Pricing and Consumption in Subscription Settings*

Pedro M. Gardete, Daniela Schmitt, Florian Stahl

February 2025

Abstract

This paper examines how subscription pricing impacts usage intensity, a critical factor for firms relying on subscription business models. Our analysis focuses on data from an online news publisher, a context where promotional pricing is widely used as a way to attract new customers, although its overall impact is unclear. Traditional analyses suggest that lower prices reduce 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 behaviors. 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 corresponds to a negative correlation between subscribers' consumption values and their 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). Our analysis finds that understanding the impact of subscription pricing on future consumption can unlock significant economic value for firms.

*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, food delivery providers or 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 actual consumption are especially important, since firms often incur costs and/or derive value from the intensity of product consumption.

The relationship between subscription fees and subsequent usage is nuanced and can involve various mechanisms. One 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 for how much they would like to consume. The traditional prediction from economic theory is 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' consumption intensities. When price decreases, for example, 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 its current customers do. 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, Hayashi, 2008), suggesting that consumption is more likely to be determined by preferences than by income effects induced by budget constraints and diminishing marginal utility of consumption of the outside good.[1] In effect, lower prices can lead to lower consumption across various contexts. For example, in the contexts of a physical newspaper subscription and an online

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

grocer, Lewis (2006) finds that discount subscribers exhibit lower repurchase rates. Just and Wansink (2011) find that a 50% price discount to an all-you-can-eat meal led participants to consume 28% less pizza, and attribute this behavior to the sunk cost fallacy. In line with these findings, Datta, Foubert, and Van Heerde (2015) find that free-trial consumers of a digital television service are significantly less valuable than regular customers.

While these focal predictions are admittedly precise, they are by no means general, as there also exist economic rationales predicting that marginal subscribers will consume more than existing ones. For example, consumers who take advantage of promotions may belong to specific populations (e.g., retired or unemployed) who feature both a lower willingness to pay and a lower opportunity cost of time. In such cases, a firm may find that subscribers acquired through a price promotion may consume more than those attracted through its regular price. Relatedly, in many contexts (i.e. streaming services, online news) price promotions may primarily attract those consumers who do not yet subscribe to competing products in the category. Subscribers of competitive products may be less responsive to price promotions due to switching costs or even costs of maintaining and monitoring multiple accounts, i.e., multihoming costs. Low prices, therefore, may appeal to new subscribers who lack multihoming opportunities and so are willing to consume more from the focal firm.

Given the complexity of mechanisms at play, empirical analysis is required to understand the dynamic relationship between subscription prices and subsequent consumption levels. In this paper, we investigate the consumption patterns of news of a leading European online news publisher offering both free and premium articles, the latter accessed 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 articles than consumers who subscribed a few days before the promotion at the regular price.

Two striking behavioral patterns are present in our data. First, the higher consumption by new subscribers is partially explained by their lower churn rate, yet the higher consumption level persists even after conditioning only on active users: Conditional on remaining active, promotion subscribers consume more news than their regular-price counterparts. Second, the difference in consumption levels remains stable over time, 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 service.

We believe there already exists evidence in the Marketing field – albeit rarely made explicit – supporting the claim that changes in subscription prices may not always have traditionally defined impacts on consumption patterns (i.e., lower prices leading to less consumption per customer). For example, Danaher (2002) manipulates subscription prices of a telecommunications service and finds that individual usage on average decreased as a result of an increase in the subscription price of the service. This results persists even after controlling for customer churn: 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, which may be determinant for their final behavior. Relatedly, Albuquerque, Pavlidis, Chatow, Chen, and Jamal (2012) and Runge, Levav, and Nair (2022) find that while price promotions tend to boost subscription rates, there is little or no evidence of changes to other activities, including product usage/consumption. The diversity of these results highlights the importance of developing a flexible empirical framework that is able to account for an arbitrary correlation between subscription behaviors and subsequent consumption.

The subscription and usage patterns we find are especially important in our setting since the online publisher accrues revenue from both subscriptions and advertising. To quantify these effects and consider the impact of different pricing policies, we develop and estimate a flexible model of subscription and consumption behavior. The model accounts for the fact that the utility a customer derives from product consumption may not be enough to characterize consumption and explain the patterns in the data. In addition to preferences, we allow consumers to hold a willingness-to-pay concern, which may be induced from various unobservable mechanisms, such as switching and multihoming costs as well as heterogeneous opportunity costs of time. As we elaborate later, the fundamental aspect is not merely introducing a willingness-to-pay concern, but rather allowing it to be correlated with preferences for consumption, and using the patterns in the data to estimate the extent of this correlation.

Our analysis reveals that failing to account for these unobservable forces and relying solely on preferences would render very biased model parameter estimates. This is significant because most empirical work relies on a single dataset of rich consumer behaviors typically sourced from a single organization, and rarely contemplates all activities that may compete, directly and/or indirectly, for the consumers' time and resources. Gathering data about alternative activities is unrealistic in most empirical settings. This paper shows that, even in single-source datasets, it is possible to utilize the variation in subscription prices to take into account unobserved alternatives and recover unbiased parameters. The reason is that subscription prices affect subscription rates directly, but are sunk at the time of consumption. Hence, they affect consumption indirectly, only through the types of consumers who select into subscribing. This asymmetry allows us to flexibly recover the relation between subscription rates and consumption levels.[2]

Marketing scientists often consider settings in which consumers take interdependent decisions over time. A context that is close to ours is the case of retail settings, in which consumers decide whether to buy an item and, if so, how much to buy. Although there already exists a framework relying on discrete-continuous models (Hanemann (1984), Krishnamurthi and Raj (1988), Chintagunta (1993), Kim, Allenby, and Rossi (2002), Bhat (2005), Tuchman, Nair, and Gardete (2018)), we find it inappropriate to model subscription contexts for two reasons. First, in integrating consumers' sequential decisions, these models generally predict an unambiguously positive correlation between prices and (average) consumption levels, which limits the types of situations they can be applied to. Second, discrete-continuous models deal with per-unit prices, such that the total price paid depends linearly on the quantity purchased/consumed. This contrasts with subscription settings where, crucially, the quantity subscribers are allowed to consume seldom depends linearly on the subscription price.

After documenting the patterns in the data, we develop a model to characterize consumers' subscription decisions and observed consumption. The working assumption in terms of underlying mechanism is that the price promotion induces selection into the program by customers who may behave differently from the company's current customers.[3] 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 multiple mechanisms, it is nonetheless essential to produce even a reasonable fit with the data. In the counterfactual analysis section, we consider different promotional programs and find that the price promotion offered by the firm is very near the optimal promotional price if only subscription revenue is taken into account. However, once the impact of future advertising revenue from new subscribers is considered, we find the firm would have been better off pro-

[2] In line with Heckman (1979), a selection issue would arise if subscription prices were not sunk at the time of consumption, which would lead to incorrect estimates of subscribers' consumption distribution.

[3] There exists a long literature modeling the behaviors of existing subscribers in detail: see for example Ascarza and Hardie (2013). Our work is different in that we are interested in modeling how changes to the subscription price induce heterogeneous subscriber pools, which requires access to price variation in the first place. There is also a long literature on consumer expansion during stockpiling (see e.g., for example Bell, Iyer, and Padmanabhan (2002), Van Heerde, Leeflang, and Wittink (2004), Chan, Narasimhan, and Zhang (2008)), which does not apply to the context of digital subscriptions such as ours.

moting more aggressively to induce more consumption and higher advertising revenue in the future. This finding is robust even at advertising rates lower than the ones communicated to us by the management team.

We then consider a counterfactual analysis in which the firm is assumed to be able to monitor actual consumption by subscribers; this is an admittedly futuristic scenario in which consumers are not able to manipulate their real 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, that is, subscribers who consume more may effectively pay less. We find that providing discounts for high levels of (verifiable) consumption would be extremely profitable to the firm vis-à-vis the status quo for a reasonable range of advertising elasticities, owing primarily to the subsequent advertising 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 is 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 said to be widespread, and credible online news publishers face significant financial pressure. 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 present a selection of mechanisms related to the interplay of subscription and consumption behaviors. Section 3 describes the dataset patterns and some model-free patterns. Section 4 presents the model, its identification, and the estimation method. Sections 5 and 6 present the estimation results and counterfactual analyses, respectively, and Section 7 concludes with implications for managers, publishers and policymakers.

2 Subscriptions and Consumption

The goal of this paper is not to separate out specific mechanisms, but 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 we describe a few mechanisms that are less likely to impact the relationship in most analyses, including the case of wealth/income effects. Overall, the goal of this section is not to identify one or two mechanisms that will be assumed for the rest of the analysis. Rather, the goal is simply to illustrate that the correlation between subscription pricing and consumption is nuanced, potentially owing to multiple interacting mechanisms.

2.1 Positive Correlation between Subscription Prices and Consumption

Consumer Preferences. We start by explaining that simple consumer preferences inherently predict 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 extramarginal consumers (i.e., those whose valuations fall below the regular price) to select into buying, resulting in 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.

