ABSTRACT

Title of Dissertation Proposal: ESSAYS ON DIGITAL CONTENT PROVISION AND CONSUMPTION

Chutian Wang
Doctor of Philosophy, 2022

Dissertation Directed by: Associate Professor, Yogesh V. Joshi
Department of Marketing

Associate Professor, Bo “Bobby” Zhou
Department of Marketing

Consumption of digital content has become an inseparable part of consumers' lives today. As providers of digital content, media platforms continuously seek to pursue pricing and product design strategies that increase their profits. This dissertation studies media platforms' digital content provision and consumers' consumption decisions. In the first essay, we focus on the pricing of digital content and analyze the impact of consumers' endogenous content consumption on platforms' paywall strategies. Paywalls increase subscription revenues for platforms, but they also impact content consumption and thus advertising revenues. We build an analytical model that endogenizes consumers' content consumption decisions. We find that under moderate ad rates, a metered paywall under which a limited amount of content is provided for free is optimal when consumers display sufficient heterogeneity in their costs of consuming content. We also study how the amount of free content and the subscription price vary with changes in the advertising rate and

consumer preference. In the second essay, we analyze the accuracy of news reported by the news media. When consumers are seeking the truth and accurate reporting is costly, determining the optimal level of accuracy in reporting is a strategic decision for a profit-maximizing media firm. We build an analytical model to study this media firm decision. When consumers and the media firm are both initially uncertain about the true state of the world, we show that the media firm always chooses full accuracy if investigation and reporting are of low cost. However, if achieving accuracy is sufficiently costly, the media firm provides news only when consumers' priors regarding the truth are not too extreme, so that they see enough value in news consumption. Interestingly, consumers' truth-seeking and the firm's profit maximization can lead to reporting inaccuracy and exaggeration of the more likely state *a priori*. We also discuss the implications of polarization in consumers' prior beliefs and the media firm's different objectives on the accuracy of news.

ESSAYS ON DIGITAL CONTENT PROVISION AND CONSUMPTION

by

Chutian Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:

Associate Professor Yogesh V. Joshi, Chair

Associate Professor Bo "Bobby" Zhou, Co-Chair

Professor David Godes

Professor P.K. Kannan

Professor Daniel Vincent, Dean's Representative

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

From 2011 to 2018, the time spent by consumers with digital media has risen significantly, from 3.5 hours per day to 6.5 hours per day (Watson 2019). As providers in this market, the success of media platforms critically depends on their understanding of consumers' utility from content consumption. In this dissertation, we investigate media platforms' digital content provision and pricing as well as consumers' consumption decisions.

The first essay (Chapter 2) investigates platforms' pricing strategy when consumers endogenously decide the amount of content to consume. Media platforms earn revenues from subscription and advertising. A change in market conditions can change platforms' main source of revenues as well as their pricing strategies. In the U.S. newspaper market, market studies have shown that a decline in advertising revenue can lead newspapers to seek consumer payments and embrace paywalls (Arrese 2016; Kumar et al. 2013; Media Insight Project 2017). Previous studies provide several theories on platform's choice of paywalls, including consumers' need for learning about content quality (Halbheer et al. 2014; Li et al. 2019), intertemporal fluctuation in demand (Lambrecht and Misra 2017), and platforms' management of ad space supply (Halbheer et al. 2014). In the first essay (Chapter

2), we build an analytical model to study a platform's optimal paywall design when consumers endogenously decide their consumption. We find that when ad rates are moderate, a metered paywall is optimal when consumers display sufficient heterogeneity in their costs of consuming content. We derive three key insights. First, when ad rates increase, one typically expects platforms to offer more free content and thus generate higher ad revenues. Instead, we find that sometimes offering less free content and lowering the price can attract new subscribers, which generates even higher revenues. Further, the optimal amount of free content may decrease or increase in consumers' valuation for content, depending on ad rates. As consumers' valuation increases, their willingness-to-pay for subscription and their desire for content consumption increases. The former effect dominates under low ad rates and incentivizes the platform to provide less free content, but the opposite is true under high ad rates. Thirdly, we also find that total content consumption can increase with the proportion of light users, because the platform may strategically raise its meter limit to increase ad revenues from non-subscribers.

The second essay (Chapter 3) studies news media platforms' product design strategy. More specifically, we study news media platforms' strategic choice of reporting accuracy. News consumers are seeking information and updating their beliefs about the true state of the world which is random in nature. While the news does not always accurately reflect the truth, it can still improve consumers' understanding of the truth. The accuracy of news influences both consumers' willingness to pay and the production cost for the media platform. Previous theoretical research suggests that the existence of news inaccuracy when consumers obtain utility from conforming ideas (Mullainathan and Shleifer 2005; Gal-Or et al. 2012; Xiang and Sarvary 2007; Zhu and Dukes 2015); or when media platforms get a higher evaluation by confirming consumers' prior (Gentzkow and Shapiro 2006). In the second essay (Chapter 3), we suggest that reporting inaccuracy can result from a platform's unbalanced resource allocation. We analyze the provision of news when information provision is costly and consumers are truth-seeking. We model a media platform's news provision strategy as well as consumers' news consumption decisions when they have common priors. We find that consumers value news more under high prior uncertainty. If investigation and reporting are of low costs, the media firm always offers news with full accuracy. If a higher level of accuracy is sufficiently costly, news provision will be profitable only when consumers' prior is not too extreme so that they see enough value in news. Importantly, although consumers only seek for truth, the cost of news provision can still lead to unbalanced investigative resource allocation. In particular, the media firm would allocate more resources on accurately reporting the more likely state to increase its probability of presenting the truth. In equilibrium, such allocation of resources will exaggerate the likelihood of the state which is more likely a priori. We extend the model to study the impact of polarization in consumers' prior beliefs on news accuracy. We find that polarization makes reporting less accurate when the cost of news provision is moderate, but has no impact when the cost is low or high. The reason is that although polarization decreases the firm's incentive to provide accurate reporting, the significance of its impact varies across cost levels. Interestingly, polarization can make news more accurate when the media firm considers improving consumers' probability of being correct about the truth in addition to earning profits. Polarization means consumers initially show less interest in learning from news, which then incentivizes the firm to attract them with more accurate reporting.

Finally, in Chapter 4, we summarize our findings and highlight our contributions to literature. We also identify several future directions for this research.

Chapter 2: Endogenous Consumption and Metered Paywalls[1]

2.1 Introduction

Ever since content started going digital, platforms have implemented pricing mechanisms to generate revenues from such content. In the news industry, *The Wall Street Journal* (WSJ) has been one of the early platforms to develop along with its digital presence a digital pricing strategy. In 1996, WSJ developed an online version of its print newspaper, *The Wall Street Journal Interactive Edition*, and charged readers for access to this digital content.[2] At the time, other newspapers wondered whether this was a good idea, and questioned the willingness of consumers to pay for access to digital content. The dominant model at the time was to build a digital news platform but generate subscription revenues primarily from the print versions of their products. Over time, as consumer reading habits have evolved, demand for printed newspapers has shrunk, and consumption has shifted dominantly to digital products (Pew Research Center 2011). To adapt to this shifting landscape, in 2011 *The New York Times* (NYT) launched a paywall on its website (Kumar et al. 2013). With the paywall in place, a non-subscriber could still access any

[1]This research is conducted with Bo Zhou and Yogesh V. Joshi.
[2]https://www.wsj.com/articles/SB849581696318836500

article on the newspaper's website, but was limited in the total number of articles accessible per month.

The product design and pricing tactics of the WSJ and NYT and their subsequent market successes were closely observed by other media firms. By May 2018, 77% of newspapers in the U.S. had launched digital paywalls (Lewis 2018). Similarly, 63% of newspapers in Europe have implemented paywalls.[3] Besides newspapers, magazines (e.g., *The New Yorker*, *The Atlantic, Wired*) and digital news media (e.g., *Medium, ESPN*) have also implemented paywalls.[4]

Today, while media platforms have adopted paywalls widely, they differ in their specific product designs. These paywall strategies can be categorized into four types, depending on the amount of content given away for free: [i] *No paywalls*, where all content is provided for free; [ii] *Hard paywalls*, where there is no free content; [iii] *Freemium paywalls,* where access to a pre-determined part of the total content is provided for free, and consumers have to pay to access the premium or "exclusive" part of the content (Kumar et al. 2013; Lambrecht and Misra 2017); and [iv] *Metered paywalls,* where the platform sets a quota on the content that consumers can select to access for free (e.g., 20 free articles per month), and consumers have to pay for accessing the rest of the content (Chiou and Tucker 2013). In other words, the distinction between freemium and metered paywalls is that in the latter, the specific content provided for free is not pre-determined by the platform. In May

[3] https://reutersinstitute.politics.ox.ac.uk/our-research/pay-models-online-news-us-and-europe-2019-update

[4] *The Atlantic*: https://tinyurl.com/y3u8eqwv. *The New Yorker* and *Wired*: https://tinyurl.com/y2xgmhnw. *Medium*: https://tinyurl.com/yyaasdpp. *ESPN*: Lambrecht and Misra (2017).

2018, 20% of newspapers in the U.S. had no paywalls, 5% used freemium paywalls, 72% used metered paywalls, 0.4% used hard paywalls, and 2.5% used other strategies (Lewis 2018). Given the prominence of the use of metered paywalls, in this research, we focus on the optimality of metered paywalls as an effective product design and pricing strategy for digital content platforms.

Consumer behavior in the markets for digital content reveals significant heterogeneity in terms of consumers' valuation for content and the amount of content consumed. The Media Insight Project (2014) reports that 55% of Americans enjoy keeping up with news a lot, 33% only enjoy it some, and 12% don't enjoy it much or at all. This survey also finds that 21% of consumers go in-depth on both breaking and non-breaking news, 49% go in-depth on only one of these two categories, and 30% on neither. A study from *MinnPost*, an online newspaper in Minneapolis with no paywall, shows that 29% of its visitors accounted for 68% of the website visits, while the top 1% accounted for 17%.[5] In a recent survey, 15 metro area publishers reported that 91% of their website visitors view no more than five articles in a thirty-day period, with 68% viewing only one article (Mele et al. 2019). Similarly, a manager in Tronc, Inc. revealed that only 3-5% of the *Los Angeles Times*' unique visitors hit the five-article limit.[6] [7]

The aforementioned heterogeneity can be attributed to several important characteristics of consumers' digital content consumption. First, consumers have hetero-

[5]https://www.minnpost.com/inside-minnpost/2009/10/appreciating-and-counting-loyal-readers.

[6]Tronc, Inc., now Tribune Publishing Company, was the owner of the *Los Angeles Times* between 2007 and 2018.

[7]https://www.thestreet.com/opinion/l-a-times-tops-100-000-in-digital-subscriptions-14315543.

geneous valuation for content, and those with higher valuation are likely to consume more. Second, consumers are heterogeneous in terms of the amount of time they allocate for content consumption since their time is spread across different activities.[8] Third, the breadth of consumers' content interests also influences the amount of content consumed. Consumers with broader news interests tend to consume more news than others. Finally, consumers' marginal utility from content consumption is often diminishing, which also impacts the amount of content consumed. Previous studies have shown that decreasing enjoyment impacts content consumption in various contexts, such as news (Huang 2009), music (Ratner et al. 1999), television shows (Nelson et al. 2009), and mobile apps (Han et al. 2016). For instance, Nelson et al. (2009) find that consumers exhibit different rates of satiation with content, and their pleasure from consumption declines as they consume more. When the marginal utility from content consumption drops below a threshold, consumers switch to other activities that provide higher marginal utility (Herrnstein and Prelec 1991).

Motivated by the above field observations and research findings, in this paper we model consumer heterogeneity along two dimensions: [i] consumers' valuation for content, and [ii] consumers' marginal cost of consumption. For the consumption of news, consumers' valuation for content can vary for a variety of reasons such as their perceived relevance of news, income levels, etc. Consumers with higher valuation for content are more likely to become news subscribers. A high marginal cost for news consumption could originate from lack of leisure time, narrow interest in news

[8] https://www.bls.gov/news.release/pdf/atus.pdf

topics, or a high satiation rate with news. Consumers with higher costs of news consumption are likely to consume less news. Note that previous research on paywalls has typically assumed that consumption is exogenous and all consumers consumed all available content (Halbheer et al. 2014; Lambrecht and Misra 2017). We relax this simplifying assumption and endogenize consumers' consumption decisions based on their individual valuation and marginal cost of consumption. This allows our model to capture consumers' heterogeneity in both the valuation for content and the amount of desired content.

These heterogeneities have important implications for the two distinct revenue sources, advertising and subscription, for digital content platforms. On the advertising side, heterogeneity in the amount of content consumption implies variation in the number of ad impressions and hence monetization per consumer. On the subscription side, heterogeneity in both the valuation for content and the amount of desired content along with the platform's paywall strategy will determine the amount of subscription revenue. By capturing consumers' heterogeneity and endogenizing their consumption decisions, our research sheds new light on platforms' optimal design of paywalls. Additionally, our research shows that consumers' endogenous consumption coupled with the platform's paywall design has significant implications for the information transmitted in the market for news (i.e., the total amount of news consumption).

Given that a media platform has to balance revenues from both advertising and subscription, we investigate the following research questions in this paper: How does heterogeneity in consumers' valuation and their costs of consumption (hereinafter, consumption cost) influence a digital platform's paywall strategy? What fraction of content should be provided for free? How does the endogenous consumption of digital content vary with changes in the marketplace? And finally, what are the implications of these decisions on the total amount of content consumed?

To answer these questions, we build and analyze a theoretical model of a digital media marketplace. The firm operates as a typical two-sided platform in this marketplace, earning revenues from consumer payments for subscriptions on one side and advertiser payments for ad impressions on the other side. The firm makes two decisions: the subscription price that it should charge, and the fraction of content that it should provide for free. Consumers observe the firm's decisions and choose whether to subscribe and how much content to consume. Consumers are heterogeneous in their valuation for content as well as their marginal consumption cost. Modeling consumption costs explicitly enables us to endogenize a consumer's content consumption decisions in this model.

Our analysis shows that consumer heterogeneity in consumption costs alongside valuation can explain the existence of metered paywalls. Specifically, a metered paywall is optimal when ad rates are moderate and two consumer segments have a sufficiently large difference in terms of their segment level consumption costs. Under such conditions, free content can be more effective than a price reduction in increasing profits earned from low-valuation consumers.

Including consumption costs in our model allows us to generate four novel insights regarding the media firm's paywall strategies and the corresponding market outcomes. First, as the ad rate increases, we find that the firm may reduce its provision of free content. This goes against conventional wisdom and previous research which suggest that with a strong ad market, more free content should be provided to consumers (Halbheer et al. 2014; Lambrecht and Misra 2017). When the difference in consumption costs between consumer segments is moderate, it becomes important for the firm to balance the number of ad impressions generated within each segment. As the ad rate increases, offering less free content and lowering the subscription price can generate more ad impressions than simply giving away more free content. The reason is that these two decisions increase subscriptions among two groups of consumers: those whose valuation for content is moderate, but usage is heavy (i.e., users with low consumption costs) and those whose valuation for content is high but usage is light (i.e., users with high consumption costs).[9] If the firm implements a metered paywall, the former group prefers to consume a lot of content but their intermediate willingness-to-pay (WTP) prevents them from subscribing; whereas the latter group is satisfied with the amount of free content so they do not subscribe either. If the firm launches a hard paywall and charges a lower price, both these consumer groups would pay for unlimited access and contribute more to ad impressions, and as a result increase the firm's profits.

