readership as a Function of Promotional Price



Note: Above, simulated average readership levels over the course of a year after subscribing as a function of promotional price level, owing to selection into subscribing.

Figure 9 shows counterfactual average readership levels for both subscriber groups. Across the board, readership levels increase with the promotion magnitude, as a result of the higher uptake of avid readers. If consumer welfare is to be measured by consumption intensity, then there is no doubt that higher promotions lead to gains in welfare.

The figure above also reveals that the difference in readership levels increases at an increasing rate as the promotional price decreases. The reason is that a lower promotion increases the subscriber base while simultaneously attracting customers each of which consume more, on average. Because each news article consumed is positively associated with advertising revenue, a conclusion from this analysis is that the promotional segment drives more advertising revenue than the regular one.[7] The extent to which the promotional segment generates higher ad revenues depends on the price promotion level practiced by the firm, in line with the results in Figure 9.

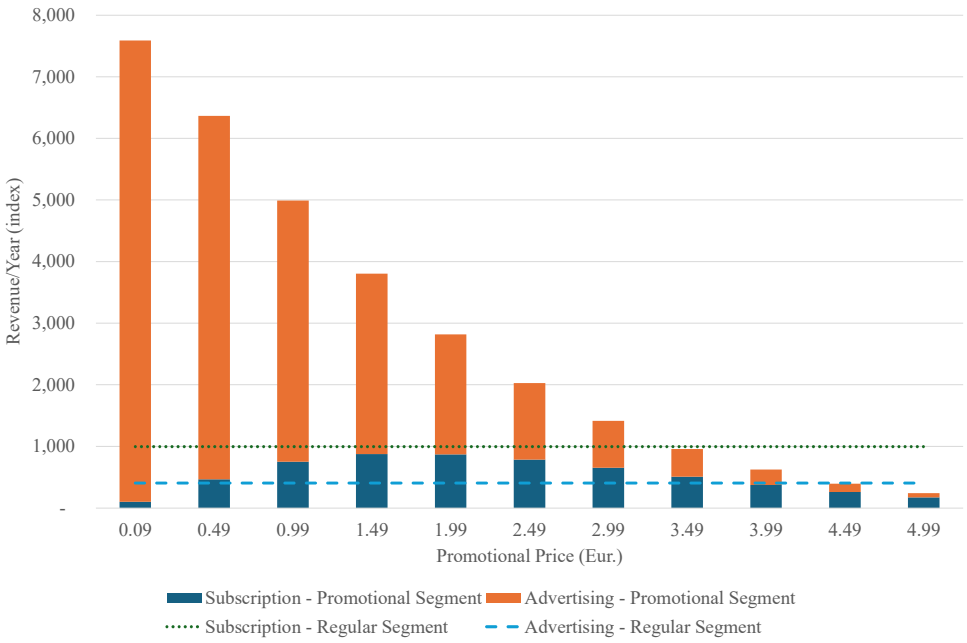
[7]Note that advertisers on this publisher pay per impression and not per click, so the critique of different unobserved ad-click behaviors by different segments as a result of the price promotion has no bearing in our setting.

6.3 Revenue from Increased Consumption via Advertising

In contrast with traditional retail markets, many firms in digital domains accrue rents via consumption, either directly (e.g., software as a service) or indirectly (e.g., advertising income). We have already shown that price promotions alter the composition of consumers, leading to different subscription and consumption patterns. The goal of this section is to quantify the extent to which that composition affects the firm's performance. We focus on subscription and advertising revenues generated within one year of individuals' subscription decisions in the sample period, or from earlier dates for consumers who churn before that time frame.

Advertising rates vary wildly across countries, industries, and types of publishers. The counterfactual scenario we consider here assumes that the advertising revenue associated with the consumption of a news article remains constant across price promotion levels. The managerial team of the news publisher indicated the incremental revenue of €0.005 per article consumed, which we will use as a focal value for our analysis. This value allows us to add the subscription and advertising revenues in order to understand the combined effect of price promotions.