Sunk Costs. The theory of sunk costs predicts that costs incurred today affect future behaviors despite their payoff irrelevance. The relationship with product usage is 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. Along these lines, the theory predicts that a price increase will lead consumers to value their past investment more later, thus increasing consumption. Overall, the sunk cost fallacy predicts that subscribers will distort their subsequent consumption decisions in the direction of the change in subscription price.

Mental Accounts. The idea that consumers may hold separate mental accounts has received significant attention (e.g., Thaler (1985)). It is likely that consumers who appreciate certain activities more are more likely to keep larger mental budgets for those activities. Similarly, consumers with higher wealth levels are more likely to hold larger mental accounts. Thus, the presence of mental accounts results in higher subscription prices driving more consumption due to willingness-to pay-selection of consumers. This is an intuitive result, and in the next section we explain that mental accounts, when joined by other factors, can induce negative correlation between subscription prices and consumption.

2.2 Negative Correlation between Subscription Prices and Consumption

The well-documented phenomenon of switching and multihoming costs makes it costly for consumers to switch between or maintain multiple offers (e.g., Klemperer (1987), Hartmann and Viard (2008), Villas-Boas (2015)). Consumers are also known to 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 instance, 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. Similarly, a price promotion of a streaming service may attract consumers who are opportunistic and/or "movie buffs," but it may fail to attract more mainstream consumers who would find it challenging to manage several streaming subscriptions. In the context of vertically-differentiated markets, the existing literature has documented the fact that "high-types" may respond less to promotions and other offers than "low-types," due to already subscribing to higher-quality – but also more expensive – alternatives (see for example Villas-Boas and Schmidt-Mohr (1999); Stole (2007); Gardete (2013)). In other words, switching and multihoming costs, along with opportunity costs of time, can sort consumer types across market offerings. These barriers may keep high-value consumers loyal to premium alternatives while making price promotions more appealing to budget-conscious consumers.

Switching and multihoming costs may be bolstered by other factors as well, such as mental accounts and heterogeneous time constraints (e.g., Thaler (1985) and Kivetz (1999)). 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 or other reasons. Also, mental accounts explain why such consumers had not subscribed to the competitive streaming service to start with, but they may be attracted by a price promotion since it gives them license to opt into a subscription program without having to exceed the mental account for the category. Having briefly outlined a few mechanisms operating in subscription contexts, we now turn to the data.

3 Data Patterns

To analyze the relationship between subscription pricing and consumers' usage decisions, we utilize a dataset that 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. Whereas 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 for 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. We learned in meetings with the management team that the marginal cost for the publisher to accommodate additional readership is considered irrelevant.

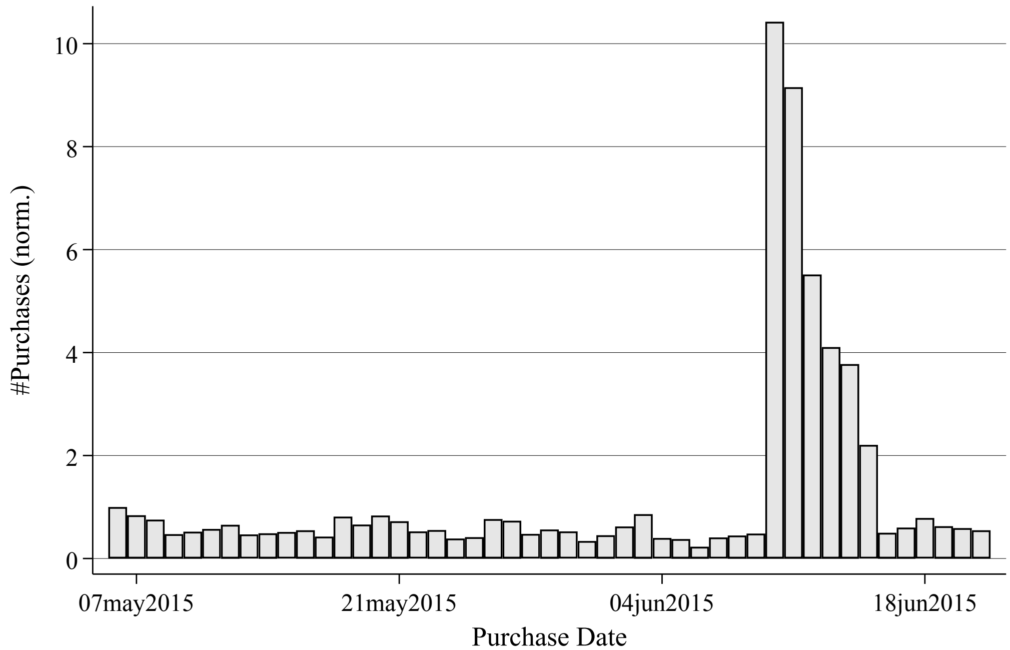
Our dataset focuses on consumers who subscribed between May 6^{th}, 2015, and June 15^{th}, 2015, providing detailed individual-level data on the browsing and subscription behavior of more than 10,000 new subscribers. For each subscriber, the dataset captures the timing and frequency of website visits,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

After two years of operating its premium content service, the publisher introduced its first price promotion in mid-June 2015. The six-day promotion ran from June 10^{th} to June 15^{th} and temporarily reduced the monthly subscription price from €4.99 to €2.00. It was not pre-announced in any channel and was available exclusively to first-time subscribers, who 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 after the promotion, which are ignored in the main analysis). The data reveal a pronounced spike in new subscriptions during the six-day promotional period, prior to which subscription activity was relatively stable, with no detectable anticipation effects.[4] This momentum diminished gradually over the promotional period and is undetectable after the promotion ended, reflecting the well-defined impact on new subscriptions.

[4]Indeed, the management team confirmed that the promotion was not advertised anywhere other than on the provider's website.

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 reasonable given the magnitude of the promotion: The discount generated more than ten times the number of new subscribers the publisher had attracted in any day before the promotion. However, the effects of the promotion are clearly diminishing over time. The decreasing pattern 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 price promotion is its impact on consumption patterns. To explore these patterns, 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 the first year following each customer's subscription decision. On average, promotional subscribers consumed almost three times as many articles as their regular counterparts. Figure 2 presents consumption levels for promotional subscribers vs. 'regular consumers' who subscribed up to one week before the promotion started, to maximize comparability between groups. When we aggregate all articles consumed by subscribers over a maximum subscription period of 12 months, we see that promotional subscribers demonstrated significantly higher consumption levels compared to the group of regular subscribers (10,677 vs. 3,697 articles consumed in the first year, p<0.001).

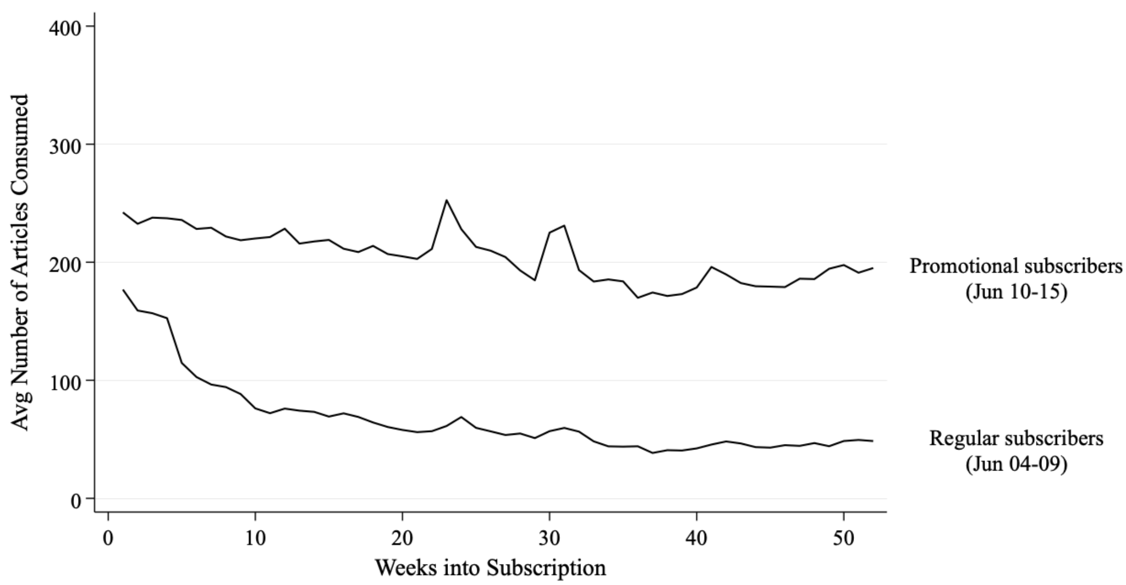
Table 1: Summary Statistics of News Articles Consumed Per Week

	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 promotional 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.

Across the 52-week period depicted in Figure 2, promotional subscribers demonstrate consistently higher consumption levels than regular subscribers. Consumption rates begin with a large gap and this difference stabilizes, with the two groups maintaining roughly parallel trends over time. This persistent disparity suggests that promotional subscribers exhibit distinct engagement behaviors that may not be solely attributable to short-run effects of the 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. An unexpected and striking result is that 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 promotional 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.