Second, we find that under a metered paywall, the impact of consumer valuation on the amount of content provided for free critically depends on the ad rate. As consumer valuation increases, less content should be offered for free under low ad rates, while the opposite is true under high ad rates. This pattern occurs because a higher valuation translates to a higher WTP for paywalled content as well as a

[9]Throughout the paper, the terms "user", "consumer" and "reader" are used interchangeably.

preference to consume more content. The former benefits the firm more when ad rates are low, incentivizing the firm to lock more content behind its paywall. By contrast, the latter matters more when ad rates are high, leading the firm to offer more free content and obtain ad impressions from non-subscribers.

Third, we find that the total amount of content consumed in the market may increase in the proportion of light users who have high consumption costs. In general, when consumers' consumption costs are higher, they consume less content. Thus one might expect that with more light users, the total amount of consumption should decrease (i.e., a decrease in the amount of information flow). This logic indeed would hold true if one did not account for the firm's strategic response to changes in segment sizes. When the light user segment is small, the majority of users (heavy users) prefer a lot of content, but the firm only provides a small amount of free content in order to charge a high price from the heavy users who subscribe. In this case, many non-subscribers hit the paywall. As there are many paywall-hitting users, even a little extra free content can lead to a large increase in ad impressions, which gives the firm an incentive to raise the meter limit. Hence, the amount of free content is highly sensitive to a change in segment sizes. Consequently, when the proportion of light users increases, the firm may find it optimal to offer a lot more free content, and with many paywall-hitting users consuming this content, leading to an increase in the amount of information flow.

Finally, we also identify an interesting relationship between consumers' consumption cost and the firm's subscription price. Intuitively, one would expect that when consumers' consumption cost increases, their WTP for content decreases, which should incentivize the firm to cut its price. Indeed, this is the case when there is a small difference between the consumption costs of light and heavy users. Under such conditions, the firm implements a hard paywall that targets both light and heavy users. If the consumption costs of light users increases, as expected, the firm finds it optimal to reduce the price to maintain subscriptions. However, this is no longer the case if the difference in consumption costs between heavy and light users is sufficiently large. Here, the firm implements a hard paywall that targets only heavy users, and ignores light users. Thus, price becomes invariant in the consumption cost of light users. In addition, when the firm operates a metered paywall, the price would even increase in the consumption cost of light users. A higher consumption cost implies less consumption by light users, so their demand for content declines. In response, it is optimal for the firm to provide less free content and raise its price.

Our results offer several managerial implications for media platform managers. First, offering more free content is not always the optimal choice when the ad rate increases. When the difference in consumption costs across consumer segments is moderate, it can be optimal for the platform to reduce the amount of free content under a higher ad rate. Second, the platform's best response to an increase in consumer valuation depends on the ad market. When the ad market is weak, it should offer less content for free in order to convert more consumers to subscribers. When the ad market is strong, it should offer more content for free to take advantage of consumers' appetite for content.

Our paper also contributes to the understanding of paywalls theoretically.

Our paper is the first analysis of paywalls that incorporates endogenous content consumption. Consumers endogenously decide the quantity of content to consume and obtain utility only from the actual consumption, so their utility may not increase in the total quantity provided by the platform. This is different from the traditional product line design, where consumers can only choose among the quality levels provided to them. By endogenizing consumers' decisions on the amount of content to consume, our model not only captures richer consumer behaviors in digital content consumption but also allows for a more accurate measure of each consumer's contribution to ad revenue. We find that a rise in consumers' valuation for content can lead to more free content being provided, which does not happen under exogenous consumption. Besides, we explicitly study heterogeneity in consumers' consumption costs, which helps us provide a novel explanation for media platforms' paywall strategies and generate new insights into the optimal paywall design. When consumers have heterogeneous valuation as well as heterogeneous consumption costs, a metered paywall can be the most profitable strategy even in the absence of uncertainty and quality learning, which have been the focus of prior studies (e.g., Halbheer et al. 2014; Li et al. 2019). We also find that an increase in ad rates can result in a reduction in the amount of free content, which is different from the findings in previous research (Halbheer et al. 2014; Lambrecht and Misra 2017).

The rest of the paper is organized as follows. Section 2.2 reviews the relevant literature. Section 2.3 introduces the model and Section 2.4 presents the analysis. Section 2.5 presents the main results. Section 2.6 extends the analysis in our main model. Finally, Section 2.7 concludes and discusses future research opportunities.

2.2 Related Literature

Our work contributes to research on revenue models for media platforms. There are three distinct models: [i] "free": content is free and revenue is completely generated via ads; [ii] "paid": content needs to be paid for, regardless of the existence of ads; and [iii] "free-plus-paid": there is a free version as well as a paid version, and ads may be present in both versions.[10] In the rest of this section, we review the previous research on each of these revenue models, and discuss our contributions relative to past work.

The "free" model where a firm charges no fee from its audience and earns revenues via advertising sees widespread use within the media industry. Radio and television programs often operate on the free model (Steiner 1954; Spence and Owen 1977). Gabszewicz et al. (2001) has found that media firms would set a zero price and provide content for free under a strong demand for advertising. When viewers have a distaste for ads, media firms may choose lower levels of ads to compete for them (Anderson and Coate 2005; Gal-Or and Dukes 2003). Additionally, competition among pure ad-supported media may also lead to lower program quality levels (Liu et al. 2004). Our analysis shows that a media platform chooses the free model when advertisers' WTP for a consumer's impression exceeds each consumer's WTP for content. Otherwise, it would launch a paywall and charge a price for the

[10] "Freemium" models are a subset of "free-plus-paid" models. We discuss the distinction between these models subsequently.

subscription.

The "paid" model is another popular model among media firms. Under this model, consumers have to pay a price to access the content. A media can choose a paid-without-ads model (e.g., HBO, Netflix) or a paid-with-ads model (e.g., printed newspapers and magazines, digital newspapers with hard paywalls). When consumers incur disutility from ads, pure ad-supported media firms can reduce ad levels and adopt a paid-with-ads model to earn higher profits (Anderson and Coate 2005). If consumers' ad disutility is sufficiently high, media firms may even operate on a paid-without-ads model (Amaldoss et al. 2020; Tåg 2009). Paid-without-ads models are analyzed by previous papers with the focus on media bias (Mullainathan and Shleifer 2005; Xiang and Sarvary 2007; Yildirim et al. 2013) and preview provision (Xiang and Soberman 2011). In those cases, the media can be viewed as a one-sided firm rather than a two-sided platform (Anderson and Jullien 2015). More theoretical papers have studied paid-with-ads models, with the focus on the impact of customer loyalty (Chen and Xie 2007), content substitutability (Godes et al. 2009; Kind et al. 2009), advertiser heterogeneity (Gal-Or et al. 2012; Lin 2020; Prasad et al. 2003), and ad externality (Amaldoss et al. 2020; Chatterjee and Zhou 2020; Godes et al. 2009; Kind et al. 2009; Lin 2020). In this stream of research, several papers study media firms' product line design, where products with different levels of ads are offered (Appel et al. 2020; Casadesus-Masanell and Zhu 2010; Lin 2020; Prasad et al. 2003; Tåg 2009). In a recent paper, Lin (2020) finds that when consumers display heterogeneous ad disutility and advertisers possess heterogeneous ability to convert ad viewers, a media platform can use versioning to screen consumers and advertisers on both sides of the market. By contrast, our paper focuses on a platform's provision of free content. The platform may choose a paid-with-ads model (using a hard paywall), but under certain conditions, it is more profitable with a free-plus-paid model (using a metered paywall). In this product line, the free and the paid versions do not differ in the intensity of ads, but only in the amount of content. Our analysis reveals the conditions under which a metered paywall is favored by the platform.

Our paper also adds to the research on paid models by incorporating consumers' heterogeneous costs of consumption. The cost of consumption has been considered in this stream of research. Chellappa and Mehra (2018) analyze the impact of homogeneous consumers' usage cost on the versioning of information goods, where the firm runs a paid-without-ads model. Three papers have incorporated homogeneous consumption costs in the analysis of paid-with-ads models. Godes et al. (2009) and Kind et al. (2009) model homogeneous consumption costs when they study media firms' competition and content substitutability. Guo et al. (2019) study the implication of consumer ad disutility on mobile apps' reward advertising strategy in the presence of homogeneous consumption costs. Different from their papers, our paper focuses on a two-sided media platform's paywall strategy. We explicitly model both the level of consumption cost and the heterogeneity of consumption costs across consumer segments. By including heterogeneous consumption costs in the analysis, we find that an increase in some consumers' consumption costs can lead to a higher subscription price. This is opposite to the finding under homogeneous consumption costs, where the price decreases in the consumption cost. We further discuss this finding in Section 2.5.3.

Besides the free model and the paid model, a media platform may choose a free-plus-paid model, where it offers a product line with both the free and the paid versions of its product. Classic papers in product line design examine physical products with positive marginal costs, which are rarely provided for free (Maskin and Riley 1984; Moorthy 1984; Mussa and Rosen 1978). By contrast, digital products can have zero marginal cost, so they can be offered for free (Lambrecht et al. 2014).

Under a free-plus-paid model, the free version and the paid version may differ in their quality, functionality, and quantity. For example, under a typical "freemium" model, the firm provides a paid version of a product and a free version with lower quality or less functionality (Li et al. 2019). Previous research has shown that firms may choose freemium models in the absence of advertising. This is because a freemium model can help consumers learn quality (Li et al. 2019; Niculescu and Wu 2014), generate word-of-mouth (Niculescu and Wu 2014; Kamada and Öry 2020), or attract free users who bring benefits to paid users through network effects (Shi et al. 2019). In addition to these benefits, two recent papers have analyzed freemium models in the presence of advertising. Lambrecht and Misra (2017) show that platforms will have more flexibility when dealing with intertemporal fluctuations in the demand for content if they launch freemium paywalls rather than hard paywalls or no paywalls. Appel et al. (2020) suggest that the firm provides two versions, an ad-supported free version and an ad-free paid version, of its product when consumers have disutility for ads and uncertainty about product fit.[11] In our

[11] Appel et al. (2020) also consider the impact of satiation on the firm's strategy. In their

paper, we analyze metered paywalls in absence of the above-mentioned benefits of free products. To design a metered paywall, a platform needs to consider consumers' endogenous choice of the amount of content to consume. We elaborate on this in the following discussion.

Under a metered paywall, the free and paid versions have the same quality and functionality, and they only differ in the quantity provided (Chiou and Tucker 2013). Consumers can endogenously choose a fraction of the available content to consume based on their valuation and costs of consumption. Once consumers have reached their personal optimal level of consumption, additional content, though available, does not increase their utility. In other words, a larger quantity of content may not directly translate into higher consumers' WTP. If the firm chooses a hard paywall, its subscription price needs to be sufficiently low to attract consumers who are interested in a relatively low amount of content. In contrast to a hard paywall, a metered paywall allows the firm to obtain impressions from those consumers while maintaining a relatively high subscription price. Our paper shows that consumers' endogenous choice of consumption quantity plays an important role in a platform's paywall design.

Metered paywalls have been studied by several recent papers. Empirical evidence has shown that a metered paywall can bring more digital subscription revenues to the newspaper (Aral and Dhillon 2020; Pattabhiramaiah et al. 2019) but also has negative consequences: a declined number of unique website visitors, reduced con-

model, they assume that satiation decreases consumers' utility but does not influence the amount of consumption. By contrast, the consumption cost in our model directly influences the amount of consumption.

sumer engagement, and less word-of-mouth on social media (Aral and Dhillon 2020; Cook and Attari 2012; Oh et al. 2016; Pattabhiramaiah et al. 2019). Halbheer et al. (2014) have provided two theoretical explanations for the optimality of metered paywalls. First, similar to the explanation in Li et al. (2019) and Niculescu and Wu (2014), a metered paywall can help consumers learn the quality of content, thus will be the optimal choice of a platform. Second, a platform may use the amount of free content to adjust its supply of ad space, which may enable it to charge a higher price from advertisers. In addition to these explanations, we find that consumer heterogeneity in consumption costs also explains the optimality of metered paywalls. We show the optimality of metered paywalls in the absence of the two factors in Halbheer et al. (2014). Note that previous research suggests that the amount of free content should increase when the demand for advertising increases (Halbheer et al. 2014; Lambrecht and Misra 2017). By contrast, we find the opposite can happen. Compared with giving more free content, reducing free content and cutting the price can lead to more subscriptions as well as consumption under certain conditions. In the next section, we describe our model setup.

2.3 Model

In our model, a digital media platform provides content and sells ad space. To facilitate exposition, we denote this media platform as a digital newspaper. We assume the newspaper to be a monopolist in the content market and a price taker in the ad market. The price taker assumption can be supported by the fact that a digital newspaper competes with other media forms (e.g., TV, radio, and magazines) for advertisers and thus has limited market power. The newspaper chooses its paywall strategy to maximize profits. Consumers maximize their utility from reading news by deciding on how much to consume and whether to subscribe. In the rest of this section, we describe the paywall strategies of the media firm, the decisions of consumers, and the timing of the game.

2.3.1 The Media Platform

The total number of articles in the digital newspaper is normalized to be one. The cost of content production is assumed to be zero so that we can focus on the newspaper's decision on the pricing stage. Taking the content as given, the newspaper can offer two versions of newspaper to its readers: a free version with a fraction $f \in [0, 1]$ of the newspaper that can be accessed to with no charge, and a paid version with full content that requires a subscription fee of $p \in [0, \infty)$.[12] Consistent with the industry practice (e.g., *The New York Times*), we assume that each piece of the article includes one piece of display ad, regardless of the version it is in. In the rest of the paper, we use "consumers" and "readers" interchangeably, as well as "the firm" and "the newspaper."

Three types of paywall strategy are defined by the amount of free content provided. A *hard paywall* strategy means offering nothing for free but only a paid version ($f = 0$ and $p > 0$). A *metered paywall* strategy means providing a free version with limited content and a paid version with all the content ($0 < f < 1$ and

[12] We discuss the implication of negative prices in Section 2.6.1.

$p > 0$). A *no-paywall* strategy means making all the content available for free and earning all revenues from advertising ($f = 1$ and $p = 0$). The firm's problem can be written as the following:

$$\max_{f,p} \pi(f, p; \theta) = p \cdot D(f, p; \theta) + A \cdot I(f, p; \theta) \tag{2.1}$$

In Equation (2.1), θ represents parameters of consumer preference, which we discuss subsequently. D is the demand for subscription. A is advertisers' valuation for an ad impression, which is also the ad price per ad impression. I is the total number of ad impressions. Given consumer preference and the ad price, the newspaper simultaneously decides the amount of free content and the subscription price to maximize its profits, π.