Figure 10: Yearly Revenue Decomposed by Segment and Source

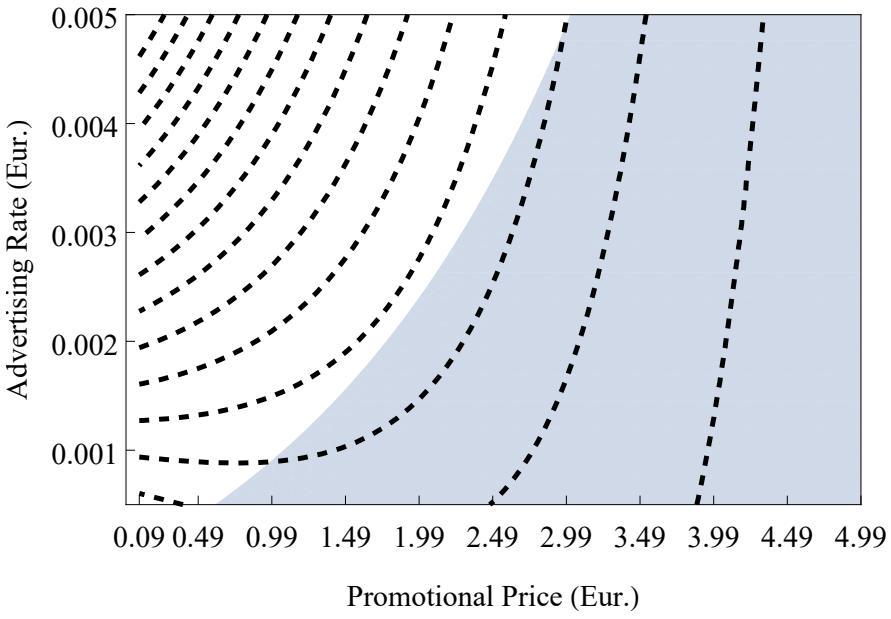


Note: Above, subscription revenue of promotional segment at the €0.09 price level normalized to 100. The results above pertain to the period of one year after individual subscriptions.

Figure 10 decomposes revenues across segments and sources (advertising and subscription), by segment (promotional and regular). The most striking result is the fact that advertising revenues make up a small portion of revenue for the promotional segment when the promotional price sits above €2.99. However, at lower promotional prices the advertising revenue of the promotional segment increases at an increasing rate, rapidly outweighing the other revenue sources. We would like to de-emphasize results obtained at very low promotional values for two reasons. First, while the merit of structural models is to allow extrapolation over the original sample support, a promotional price of €0.09 may nonetheless sit too far from the original support of the data to produce reliable prediction. Second, while we were provided with a constant monetary value that may remain valid close to the observed data, but may no longer apply when the subscriber base increases substantially, since it would also be associated with a much larger ad impression stock. These two reasons require prudence judging the scenarios with very low promotional prices. Nonetheless, the promotional price of €0.09 is observed in the data, and at that level we can observe that the revenues generated by the promotional segment already dominate the regular segment. It is clear that the promotion generated value, primarily generated through advertising revenue.

We now perform a sensitivity analysis to the advertising rate provided by the managerial team by considering situations in which the ad rate is lower than €0.005 rate used in the previous analysis. We consider the interval [0.0005,0.005] euros per news article read for the analysis. In particular, we investigate the robustness of the result that the customers attracted via the price promotion are more valuable due to their advertising rents rather than their subscription payments.

Figure 11: Revenue Decomposition as a Function of Promotional Price and Advertising Rate



Note: Above, shaded area corresponds to the combinations of promotional price and advertising rate where revenues from subscriptions dominate ad revenue. Contour lines refer to total revenue (subscription plus advertising), with higher values located to the top and left. Advertising rates range from €0.0005 to €0.005 per news article read.

Figure 11 plots the region in which subscription revenues dominate advertising revenue, as a function of promotional price and advertising rates. In addition, it overlays indifference curves for the firm, with firms to the left being associated with higher revenues. Across the range of advertising rates considered we find that at a low enough promotional price, advertising rents will always dominate the revenue from subscriptions. However, given the caveats related with low promotional prices, we believe it is more informative to focus nearer to the promotional price observed in the data of €1.99. In this neighborhood, it is clear that the question of which source of revenue dominates depends crucially on the advertising rate. At approximately €0.0025 per article read the gains from the promotion are approximately evenly split between subscription and advertising revenues. This means that even at half the estimate provided by the managerial team for the advertising rate, rents from advertising sourcing from the price promotion are extremely significant for the company from a revenue point of view. It is clear that a central effect of the price promotions is to boost advertising revenue through increased future consumption. In the following section we incorporate the

effect of subscription levels on advertising rates explicitly.

6.4 Charging Less for More Consumption

Subscription pricing is often complemented by second-degree pricing practices, that is, the existence of different menus with combinations of price-quantity/quality offers that consumers select into. Second-degree pricing can be profitable specifically when a firm is able to effectively separate consumers by ensuring incentive compatibility. This idea is more complicated in the setting we analyze, because the 'best consumers' (i.e., those who would like to consume the most) are not necessarily the 'best customers' (i.e., those who are willing to pay the most) from a subscription revenue standpoint. For example, offering a menu that includes more quantity at a higher price may fail to attract enough consumers due to the negative correlation between consumption value and willingness-to-pay.