While the different consumption levels above are suggestive of the presence of heterogeneous consumers selecting into subscribing at different prices, they could also reflect that promotional subscribers have less of an incentive to churn, as doing so would forfeit their promotional price. In other words, the lower churn rates of promotional subscribers may be enough to explain their higher consumption levels, a consequence of straightforward strategic behavior. To examine this possibility, in Figure 3 we plot consumption levels for the two subscriber groups only during periods in which each subscriber was active. Notably, conditional on not churning, promotional subscribers still display substantially higher engagement levels, consuming approximately 50% more news articles than regular subscribers.

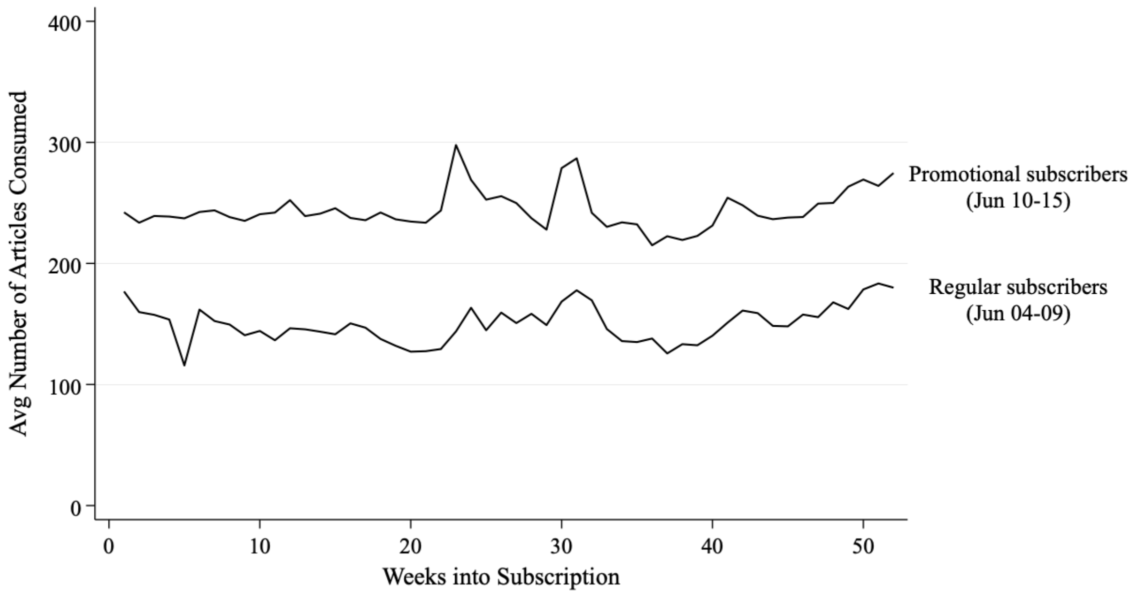
Overall, the data patterns above point to the price promotion having attracted fundamentally different types of consumers into the publisher's subscriber pool. First, the consumption levels of the promotional group are higher despite the fact that subscription costs are sunk at the time of consumption. Second, the effect of the promotion on subscription rates is significant and decays quickly, pointing to a well-defined segment of consumers who were 'not in the market to subscribe' at the regular price, but 'rushed into subscribing' once the promotion was offered. Third, promotional subscribers consume more articles than their regular counterparts even after we control for consumer churn and after a year has elapsed since the initial subscription. Finally, while habit formation may play a role in these patterns, the parallel trends in Figure 3 suggest a steady difference in consumption levels between the two groups from the original subscription moment on, leaving little space for dynamics related to habit formation.[5]

Overall, 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

[5]Our analysis emphasizes the mechanism of selection and, under this assumption, the counterfactual analyses suggest that the publisher could have gained significantly by adopting a more aggressive promotional strategy. If habit formation indeed takes place, then its compounding effect would imply that the benefits for the publisher could have been even greater, making our estimates a lower bound for the full potential value of promotional activities for the publisher.

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).

Building on these descriptive findings, we introduce an empirical model in the next section that allows us to explain the subscription and consumption patterns above, and investigate the consequences of counterfactual pricing policies by the seller.

4 Model

We propose a simple model of economic behavior that recovers the fundamentals driving the observed behaviors in the data. After estimating the model, we are able to consider the subscription/advertising revenue balance faced by the seller as a function of the subscription price, as well as investigate the effects of offering different pricing menus on subscription and consumption levels. The model contemplates two-dimensional heterogeneity: Each consumer type is characterized by its optimal consumption quantity (conditional on subscribing) and willingness to pay. The first dimension, noted as v_i, is the quantity that customer i would like to consume once the decision to subscribe has been taken. This feature is especially appealing because, at the time of consuming, subscription costs are sunk and so have no bearing on consumption intensity. The working assumption that v_i is consumer i's own type means that, for counterfactual analyses, subscription prices affect consumption levels via subscriber selection. In effect, only the consumption levels v_i of individuals who effectively subscribed are observed, much in the spirit of Heckman (1979).

Consumers also hold a w.t.p. w_i dimension that depends on forces unrelated to direct consumption utility. This dimension pertains to unobservable forces (such as unobserved competitors or opportunity costs of time) that affect how much consumers are willing to pay. Crucially, the model allows the consumption preference v_i and the willingness to pay (w.t.p.) w_i to be arbitrarily correlated. This is made possible, as we discuss later, by the fact that the subscription price is sunk once at the time of consumption.

For simplicity, we assume each consumer has a single chance to subscribe during the sample period, at the posted subscription price p_i (this is a common assumption in empirical work; see Waisman (2021)). Those who 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, and v_i is consumer i's optimal news consumption, conditional on subscribing. Shocks ε_i^1 and ε_i^0 are extreme-value type 1 distributed, p_i is the price observed by consumer i upon arrival, and w_i is consumer i's willingness to pay. For simplicity, the model abstracts away from consumers' subscription renewal decisions, as we discuss later.

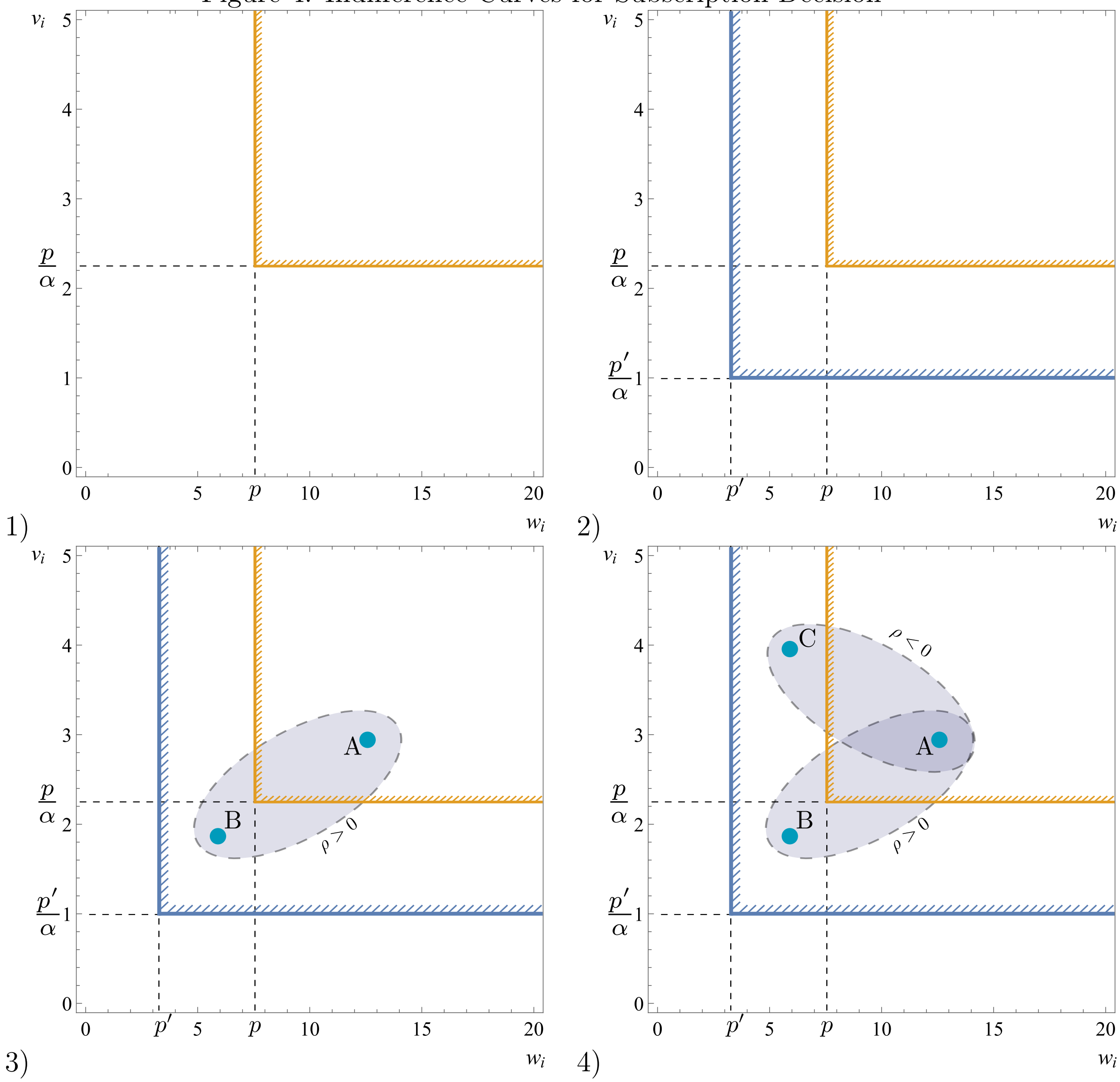
Figure 4 provides a visual representation of this simple model. Panel 1 depicts the space of consumption v_i and willingness to pay w_i that justifies subscribing. The top-right area, outlined in orange, reveals that only consumers who are both interested in consuming enough news ($v_i \geq p/\alpha$) and are willing to pay enough for that service ($w_i \geq p$) will subscribe (notice that 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 downward and to the left, a combination of the preference effect and the w.t.p. effect. Panel 3 introduces the support of a possible distribution of individuals' $\{w_i, v_i\}$ pairs, with a positive correlation parameter ρ. It is now possible to see 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 between the blue and orange lines had not subscribed at the regular price, either because they exhibited relatively weak preferences for consumption or because they held a low willingness to pay related to other factors. Moreover, assuming a uniform distribution over the support, it is clear that the incremental customers (in between the blue and orange regions) will consume less per capita than the ones who are willing to subscribe at the regular price (orange region). Finally, Panel 4 depicts the analogous case of a price decrease when the correlation of consumption and willingness to pay is negative. The price reduction captures consumers who tend to take great interest in reading the news, but other factors (such as being locked into unobservable competitive offers) discourage them from subscribing at the regular price. For clarity, consider the mass points A and B in Panel 3: We see 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. From the above, the connection between the correlation of preferences and w.t.p. with the effects of pricing on consumption is now straightforward and readily interpretable. This formulation also helps illustrate the central reason why we are able to identify the correlation parameter ρ: the fact that price variation affects consumption levels only through selection into 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 consider the parameterization of the joint distribution of v and w. We start by noting that the distribution of w is unobservable to researchers. Indeed, being able to account for this unobserved latent variable is one of the contributions of the paper. As for the levels of consumption v_i, they are 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 belongs 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). To be precise, we define v_i as the consumption level by consumer i over a 12-month