2.3.2 Consumers

Consumers obtain utility from content consumption. Their marginal utility from consuming news decreases in the number of articles they already read. When the marginal utility declines to zero (the utility from one's outside option), a consumer would stop reading. Consumers' utility function is given by Equation (2.2),

$$u(n; v, c) = vn - \frac{cn^2}{2} \tag{2.2}$$

In Equation (2.2), $n \in [0, 1]$ is the amount of articles read by a consumer. v is a consumer's valuation for each piece of article and is heterogeneous across consumers: $v \sim U[0, V]$, in which $V > 0$ is the upper-bound of the uniform distribution.[13] c

[13]If we consider consumers' disutility from seeing ads, then we can interpret v as a consumer's

measures the rate at which a consumer's marginal utility from content decreases. A larger c implies a higher consumption cost (i.e., reading cost) due to the lack of time, a narrow interest, or satiation. In reality, some consumers have a high cost and only read a small fraction of the newspaper, while others can spend hours reading news and are thus being called "news junkie." We capture such heterogeneity by assuming that there are two groups of consumers: light readers (L type) with a high cost, $c_L > 0$, and of size $\alpha \in [0, 1]$, and heavy readers (H type) with a low cost, $c_H = 0$, and of size $(1 - \alpha)$. In the rest of the paper, we omit the subscript and use c to represent light readers' consumption cost. c can also be interpreted as the difference between the two segments' consumption costs. We also assume that a consumer's valuation for content is independent of her consumption cost.[14]

Consumers simultaneously make two decisions: how much to read and whether to subscribe. If they choose the free version, the number of articles is capped by f. If they pay a price of p, they will have unlimited access to all the content. If the newspaper offers everything for free, then they only decide on how much to read.

2.3.3 Timing of the Game

The timing of the game is summarized as below.

1.*Paywall strategy*: given consumer preference ($\theta = (V, c, \alpha)$) and ad market condition (A), the firm simultaneously decides the fraction of free content and the subscription price.

---

utility from a unit of content net of the disutility from a piece of display ad.

[14]In Section 2.6.2, we analyze an extension where v and c are correlated.

2. *Subscription and consumption*: each consumer simultaneously decides on the amount of content to consume and whether or not to subscribe.

3. *Revenue realization*: the firm obtains subscription and advertising revenues.

2.4 Model Analysis

2.4.1 Consumers' Decisions

We derive the equilibrium through backward induction and start with consumers' decisions. Consumers choose the number of articles to read to maximize their utility. In our model, a heavy reader with $c_H = 0$ would read as much as she can. If she doesn't subscribe, she would finish all the free articles and obtain a utility of $u(f; v, 0)$. If she pays p to subscribe, she would read the whole newspaper and obtain a utility of $u(1; v, 0) - p$. We denote $\overline{v}_H$ as the valuation of the heavy reader who is indifferent between the free version and the paid version. As Figure 2.1 shows, anyone whose $v \in [0, \overline{v}_H]$ will only read the free articles and hit the meter limit, while those with $v \in (\overline{v}_H, V]$ will pay to subscribe.[15]

Insert Figure 2.1 about here

For a light reader, the optimal amount to read is $\frac{v}{c}$, so light readers with higher valuation tend to consume more news. If $\frac{v}{c}$ is smaller than f, this reader

[15]If the newspaper offers some free articles but charges a price that is too high, it is possible that $\overline{v}_H > V$, which implies that the entire heavy reader (H) segment chooses the free version. We analyze this case in Appendix A.

will be fully satisfied with the free version without hitting the paywall. However, if her valuation v is sufficiently high, she would like to read more than f. In this situation, she either stays as a free user and uses up her free quota, or pays to subscribe and reads articles until her marginal utility declines to zero. If $\frac{v}{c}$ is greater than 1, then she would finish all the articles. The condition for a light reader to subscribe is $u(f; v, c) \leq u(\min\{\frac{v}{c}, 1\}; v, c) - p$. Define $\overline{v}_L$ as the valuation of the light reader who is indifferent between the free version and the paid version. Given the price and the amount of free content, there can be three groups of light readers with different behaviors, as Figure 2.1 indicates. Readers with a sufficiently low valuation ($v \in [0, fc)$) read $\frac{v}{c}$, while those with an intermediate valuation ($v \in [fc, \overline{v}_L]$) read f and hit the paywall. Those with a sufficiently high valuation ($v \in (\overline{v}_L, V]$) become subscribers and read $\min\{\frac{v}{c}, 1\}$. The minimum function restricts the amount of reading to be weakly less than one, which is the total amount of content of the newspaper.

2.4.2 The Media Platform's Decisions

Based on consumers' decisions described above, we can calculate the newspaper's subscription demand and total ad impressions. For example, when the firm obtains subscriptions from both heavy and light reader segments ($\overline{v}_H, \overline{v}_L \in (0, V)$), the demand is Equation (2.3) and the total ad impressions are Equation (2.4).

$$D(p, f; \theta) = D_H + D_L = (1 - \alpha)\frac{V - \overline{v}_H}{V} + \alpha\frac{V - \overline{v}_L}{V} \tag{2.3}$$

$$I(p, f; \theta) = I_H + I_L, \text{ where} \tag{2.4}$$
$$I_H = (1-\alpha)(\underbrace{\int_0^{\overline{v}_H} \frac{f}{V} dv}_{\text{H,hitting}} + \underbrace{\int_{\overline{v}_H}^{V} \frac{1}{V} dv}_{\text{H,paid}}),$$
$$I_L = \alpha(\underbrace{\int_0^{fc} \frac{v}{cV} dv}_{\text{L,not hitting}} + \underbrace{\int_{fc}^{\overline{v}_L} \frac{f}{V} dv}_{\text{L,hitting}} + \underbrace{\int_{\overline{v}_L}^{V} \frac{\min\{\frac{v}{c}, 1\}}{V} dv)}_{\text{L,paid}}$$

With the demand and the ad impressions, we can solve for the firm's optimal decisions. When the firm obtains subscription only from one segment, the demand and the ad impression functions will change, which leads to different solutions to the profit maximization problem.[16] We compare the profits the newspaper can possibly generate across all strategies to find its equilibrium strategy.

To understand the firm's trade-off behind its free content provision, it is helpful to highlight two contrasting effects of free content on the firm's profits: cannibalization effect and consumption-expansion effect. We describe these two effects in detail using Equation (2.5), where $p^*(f)$ is the solution to the first-order condition, $\partial\pi(p, f)/\partial p = 0$.

$$\frac{d\pi(p^*(f), f)}{df} = p^* \underbrace{\left(\frac{\partial D}{\partial f} + \frac{\partial D}{\partial p}\frac{\partial p^*}{\partial f}\right)}_{\text{decrease in demand}} + D \underbrace{\frac{\partial p^*}{\partial f}}_{\text{price cut}} + A \underbrace{\left(\frac{\partial I}{\partial f} + \frac{\partial I}{\partial p}\frac{\partial p^*}{\partial f}\right)}_{\text{increase in impression}} \tag{2.5}$$
$$= p^* \frac{\partial D}{\partial f} + A \frac{\partial I}{\partial f} \tag{2.6}$$

In Equation (2.5), the *cannibalization effect* of free content on profits consists of two parts: the lost profit due to a decrease in subscription demand and the lost profit

[16] This case is analyzed in Appendix A.

due to a price cut (the first and second terms in Equation (2.5), respectively). First, when more free content is provided, marginal subscribers would churn and choose the free version (captured by $\frac{\partial D}{\partial f}$). Faced with this weaker demand for subscription, the firm cuts its price to retain some subscribers ($\frac{\partial D}{\partial p}\frac{\partial p^*}{\partial f}$), which attenuates the negative impact of subscriber churning. Second, the price cut results in subscribers paying less for their access to content, so subscription revenues drop ($D\frac{\partial p^*}{\partial f}$). The remaining part of Equation (2.5) captures the *consumption-expansion effect* of free content on profits. More free content allows some non-subscribers to consume more, but it also makes some subscribers churn and consume less (both included in $\frac{\partial I}{\partial f}$). Besides, more free content also indirectly influences consumption through the price ($\frac{\partial I}{\partial p}\frac{\partial p^*}{\partial f}$). As we discussed above, the price cut caused by free content provision can attract some subscriptions. With a subscription, these readers are no longer limited by the paywall and would consume more content.

We can simplify Equation (2.5) by removing the terms representing the indirect effects of free content that sum up to zero.[17] This is because given any change in the amount of free content, the firm adjusts its price accordingly such that the revenues from the additional subscribers ($p^*\frac{\partial D}{\partial p}\frac{\partial p^*}{\partial f}$) and their ad impressions ($A\frac{\partial I}{\partial p}\frac{\partial p^*}{\partial f}$) offset the loss due to the price cut ($D\frac{\partial p^*}{\partial f}$). After removing those indirect effect terms from Equation (2.5), the terms in Equation (2.6) represent the direct effects of cannibalization (the first term) and consumption-expansion (the second term). In the discussion of the results, we will refer to these two equations to elaborate on the

[17]In Equation (2.5), $p^*\frac{\partial D}{\partial p}\frac{\partial p^*}{\partial f}+D\frac{\partial p^*}{\partial f}+A\frac{\partial I}{\partial p}\frac{\partial p^*}{\partial f}=\left(p^*\frac{\partial D}{\partial p}+D+A\frac{\partial I}{\partial p}\right)\frac{\partial p^*}{\partial f}=\frac{\partial \pi}{\partial p}\frac{\partial p^*}{\partial f}=0$, because $\frac{\partial \pi}{\partial p}=0$ when $p=p^*(f)$. This result can also be obtained by applying the Envelope Theorem to Equation (2.5).

intuition behind the firm's decisions.

Next, we study the firm's equilibrium strategy and discuss how these two effects guide the firm's provision of free content.

2.4.3 Equilibrium Strategy

As described in Section 2.4.2, we compare the profits across all strategies and derive the equilibrium strategy. Specifically, we analyze the game in the following range of parameters: $V > 0$, $c > 0$, $0 \leq \alpha \leq 1$, and $A \geq 0$. In this subsection, we first discuss the case where only heavy readers exist in the market ($\alpha = 0$), then analyze the equilibrium strategy when both segments are in the market ($0 < \alpha < 1$), and finally present the case where only light readers are in the market ($\alpha = 1$). Lemma 1 below presents the firm's equilibrium strategy in a market full of heavy readers.

Lemma 1. *When $\alpha = 0$, the firm maximizes its profits by choosing a hard paywall strategy targeting the H segment (denoted by subscript "hard, H") with $f_{h,H} = 0$ and $p_{h,H} = \frac{V-A}{2}$ when $A < V$; or choosing a no-paywall strategy with $f_n = 1$ and $p_n = 0$ when $A \geq V$.*

When all consumers are heavy readers ($\alpha = 0$), the firm chooses either a hard paywall strategy or a no-paywall strategy. Under low ad rates, it launches a hard-H paywall and provides no free content. This hard paywall strategy targets the heavy reader (H) segment, which means that all subscriptions are obtained from that segment. Though free content can generate ad revenues from non-subscribers, it can also lead to subscriber churning as well as a price cut, resulting in a larger loss in subscription revenues. When the ad rate increases, the firm has a stronger incentive to earn ad impressions, so it cuts the price to attract subscribers to consume its content. This negative impact of ad rates (or advertisers' valuation for impressions) on subscription price is well-documented in the research on media markets (Gal-Or et al. 2012; Godes et al. 2009; Weyl 2010). When the ad rate exceeds the highest consumers' WTP for content, the subscription price is cut to zero. As monetizing eyeballs becomes more profitable than charging for content, the firm would give all content for free and let readers read as much as they like, i.e., choosing a no-paywall strategy. This is usually the case for news outlets offering popular articles without any in-depth analysis.

Next, we summarize the equilibrium outcome when both types of consumers exist in the market.

Lemma 2. *When $0 < \alpha < 1$, the firm's equilibrium strategy is summarized in Table 3.1, where* $p_{h,HL2} = \frac{2V-2A-c\alpha}{4}$, $p_{h,HL} = \frac{V-A}{2(1-\alpha)} + \frac{9c\alpha^2-3\alpha\sqrt{16c(1-\alpha)(V-A)+9c^2\alpha^2}}{16(1-\alpha)^2}$, $f_m = \frac{V}{c} - \frac{(1-\alpha)(V-A)^2}{4c\alpha A}$, *and* $p_m = \frac{V-A}{2}\left(1-\frac{V}{c}+\frac{(1-\alpha)(V-A)^2}{4c\alpha A}\right)$.[18]

Insert Table 2.1 about here

When the market consists of a mixture of both consumer segments ($0 < \alpha < 1$), a metered paywall strategy is the equilibrium strategy when ad rates are moderate ($\underline{A} < A < V$) and consumers display sufficient difference in consumption costs ($c > \underline{c}$). Figure 2.2 illustrates the conditions for different equilibrium strategies.

[18]The (implicit) expression of $\underline{c}$ is given in Appendix A. $\underline{A} = \frac{(1-\sqrt{\alpha})^2 V}{1-\alpha}$.

Insert Figure 2.2 about here

It is clear that with a sufficiently high ad rate ($A \geq V$), the media firm will completely rely on ad revenues and adopt a no-paywall strategy by offering all the content for free. In the rest of the discussion of Lemma 2, we focus on the firm's paywall strategies under low ($A \leq \underline{A}$) and intermediate ($\underline{A} < A < V$) ad rates. Under those conditions, the consumption cost plays an important role in the firm's decisions. First, it determines how different the two consumer segments are. This influences the firm's choice of target subscribers and thus its price. Second, it also affects the magnitudes of the two effects of free content discussed in Section 2.4.2, which in turn determines the provision of free content and therefore the type of paywall.

The firm always chooses a hard paywall when the difference in consumption costs is small ($c \leq \underline{c}$). In this case, even the light readers desire a large amount of content, especially those with a high valuation for content. As both light and heavy readers want a similar amount of content and have similar WTP, the firm can earn more profits by attracting subscribers from both segments than from only one segment (of heavy readers). In this case, providing free content will conflict with the firm's effort in incentivizing light readers to subscribe, because these readers obtain most utility from the consumption of the first several units of content. Most of them would choose the free version if it is available. Therefore, the firm offers only a paid version targeting both segments (a hard-HL paywall or a hard-HL2 paywall), which can be interpreted as setting up a "mass market" hard paywall. In the real world, this corresponds to the cases where all the consumers of the firm have similar interests or tastes for content. When ad rates increase, it would decrease the price to attract more subscribers, similar to the case in Lemma 1.