We consider the introduction of a different type of second-price discrimination that includes two menus: one with unlimited consumption and another one with a capped consumption, both sold at different prices. The unique feature of this plan is that the unlimited consumption plan is accessible at a lower price – via a discount – than the capped plan, provided consumers reach the minimum consumption threshold. Offering a plan that offers more quantity at a lower price can only be effective if the firm can detect true consumption (e.g., consumers reading articles) from fake consumption (e.g., consumers using bots to 'read' articles to obtain a discount). Given the recent trends in consumer monitoring technology (see for example Jiang, Li, Chen, and Wang (2018) and Delouya (2024)), we believe such a future is viable and worth contemplating.[8]

In the plan described above, consumers all start at a regular subscription price, and obtain a discount if they consume more than a given number of articles. This model makes sense to firms operating subscription products whose usage brings additional revenue to them, as in our context. To emphasize the uniqueness of the plan we inspect in this section, the central idea is that consumers pay a lower *total* price for reading more articles than a preset threshold, which is different from classical unit-discount settings in which more quantity is associated with marginal positive prices. In conclusion, the plan we consider applies when it is costly or impossible for consumers to artificially boost product consumption, and when the marginal cost of consumption is negligible for the firm.

We revisit the promotional campaign of the firm and search for the optimal minimum consumption threshold t^* that earns consumers an amount a^*. The idea is that customers who consume a number of articles superior to t^* over the span of one year earn a discount,

[8]Relatedly, in the mobile games industry, for example, apps often provide rewards for intensive and/or consecutive usage in part to bolster advertising revenue.

be it through a direct monetary transfer or by earning a discount towards a renewal of their subscription, for example. For the counterfactual analysis, we maintain the promotional schedule and subscription prices of the original dataset and search for the optimal levels of t^* and a^* that maximize the firm's joint subscription and advertising revenues. For those consumers with consumption level $v_i \geq a^*$, we change the willingness-to-pay from w_i to $w_i + t^*$. We keep the indirect utility constant under the interpretation that the 'windfall' t^* keeps the utilitarian component constant, but may affect consumers' willingness-to-pay via a mental account, for example. While this approach focuses on the w.t.p. mechanism, ignoring the utility component may lead to an underestimation of the total profits from the price discrimination scheme.

The introduction of a minimum consumption threshold is analogous to the different problem of motivating salespeople with variable compensation schemes. In that context, the existing literature has documented that salespeople whose performance falls just below a given compensation cutoff often exert more effort to surpass it (see for example ... and ...). Our analysis abstracts away from this behavioral pattern, which most likely would also apply to consumption in subscription settings, as it is better suited for dedicated future research efforts. Hence, as before, we maintain the assumption that consumption levels are intrinsic to customers' own types.

We assume a constant-elastic relationship between the advertising price and overall readership:

$$r(V) = r_0 \left(\frac{V_0}{V} \right)^{\eta} \tag{20}$$

where r_0 is the base unit advertising price (€0.005 per article read), V is the number of articles read over the duration of a year, V_0 is the number of articles read at the regular price (from the data), and finally η is the advertising elasticity. We consider the range $\eta \in [0.12, 0.2]$ for the elasticity parameter, which spans the estimates of the meta analyses by Sethuraman, Tellis, and Briesch (2011) and Schöndeling, Burmester, Edeling, Marchand, and Clement (2023). This specification takes into account that different levels of readership are associated with different per-unit prices of advertising. In other words, we take into account the fact that large changes to readership are necessarily accompanied by changes in advertising rates.

The revenue of the firm net of the program cost is given by:

$$\begin{aligned}
\pi = {} & p\left(D_{disc}\left(p, a^*, t^*\right) + D_{regular}\left(p, a^*, t^*\right)\right) && \text{(Subscription Revenue)} \\
& + r\left(\overline{V}_{all}\right).\left(S_{disc}\left(p, a^*, t^*\right) + S_{regular}\left(p, a^*, t^*\right)\right) && \text{(Advertising Revenue)} \\
& - t^* D_{disc}\left(p, a^*, t^*\right) && \text{(Discount Payments)}
\end{aligned}$$

where p is the regular price of €4.99, $D_{(\cdot)}$ denotes the total number of subscription payments made by the regular and discount segments, S_{disc} denotes the number of new subscribers of those same segments, and parameter $\overline{V}_{all}$ is the average readership across subscribers. For clarity, we present the demand $D_{disc}(\cdot)$ of the consumers who obtain the discount:

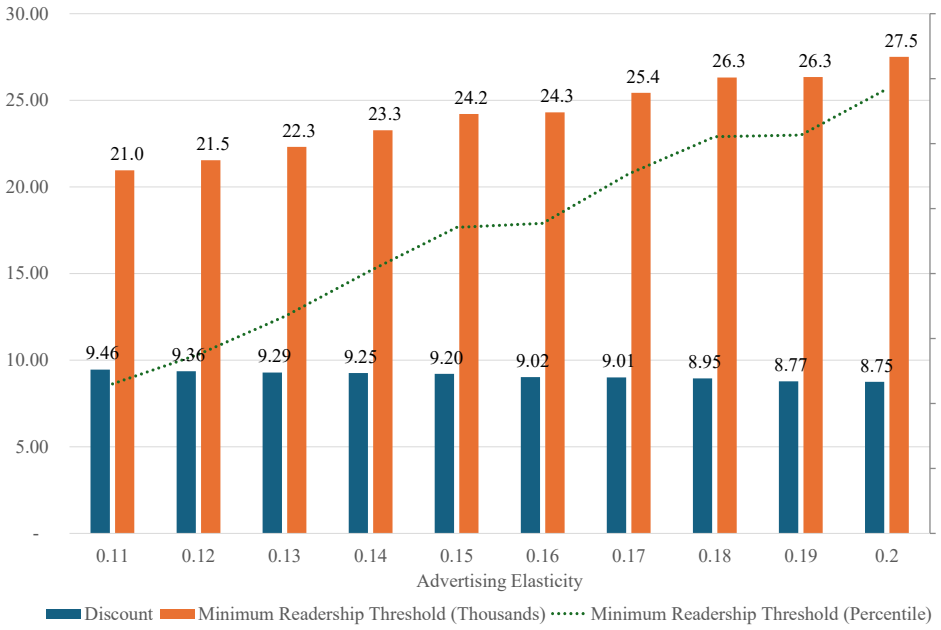
$$D_{disc}(p, a^*, t^*) = \underbrace{M\gamma}_{Potential\ Market} \underbrace{.\frac{1}{R}\sum_{i=1}^{R} 1\left(v_i \geq a^* \wedge \alpha v_i + \varepsilon_i^1 \geq p_i + \varepsilon_i^0 \wedge w_i + t^* \geq p_i\right).}_{Market\ Share} \underbrace{n_i}_{N.\ Payments} \tag{21}$$

where R is a preset large number of simulations. Similarly, the number of new subscribers in the discount plan is given by

$$S_{disc}(p, a^*, t^*) = \underbrace{M\gamma}_{Potential\ Market} \underbrace{.\frac{1}{R}\sum_{i=1}^{R} 1\left(v_i \geq a^* \wedge \alpha v_i + \varepsilon_i^1 \geq p_i + \varepsilon_i^0 \wedge w_i + t^* \geq p_i\right)}_{Market\ Share} \tag{22}$$

For each advertising elasticity level η we maximize the publisher's profit π with respect to the minimum consumption threshold a^* and the discount level t^*. Figure 12 shows the optimal discount rate d^* and minimum readership threshold a^* as a function of advertising elasticity.

Figure 12: Discount Pricing and Consumption Threshold as a Function of Advertising Elasticity

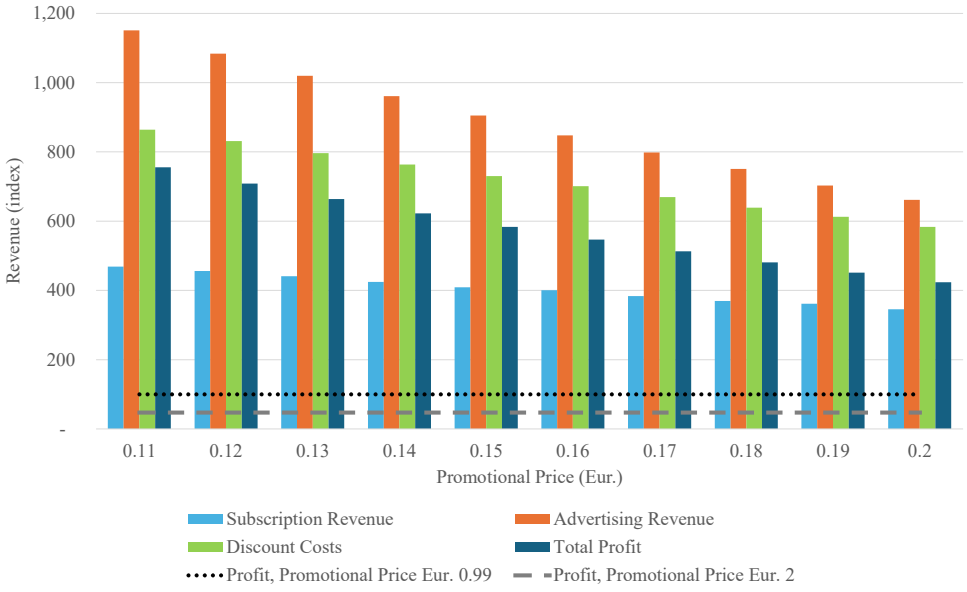


Note: The Y axis above is used to simultaneously represent Euros (short blue bars) and thousands of articles consumed (tall orange bars). The dotted line represents the same data as the orange bars but in percentile terms, and should be read on the right Y axis.