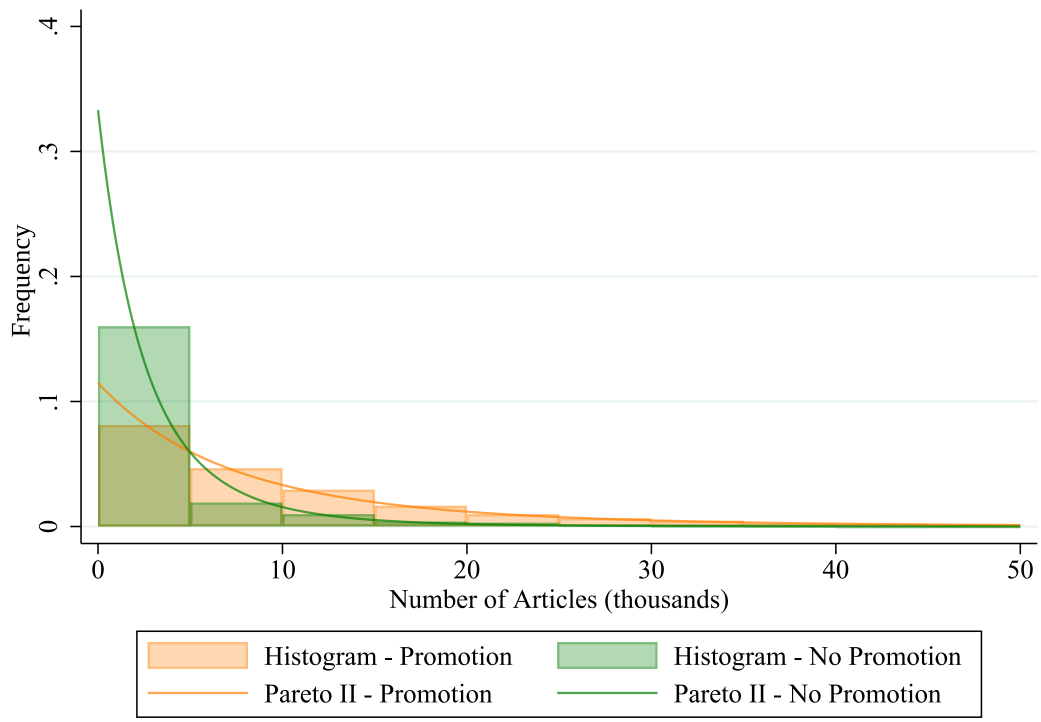
Figure 4: Indifference Curves for Subscription Decision



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

period after the initial subscription decision.[6]

Figure 5: Histograms of the Number of Articles Consumed up to One Year after 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 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:
Subscribed at Promotional Price	9.635	9.672	0.001	51.104	6,480
Subscribed at Regular Price	3.066	5.116	0.001	30.73	3,569

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

[6]We focus on the sample of consumers who, within each price group, fall within three standard deviations from the mean in terms of news readership. This removed a few very high values that may be associated with non-human activities such as bots, scraping tools, etc.

The fact that the standard deviation of consumption exceeds its average is an especially striking feature of Table 3's statistics, given that these distributions are 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 allow the standard deviation to surpass its mean. Indeed, its mean and standard deviation are equal to $\frac{1}{\lambda}$, and in cases where v is truncated from below at scalar $a > 0$, the standard deviation falls strictly below the mean ($\mathrm{E}\left(v \mid v > a\right) = a + \frac{1}{\lambda}$; Std.Dev.$(v \mid v > a) = \frac{1}{\lambda}$). 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 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 fits the observed news-consumption data quite well. The specification of the density of v is given by:

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

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

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

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

The interpretation of this specification is that there exists a mass of consumers of size $1-\gamma$ willing to pay nothing to access premium news.[7] This assumption is inspired by 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 (e.g., Reshadi and Fitzgerald (2023)). Finally, the normal specification is focal among distributions and can be justified by a central limit motivation, because willingness to pay can be thought of as aggregating over many articles, time units and/or multiple benefits of subscribing.

The fundamental relationship 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

[7] 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.

a statistical copula to characterize their relationship. This allows us to assume a flexible bivariate density of v and w, approximated by a Gaussian copula with density

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

where

$$u_1=F_w\left(w\right)$$
$$u_2=F_v\left(v\right)$$

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

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

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

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 $\left(p^r\right)$ 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):

$$obj\left(\theta\right)=\sum_{i=1}^{6}m_{i}^{'}\left(\theta\right)^{2} \tag{12}$$

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,

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

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% versus an actual market share of 4% is penalized as much as an average consumption of 9,900 articles 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. The main object to recover through estimation is the correlation between the unobserved w.t.p. levels and the partially observed (due to selection) consumption levels. This requires us to jointly simulate w_i and v_i for each guess of parameter ρ, and then simulate predicted moments that are matched with the ones in the data. In each iteration of the set of parameters θ, the model is used to simulate consumption levels, w.t.p. levels, and subscription decisions. We follow this procedure:

1. Consider some guess of parameters, θ.

2. Take K draws of $\{w_k, v_k\}$ pairs via the Gaussian copula, which incorporates candidate cumulative distribution functions F_w and F_v.

3. For each price level p^t, $t\in\{r,d\}$, calculate whether each simulated consumer satisfies the utility and w.t.p. conditions, i.e., $I_{t,k}^{v}=\alpha v_k+\varepsilon_k^1\geq p_t+\varepsilon_k^0$ and $I_{t,k}^{w}=w_r\geq p_t$. For example, $I_{r,k}^{v}=1\wedge I_{r,k}^{w}=1$ means that consumer k is willing to subscribe at the regular price, whereas $I_{r,k}^{v}=0\wedge I_{r,k}^{w}=1$ means that the consumer is unwilling to subscribe at the regular price due to insufficient consumption value.

4. 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 $\left\{v: I_{r,k}^{v}=1\wedge I_{r,k}^{w}=1\right\}$. 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, and multiplied by the market share parameter γ .

5. 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 of consumption (F_v) 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 in our simulations with randomly generated data. 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.

Taking draws of v and w pairs is relatively simple, and it 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 the 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.,

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

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.

We normalize two parameters of the marginal distributions of v and w. For the distribution of v, we normalize the shape parameter $\alpha_w = 5$. This 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. Whereas α_w is identified from the higher-order moments of v in the data, strictly speaking (via functional form), those moments exist only if the true value of α_w is indeed high enough. Rather than run the risk of employing an incorrect estimator, we 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.

The other parameter we normalize is the value of the mean of w_i, μ_w. We 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 accomplished by parameter γ, which has a much more direct interpretability.

Overall, we prefer to normalize these two parameters because 1) the parameters we do include already have focal roles in matching moments in the data, and 2) experimentation revealed very little effect of these parameters in terms of their ability to improve model fit. The parameters of our model are summarized in Table 4.