A metered paywall can be the equilibrium strategy under a sufficiently large difference in consumption costs ($c > \underline{c}$). Under this condition, light readers want much less content than heavy readers so that their willingness-to-pay for a subscription will be substantially lower. As a result, the firm only targets heavy readers as its potential subscribers, shifting from the mass market appeal to a niche market focus. When ad rates are low ($A \leq \underline{A}$), the firm operates a hard paywall targeting heavy readers (hard-H), offering no free content. Because low ad rates provide little incentive for the firm to capitalize on non-subscribers' impressions. This changes when the ad rate falls in an intermediate range ($\underline{A} < A < V$). Higher ad rates make non-subscribers' impressions more valuable. The firm would like to attract those readers, but it is not worthwhile to lower the price to convert them to subscribers, especially the light readers. Instead, the firm can expand its product line using a metered paywall strategy. It offers a paid version to get subscriptions from its core consumers (heavy readers with high valuation) and a free version to obtain ad impressions from all other readers. On the consumer side, the price is higher than that under a hard-HL paywall, because the firm no longer needs to induce light readers to subscribe. On the advertising side, more ad impressions are generated than those under a hard-H paywall, because non-subscribers would consume the free content. To summarize, when consumers display sufficient difference in consumption costs, a metered paywall dominates hard paywalls under moderate ad rates. Note that we need to compare the ad rate with the highest valuation to determine whether it is low or moderate. In the example of WSJ, the price of an ad impression is relatively low compared with consumers' valuation for the content, because consumers rely on the information in articles to make investment decisions. By contrast, the ratio between the ad rate and consumers' valuation for content is larger at NYT. As a result, a hard paywall targeting heavy readers is more favorable to WSJ, while a metered paywall is better for NYT.

It is worth noting that the difference in consumption costs across segments can be viewed as a difference in segment-level price elasticities. The subscription demand in the light reader segment is more elastic than that in the heavy reader segment. Light readers' higher consumption costs reduce their valuation for each unit of content and make them endogenously want less content. As a result, obtaining subscriptions from light readers is optimal only when their subscription demand is not much more elastic than heavy readers (i.e., c is small). By contrast, the firm would either focus on the segment with a less elastic demand (i.e., heavy readers) under low ad rates or offers different versions to serve both segments under moderate ad rates if the difference in segment-level price elasticities becomes sufficiently large (i.e., c is large).

Finally, we present the equilibrium outcome when the market is full of light readers.

Lemma 3. *When $\alpha = 1$, the firm maximizes its profits by choosing a hard* *paywall strategy with* $p_{h,HL2}$ *when* $0 \leq A < V$ *and* $0 < c \leq \frac{2(V-A)}{3}$*; or a hard paywall strategy with* $p_{h,L} = \frac{2(V-A)^2}{9c}$ *when* $0 \leq A < V$ *and* $c > \frac{2(V-A)}{3}$*; or a no-paywall strategy when* $A \geq V$*.*

When all consumers are light readers ($\alpha = 1$) and have homogeneous consumption costs, the firm should choose a hard paywall under low ad rates and no-paywall under high ad rates. As analyzed in the discussion of Lemma 2, a free version can make it harder to drive light readers to subscribe, so the firm would use hard paywall strategies under low ad rates, serving only readers with a high valuation. The price decreases with ad rates and becomes zero under sufficiently high ad rates ($A \geq V$).

2.5 Results

Our analysis reveals several interesting insights into the design of paywalls and its implications for the market outcomes. In this section, we present the most important findings and relegate other findings to Appendix A. Specifically, we first discuss the optimal provision of free content, and then further discuss its implication for the information flow from the media to its readers. Finally, we elaborate on the optimal subscription price under different types of paywalls and how it changes in the consumption cost. Given our primary interest in metered paywalls, we focus on the case where metered paywalls can be an equilibrium strategy, i.e., both segments exist in the market ($0 < \alpha < 1$).

2.5.1 Free Content Provision

By endogenizing consumers' consumption of content, our model offers novel insights into the firm's response to a change in market circumstances. For example, conventional wisdom suggests that the firm provides a larger fraction of content for free when the ad market is strong because that should attract more eyeballs and increase ad revenues. Previous theoretical research, assuming consumers' exogenous consumption of content, also makes similar recommendations (Halbheer et al. 2014; Lambrecht and Misra 2017). In our model, we consider endogenous consumption and obtain an unexpected result: When the ad rate increases, the firm can earn more by reducing the amount of free content. We summarize this result in the following proposition.

Proposition 1. *As the ad rate* (A) *increases, the firm can increase its profits by reducing the amount of free content. This happens under moderate ad rates and moderate differences in the consumption costs between segments* ($A = A(\underline{c})$ *and* $\underline{c}(\underline{A}) < c < \max_{A \in (\underline{A}, V)} \underline{c}(A)$)*.*[19]

Proposition 1 shows that a higher ad rate can lead to a lower amount of free content. More specifically, under a moderate ad rate and a moderate difference between light and heavy readers' consumption costs, an increase in the ad rate can induce the firm to shift from a metered paywall to a hard paywall, decreasing its amount of free content. For example, when c is slightly larger than the threshold $\underline{c}(\underline{A})$ (e.g., $c = 2$ in Figure 2.2), as the ad rate increases, the firm first shifts from a

[19] $A(\underline{c})$ is the inverse function of $\underline{c}(A)$.

hard paywall targeting heavy readers (hard-H) to a metered paywall, then to a hard paywall targeting both segments (hard-HL), next to a metered paywall again, and finally to no-paywall. The changes in the firm's decisions as A increases from 0 to V are shown in Figure 2.3.

Insert Figure 2.3 about here

Next, we explain the intuition behind these strategy shifts. When the ad rate is low ($A \leq \underline{A}$), the firm chooses a hard-H paywall. As the ad rate increases, the firm cuts the price to attract more subscribers. When the ad rate exceeds the threshold, $\underline{A}$, the consumption-expansion effect (discussed in Section 4.2, after Equation (2.5)) starts to dominate the cannibalization effect so that the firm switches to a metered paywall. Further increases in ad rates induce the firm to encourage more content consumption with more free content and lower prices until the metered paywall is dominated by a hard-HL paywall.

The shift from a metered paywall to a hard-HL paywall under moderate ad rates is more nuanced. Under a metered paywall, two groups of readers hit the meter limit. The first group consists of heavy readers with relatively low valuation, who thirst for more content but would not pay much for it. The second group consists of light readers with high valuation, who value the content but only consume a little. These two groups read only the f_m free articles and would not pay the subscription fee. However, if the firm cuts the price as well as the amount of free content, it can incentivize some of those readers to subscribe. Heavy readers with intermediate valuation find the low price attractive, and light readers with high valuation no longer have the free option. The removal of the free version is very important because light readers enjoy the first several units of content more than the subsequent units. As a result, the appeal of a subscription increases so that these two groups of readers are willing to become subscribers. Upon subscribing, their ad impression will be 1 (for heavy readers) or $\frac{v}{c}$ (for light readers), much higher than f_m under a moderate ad rate. As a result, when the firm shifts from a metered paywall to a hard-HL paywall, the additional impression gained from these new subscribers outweighs the impression loss from low-valuation readers in both segments.[20]

Under a hard-HL paywall, the firm further reduces the price as the ad rate increases, but it will provide free content again when the ad rate is sufficiently high. There are two reasons why price reduction will be less effective in generating revenues than free content provision under these circumstances. First, the price has to be very low to attract new subscribers. Such a low price means leaving a large amount of surplus to high-WTP consumers, especially those in the heavy reader (H) segment. Second, the ad impressions generated by the marginal subscriber also decline because the light readers who haven't subscribed yet are those who have

[20]To elaborate on the intuition, we provide a numerical example for this strategy shift and compared it with the case where only price reduction or free content provision is considered. When $V = 1$, $\alpha = 0.25$, $c = 2$, the firm chooses a metered paywall at $A = 0.38$, with $f_m = 0.12$ and $p_m = 0.27$. In this situation, 69% of heavy readers subscribe, while 31% of heavy readers and 25% of light readers hit the paywall. When the ad rate increases to $A = 0.40$, the firm shifts to a hard-HL paywall, with $f = 0$ and $p = 0.18$. As a result, 89% of heavy readers and 34% of light readers subscribe. The shift of strategy converts most of the paywall-hitting readers to subscribers. If we only consider a price reduction (holding f constant), then at $A = 0.40$, the price will be cut to $p = 0.26$. This results in 70% of heavy readers subscribing and no light readers subscribing. Alternatively, if we only consider free content provision (holding p constant), then at $A = 0.40$, the amount of free content will be increased to $f = 0.13$. The increase in profits is smaller than that from jointly optimizing p and f.

a relatively low valuation for content. Even if they are converted to subscribers, they would only read a little. Combining these two factors above, the benefit from a further price reduction becomes smaller under a high ad rate. Compared with a hard paywall, launching a metered paywall not only captures the surplus from the high-WTP readers but also obtains the ad impressions from low-WTP readers. It is very difficult to convert the latter group to subscribers with a hard paywall, but a metered paywall can obtain those ad impressions without sacrificing too much on the price.

Finally, an ad rate as high as the highest valuation makes the no-paywall strategy optimal, where the consumption of news articles is fully subsidized by advertising. When consumers' valuation for content is lower than advertisers' payment for their impressions, being purely ad-supported is the most profitable choice for the media firm.

Besides studying the impact of advertisers' valuation on free content provision, we also analyze the impact of consumers' valuation on the provision of free content. Interestingly, depending on the ad rate, an increase in consumers' valuation for content can either increase or decrease the amount of free content, as summarized in the proposition below.

Proposition 2. *As the upper-bound of consumers' valuation for content* (V) *increases, the firm that operates a metered paywall should decrease the amount of free content under low ad rates* ($A < \frac{(1-\alpha)V}{1+\alpha}$)*, but increase it under high ad rates* ($A > \frac{(1-\alpha)V}{1+\alpha}$).

First, note that an increase in the upper-bound of consumers' valuation means on average consumers have a higher valuation for content. Intuitively, when consumers' WTP increases, the firm should always charge a higher price. However, its decision on the amount of free content is less straightforward, as it depends on the ad rate. We explain the specifics below.

In Section 2.4.2, we discussed the cannibalization and consumption-expansion effects of free content on profits. The optimal amount of free content balances these two driving forces. Now consider the impact of an increase in the upper-bound, V, on each of these effects. On the one hand, the cannibalization effect becomes stronger. Because when readers have a higher valuation for content, the firm can charge a higher price for the paywalled content ($p_m = \frac{(1-f_m)(V-A)}{2}$). Consequently, giving content away for free means a larger loss. On the other hand, one part of the consumption-expansion effect, the impression gain from light readers, also becomes larger, since overall they want to read more content with a higher V (recall that they read $\min\{\frac{v}{c}, 1\}$). To compare these impacts of consumers' valuation on the two effects of free content, we use Equation (2.7) to illustrate the firm's trade-off.

$$\begin{aligned} \frac{d\pi_m(p_m(f), f)}{df} &= p_m \underbrace{\frac{\partial D_H}{\partial f}}_{\text{demand decrease}} + A \underbrace{\frac{\partial I_L}{\partial f}}_{\text{increase in impression, L}} \\ &= \underbrace{-\frac{(1-\alpha)(V-A)^2}{4V}}_{\text{loss, H}} + \underbrace{\frac{A\alpha(V-f_m c)}{V}}_{\text{gain, L}} \end{aligned} \tag{2.7}$$

Equation (2.7) is a special case of Equation (2.6) in a metered paywall equilibrium, and it divides the two effects of free content by segments. In Equation (2.7), the first term in the first row captures the impact of free content on subscription revenues. Specifically, more free content leads to a drop in subscription demand, which reduces the subscription revenues from heavy readers, who are the target subscribers under a metered paywall. The second term captures the fact that more free content gives light readers more content to consume, generating additional advertising revenues. The second row in Equation (2.7) summarizes these impacts by segments, where the first term is the loss from heavy readers and the second term is the gain from light readers.

Whether the firm should offer more or less free content depends on which force is enlarged more by an increase in the upper-bound of consumers' valuation. As V increases, based on Equation (2.7), the loss is enlarged by a size that is proportional to the unit price of paywalled content ($\frac{p_m}{1-f_m} = \frac{V-A}{2}$). Meanwhile, the gain is enlarged by a size that is proportional to the ad rate (A). When the ad rate is low, the firm should reduce its free content when V increases. The reason is that under low ad rates, the main impact of an increase in consumer valuation is enlarging the lost subscription revenues resulting from free content provision. To avoid the loss, the firm should lock more content behind the paywall and charge a higher price. By contrast, when the ad rate is high, it should offer more free content when V increases. The reason is that such an increase in valuation enlarges the gain from light readers' consumption more than it enlarges the loss from heavy readers. As a result, the firm should offer more free content to earn greater ad impressions, taking advantage of the strong ad market. Figure 2.4(b) illustrates these results. In a metered paywall equilibrium, when the ad rate is relatively low compared with the upper-bound of valuation (i.e., in Figure 2.4(b), the "low A" condition), the amount of free content decreases in V. When the ad rate is relatively closer to the upper-bound of valuation (i.e., in Figure 2.4(b), the "high A" condition), the amount of free content increases in V.

Insert Figure 2.4 about here

To highlight the role of endogenous consumption in the finding of Proposition 2, we now consider what would happen under exogenous consumption. When consumption is exogenous, consumers' desired amount of content does not change with V. In this case, the loss from offering free content still increases in V, because consumers are willing to pay more for paywalled content. However, the gain from offering free content does not increase in V, because under exogenous consumption, the amount of consumption is not influenced by V. Consequently, a higher consumer valuation always leads to the dominance of the loss side, thus driving the firm to reduce its free content. In our model, we consider the impact of valuation on consumption by endogenizing consumers' decisions. This enables us to uncover the new insight that when V increases, the firm may provide more free content.

Our framework also allows us to speak to the impact of the light reader segment size. Recall that light readers (L) consume less compared to heavy readers (H). When the proportion of light readers increases, one may intuit that the firm would provide less content for free. After all, consumers overall want to read less content. However, the opposite may be true, as summarized in Lemma 4 below.

Lemma 4. *As the light reader segment size* (α) *increases, more free content will be provided under a metered paywall.*

In the discussion of Proposition 2, we use Equation (2.7) to illustrate the impacts of free content on profits in a metered paywall equilibrium. The second row of Equation (2.7) can also help us understand the impact of a larger L segment on the firm's choice of its meter limit. On the one hand, a larger L segment implies a smaller H segment, which decreases the subscription revenue loss from offering more free content. On the other hand, a larger L segment increases the ad impression gains from providing more free content. Therefore, more free content should be provided in equilibrium, and the price should be decreased accordingly. In the next subsection, we continue our discussion about the light reader segment size, analyzing its impact on the information flow.

2.5.2 Information Flow

In our model, the total ad impression, I, also measures the total amount of content consumed by consumers. If we assume that each unit of content presents one unit of information, we can also interpret I as the size of the "information flow" from the media to its readers. Given that one of the most fundamental functions of a media firm is to transmit information, we next analyze how the information flow changes when market circumstances change. Indeed, the firm's optimal paywall strategy in response to environmental changes can lead to unexpected changes in the size of the information flow. For example, when the proportion of light readers increases, consumers' overall desire for content declines, so one might expect a drop in the information flow. However, we find that the information flow can increase, as shown in the following proposition.