As the willingness to pay by advertisers becomes more sensitive to the size of the subscriber pool, the publisher has an incentive to increase the minimum consumption threshold and decrease the subscription discount. A striking result is the fact that the per-subscription discount is higher than the regular price (€4.99); in other words, if possible, the publisher is better off paying a few consumers to be part of its reading pool. As we show later, the loss incurred through the program cost is offset by the incremental gains in advertising revenue. Finally, the dotted line presents the minimum consumption threshold in terms of readership percentiles to provide better interpretability about the required readership levels among the population necessary to opt into the discount.

Figure 13 presents normalized profits of the discount program across advertising elasticity values. It also disaggregates profits by subscription revenue, advertising revenue, and program cost. In addition, the last bar of each cluster presents normalized profits of the program, which can be compared with the status quo represented in dotted and dashed lines. The status quo scenarios represent the firm's profit in the original promotional program of the data, with promotional prices of €0.99 (dotted line) and €2 (dashed line), as in the results presented in Figure 10.

Figure 13: Profit Decomposition of Discount Program



Note: Above, shaded area corresponds to the combinations of promotional price and advertising rate where revenues from subscriptions dominate ad revenue. Contour lines refer to total revenue (subscription plus advertising), with higher values located to the top and left. Advertising rates range from €0.0005 to €0.005 per news article read.

Inspecting each cluster above, we find that subscription revenues tend to represent less than half of advertising revenues. Moreover, the cost of the discount program is very significant, surpassing the magnitude of subscription revenues. The overall effect is depicted in the last bar of each cluster. As expected, the total profit of the discount program decreases with advertising elasticity. However, it is extremely profitable: It generates approximately between 9 and 16 time more profit than the discount program that the publisher ran during the sample period. The stark difference is impressive but perhaps not too surprising due to the fact that the discount program constitutes third-degree price discrimination. In the current scenario consumers do not maintain information rents and put pressure on the seller through menu selection due to the assumption that the publisher can effectively monitor real consumption by it subscribers, effectively engaging in third degree price discrimination. While not fully realistic, this counterfactual analysis shows that efforts in the direction of monitoring consumer relationships and rewarding them accordingly can be a great opportunity for firms managing subscription programs.

7 Concluding Remarks

We have investigated the relationship between subscription prices and consumption behaviors in the context of an online news publisher. In addition to analyzing simple behavioral patterns, we developed a model that allows for a flexible relationship between subscription behaviors and consumption, and used it to find significant opportunities for the firm via different pricing policies.

Our model takes into account substitution patterns that counter traditional economic theory, and its flexibility allows us to characterize consumer behavior that most likely was generated by competitive forces in the context of single-source datasets.

It is worth noting that the problem of linked decisions is pervasive in Marketing and Economics. In consumer search, consumers form expectations that link search decisions with purchase decisions. In discrete-continuous settings, consumers maximize their utilities subject to a budget constraint, which links their propensity to buy with their optimal quantity. Similarly, in subscription settings consumers' decisions of purchasing correlated with their subsequent usage. We believe that investigating the extent to which classical predictions from Economic theory regarding the extent to which consumers' decisions are linked, through empirical exercises, may allow researchers to advance empirical methods and also to revisit long-standing assumptions of consumer behavior.

8 Appendix

8.1 Illustrative Model of Subscriptions and Consumption

Consider the following simple illustrative model: A focal firm offers a number of news articles v at subscription price p. This means that, upon paying p, its subscribers can consume up to the v available articles free of charge. There also exists an "outside option:" A firm offering $v_o > v$ articles at price p_o (we use the case of a competitor news firm, but it could provide similarly entertaining content, such as an online games publisher). Each firm may already have a readership base, which we do not model here. We consider only two consumers, L and H, with different budgets and different readership preferences. Specifically, consumer H has a budget of w_H and derives utility from reading up to v_H articles (this is a simplified assumption in lieu of decreasing marginal utility). Consumer L has a budget of $w_L < w_H$ and derives utility from reading up to $v_L < v_H$ articles. Consumer i's utility from reading v_i articles is given by v_i. We assume $w_H > p_o > p > w_L$, such that only the first consumer can afford subscribing either news service. For example, consumer L may live from paycheck to paycheck, often reaching the end of the month with difficulty to pay his credit card balance. In the status quo case, we assume consumer H subscribes *only* the outside option, by assuming that $v_o - p_o > v - p$ and

$$\underbrace{v_H - v_o}_{additional\ news\ consumption} < p \tag{23}$$

The condition above means that the additional utility obtainable by consumer H from multihoming, $v_H - v_o$, is not enough to offset the subscription price p. While it is true that the focal firm offers v articles, due to diminishing returns, it turns out that $v_H < v$, such that consumer H will only draw utility from the extra $v_H - v_o$ articles offered by the focal firm.