Table 4: Model Parameters

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

Identification. We discuss how the parameters above are identified by the estimation moments. 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 has a direct effect on market share, regardless of subscribers' consumption levels, for example. As a result, it matches the overall market share across pricing regimes. As for parameter α, it weighs how much consumption, as captured by v_i, translates into utility, and thus, to the act of subscribing. It follows that parameters α and γ can be used to match the market share during the promotional and the regular price periods.

As we explained before, the correlation parameter ρ is identified by the difference in average readership across price levels. This follows from the discussion of Figure 4: The impact of price changes on consumption rates informs the correlation between w.t.p. and consumption.

Parameters λ_v and σ_v match the means and standard deviations of consumption. 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. Like in the case of parameters α and γ with market shares, it does not suffice to match the moments overall; they must also match within each price condition. Parameter σ_w captures the decay of the mean and standard deviation of v as the price changes, following the assumed curvature of the normal distribution. Referring back to Figure 4, it is clear that when w is very spread out, the moments of v vary less across price levels than if w is concentrated near zero (i.e., a low value of σ_w), ceteris paribus. Hence, parameters λ_v and σ_v play a fundamental role in matching the consumption moments across price levels.

Additional Modeling Assumptions. For simplicity, the model above abstracts away from consumers' subscription renewal decisions, focusing only on the first subscription decision and 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 without producing an obvious benefit. In Section 6.2 we explain that consumption levels (which we model) are highly correlated with contract duration, and so it suffices to assume that the relationship between these two constructs is stable. We model this relationship explicitly in the counterfactual analysis section later in the paper.

Second, we define v_i as the consumption of all types of articles by consumers, rather than just premium ones. This is useful for counterfactual analyses, namely whenever advertising revenue is relevant. Since different types of articles accrue the same revenue each time they are read, accounting for the overall consumption of each subscriber takes into account the total 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 due to it increasing the consumption of premium articles, while potentially ignoring advertising revenue lost due to a substitution from consumption of free articles.

We now turn to the estimation results and present measures of fit with the moments in the data.

5 Results

We start by reporting the model estimates, and providing brief interpretations of their magnitudes. We then present measures of fit and document the bias that arises from ignoring the correlation parameter ρ.

Table 5 summarizes the maximum likelihood estimates of the model:

Table 5: Model Estimates

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

Note: Standard errors in parentheses. Significance levels: $^{\dagger}p \leq 0.1$, $^{*}p \leq 0.05$, $^{**}p \leq 0.01$. 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 w.t.p. for online news. Although this number appears small, when multiplied by the multiple millions of potential consumers in the market, it still results in hundreds of thousands of potential customers. This estimate is reasonable; for instance, the equivalent figure for The New York Times is at least 0.5%.[8]

We now ask to what extent are subscriptions driven by consumption value alone? The value of parameter α translates the number of articles consumed to monetary units. We

[8]The New York Times specifies "**nearly** 2 Billion readers" and "**more than** 10 million paid subscribers" (https://advertising.nytimes.com/audience-and-insights/).

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 w.t.p. effects. In this case, the ratio of demand between the discounted price and the regular price would be given by the expected ratio of logit expressions:

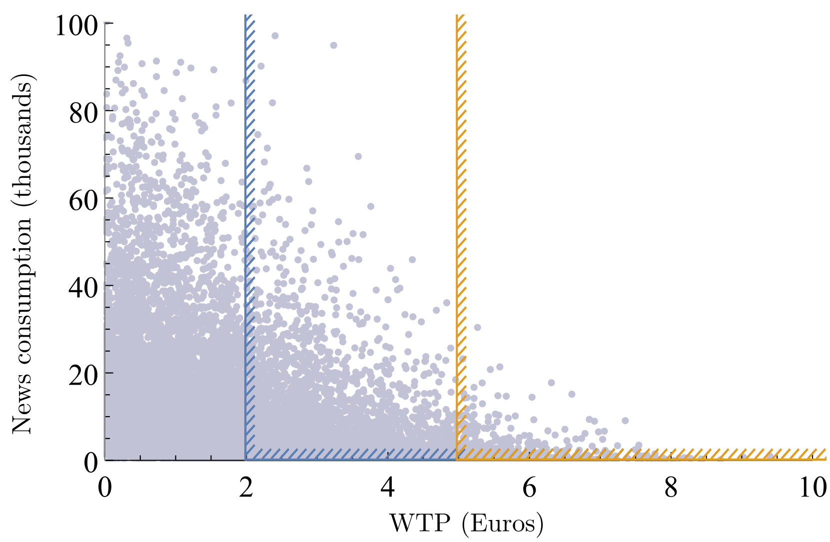
$$E_v\left(\frac{\frac{1}{1+\exp\left(-\left(\widehat{\alpha}v-p^d\right)\right)}}{\frac{1}{1+\exp\left(-\left(\widehat{\alpha}v-p^r\right)\right)}}\right) = \int_0^\infty \frac{1+\exp\left(-\left(\widehat{\alpha}v-p^r\right)\right)}{1+\exp\left(-\left(\widehat{\alpha}v-p^d\right)\right)} f_v\left(v;\widehat{\lambda_v}\right)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

$$= 1.0108 \tag{18}$$

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%. This is a relatively low number given the price discount of more than 50% (implied elasticity equals -0.007), and points to the possibility that the w.t.p. concern plays a more important role than direct consumption utility.

The main parameter estimate of the model, ρ, can be interpreted as a relatively strong negative correlation between news consumption and willingness to pay. Figure 6 plots simulations of $\{v, w\}$ pairs 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 w.t.p., in line with the negative estimate of parameter ρ. Although this is reassuring, the negative value of ρ is necessary but not sufficient for a price discount to induce an increase in the

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.

average news consumption level. 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. Table 6 shows the data moments and 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 deviations 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 the two most fundamental patterns in the data: First, both standard deviations of readership exceed the segment means and, second, the average readership is higher for the promotional segment than the regular one. Matching these two patterns is essential to capture the empirical patterns and the underlying economic forces in play.

6 Counterfactuals

We now turn to evaluating how different pricing policies interact with the demand we have recovered. All counterfactual analyses below maintain the sample period constant and the regular subscription price unless otherwise stated. In particular, the negative trend observed for new subscribers during the promotional period (see Figure 6) motivates us to not extrapolate promotional effects to longer periods.

6.1 Effect of correlation parameter ρ

Our first goal is to understand the effect of the correlation between preferences and w.t.p. on the demand system. We investigate this effect by ignoring/allowing for the correlation parameter ρ in estimation, namely since this parameter is typically absent from discrete-continuous model specifications.

Table 7: Model Estimates

Parameter	Original Estimate	Estimate with $\rho = 0$
α	141.5*	174.779**
	(63.124)	(5.481)
γ	0.01**	0.011**
	(0.000)	(0.000)
λ_v	124.04**	16.128**
	(6.484)	(0.524)
ρ	-0.67**	–
	(0.016)	–
σ_w	2.444**	2.386**
	(0.017)	(0.02)
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 presents side-by-side estimates of the model parameters with and without fixing parameter ρ at zero. When preferences and w.t.p. are assumed to be uncorrelated, 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 in the news-consumption predictions. These

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.

differences are relatively modest for consumers subscribing at the regular price, but the constrained model 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 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.[9] In summary, failing to account for the correlation between preferences and w.t.p. would not only produce biased coefficient estimates, but it would also absolutely prevent the model from explaining the moments in the data.

6.2 Effect of Promotional Price Level on Profits

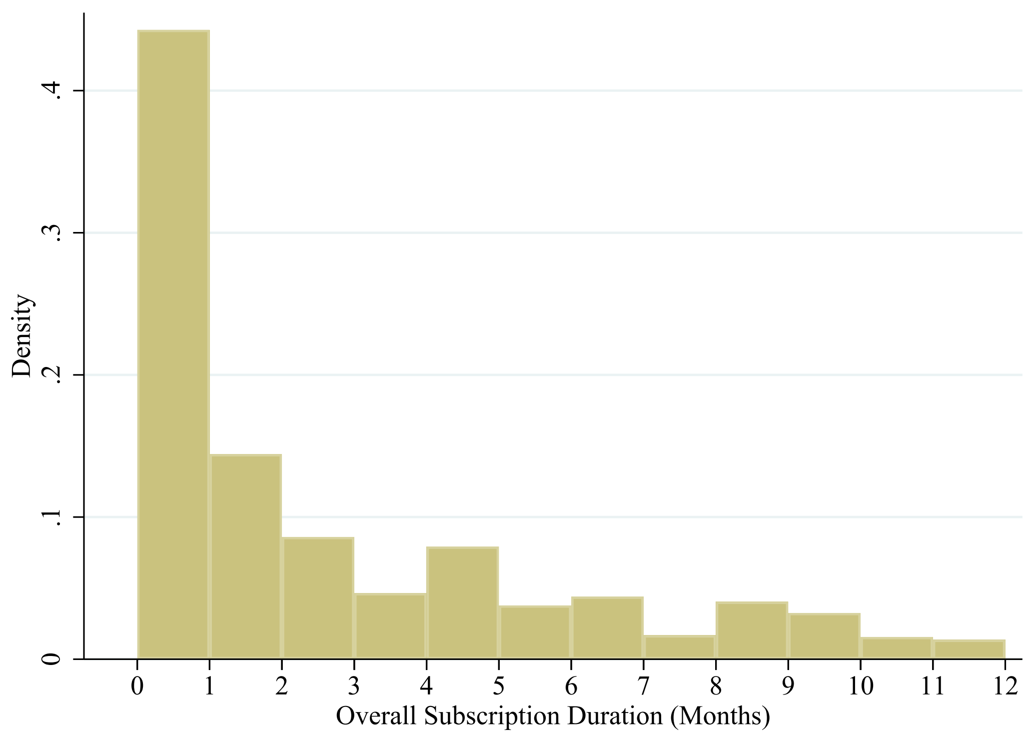
We now simulate firm profitability at different price promotion levels. In line with the data, we consider the consumers acquired during the estimation period of the data (May 6^{th} until June 15^{th} 2015), with the regular price of €4.99 until June 9^{th}, followed by a promotional price – which we vary in the counterfactual analyses – starting June 10^{th} through June 15^{th}. We first consider the effect of different promotional prices, which are held constant over the course of a year, like in the dataset. 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 counterfactual scenarios in which we change the

[9]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.

promotional price offered in the period from June 10^{th} to June 15^{th}, and we analyze the firm-profitability effects on customers up to one year after their subscription starts.