Proposition 3. *Under a metered paywall, as the light reader segment size* (α) *increases, the size of the information flow increases if both the light reader segment size and the difference in consumption costs are small* ($\alpha < \widetilde{\alpha}$ *and* $c < \widetilde{c}_\alpha$)*, and weakly decreases otherwise.*[21]

Proposition 3 suggests that under a metered paywall, a larger light reader segment can result in more content being consumed. This counterintuitive result happens when a metered paywall is the equilibrium strategy, and both the L segment size and the difference in consumption costs are small. Figure 2.5 below illustrates the impact of the light reader segment size on the price, the meter limit, and the information flow under the condition specified in Proposition 3.

Insert Figure 2.5 about here

Next, we elaborate on the intuition behind Proposition 3. An increase in α has both a direct and an indirect effect on the size of the information flow. The direct effect is the change in information flow when some heavy readers are replaced by some light readers. This impact on the information flow is always negative because an average heavy reader reads more than an average light reader does. Specifically,

[21] The (implicit) expression of $\widetilde{\alpha}$ is given in Appendix A. $\widetilde{c}_\alpha = \frac{V(V+A)^2}{2(V-A)A}$.

this direct effect consists of a reduction in I due to the decrease in the size of the heavy reader segment (the first term in Equation (2.8)), and a lift due to the increase in the size of the light reader segment (the second term in Equation (2.8)).

$$\frac{\partial I_m}{\partial \alpha} = \underbrace{-\left(1-\frac{p_m}{V}\right)}_{\text{smaller H segment size (-)}} + \underbrace{\frac{f_m(2V-f_m c)}{2V}}_{\text{larger L segment size (+)}} + \underbrace{\left(\frac{(1-\alpha)(V-A)}{2V}+\frac{\alpha\,(V-f_m c)}{V}\right)\frac{\partial f_m}{\partial \alpha}}_{\text{paywall-hitting readers get more free content (+)}} \tag{2.8}$$

Besides the direct effect, a change in α also has an indirect effect on the information flow, which is always positive. This is because of the firm's strategic response: when there are more light readers, more free content will be provided (Lemma 4). The readers who used to hit the paywall would read more, leading to an increase in total consumption. The magnitude of this indirect effect depends on the size of paywall-hitting readers ($\frac{(1-\alpha)(V-A)}{2V}+\frac{\alpha(V-f_m c)}{V}$) and how sensitive the amount of free content is to a segment size change ($\frac{\partial f_m}{\partial \alpha}$).

When the L segment size and the difference in consumption costs are both small, the positive effect of more free content outweighs the negative effect due to the decrease in the H segment size ($(1-\alpha)$ decreases). The intuition can be seen from the following. First, when the L segment size is small, the size of paywall-hitting readers is large. In this case, most readers are heavy readers who want a lot of content, but the firm would offer only a little content for free in order to charge a high price from the heavy readers who subscribe. Consequently, more non-subscribers would hit the paywall. Second, with a large group of paywall-hitting readers, even a small increase in the amount of free content can generate a lot of additional ad impressions, especially when light readers also want to consume a lot of content (which is the case under a small c). This gives the firm a strong incentive to provide more free content. As a result, the amount of free content is highly sensitive to a change in the consumer segment size (in Equation (2.8), $\frac{\partial f_m}{\partial \alpha}$ is large). Combining the two factors above, the indirect effect dominates so that the size of the information flow increases in α.

However, in the case of a large light reader segment size or a large difference in consumption costs, there are fewer paywall-hitting readers. In addition, the amount of free content becomes less sensitive to a segment size change. Both factors lead to a small indirect effect. Furthermore, when light readers account for a large share of the market, more content is available for free. In this case, the subscription price is low, so the losses from the first term in Equation (2.8) become larger. The gain from a larger L segment (the second term in Equation (2.8)) is small in magnitude compared with the loss. As a result, the firm's strategic response does not cancel out the direct effect of an increase in α, and the information flow decreases.

After analyzing the free content provision of the media firm and its implications for the information flow, we next examine the firm's pricing strategy under a change of market circumstances. Under different types of paywall, the relationships between the price and consumers' consumption costs are different. In the next subsection, we highlight those differences and discuss the intuition behind them.

2.5.3 Subscription Price

Recall that incorporating consumption costs and thus endogenizing consumers' content consumption is a key element in our framework. In fact, they have a significant impact on the firm's subscription price. Next, we analyze how consumption costs and endogenous consumption influence the firm's pricing decision. Usually, the price is set based on the distribution of consumers' WTP. If there is a decrease in WTP, the firm should lower its price as a response to the weaker demand. Hence, one might expect the price to drop when some consumers experience an increase in their consumption costs and want to consume less. This is indeed the case when consumers have homogeneous consumption costs (Godes et al. 2009). However, the result is more nuanced in our model, as presented in the following proposition.

Proposition 4. *As the consumption cost of light readers* (c) *increases, the subscription price decreases under a hard-HL or a hard-HL2 paywall, stays unchanged under a hard-H paywall, and increases under a metered paywall.*

Under a hard-HL or a hard-HL2 paywall, the price decreases in the consumption cost of the light reader segment. This happens when light readers' consumption costs are close to that of heavy readers. As expected, when the firm targets both segments, it has to lower the price to keep light readers' subscriptions when their WTP decreases. Different from these two specific types of hard paywall, the price does not depend on c under a hard-H paywall, which is chosen under high differences in consumption costs and low ad rates. Under those conditions, the firm would only target heavy readers with high valuation and set the price based on their WTP.

Light readers thus do not influence the firm's pricing decision. Finally, under a metered paywall, the price increases in c. This occurs because the firm provides two versions to target different consumers. Light readers are not the target of the paid version but the target of the free version. When their consumption cost increases, they consume less content. Thus the firm would see less benefit in maintaining its current meter limit. As a result, it would lock more content behind the paywall and raise the price.

2.6 Extension

In this section, we consider the non-negative-constraint on price and discuss the implication of relaxing it. We also incorporate another pricing strategy (price-per-unit) into our analysis and compare its performance with paywall strategies.

2.6.1 Negative Price and Price-Per-Unit Strategy

In our model, the firm charges $p \geq 0$ for access to the full version of its product. Once a consumer has access, the price does not influence her decision on how much to consume. She would consume $\min\{\frac{v}{c}, 1\}$ if she is a light reader, or consume 1 if she is a heavy reader. This is because a subscription grants this consumer access to all the content, but does not regulate the amount of content to be consumed. In this case, the firm cannot enforce the consumption of content (and viewing of ads), so charging $p < 0$ does not generate more ad impressions than $p = 0$. Therefore, it is not optimal to choose a negative price for the subscription.[22]

Now, we consider what happens if the firm grants access to its content in a different way. Compared with traditional media firms, digital media firms have more flexibility in selling their content separately. Suppose the firm chooses a *price-per-unit* (PPU) strategy and charges r for each unit of content.[23] When $r < 0$, the firm pays $|r|$ to the consumer for each unit of content consumed (and earns A from advertisers for the ad impression). In this case, a consumer solves $u(n; v, c) - rn$ to decide her optimal amount of consumption. For light readers, this utility function maximizes at $n = \frac{v-r}{c}$.

Following the same approach in Section 2.4, we can obtain the firm's optimal price when it chooses a price-per-unit strategy, as presented by the lemma below.

Lemma 5. *When the media firm chooses a price-per-unit strategy, it sets a positive price* ($r = \max\{\frac{2V-2A-c\alpha}{4}, \tilde{r}\} > 0$) *under low ad rates* ($A < \underline{A}_{PPU}\}$)*, a negative price* ($r = max\{\frac{\sqrt{(A-c)^2+6Vc/\alpha}-A-2c}{3}, \frac{V-2A-2c(1-1/\alpha)}{4}\} < 0$) *under high ad rates* ($A > \overline{A}_{PPU}$)*, and a price of zero* ($r = 0$) *under moderate ad rates* ($\underline{A}_{PPU} \leq A \leq \overline{A}_{PPU}$).[24]

The firm would charge a positive price per unit under low ad rates, a price of

[22] Another reason behind the non-negative subscription price is the assumption that consumers obtain zero utility from their outside option. With this assumption, free content is sufficient to attract all non-subscribers, thus the firm does not need to pay consumers. If the outside option provides positive utility, consumers with low valuation will prefer the outside option to the free content, and the firm would choose $p < 0$ to attract those consumers under sufficiently high ad rates.

[23] A hard paywall can be viewed as a pure subscription tariff (Armstrong 2006a), while a PPU strategy can be viewed as a pure transaction tariff (Rochet and Tirole 2003).

[24] $\tilde{r} = \frac{(2V-A)\alpha+2c(1-\alpha)-\sqrt{((V+A)\alpha+2c(1-\alpha)^2)-2c(V+A)\alpha(1-\alpha)}}{3\alpha}$ when $\alpha \neq 0$; $\underline{A}_{PPU} = \max\{\frac{2V-c\alpha}{2}, \frac{V(V\alpha+2c(1-\alpha))}{2(V\alpha+c(1-\alpha))}\}$, and $\overline{A}_{PPU} = \max\{\frac{2V-c\alpha}{2\alpha}, \frac{V\alpha+2c(1-\alpha)}{2\alpha}\}$ when $\alpha \neq 0$. When $\alpha = 0$, the firm always chooses $r \geq 0$.

zero under moderate ad rates, and a negative price under high ad rates. This, again, reflects the subsidy from the advertising side to the consumer side. A sufficiently strong ad market incentivizes the firm to pay its consumer for their consumption. Next, we compare the profit obtained from PPU strategies with those from paywall strategies and summarize the result in the following proposition.

Proposition 5. *Compared with a price-per-unit strategy, a paywall strategy (i.e., either hard or metered) can be more profitable under moderate ad rates* ($\tilde{A}_{PPU} < A < min\{V, \overline{A}_{PPU}\}$).[25]

Figure 2.6 illustrates the firm's equilibrium strategies using a numerical example. Figure 2.6(a) displays the optimal paywall strategies (as described in Lemma 2), while Figure 2.6(b) incorporates PPU strategies into the analysis. It is worth noting that some of the paywall strategies earn the same profit as a PPU strategy. First, it is clear that a no-paywall strategy is identical to a PPU with $r = 0$. Second, according to Lemma 2, when both A and c are low ($A < V$ and $c \leq \frac{2(V-A)}{2+\alpha}$), a hard-HL2 paywall with $p_{h,HL2} = \frac{2V-2A-c\alpha}{4}$ is chosen. In this case, the optimal PPU strategy also charges $r = \frac{2V-2A-c\alpha}{4}$ and earns the same profit as a hard-HL2 paywall.[26]

Proposition 5 suggests that paywalls can be more profitable than PPU strategies when ad rates are moderate. First, when the consumption costs of light readers

[25]The (implicit) expression of $\tilde{A}_{PPU}$ is given in Appendix A.

[26]Although $r = p_{h,HL2}$ in this case, the consumption behaviors of light readers under a hard-HL2 paywall are different from those under a PPU strategy. Under the hard-HL2 paywall, light readers with $v < \bar{v}_L$ do not read, while those with $v \geq \bar{v}_L$ read all the content. Under the PPU strategy, light readers with $r \leq v < r + c$ reads $\frac{v-r}{c} < 1$, while those with $r + c \leq v \leq V$ reads all the content.

and ad rates are both moderate, a hard-HL paywall is the most profitable strategy. Compared with a PPU strategy, a hard-HL paywall can be viewed as a special case of a two-part tariff, with a fixed charge of $p_{h,HL}$ and a marginal price of zero (which equals to the marginal cost). On the one hand, a hard-HL paywall earns less revenue from consumers than a PPU strategy. Previous research on price discrimination has shown that a two-part tariff is better than a uniform price in extracting an individual consumer's surplus and encouraging consumption (Lewis 1941; Armstrong 2006b). However, in our setting, the firm cannot set individualized fixed charges for different consumers because it has no information on each consumer's valuation and consumption costs. As a result, its hard-HL paywall can only fully extract the surplus of the marginal subscribers, while leaving a larger amount of surplus to other subscribers than a PPU strategy. On the other hand, a hard-HL paywall earns more ad revenue than a PPU strategy. Under a hard paywall, the zero marginal price encourages subscribers to consume as much as they can. By contrast, under a PPU strategy, the firm needs to set the price per unit above the marginal cost to make a profit, which discourages consumption. Consequently, a hard-HL paywall generates more ad revenues because more content is consumed. When the ad rate is sufficiently high, a hard-HL paywall's advantage on the advertising side outweighs its disadvantage on the consumer side, which makes it dominate a PPU strategy.

Insert Figure 2.6 about here

When c is sufficiently high and A is low (i.e., the bright yellow region in 2.6(b)), a PPU strategy becomes more profitable than a paywall strategy. First, a higher c enlarges a hard-HL paywall's disadvantage in earning consumer payment, because it needs to cut price to keep the subscribers from L segment, leaving an even greater amount of surplus to other subscribers. Second, a hard-HL paywall's advantage in ad revenues shrinks when the ad rate becomes lower. Both factors favor a PPU strategy over a hard-HL paywall. Thirdly, compared with a hard-H paywall, a PPU strategy is more profitable because of its flexibility in accommodating the demand of light readers with high valuation. Under high consumption costs, those consumers are only interested in a small proportion of content, thus not willing to pay a high subscription fee. A PPU strategy allows them to pay only for the part that they want, which not only makes those consumers better off but also generates extra revenues for the firm.

When c is high and A is moderate, a metered paywall is preferred to a PPU strategy. The intuition behind this result is similar to that behind the shift from a hard-HL paywall to a metered paywall (as in the discussion following Proposition 1). When ad rates increase, the firm cuts the price per unit to encourage consumption. Note that the drop in the price per unit only influences light readers' consumption, but has no impact on heavy readers' consumption, so the firm is attracting eyeballs at the cost of its revenue from consumer payment. By contrast, a metered paywall can attract consumers with low valuation with free content, which has a relative low impact on the revenue from the consumer side.

Finally, Figure 2.6(b) also shows that when the ad rates become sufficiently high ($A > \overline{A}_{PPU}$), the firm may choose to pay consumers for their eyeballs. Under this condition, the price per unit changes non-monotonically with light readers' consumption costs. As c increases, the firm first decreases its price from $r = 0$ to $r < 0$, then increases it until the price per unit becomes zero again. This is because when light readers have a strong desire for content (very low c), most of them would finish all the content even at $r = 0$. A negative price only incentives a small portion of consumers to consume more, so the benefit of paying consumers is little. When light readers experience very high costs of consumption, though a negative price can drive all of them to consume more, the marginal impact on total consumption is small, because consumption is less sensitive to price under high c.[27] To summarize, a negative price is most effective when c is at a moderate level.[28]

2.6.2 Correlation Between Valuation and Consumption Cost

In our main model, a consumer's consumption cost (c) and valuation for content (v) are assumed to be independent. In reality, c and v may be correlated, and the correlation between them depends on the interpretation of c. First, c and v can be *negatively correlated* if c is the cost of processing information. A consumer has lower costs of processing information if she has expertise in related areas (e.g., knowing the jargon and context). On the other hand, the expertise in an area increases the value of consuming content (e.g., thorough understanding of the content), which makes c and v negatively correlated. Second, c and v may have *no clear correlation* if c represents the (inverse of) breadth of interest or the rate of satiation. Previ-

[27]Specifically, in $n = \frac{v-r}{c}$, the marginal impact of r is $-\frac{1}{c}$, which gets closer to zero as c increases.