Now, suppose that the focal firm reduces its price from p to p', such that $p' < w_L$. In this case, both consumers will subscribe if

$$v_H - v_o \geq p' \tag{24}$$

$$v_L \geq p' \tag{25}$$

In addition, for consumer L to consume more articles from the focal firm than consumer H, it suffices that

$$v_L > v_H - v_o \tag{26}$$

that is, the utility for additional news by the low type consumer is greater than that of the

high-type consumer.

For the effect above to take place, it is crucial that the low type consumer cannot or will not access the outside option at the same rate as the high type consumer. The result is that, in equilibrium, the marginal utility of the high-type consumer for the news of the focal firm is lower than that of the low type. From the focal firms' point of view, the result is a negative correlation between wealth ($w_H > w_L$) and consumption activity ($v_H - v_o < v_L$). This does not mean that the firm offers an inferior good. Rather, the negative correlation it observes is an equilibrium effect that potentially results from multiple forces.

In the model we incorporate the effects above by introducing an additional willingness-to-pay constraint, so that different consumers may be willing to pay for the subscription at different rates. More importantly, we allow consumers' willingness-to-pay to be correlated with consumption utilities, so that the relationship remains flexible and can be identified from the data.

8.2 Proof of Proposition 1

We start adding detailed mathematical structure to the proposition (omitted in the main text for readability):

Proposition 1 (with mathematical detail): *Let $u(\cdot)$ be a strictly increasing differentiable utility function and v be a continuous random variable with continuous support with defined first moment, and let p be a scalar. Then, $-\frac{\partial}{\partial p}P(u(v) \geq p) > 0$ and $-\frac{\partial}{\partial p}E(v|\, u(v) \geq p) < 0$.*

- **Purchase:** The effect of price on demand is given by

$$\begin{aligned}\frac{\partial}{\partial p}P(u(v) \geq p) &= \frac{\partial}{\partial p}\left(1 - F_v\left(u^{-1}(p)\right)\right) \\ &= -\underbrace{f_v\left(u^{-1}(p)\right)}_{(+)}\underbrace{\left(u^{-1}(p)\right)'}_{(+)} < 0\end{aligned}$$

  where the first positive sign follows from the fact that cumulative distribution functions are increasing, and the second positive sign results from the fact that the derivative of a function is equal to the reciprocal of the derivative of the corresponding inverse function.

- **Expected Consumption:** The expected consumption is given by $E(v|\, Subscribe)$.

The effect of a small price increase is given by

$$\begin{aligned}\frac{\partial}{\partial p}E\left(v|\,u\left(v\right)\geq p\right) &= \frac{\partial}{\partial p}\int_{u^{-1}(p)}^{\infty} vf_{v|u(v)\geq p}\left(v\right)dv = \frac{\partial}{\partial p}\frac{\int_{u^{-1}(p)}^{\infty} vf_v\left(v\right)dv}{P\left(u\left(v\right)\geq p\right)} \\ &= \frac{\partial}{\partial p}\frac{\int_{u^{-1}(p)}^{\infty} vf_v\left(v\right)dv}{P\left(u\left(v\right)\geq p\right)} = \frac{\partial}{\partial p}\frac{g\left(u^{-1}\left(p\right)\right)}{G\left(u^{-1}\left(p\right)\right)} \\ &= \frac{\left(u^{-1}\left(p\right)\right)'\left[G\left(u^{-1}\left(p\right)\right)g'\left(u^{-1}\left(p\right)\right)-g\left(u^{-1}\left(p\right)\right)G'\left(u^{-1}\left(p\right)\right)\right]}{G\left(u^{-1}\left(p\right)\right)^2} \\ &\propto G\left(u^{-1}\left(p\right)\right)g'\left(u^{-1}\left(p\right)\right)-g\left(u^{-1}\left(p\right)\right)^2\end{aligned}$$

where $g\left(u^{-1}\left(p\right)\right)\equiv\int_{u^{-1}(p)}^{\infty} vf_v\left(v\right)dv$ and $G\left(u^{-1}\left(p\right)\right)\equiv P\left(u\left(v\right)\geq p\right)$. The last proportionality relation above captures the fact that the sign of the derivative depends only on the last expression. Note that by Leibniz rule, $g'\left(u^{-1}\left(p\right)\right)=-u^{-1}\left(p\right)f_v\left(u^{-1}\left(p\right)\right)$ and that $G'\left(u^{-1}\left(p\right)\right)=-f_v\left(u^{-1}\left(p\right)\right)$, such that the last expression can be rewritten as

$$\underbrace{f_v\left(u^{-1}\left(p\right)\right)}_{(+)}\left[-u^{-1}\left(p\right)G\left(u^{-1}\left(p\right)\right)+g\left(u^{-1}\left(p\right)\right)\right] \tag{27}$$