So far, we have opted to model initial subscription decisions and consumption levels over the period of a year explicitly. To incorporate subscription revenues over the course of a year as well, we forecast consumer churn via a prediction model. The working assumption is that the relationship between consumer churn and number of articles consumed remains stable across counterfactual analyses. Indeed, the correlation between the two components in the data is 0.42, which is to be expected since consumption decisions are intimately linked with churn decisions.[10] We plot the histogram of subscription durations in Figure 7, to determine whether additional factors need to 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 reveals a monotonously decreasing churn risk over time since the initial

[10]Note that in the data there exists heterogeneity in terms of how rapidly 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 of the total profitability effects.

Table 9: Cox Estimated Hazard Rates

Regressor	Parameter Estimate	Standard Error	
Number of Articles ($\times$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$\leq$0.10, * p$\leq$0.05, ** p$\leq$0.01. Figures above rounded to three decimal places with trailing zeros omitted.

subscription. The monotonic decreasing trend can be described 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 and advertising revenue, ceteris paribus).

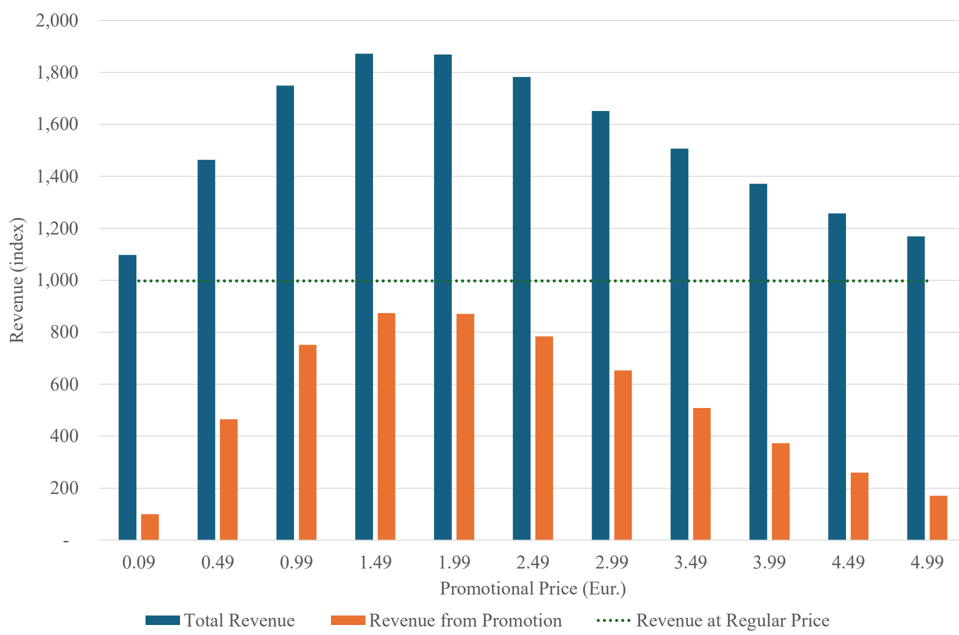
We regress a Cox survival model on the churn data, including as predictors the total number of articles read by a consumer (i.e., the consumer type) as well as dummy variables to capture potentially higher churn risk near exact month thresholds. Table 9 reports the results of the Cox regression model as well as the baseline survival probabilities. As expected, the parameter estimate associated with the number of articles read is negative, meaning 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 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 months to count fractional month occurrences as a single payment and to include the initial subscription 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.[11]

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.

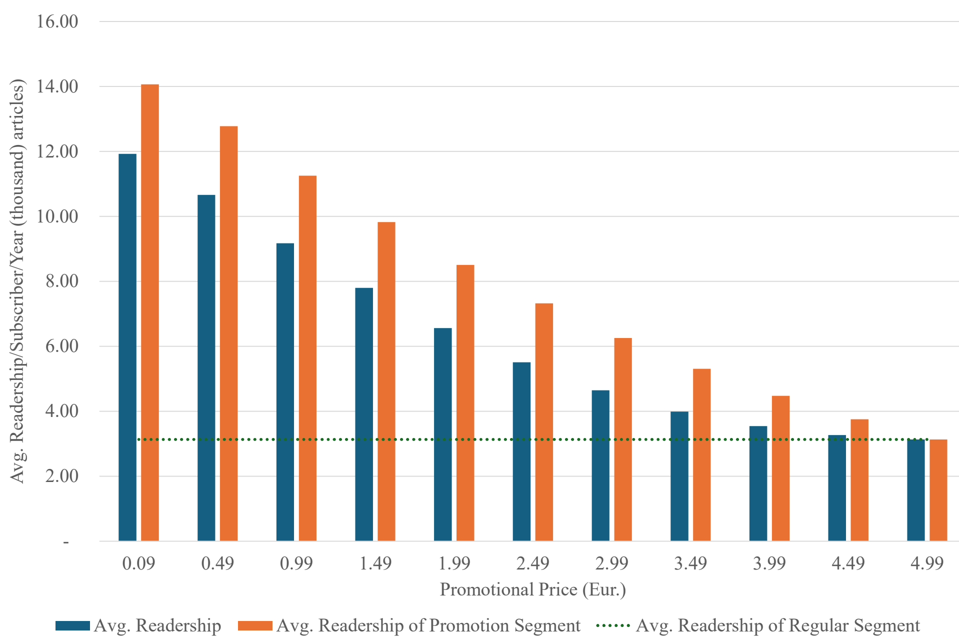
As illustrated in Figure 8, revenue is concave in price, achieving its maximum at the actual promotional price observed in the data. We remind the reader that there is no structural or theoretical reason that imposes this result. It seems that, by coincidence, the firm's promotional price was set very close to the empirical 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 close to €1.99. Note that when no promotion is introduced (rightmost bars), the difference between promotion revenue and the dashed lines is 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 preferable to overdiscount by a little than to underdiscount. However, when larger deviations from €1.99 are considered, we notice that revenue decreases more slowly with underdiscounting than overdis-

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

counting. So, from a managerial perspective, it is best to make small overdiscounting errors or large underdiscounting 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 overdiscounting the optimal price. In contrast, when they are less experienced or their information is very poor, they are better off underdiscounting.

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 consumes 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.[12] The extent to which the promotional segment generates higher ad revenues depends on the price promotion level set by the firm, in line with the results in Figure 9.

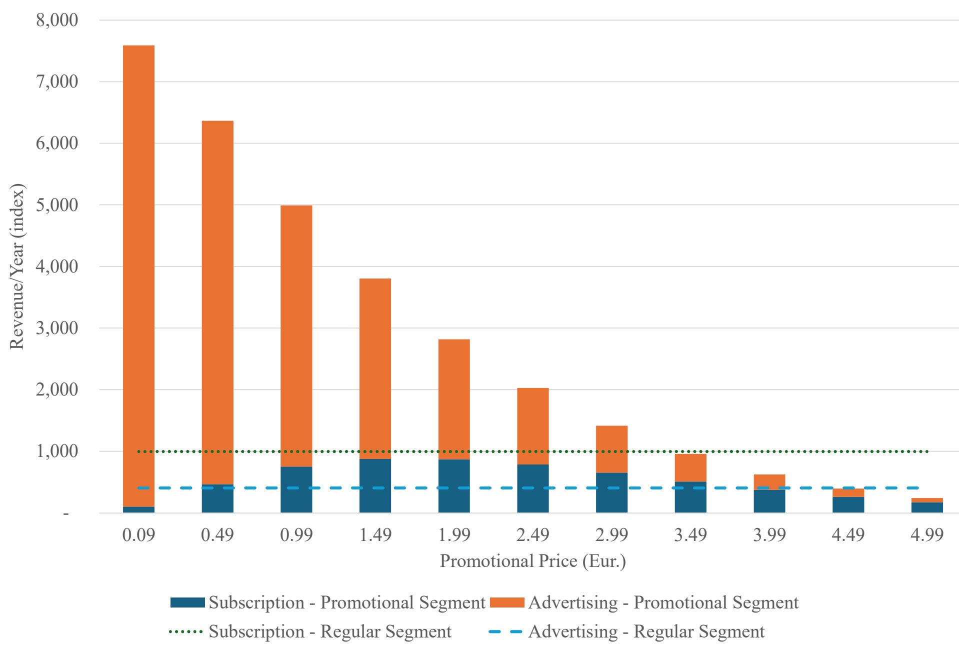
6.3 Revenue from Increased Consumption via Advertising

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 up to earlier dates for consumers who churn before that period ends.