[28]This result slightly changes under very high ad rates ($A > \frac{V}{\alpha}$), where a negative price is optimal when c is low ($c < \frac{(2A-V)\alpha}{2-2\alpha}$).

ous research finds that the satiation rates of individuals depend on factors like age and trait self-control (Nelson et al. 2009). These factors are not clearly correlated with the value of consuming content. Thirdly, c and v can be *positively correlated* if c measures the tightness of time constraint. People with higher levels of education/salaries obtain more value from consuming news (because their decisions based on news involves higher monetary value), but also have tighter constraint on leisure time (i.e., higher c).

In this subsection, we incorporate the correlation between c and v by analyzing two cases. In both cases, consumers' valuation (v) distributes uniformly between 0 and V. We assume that light readers' consumption costs are correlated with their valuation, while maintaining the assumption that heavy readers have zero consumption cost. In the first case, light readers' consumption costs increase in their valuation: $c = kv$, where $k > 0$ is a constant. A larger k corresponds to a larger difference between the consumption costs across segments. Since $k > 0$, c and v are positively correlated.[29] The equilibrium result can be obtained by following the approach in Section 2.4. However, this extension does not generate closed-form solution for all values of α. In the following, we discuss the result of a special case, $\alpha = \frac{1}{2}$.[30]

Figure 2.7(a) illustrates the firm's equilibrium strategy when $\alpha = \frac{1}{2}$. A metered paywall is optimal when the ad rate is moderate and the difference in the

[29]The correlation coefficient, $\mathrm{corr}(c, v) = \frac{\mathrm{cov(c,v)}}{\sigma_c \sigma_v} = \frac{\alpha}{\sqrt{-3\alpha^2+4\alpha}}$, increases in α, where $\mathrm{cov}(c, v) = \frac{\alpha k V^2}{12}$, $\sigma_c = kV\sqrt{\frac{-3\alpha^2+4\alpha}{12}}$ is the standard deviation of c, and $\sigma_v = \frac{\sqrt{3}V}{6}$ is the standard deviation of v.

[30]In Appendix A, we discuss another special case where $\alpha = 1$, which means c and v are perfectly positively correlated for all consumers.

consumption costs between segments is sufficiently large. When $k \leq 1$, all readers want more content than the total amount of available content (i.e., $\frac{1}{k} \geq 1$). Under this condition, the cannibalization effect of a free version is greater than its consumption-expansion effect, so the firm chooses either a hard paywall or no paywall. When $k > 1$, the amount of free content weakly increases in the ad rate. Similar to the main model, an increase in the ad rate makes the firm shift from a hard-HL paywall to a metered paywall when the difference between segment-level consumption costs is moderate. However, further increases in the ad rate do not make the firm shift back to a hard paywall. This is because under this metered paywall, the amount of free content provided equals to the amount that light readers want ($f^* = \frac{1}{k}$), so no light reader hits the limit. Meanwhile, heavy readers with low valuation also get a large portion of content for free when k is close to 1. In this case, shifting back to a hard-HL paywall does not generate much additional impressions. The strategies under higher differences in segment-level consumption costs are similar to those in the main model.

Insert Figure 2.7 about here

In the second case, light readers' consumption costs decrease in their valuation: $c = k(V - v)$, where $k > 0$ is a constant. This implies a negative correlation between c and v.[31] Note that in this case, a hard-H paywall does not exist, because the light

[31] The correlation coefficient, $\mathrm{corr}(c, v) = \frac{\mathrm{cov(c,v)}}{\sigma_c \sigma_v} = -\frac{\alpha}{\sqrt{-3\alpha^2+4\alpha}}$, decreases in α, where $\mathrm{cov}(c, v) = -\frac{\alpha k V^2}{12}$, $\sigma_c = kV\sqrt{\frac{-3\alpha^2+4\alpha}{12}}$, and $\sigma_v = \frac{\sqrt{3}V}{6}$.

reader with the highest valuation (V) has zero consumption cost, thus her WTP is the same as her counterpart in the heavy reader segment. As a result, any hard paywall will obtain subscriptions from both segments. We analyze the model following the approach in Section 2.4. Figure 2.7(b) provides a numerical example of the firm's equilibrium strategies. Similar to the main model, a metered paywall is chosen under moderate ad rates and sufficiently large differences in segment-level consumption costs. As k increases, the area for a metered paywall equilibrium expands. A higher k means that the light readers with low valuation have even higher consumption costs and become even more reluctant to pay for a subscription. Under this condition, a metered paywall can keep obtaining impressions from those consumers, while a hard paywall loses both the subscriptions and impressions from those consumers, so a metered paywall becomes more likely to be chosen.

2.7 Concluding Remarks

In this paper, we study the paywall strategies of a digital media platform. The platform decides on the amount of free content and the subscription price to maximize its profits. It does not have information on each consumer's characteristics but does know the distribution of consumers' valuation for content and the distribution of their consumption costs. Consumers optimize their content consumption decisions and their subscription decisions. Our analysis demonstrates that the firm's optimal strategy depends on consumers' valuation for content, the proportions of light versus heavy readers, the difference between their consumption costs, and the advertising rate.

Our model and findings contribute to the theoretical analysis of paywalls. First, our paper introduces endogenous consumption into the analysis of paywalls. We explicitly model consumers' decisions on the quantity of content consumption, which have important implications for platforms' product line design. Unlike that of the quality level, platforms have limited control over the quantity levels of consumption. As a result, understanding the amount of content that consumers will choose to consume is important for optimally designing a paywall. Second, we provide a novel explanation for metered paywalls in the absence of uncertainty and quality learning. Specifically, in the presence of heterogeneous valuation as well as heterogeneous consumption costs, we find that a metered paywall is the optimal strategy when the ad rate is moderate and the difference in consumption costs across consumer segments is sufficiently high.

Our research identifies several interesting findings regarding the optimal provision of free and paid content by a platform. First, we find that the fraction of free content may decrease under higher ad rates. In fact, there is a non-monotonic relationship between the ad rate and the firm's amount of free content. A higher ad rate typically incentivizes the firm to obtain more impressions by offering more content for free. However, under a moderate ad rate and a moderate difference in consumption costs, it is profitable to attract two groups of consumers to subscribe: the heavy users with intermediate valuation and the light users with high valuation. To convert these two groups of consumers to subscribers, the firm lowers both the amount of free content and the price. When the ad rate becomes even higher, the firm increases the amount of free content again to obtain impressions from low-valuation consumers.

Second, when consumers' valuation for content increases, the platform should offer less content for free under low ad rates, but more content for free under high ad rates. This is because of two implications of a higher valuation. On the one hand, a higher valuation means an increase in consumers' WTP for paywalled content. On the other hand, it implies consumers' desire for more content. When ad rates are low, the firm relies on consumer payments and locks more content. Under high ad rates, it encourages more consumption and thus offers more free content.

We also discuss the implications of paywalls on the amount of information flow. When the proportion of light users increases, although the average consumption cost increases, the total amount of content consumed may increase. This happens because of the platform's strategic response to changes in segment size. When the light user segment is small, consumers overall have a strong desire for content, but the platform offers only a little content for free in order to charge a high subscription price. Hence, many consumers hit the paywall. Under this situation, increasing the meter limit can significantly increase ad revenue. The incentive for the platform to raise the limit is even stronger when light users want to consume a lot (i.e., they have low consumption costs). As a result, the amount of free content can be highly sensitive to a change in segment size. With a large number of paywall-hitting consumers, this highly sensitive free content provision can increase the total consumption, leading to a larger amount of information flow. This result cautions us against the conventional wisdom that a higher average level of consumption cost amongst all consumers will result in a reduction in total information flow.

This paper offers several important managerial insights. First, when the ad rate increases, a media platform should jointly optimize its subscription price and its free content provision to maximize profits. Under certain conditions, it can benefit from offering less free content even if ad rates increase. The reason is that when different consumer segments have moderately different consumption costs, acquiring subscribers with a price reduction can be more effective in generating additional ad impressions than attracting low-valuation readers with free content. Second, when there is an increase in consumers' valuation for content, the media platform should paywall more content if the ad rate is low, but offer more free content if the ad rate is high. The key is that the firm's paywall decision depends on the interplay of both sides of the market.

Our model focuses on the impact of consumer heterogeneity and endogenous content consumption on media platforms' paywall strategies. Additional factors may also influence consumer behaviors and platform strategies. First, consumers may have disutility from seeing ads alongside the content, which we do not explicitly model in consumers' utility function. However, if we interpret v as a consumer's utility from a unit of content net of the disutility from a piece of display ad, the results of our model still hold. In other words, consumers' ad disutility is implicitly captured in our model.

Second, we assume the platform to be a price taker in the ad market. Given its monopoly power in the content market, it is possible that its decisions, which determine the supply of ad impressions, can influence the price charged to advertisers. This mechanism, proposed by Halbheer et al. (2014), can also explain the optimality of metered paywalls. In this paper, we assume exogenous advertising rates and abstract away from the platform's influence on the ad market in order to focus on its impact on the content market. However, even if the advertising rates are influenced by the platform, the analysis should lead to similar conclusions as long as the ad price elasticity is not too high. In reality, both mechanisms drive media platforms to consider a metered paywall strategy, and our research enriches the understanding of the consumer side of two-sided media platforms.

Third, our model focuses on the consumption decisions made in a single period, so the changes in consumers' valuation for content across time are not incorporated. Previous research has studied the temporal fluctuation in consumer valuation and found it has an impact on a firm's paywall strategy (Lambrecht and Misra 2017). Although the dimension of time is not explicitly modeled in our paper, the comparative statics analysis can shed some light on the implications of changes in valuation. We find that a rise in valuation can lead to either more or less free content, depending on the ad rate. If we incorporate the changes in valuation by modeling multiple periods, our result should hold as long as the platform optimizes its paywall design at the beginning of each period. If the platform needs to commit to a paywall design, we expect it to be more likely to choose a metered paywall than in the case with no commitment. Without commitment, the optimal strategy may be a hard paywall for the high-demand season and no paywall for the low-demand season (or the reverse if consumer valuation changes in the way described by Lambrecht and Misra (2017)). With commitment, the platform would design its paywall based on the average valuation across time, which makes a metered paywall more preferred.

Fourth, consumers' perception of content value may depend on the amount of free content. One mechanism for that is consumers' learning of quality (Halbheer et al. 2014). In our model, consumers' valuation is independent of the firm's meter limit. Relaxing this assumption can lead to an expansion or shrinkage of the area for a metered paywall equilibrium. If giving more content for free leads to a decline in valuation, the consumption of content and the subscription demand will decrease. Hence, the firm will have less incentive to provide free content. In this case, the parameter range for a metered paywall to be optimal shrinks. If giving more free content raises consumer valuation, the opposite should happen. It merits further exploration whether the amount of free content under a metered paywall will increase or decrease, because the ad rate will also influence the firm's decision, as in Proposition 2.

Additionally, some media platforms have a "leaky" paywall, which allows consumers to access more articles than the meter limit by deleting cookies or using "private" browsing.[32] Though we didn't consider this in the model, leaky paywalls could shape platforms' strategies and profits in the following fashion. Consumers who take advantage of a leaky paywall are either light readers with a high valuation or heavy readers. For those consumers, a leaky paywall allows them to read more and generate more ad impressions. However, a leaky paywall can also lead to a loss in subscriptions. If the impression gain dominates, the platform is more likely to choose a metered paywall (compared to a hard paywall) and make it "leaky." If

[32] https://www.cjr.org/business_of_news/news-paywalls-new-york-times-wall-street-journal.php

instead, the subscription loss is more significant, the platform would choose a hard paywall or make its metered paywall less porous.

Finally, we assume that the platform knows only the distribution of consumers' preferences but not any individual consumer's preference. As the technology of big data and analytics improves, platforms may know more about their consumers, which could allow for more "individualized" paywalls. On the one hand, the platform can capture more consumer surplus with more sophisticated price discrimination. On the other hand, it should also be cautious about potential backlash from upset consumers who face tighter meters or higher prices. The latter can make a metered paywall more favorable than an "individualized" paywall.

There are quite a few directions for future research. In this paper, we assume that each consumer has the same valuation for all content. This assumption can be relaxed to capture each consumer's heterogeneous tastes for different types of content (e.g., political news, entertainment news). As for the market structure, the media firm in our model is a monopolist. It would be interesting to study how competition between media firms will influence their paywall strategies. For the advertising side of the platform, because the readers of the free and paid versions of the newspaper are different, advertisers may have different WTP for the impressions obtained from different readers. Additionally, a consumer's cost of content consumption is assumed to be exogenous in this paper. Previous behavioral research suggests that this cost can be influenced by interventions (Redden 2008). Incorporating firms' strategic actions to influence consumers' consumption costs can offer further insights into the optimal design of paywalls.

Parameter Range	Paywall Strategy	Source of Subscribers	Readers' Consumption Behavior	Meter Limit (f)	Subscription Price (p)
$A < V$ and $c \leq \frac{2(V-A)}{2+\alpha}$	Hard-HL2	H, L	All subscribers read the whole newspaper	0	$p_{h,HL2}$
$A < V$ and $\frac{2(V-A)}{2+\alpha} < c \leq \underline{c}$	Hard-HL	H, L	Some subscribers only read part of the newspaper	0	$p_{h,HL}$
$A \leq \underline{A}$ and $c > \underline{c}$	Hard-H	H		0	$p_{h,H}$
$\underline{A} < A < V$ and $c > \underline{c}$	Metered	H		f_m	p_m
$A \geq V$	No-paywall	-		1	0

Note: In this table, we highlight the difference in readers' consumption behavior under hard-HL2 and hard-HL paywalls. We discuss the consumption behaviors under other types of paywall in the appendix.