Finally, note that

$$\begin{aligned}g\left(u^{-1}\left(p\right)\right) &= \int_{u^{-1}(p)}^{\infty} vf_v\left(v\right)dv > \int_{u^{-1}(p)}^{\infty} u^{-1}\left(p\right)f_v\left(v\right)dv \\ &> u^{-1}\left(p\right)\int_{u^{-1}(p)}^{\infty} f_v\left(v\right)dv = u^{-1}\left(p\right)G\left(u^{-1}\left(p\right)\right)\end{aligned}$$

so that $g\left(u^{-1}\left(p\right)\right) > u^{-1}\left(p\right)G\left(u^{-1}\left(p\right)\right)$, which in turn implies that (27) is indeed positive, proving the proposition.

8.3 Mechanisms

Below we consider mechanisms that are unlikely to play a central role in typical analysis pertaining to subscription settings.

Wealth Effects. These effects take place, formally speaking, when a focal purchase of a good affects the marginal utility of an outside good via a budget constraint (rather than through complementary utility, for example). The outside good is often assumed to be a composite good of all other consumption alternatives by the individual. Consider the following form:

$$\alpha v_i + U\left(w_i - p_i\right) \geq U\left(w_i\right) \tag{28}$$

For simplicity, function $U(\cdot)$ captures the utility of the outside good; it is an increasing concave utility function, such that $U''(\cdot) < 0$. One may also impose the condition $U(x) = -\infty$ when $x < 0$ to ensure that utility is only relevant for choice whenever the good sold is affordable, that is, $w_i > p_i$. Consumer i subscribes if and only if expression 11 is satisfied. As before, v_i is consumer $i's$ optimal news consumption level conditional on subscribing. The decision of whether to subscribe therefore rests on the utility gain from subscribing dominating the loss in outside utility, $U(w_i) - U(w_i - p_i)$.

A popular specification for the utility function is

$$U(x) = \frac{x^{1-\tau} - 1}{1 - \sigma} \tag{29}$$

where τ is the correlation of relative risk aversion, which can be sourced from the literature (cite, cite). This quick exercise shows that incorporating income effects to our application is straightforward and, as in our application, the results can be explained by allowing v_i and w_i to be negatively correlated.

There exist, however, reasons to believe that income effects are not involved in the patterns we observe. First, while formally appealing, it is challenging, even for a consumer, to determine her own budget constraint, which in addition to bank accounts and assets, may also include borrowing from friend and family, using credit cards, etc. Second, there exist theoretical results that show income effects are expected to play a minor role for small purchases. One stream, based on the analysis of classical demand functions (see the discussion in Vives, 1987 as well as Hayashi, 2008), concludes that income effects decrease rapidly as the price of the good decreases as well. As a result, they are often negligible when the good in question represents only a small share of the individual's income. A different criticism is provided in Rabin (2000), who explains that a degree of risk aversion at small stakes that is deemed reasonable implies an unreasonably-high degree of risk aversion at high stakes, implying that individuals must refuse larger bets regardless of how favorable they are. This result also implies that risk aversion, or income effects as captured by the curvature of the utility function, should not play a discernible role in small purchases.

Reference-dependent preferences. A possible alternative explanation is the role of loss aversion, which may indeed act even in small stakes (Tversky and Kahneman, 1991). It is indeed possible that consumers in this market were used to a reference price of €4.99, and realized an extra-monetary gain by observing the promotional price of €2. Their decision to

subscribe may be formalized as

$$\alpha v_i - p^d + \lambda \left(p^r - p^d\right) \geq 0 \tag{30}$$

where $\lambda \left(p^r - p^d\right)$ is a reference-dependent utility component with loss-aversion parameter $\lambda > 0$. While this effect can indeed play a role in subscription decisions, it is unclear whether it would imply that price promotions would attract consumers who would like to consume more news. Rather, one would need to make λ_i individual-specific, and have it correlate with v_i appropriately. For example, we would need to state that consumers who would like to consume more news also tend to perceive higher gains from price promotions. While we cannot reject this hypothesis, the lack of empirical evidence for this fact means we lack motivation to put loss aversion forward as a main mechanism.