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 use as a focal value for our analysis (we later consider sensitivity checks). This value allows us to add the subscription and advertising revenues in order to understand the combined effect of price promotions.

[12] Note that advertisers in this publication 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.

Figure 10: Yearly Revenue Decomposed by Segment and Source

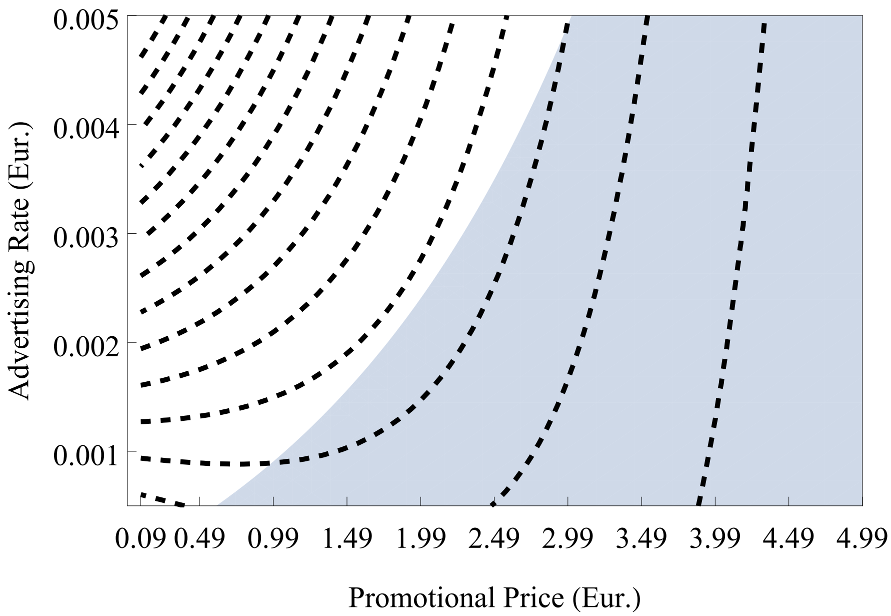


Note: Above, subscription revenue of promotional segment at the €0.09 price level normalized to 100. The results above pertain to the period of one year after individual subscriptions.

Figure 10 decomposes revenues across segments and sources (advertising and subscription), by segment (promotional and regular). The most striking result is that advertising makes up a small portion of revenue for the promotional segment when the promotional price sits above €2.99. However, at lower promotional prices it increases at an increasing rate, rapidly outweighing the other revenue sources, all the way to the lowest promotional price. Although we include results obtained at very low promotional values, we would like to de-emphasize them for two reasons. First, while the merit of structural models is allowing 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, we were provided with a constant monetary value of ad impressions that may no longer apply when the subscriber base increases substantially due to very low subscription prices, since it would also be associated with a much larger ad impression stock. These two reasons require prudence when judging the scenarios with very low promotional prices. Regardless of this, the promotional price of €2 is observed in the data, and at that level we can observe that the revenues generated by the promotional segment already dominate those from the regular segment. It is clear that the promotion generated value, primarily via advertising revenue.

Later, we consider a counterfactual analysis where advertising rates vary with the number of subscribers. First, we perform a sensitivity analysis to the unique advertising rate provided by the managerial team by, independently of the subscription base, considering situations in which the ad rate is lower than the €0.005 rate used in the previous analysis. We consider the interval [0.0005, 0.005] euros per news article read. We to investigate the robustness of the previous result that the customers attracted via the price promotion are more valuable due to their advertising rents than their subscription payments.

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



Note: Above, the 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 curves to the left 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 to very low promotional prices, we believe it is more informative to focus nearer to the promotional price observed in the data of €2. In this neighborhood, it is clear that the question of which source of revenue dominates depends crucially on the advertising rate. At about €0.0025 per article consumed, 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. A central effect of the price promotion is to boost advertising revenue through increased future consumption. In the following section we explicitly incorporate the effect of subscription levels on advertising rates.

6.4 Charging Less for More Consumption

Subscription pricing is often complemented by second-degree pricing practices, that is, the provision of different menus with combinations of price-quantity/quality offers that consumers select into. Second-degree pricing can be profitable when a firm is able to effectively separate consumers by ensuring incentive compatibility. This strategy is less effective 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 a negative correlation between consumption value and w.t.p..

However, we believe it is relevant to consider a setting in which the firm is able to monitor subscribers' actual consumption values, regardless of their efforts to hide them, as we now explain. Consider a plan that includes two price packages: one with unlimited consumption and another one with a capped consumption. The unique feature of this pricing program 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 with more quantity at a lower price can only be effective if the firm can detect real 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 technologies (see for example Jiang, Li, Chen, and Wang (2018) and Delouya (2024)), we believe such a futuristic scenario is viable and worth contemplating.[13]

In the plan described above, consumers all start at a regular subscription price, and obtain a discount only 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, the central idea is that

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

consumers end up paying 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. The plan we consider applies when it is costly or impossible for consumers to artificially boost product consumption, especially 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 a^* that earns consumers an amount t^*. The idea is that customers who consume a number of articles beyond a^* over the span of one year earn a reward, be it through a direct monetary transfer or a discount on 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 a^* and t^* that maximize the firm's joint subscription and advertising revenues. For consumers with consumption level $v_i \geq a^*$, we change the w.t.p. from w_i to $w_i + t^*$. We keep the indirect utility fixed under the interpretation that the "windfall" t^* keeps the utilitarian component constant, but may affect consumers' w.t.p. via a mental account, for example. This approach focuses on the w.t.p. mechanism, but 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 literature documents that salespeople whose performance falls just below a given compensation cutoff often exert more effort to surpass it (e.g., Misra and Nair (2011) and Chung, Steenburgh, and Sudhir (2014)). We leave this analysis for future research efforts.

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 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 note 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. The discount payments above (last line of the firm revenue) are formulated so that t^* is interpreted as a per-month payment to customers. This specification provides clarity, and is equivalent to a one-time payment since we abstract away from time discounting. For illustration purposes, we present the number of subscription renewals by consumers who obtain the discount, $D_{disc}(\cdot)$:

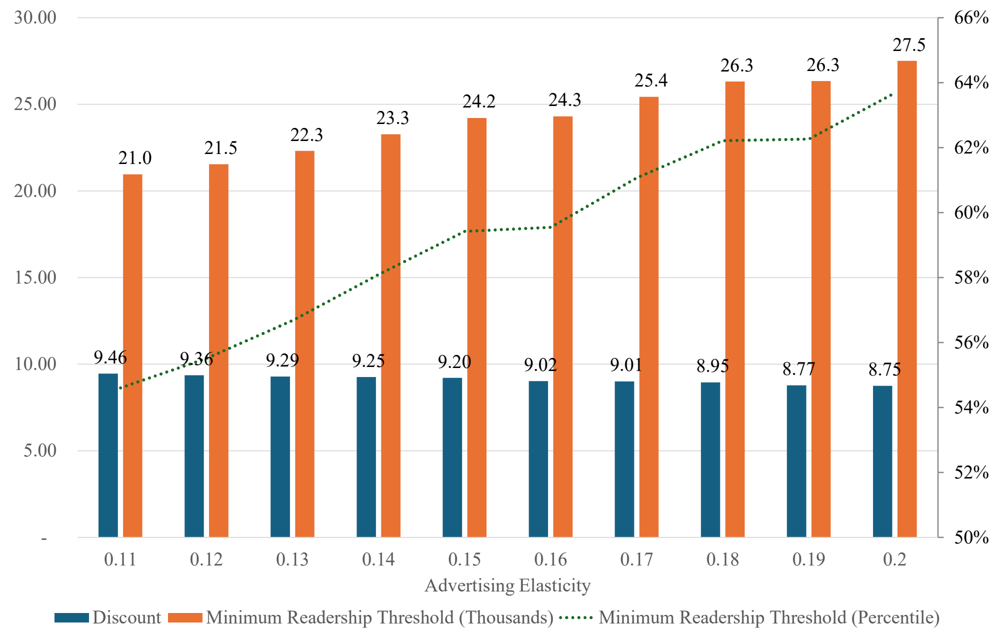
$$D_{disc}\left(p,a^*,t^*\right)=\underbrace{M\gamma}_{Potential\,Market}\cdot\underbrace{\frac{1}{R}\sum_{i=1}^{R}1\left(v_i\geq a^*\wedge\alpha v_i+\varepsilon_i^1\geq p_i+\varepsilon_i^0\wedge w_i+t^*\geq p_i\right)}_{Market\,Share}\cdot\underbrace{n_i}_{N.\,Payments} \tag{21}$$

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

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

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 level t^* and minimum readership threshold a^* as a function of advertising elasticity.