Table 2.1: Equilibrium Strategy When $0 < \alpha < 1$

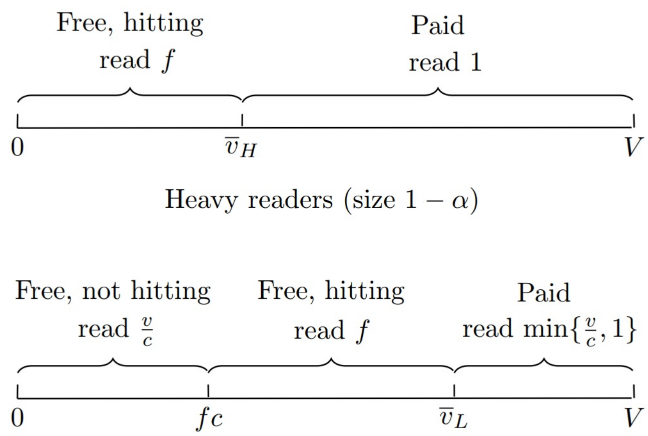
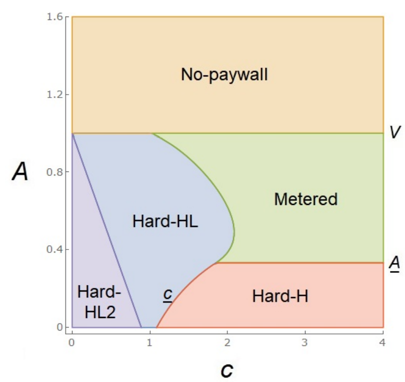
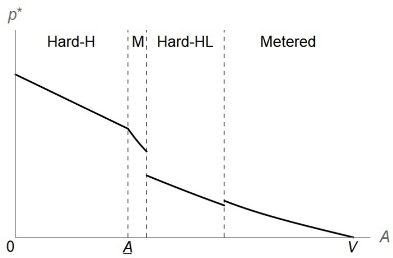


Figure 2.1: Consumers' Subscription and Reading Decisions



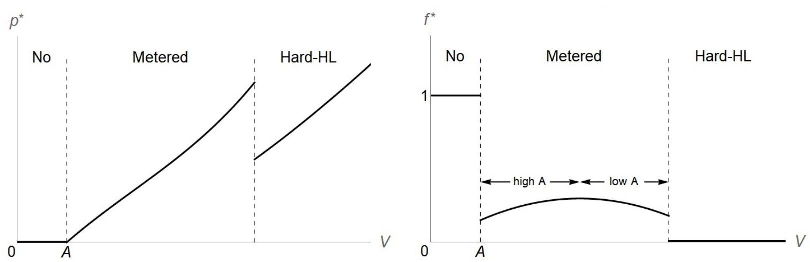
Note: In this figure, $V = 1$ and $\alpha = 0.25$. $\underline{c} = 1.08$ when $A = 0$.

Figure 2.2: Equilibrium Strategy ($0 < \alpha < 1$)



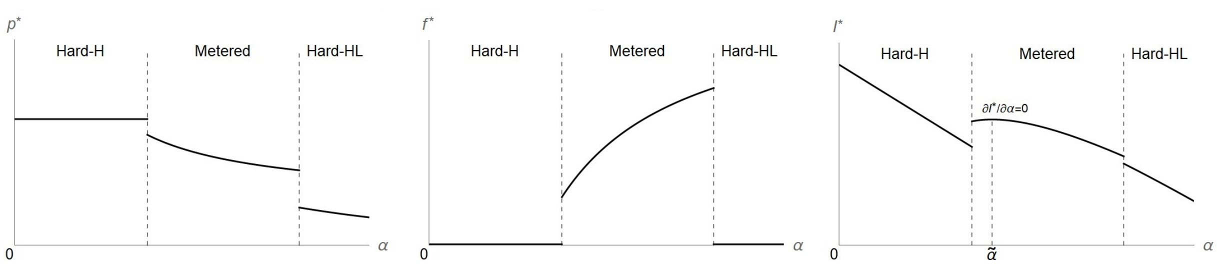
(a) Price (b) Meter Limit

Figure 2.3: Impact of the Ad Rate (A) on the Firm's Decisions When $\underline{c}(\underline{A}) < c < \max_{A \in (\underline{A}, V)} \underline{c}(A)$



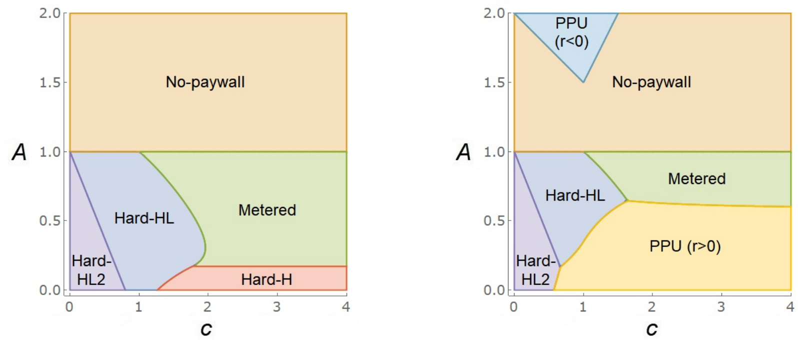
(a) Price (b) Meter Limit

Figure 2.4: Impact of the Upper-bound of Valuation (V) on the Firm's Decisions



(a) Price (b) Meter Limit (c) Information Flow

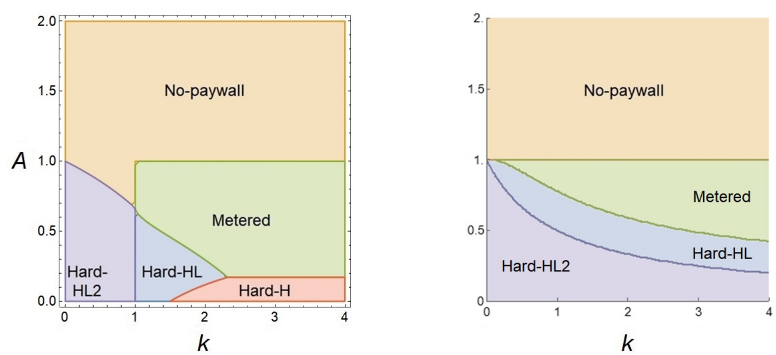
Figure 2.5: Impact of the Light Reader Segment Size (α) When $\underline{c} < c < \widetilde{c}_{\alpha}$



(a) Without PPU Strategies (b) With PPU Strategies

Note: In Panel (b), a no-paywall strategy is equivalent to a PPU strategy with $r = 0$. In the purple region, a hard-HL2 paywall is equally profitable as a PPU strategy with $r = p_{h,HL2}$. We mark the purple region as Hard-HL2 because a hard paywall should be easier to implement than a PPU strategy, thus is preferred by the firm.

Figure 2.6: Equilibrium Strategies ($V = 1$, $\alpha = 0.5$)



(a) $c = kv$ (b) $c = k(V - v)$

Figure 2.7: Equilibrium Strategies When c and v Are Correlated Among Light Readers ($V = 1$, $\alpha = 0.5$)

Chapter 3: The Accuracy of News[1]

3.1 Introduction

In April 2014, the city of Flint in Michigan changed its source of drinking water, which turned out to be the beginning of the "Flint water crisis."[2] Since August 2014, residents have been concerned about the discolored tap water and some of them even experienced unexplained illnesses.[3] People wanted to know whether the water was safe to drink. The city stated, "the Flint River presented a safe ... water source" and a review from a third-party company concluded that "the water is considered to meet drinking water requirements."[4][5] More evidence emerged later, suggesting the water contained an elevated level of lead, and city officials might have known this risk for months before they publicly admitted it.[6] Though the state of Michigan tells its residents that the water quality has met the federal standard since 2016, local officials still recommend people to drink only bottled or filtered water.[7] Since this crisis, concerns over water safety also emerged in other cities including Newark,

[1] This research is conducted with Bo Zhou and Yogesh V. Joshi.

[2] https://tinyurl.com/czj585e8

[3] https://tinyurl.com/y4xcegmc

[4] https://www.cityofflint.com/wp-content/uploads/CoF-Water-System-QA.pdf

[5] https://www.cityofflint.com/wp-content/uploads/Veolia-REPORT-FlintWater-Quality-201503121.pdf

[6] https://tinyurl.com/wny5fmk

[7] https://www.nytimes.com/2019/04/25/us/flint-water-crisis.html

New Jersey; Portland, Oregon; and Pittsburgh, Pennsylvania.[8]

Since the water crisis in Flint was covered by the media, it has attracted the attention of the entire nation. During the long process of investigation, news consumers who follow the development of the water crisis have many questions on their minds. Should the third-party company or the government officials be held accountable? Is the water crisis resolved in Flint, Michigan? Do I have to worry about my local water supply system? Individuals seek answers to these questions by consuming news from the media. They become more certain about the true state of the world after news consumption when the news is more informative. Responding to such consumer needs, media firms engage in investigative journalism and provide news. When deciding the level of accuracy in its reporting, a media firm needs to balance between consumers' willingness to pay (WTP) and the cost of investigative reporting. The cost concern of media firms in the U.S. is partially reflected by the following observations: the overall newsroom employment in the U.S. decreased by 26% from 2008 to 2020, and one third of large U.S. newspapers suffered layoffs between 2017 and 2018.[9] Many local news outlets even ended up in closure of business as their revenues could not cover costs, which raised the concern over "the death of local news" among media practitioners.[10]

In addition to the supply side issue of costs, news media firms also face major challenges on the demand side. In the U.S., the increasing polarization in public views has been a concern since the 1980s and has reached an alarmingly high level

[8] https://www.theatlantic.com/health/archive/2019/09/millions-american-homes-have-lead-water/597826/

[9] See https://tinyurl.com/45uud4pd and https://tinyurl.com/k6vmprwu

[10] https://tinyurl.com/ymauj5t3

in recent years (Abramowitz and Saunders 2008; Pew Research Center 2017). As individuals become more divided in their opinions, their consumption and media firms' provision of news will be directly affected. While the impact of news provision on polarization has been studied, there is limited discussion about the impact of polarization on news provision.

In this paper, we study a media firm's news provision strategy and consumers' news consumption decisions. More specifically, we intend to answer the following research questions: When consumers are seeking the truth, how does their prior belief influence their news consumption? How would the firm optimally choose its accuracy of news reporting? What is the impact of polarization in consumers' prior beliefs on the firm's news reporting?

We build an analytical model to study the strategic interaction between a media firm and news consumers. We assume that the underlying state of the world is a binary random variable. Neither the media firm nor the consumers know the truth definitively, but they share a common prior belief about its distribution. The firm decides on its news provision strategy, which is operationalized as the accuracy with which its report reflects the truth. This strategy is a long-term decision and is announced before the realization of the true state. In order to uncover the truth and report the news to consumers, the firm expends efforts in investigating and learning the state of the world. This process is costly, and a higher level of accuracy is typically associated with a higher cost. After this decision on its amount of investigation, the firm announces the price for its news. Then nature decides the true state, and a news report is generated according to this true state and the firm's reporting accuracy. Consumers know the firm's strategy and they want to learn the truth. They will buy and consume news if their expected benefit is greater than the price.

Our model generates several interesting results on media accuracy. First, we find both the cost of providing information and consumers' prior beliefs over the true state influence the provision of news. When improving accuracy is not very costly, the firm always provides news. When improving accuracy is sufficiently costly, the firm is profitable only when consumers' prior belief is not too extreme. In other words, the media market will fail when consumers' prior belief is relatively extreme. This happens because consumers' ex ante valuation for news can only exceed the firm's cost if their prior uncertainty is sufficiently high. Second, although consumers are indifferent between different states and obtain no extra utility from biased reporting, inaccuracies will emerge when news reporting is sufficiently costly. In order to maximize its probability of offering an accurate report, the firm will focus on accurately reporting a certain outcome, but spend relatively little resources on investigating the other possibility. As a result, in equilibrium, the probability with which the ex-ante more likely state is reported will be greater than the actual probability with which that state is realized. We label this as the exaggeration effect.

Third, we discuss the implications of different market characteristics on the level of reporting inaccuracy. When the cost of news provision increases, reporting inaccuracy increases. The impact of prior beliefs on reporting inaccuracy is also interesting. In an environment with higher uncertainty, the journalistic investigation is more costly, so one might expect higher reporting inaccuracy. However, we find that reporting inaccuracy is independent of the level of prior uncertainty as long as the firm provides news. Note that higher uncertainty implies not only a higher cost but also a higher WTP for news. We show that these two effects cancel each other out in equilibrium, resulting in their independence from the prior belief.

To study the impact of polarization, we next extend our model by incorporating the heterogeneity in consumers' prior beliefs. Heterogeneous prior beliefs lead to heterogeneous WTP for news, which makes the demand for news more sensitive to the price. This limits the firm's ability to extract consumer surplus and reduces its incentives to provide accurate news. We find that when consumers' prior beliefs become more polarized, media accuracy will decrease under moderate costs of news provision, but does not change under low or high costs. With more polarized beliefs, consumers become more certain under their priors and tend not to consume news. To encourage consumers to use news, the firm cuts its price while trying to maintain a high accuracy level. Such a strategy is effective under low costs, but higher levels of costs (i.e., moderate costs) will drive the firm to compromise and lower its accuracy. Under even higher costs, the cost-saving motivation dominates, and the firm would provide no news regardless of the degree of polarization. This result is concerning from a public policy standpoint, because a higher level of polarization, particularly in intricate and sensitive fields such as politics and healthcare, calls for more and more accurate news reporting instead of less.

As many consumers rely on news to learn about the state of the world, a news media firm may consider the public interest in addition to the commercial interests. We further extend our model to incorporate a media firm's consideration for the public interest. We start our analysis with an extreme case where the firm only focuses on improving the objective welfare of a society, which increases with consumers' probability of being correct about the truth. Consumers' expected probability of being correct increases when consumers learn about the state of the world from news, and the media firm can improve accuracy and reduce the price to encourage news consumption. We find that compared with a profit maximizer, an objective-welfare maximizer is more likely to provide news, and its accuracy is also weakly higher because it has more incentives to attract extreme consumers than a profit maximizer. When the degree of polarization increases, a profit maximizer never increases its accuracy, while an objective-welfare maximizer may increase its accuracy. Polarization decreases consumers' WTP for news, which discourages a profit maximizer from raising accuracy. However, an objective-welfare maximizer expects more gain from improving its accuracy than its profit-oriented counterpart, because the payoff gain of converting a non-consumer to a consumer does not decrease in the degree of polarization.

Finally, we analyze a more general case where the firm has a dual objective of earning profits and improving objective welfare. We find that if the firm's profit-seeking motivation is strong, it never offers news for free. Even if its profit-seeking motivation is weak, the firm offers news for free only when the cost is sufficiently high or when consumers are sufficiently polarized. This is because low polarization means most consumers have a high demand for accurate news, and low costs make providing such news relatively easy. The firm would charge a price from those high-demand consumers and ignore the extreme consumers. The size of the latter group is very small under low polarization, so the loss due to leaving them uninformed is smaller than the revenue gain. If the cost increases, the firm needs to reduce the accuracy and thus the price till the news becomes free. If the level of polarization increases, the firm tends to encourage consumption, which also leads to price cutting.

We also study a dual-objective firm's response to polarization and compare the results with those of a single-objective firm (i.e., maximizing only profits or only objective welfare). We find that in the equilibrium where the firm chooses an interior level of accuracy and a positive price, its accuracy increases in the degree of polarization only if it sufficiently values objective welfare, and the degree of polarization is sufficiently high. Greater polarization strengthens the firm's objective-welfare-improving incentives while weakening its profit-seeking incentives. The former impact dominates the latter under the specified condition and leads to higher accuracy.

The rest of the paper is organized as follows. Section 3.2 reviews the relevant literature, and Section 3.3 introduces the model and presents the analysis. Section 3.4 presents the results of the main model. Section 3.5 extends the model to study the implications of heterogeneous prior beliefs and the media firm's objective. Finally, Section 3.6 concludes and discusses future research opportunities.