Habit Formation. This mechanism does not speak directly to the correlation between subscription prices and consumption. Rather, it pertains to consumption dynamics taking place after the initial consumption periods following the subscription. The prediction that is most relevant for our context is that, over time, the consumption intensity of new subscribers will increase as a consequence of their past consumption and that, in addition, consumption differences between two groups of consumers who faced different subscription prices are expected to diverge, at least momentarily, due to behavioral reinforcement (see for example Wood and Neal (2009)).

References

Arkes, H. R., and C. Blumer (1985): "The psychology of sunk cost," *Organizational behavior and human decision processes*, 35(1), 124–140.

Bhat, C. R. (2005): "A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions," *Transportation Research Part B: Methodological*, 39(8), 679–707.

Chintagunta, P. K. (1993): "Investigating purchase incidence, brand choice and purchase quantity decisions of households," *Marketing Science*, 12(2), 184–208.

Chintagunta, P. K., and H. S. Nair (2011): "Structural workshop paper-discrete-choice models of consumer demand in marketing," *Marketing Science*, 30(6), 977–996.

Chou, C., and V. Kumar (2024): "Estimating Demand for Subscription Products: Identification of Willingness to Pay Without Price Variation," *Marketing Science*.

Conlisk, J., E. Gerstner, and J. Sobel (1984): "Cyclic pricing by a durable goods monopolist," *The Quarterly Journal of Economics*, 99(3), 489–505.

Danaher, P. J. (2002): "Optimal pricing of new subscription services: Analysis of a market experiment," *Marketing Science*, 21(2), 119–138.

Deaton, A., and J. Muellbauer (1980): *Economics and consumer behavior*. Cambridge university press.

Delouya, S. (2024): "Netflix cracked down on password sharing. The result? Millions of new subscribers," *CNN-Business (Ed.), https://edition.cnn.com/2024/04/18/business/netflix-earnings-first-quarter/index.html*.

Esteves-Sorenson, C., and F. Perretti (2012): "Micro-costs: Inertia in Television Viewing," *The Economic Journal*, 122(563), 867–902.

Greene, W. H. (2000): *Econometric Analysis*. Prentice Hall Inc.

Hartmann, W. R., and V. B. Viard (2008): "Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program," *Quantitative Marketing and Economics*, 6, 109–137.

Hayashi, T. (2008): "A note on small income effects," *Journal of Economic Theory*, 139(1), 360–379.

Heckman, J. J. (1979): "Econometrica," *Econometrica*, 47(1), 153–161.

Jiang, J.-Y., C.-T. Li, Y. Chen, and W. Wang (2018): "Identifying users behind shared accounts in online streaming services," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp. 65–74.

Kim, J., G. M. Allenby, and P. E. Rossi (2002): "Modeling consumer demand for variety," *Marketing Science*, 21(3), 229–250.

Klemperer, P. (1987): "Markets with consumer switching costs," *The quarterly journal of economics*, 102(2), 375–394.

Nelder, J. A., and R. Mead (1965): "A simplex method for function minimization," *The computer journal*, 7(4), 308–313.

Rabin, M. (2000): "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68(5), 1281–1292.

Runge, J., J. Levav, and H. S. Nair (2022): "Price promotions and "freemium" app monetization," *Quantitative marketing and economics*, 20(2), 101–139.

Schöndeling, A., A. B. Burmester, A. Edeling, A. Marchand, and M. Clement (2023): "Marvelous advertising returns? A meta-analysis of advertising elasticities in the entertainment industry," *Journal of the Academy of Marketing Science*, 51(5), 1019–1045.

Sethuraman, R., G. J. Tellis, and R. A. Briesch (2011): "How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities," *Journal of marketing research*, 48(3), 457–471.

Thaler, R. (1985): "Mental accounting and consumer choice," *Marketing science*, 4(3), 199–214.

Tuchman, A. E., H. S. Nair, and P. M. Gardete (2018): "Television ad-skipping, consumption complementarities and the consumer demand for advertising," *Quantitative Marketing and Economics*, 16(2), 111–174.

Tversky, A., and D. Kahneman (1991): "Loss aversion in riskless choice: A reference-dependent model," *The quarterly journal of economics*, 106(4), 1039–1061.

Villas-Boas, J. M. (2015): "A short survey on switching costs and dynamic competition," *International Journal of Research in Marketing*, 32(2), 219–222.

Vives, X. (1987): "Small income effects: A Marshallian theory of consumer surplus and downward sloping demand," *The review of economic studies*, 54(1), 87–103.

Wood, W., and D. T. Neal (2009): "The habitual consumer," *Journal of Consumer Psychology*, 19(4), 579–592.