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

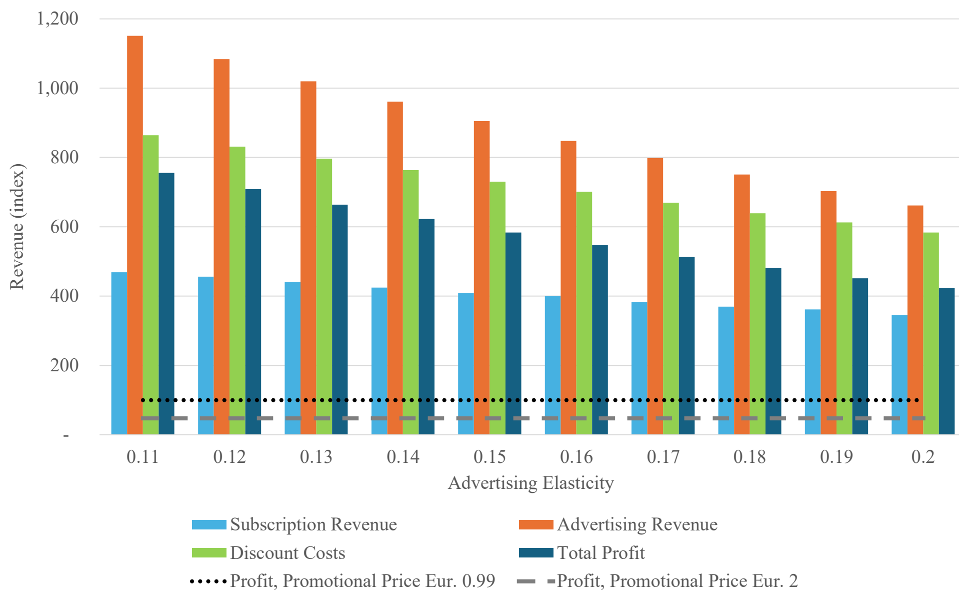


Note: The Y axis above is used to simultaneously represent euros (short blue bars) and thousands of articles consumed (tall orange bars). The dotted line represents the same data as the orange bars but in percentile terms, and should be read on the right Y axis.

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 subscribers 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. We find that the firm is never better off offering discounts to consumers below the 54^{th} consumption percentile.

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, firm revenues as a function of advertising elasticity. For comparison, profits of the regular price promotion with discounted prices of Eur. 0.99 and Eur. 2 are represented by the dashed lines.

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 throughout: It generates approximately 9 to 16 times 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 does not rely on providing information rents to enforce price discrimination. Under perfect consumption observability, the firm can effectively practice third-degree price discrimination. In the current scenario, consumers do not maintain information rents and put pressure on the seller only through menu selection due to the assumption that the publisher can effectively monitor real consumption by its subscribers. 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 to increase profits by pricing the plans accordingly.

7 Concluding Remarks

In this paper, we investigate the relationship between subscription prices and consumption behaviors in the context of an online news publisher. Beyond examining basic behavioral patterns, we propose and specify a flexible model that is able to capture the dynamics of this relationship. The model was then used to identify significant opportunities to increase revenues for the publisher through various pricing strategies. 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.

We find that failing to allow for consumption utility to be correlated with willingness to pay not only has a strong effect on the estimated model coefficients, but it also absolutely precludes the model from fitting the moments of the data. In our setting, the counterfactual analyses reveal that there is significant value from introducing price promotions, not only due to attracting additional subscription revenue but especially due to incremental advertising revenue driven by consumers who are not willing to pay the regular subscription price. Finally, we considered the case of perfect consumption monitoring, which opens up the possibility of generating significant incremental revenue through the introduction of an effective negative price on additional consumption.

Our analysis demonstrates that even small reductions in access fees can lead to large increases in news consumption, highlighting a way to improve public access to information. This has important implications to social welfare, as lower access costs could contribute to a better-informed population. Policies designed to promote access to socially beneficial services should carefully consider the effects of pricing on consumption patterns. Understanding the relationship between pricing and consumption is not only crucial for firms evaluating revenue trade-offs but essential for assessing the broader societal consequences of subscription pricing. This is especially relevant in an era of widespread misinformation and financial challenges for credible online news publishers.

For future research, 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' purchase decisions correlate 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 about 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 extra 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* to 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 of the low type consumer is greater than that of the high-type consumer.

For the effect above to occur, 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 firm's 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 by adding a 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

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

where $g\left(u^{-1}(p)\right) \equiv \int_{u^{-1}(p)}^{\infty} v f_v(v)\, dv$ and $G\left(u^{-1}(p)\right) \equiv P\left(u(v) \geq p\right)$. 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, $g'\left(u^{-1}(p)\right) = -u^{-1}(p) f_v\left(u^{-1}(p)\right)$ and $G'\left(u^{-1}(p)\right) = -f_v\left(u^{-1}(p)\right)$, such that the last expression can be rewritten as

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

Finally, note that

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

so that $g\left(u^{-1}(p)\right) > u^{-1}(p) G\left(u^{-1}(p)\right)$, 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 occur, 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 (28) 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. 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.

Yet, there are reasons to believe that income effects are not involved in the patterns we observe. First, while formally appealing, it is challenging to determine one's own budget constraint, which in addition to bank accounts and other assets, may also include borrowing from friends 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 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, indicating 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 possible that consumers in this market were accustomed 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 $\lambda > 0$. This effect can indeed play a role in subscription decisions, but it is unclear whether this implies that price promotions will attract people who want 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 people who want to consume more news also tend to perceive higher gains from price promotions. Though we cannot reject this hypothesis, no empirical evidence supports loss aversion 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 period following the subscription. The prediction 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 consumption differences between two groups of consumers with different subscription prices are expected to diverge, at least momentarily, due to behavioral reinforcement (e.g., Wood and Neal (2009)).

References

Albuquerque, P., P. Pavlidis, U. Chatow, K.-Y. Chen, and Z. Jamal (2012): "Evaluating promotional activities in an online two-sided market of user-generated content," *Marketing Science*, 31(3), 406–432.

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

Ascarza, E., and B. G. Hardie (2013): "A joint model of usage and churn in contractual settings," *Marketing Science*, 32(4), 570–590.

Bell, D. R., G. Iyer, and V. Padmanabhan (2002): "Price competition under stockpiling and flexible consumption," *Journal of Marketing Research*, 39(3), 292–303.

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.

Chan, T., C. Narasimhan, and Q. Zhang (2008): "Decomposing promotional effects with a dynamic structural model of flexible consumption," *Journal of Marketing Research*, 45(4), 487–498.

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*.

Chung, D. J., T. Steenburgh, and K. Sudhir (2014): "Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans," *Marketing Science*, 33(2), 165–187.

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.

Datta, H., B. Foubert, and H. J. Van Heerde (2015): “The challenge of retaining customers acquired with free trials,” *Journal of Marketing Research*, 52(2), 217–234.

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.

Gardete, P. M. (2013): “Cheap-talk advertising and misrepresentation in vertically differentiated markets,” *Marketing Science*, 32(4), 609–621.

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

Hanemann, W. M. (1984): “Discrete/continuous models of consumer demand,” *Econometrica: Journal of the Econometric Society*, pp. 541–561.

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.

Just, D. R., and B. Wansink (2011): “The flat-rate pricing paradox: conflicting effects of “all-you-can-eat” buffet pricing,” *The Review of Economics and Statistics*, 93(1), 193–200.

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

Kivetz, R. (1999): “Advances in research on mental accounting and reason-based choice,” *Marketing Letters*, 10, 249–266.

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

KRISHNAMURTHI, L., AND S. RAJ (1988): "A model of brand choice and purchase quantity price sensitivities," *Marketing Science*, 7(1), 1–20.

LEWIS, M. (2006): "Customer acquisition promotions and customer asset value," *Journal of marketing research*, 43(2), 195–203.

MISRA, S., AND H. S. NAIR (2011): "A structural model of sales-force compensation dynamics: Estimation and field implementation," *Quantitative Marketing and Economics*, 9, 211–257.

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.

RESHADI, F., AND M. P. FITZGERALD (2023): "The pain of payment: A review and research agenda," *Psychology & Marketing*, 40(8), 1672–1688.

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.

STOLE, L. A. (2007): "Price discrimination and competition," *Handbook of industrial organization*, 3, 2221–2299.

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.

Van Heerde, H. J., P. S. Leeflang, and D. R. Wittink (2004): "Decomposing the sales promotion bump with store data," *Marketing Science*, 23(3), 317–334.

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

Villas-Boas, J. M., and U. Schmidt-Mohr (1999): "Oligopoly with asymmetric information: Differentiation in credit markets," *The RAND Journal of Economics*, pp. 375–396.

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.

Waisman, C. (2021): "Selling mechanisms for perishable goods: An empirical analysis of an online resale market for event tickets," *Quantitative Marketing and Economics*, 19(2), 127–178.

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