3.2 Related Literature

This paper is related to the theoretical literature on consumers' news consumption and media bias. We analyze the issue of reporting inaccuracy in the media market, which is related to but also different from media bias. Media bias is defined as the difference between the news report and the media's private observation of the state of the world (Mullainathan and Shleifer 2005, hereinafter MS; Xiang and Sarvary 2007, hereinafter XS). It describes the extent to which a media firm distorts its report from its own observation. Such distortion can be a source of reporting inaccuracy, which we define as the difference between the news report and the truth. In our paper, we assume that there is no distortion between a media firm's observation and its report (i.e., no media bias). As a result, the only source for reporting inaccuracy is the difference between the firm's observation and the truth.

Previous studies suggest media bias as the source of reporting inaccuracy and provide the following explanations for those biases: biases emerge when consumers obtain utility from conforming ideas (MS 2005; XS 2007; Gal-Or et al. 2012; Zhu and Dukes 2015; Godes 2020), or when media firms get a higher evaluation by confirming consumers' prior (Gentzkow and Shapiro 2006, hereinafter GS; Godes 2020). While the former explanation captures the entertaining value of news, they do not completely reflect the value of news. Besides the entertainment value, a critical characteristic of news is its informational value, which is determined by its accuracy on truthful reporting. GS analyze a model where consumers care only about the informational value. The media firm obtains a private signal related to the truth and decides on whether to truthfully report it. A medium will be regarded as a more accurate source by consumers if its report confirms consumers' prior, and thus it has incentives to misreport. Similarly, in our paper, consumers also only care about the informational value and we also find that the media firm would confirm consumers' prior. But the mechanism in our paper is markedly different. In their paper, the firm's reporting accuracy is exogenously determined and unknown to consumers. Consumer inference for the firm's accuracy is a key driver of the firm's misreporting behavior. In our paper, reporting accuracy is endogenously chosen by the firm and is known to consumers. Reporting inaccuracy emerges because accurate reporting is costly and uncertainty may never be fully resolved. Moreover, GS shows that the media firm reports each state with the underlying probability of each state (i.e., no exaggeration in news content).[11] Instead, we find that in equilibrium, the content of news exaggerates the ex ante more likely state in consumers' prior belief. More recently, Godes (2020) extends the model in GS and study the impact of feedback on media bias. While GS suggest feedback reduces media bias, he finds that feedback can exacerbate media bias in the presence of consumers who seek conforming news. In contrast, all consumers in our model are truth-seekers, and we focus on the supply-side factors in news provision.

In contrast with the studies on demand-side factors, Baron (2006) explains the

[11] In GS, the firm's report presents each state with the underlying likelihood, regardless of being from a high-type firm (that honestly reports its observation) or a low-type firm (that pretends to be a high-type by distorting its observation in news). For example, when the underlying distribution is Pr(state L happens)=30% and Pr(state R happens)=70%, either firm would reports l in news 30% of the time, while report r 70% of the time.

emergence of media bias with supply-side factors. He suggests that journalists would bias their reports to enhance their career prospects, and a media firm needs to pay higher salaries to journalists to prevent them from generating bias. In equilibrium, the firm may tolerate some bias in its news to save on labor costs. By contrast, our paper provides a more general framework to capture the relationship between a media firm's reporting costs and its accuracy. We also explore the impact of polarization on the firm's news provision strategy. In a recent paper, assuming the accuracy to be exogenous, Papanastasiou (2020) studies the spread of inaccurate news on a social media platform as well as the platform's intervention decision. Different from that paper, the accuracy level is endogenous in our paper and we focus on the production of accurate information rather than the spread of it.

This paper also relates to the research on polarization in public views. Three types of polarization have been discussed in the literature: [i] *belief polarization*, which is also called "ideological polarization," refers to an increasing difference between the stances and policy positions of different people (DiMaggio et al. 1996; Fiorina et al. 2005); [ii] *affective polarization*, which means the tendency of people to view opposing partisans negatively and co-partisans positively (Iyengar and Westwood 2015); and [iii] *behavior polarization*, which means people becoming more divergent in their behaviors, for example, voting decisions (Jacobson 2000; Fleisher and Bond 2004). Earlier research suggests the polarization among the American public is primarily affective but not ideological (Iyengar et al. 2012), but recent studies show the existence of both and provide evidence on the connection between these two types of polarization (Rogowski and Sutherland 2016; Pew Research Center 2017; Webster and Abramowitz 2017). A stream of literature provides economic explanations for the emergence of belief polarization (Rabin and Schrag 1999; Dixit and Weibull 2007; Glaeser and Sunstein 2009; Zimper and Ludwig 2009; Baliga et al. 2013; Acemoglu et al. 2016; Nielsen and Stewart 2020). In a recent paper, Iyer and Yoganarasimhan (2021) explain the emergence of behavior polarization with people's incentive to influence group decision. Different from this stream of literature that focuses on the antecedence of belief polarization, we take belief polarization as given, and study its impact on media market outcomes.

Previous research has also investigated the relationship between polarization and news media. On the one hand, it is shown that exposure to partisan media can lead to both ideological and affective polarization (Levendusky 2013a, b; Levy 2021). On the other hand, polarization can also influence news consumption. For example, Iyengar and Hahn (2009) and Flaxman et al. (2016) find that consumers' ideological leanings influence their consumption of news. Our analysis not only captures the impact of consumers' beliefs on their news consumption decision, but also reveals the implication of polarization for news provision. We model polarization as an increase in the heterogeneity in consumers' beliefs. Previous theoretical papers find that an increase in heterogeneity has no impact on media bias (MS; Zhu and Dukes 2015). In our paper, we find polarization either has no impact or decreases media accuracy, depending on the cost of reporting. Interestingly, when the media firm intends to improve consumers' understanding of the truth rather than maximize its profits, we find it may even increase accuracy in response to polarization.

Our research also adds to a fast-growing stream of literature on Bayesian persuasion and information design (Kamenica and Gentzkow 2011; for a review, see Kamenica (2019) and Bergemann and Morris (2019)). The framework of Bayesian persuasion/information design analyzes the game between a Sender and a Receiver of information in a world with an unknown state. Sender and Receiver have symmetric information. Sender's payoff depends on both the true state and Receiver's action. However, Sender can only influence Receiver's action by providing information. Specifically, Sender designs an information structure and commits to it, from which Receiver can obtain information about the truth and act accordingly. We apply this framework to the media market. Although consumers in our setting take no explicit action and their belief determines their utility, the firm's payoff from selling news can still be written as a function of consumers' posterior belief like in Kamenica and Gentzkow (2011). Note that Bayesian updating plays an important role in consumers' interpretation of news. Consumers understand that media firms take stances and may generate inaccurate news, so they will not take the news content literally. Their prior beliefs influence their interpretation of a given news report. Consumers with a more extreme prior are less likely to change their belief over the default state (i.e., the more likely state under their prior).

Our model can be categorized as the ex ante pricing of information (Bergemann and Bonatti 2019). Bergemann et al. (2018) study the optimal menu design by a monopolist information seller. Assuming costless information provision, they characterize the properties of any revenue-maximizing menus. On the contrary, in our model, it is costly for the media firm to provide information, and this cost plays a major role in its reporting accuracy. Another related paper in terms of model setup is Jerath and Ren (2021). They study consumers' attention allocation between positive and negative product information and assume the cost of information to be incurred by homogeneous consumers. In our model, we assume the cost to be born by the information provider, as Gentzkow and Kamenica (2014). As mentioned in the introduction, the cost concern widely exists among U.S. newspapers. The cost of investigating plays an important role in how much information it can provide. In addition, we also consider the heterogeneity in consumers' beliefs and study its implication for market outcomes.

3.3 Model

In our main model, there are a monopoly media firm and a group of homogeneous consumers. There is a true underlying state of the world, t. The state is either L or R. Ex ante neither the firm nor consumers know the true state with certainty, but they have a common prior belief over the distribution of the true state. We denote $\mu = \Pr(t = R)$ as the belief that the true state is R with probability μ, and the common prior belief is denoted by μ_0.[12] In the main model, this common prior belief is assumed to be consistent with the underlying distribution of t. The media firm decides on its reporting strategy α and price p, which will be described in detail in Section 3.3.1. Consumers know α before deciding whether to purchase the news or not. If they purchase the news, they will observe the news content and update their belief about t. At the end of the game, consumers still don't know the true state with certainty (unless the news is fully accurate) and their expected utility

[12] Since no one knows the true state with certainty ex ante, $\mu_0 \in (0, 1)$.

will be determined by their belief. In general, consumers' expected utility is higher if they feel more certain about their belief, which will be elaborated in Section 3.3.2. In the rest of this section, we describe the media firm's strategy, consumers' decision and utility, the timing of the game, and then the analysis of the game.

3.3.1 Media Firm

The media firm provides a news report to consumers. Though the firm cannot directly observe the true state, it can collect information to construct an informative news report. We use s to denote the firm's observation from information gathering, and assume that this observation (s) is reported to consumers without distortion. The media firm's reporting strategy is defined as the probability with which the collected information is consistent with the true state, $\alpha = (\alpha_L, \alpha_R)$, where $\alpha_L = \Pr(s = l|t = L)$ is the probability of reporting l when the true state is L and $\alpha_R = \Pr(s = r|t = R)$ is the probability of reporting r when the true state is R. Higher values of α_L and α_R indicate higher levels of reporting accuracy. In particular, $\alpha_L = \alpha_R = 1$ means completely accurate reporting. The reporting strategy is decided by the firm before its production of news. For example, when a media firm is producing a news report about the impact of lockdown policies on people's life, the true state can be the benefits of those policies are greater than their costs (L), or vice versa (R). The firm implicitly chooses its reporting accuracy by deciding how many and which journalists to assign for this report. Specifically, it can increases α_L by assigning journalists with expertise on public health, and increases α_R by assigning those with expertise in economics.

To achieve higher levels of reporting accuracy, the firm needs to recruit qualified journalists, interview relevant parties, and collect and analyze information. This process can take a substantial amount of time and effort. In our model, we use a function, $C(\alpha)$, to capture the firm's news production costs when its reporting strategy is α. $C(\alpha)$ is proportional to the amount of information included in the firm's news report, which is measured by the expected amount of uncertainty reduction caused by this report (Gentzkow and Kamenica 2014). Therefore, C is proportional to the expected uncertainty reduction, as Equation (3.1). $c > 0$ represents the unit cost of providing information, which depends on the characteristics of the information environment and the firm (e.g., the availability of evidence in the environment, the firm's capability to collect relevant information). $H(\mu)$ is residual variance, which measures the amount of uncertainty under a belief μ (Ely et al. 2015; Augenblick and Rabin 2021).[13] μ_s is the media firm's posterior belief after observing s.

$$C(\alpha;\mu_0) = c(H(\mu_0) - E_s[H(\mu_s)]),$$

$$\text{where } H(\mu) = (1-\mu)\mu; \text{ and } \mu_s = \Pr(R|s) = \frac{\Pr(s|R)\mu_0}{\Pr(s)} = \begin{cases} \frac{\alpha_R\mu_0}{(1-\alpha_L)(1-\mu_0)+\alpha_R\mu_0} & \text{if } s = r \\ \frac{(1-\alpha_R)\mu_0}{\alpha_L(1-\mu_0)+(1-\alpha_R)\mu_0} & \text{if } s = l \end{cases}. \tag{3.1}$$

[13]We tried different specifications for the measure of uncertainty (e.g., an entropy function used by Sims (2003)). For the findings in Proposition 1-3, we obtained qualitatively similar results when using an entropy function to measure uncertainty.

This cost function has several desirable properties. First, this cost function is non-negative. Blackwell's theorem suggests that $E_s[H(\mu_s)] \leq H(\mu_0)$ for all α and μ_0 if and only if H is concave (Blackwell 1953; Gentzkow and Kamenica 2014). Second, it is a convex function of reporting accuracy ($\frac{\partial^2 C}{\partial \alpha_t^2} > 0$) and is an increasing function of accuracy under certain constraints ($\frac{\partial C}{\partial \alpha_t} > 0$ when $\alpha_L + \alpha_R > 1$).[14] Therefore, it is always more costly to pursue higher reporting accuracy conditional on any given prior belief. When $\alpha_L = \alpha_R = 1$, the report is fully accurate so that the uncertainty is reduced to zero, $H(\mu_s) = 0$, and the cost of providing such news is proportional to the prior uncertainty. Third, the cost is zero when $\mu_0 = 0$ or $\mu_0 = 1$. Under these prior beliefs, the true state is certain ex ante, so no information is needed to further reduce uncertainty. In the model, we assume that there is uncertainty under the prior, so the cost is always positive when the firm provides information.

The media firm chooses its reporting strategy α based on the prior belief μ_0, and then it chooses the price, p, for its news. The price is also chosen before the completion of the news report so that it is not influenced by the firm's observation s. The firm maximizes its profit, $p - C$, and it would provide news only if $p - C > 0$.

3.3.2 News Consumers

Consumers have the same prior belief as the firm. With the Bayesian approach, they update their prior belief of μ_0 to the posterior belief, μ_s, upon observing s in the news report. The reporting strategy of the firm, α, is known to consumers before they purchase and consume news. As a result, they can incorporate this knowledge

[14] We discuss the implication of this constraint in Section 3.3.4.

when interpreting the signal/content s in the news. For example, if a firm chooses $\alpha_L = \Pr(s = l|t = L) = 1$ and $\alpha_R = \Pr(s = r|t = R) < 1$, it always reports l when $t = L$ and would report r only when $t = R$. Therefore, consumers receiving a report of r will update their posterior belief to $\mu_r = 1$. In contrast, a report of l from this firm has lower credibility. Consumers receiving l would update their posterior belief to $\mu_l = \frac{(1-\alpha_R)\mu_0}{1-\mu_0+(1-\alpha_R)\mu_0} > 0$, which implies that consumers would discount the l report from this firm.

Consumers' objective is to seek the truth.[15] They will obtain a higher utility if the state they believe in turns out to be the truth, but a lower utility otherwise. We model the truth-seeking consumers with the utility function in (3.2). $\mathbf{I}$ is the indicator function which is 1 when the conditions in its parentheses hold and is 0 otherwise.

$$u(\mu, t) = \mathbf{I}\left(\mu < \frac{1}{2}, t = L\right) \cdot v + \mathbf{I}\left(\mu > \frac{1}{2}, t = R\right) \cdot v + \mathbf{I}\left(\mu = \frac{1}{2}\right) \cdot \frac{v}{2} \tag{3.2}$$

We denote the more likely state in consumers' beliefs as τ. When consumers believe state L is more likely than state R (i.e., $\mu < \frac{1}{2}$), then $\tau = L$; similarly, when consumers believe state R is more likely than state L (i.e., $\mu > \frac{1}{2}$), $\tau = R$. When consumers believe the two states L and R are equally likely (i.e., $\mu = \frac{1}{2}$), τ is L with probability $\frac{1}{2}$ and is R with probability $\frac{1}{2}$. According to Equation (3.2), consumers' utility is $v > 0$ when the more likely state in their beliefs is consistent with the true

[15] Based on the literature in communication and journalism, informational gain is one of the fundamental motivation for news consumption (Lee 2013). News satisfies the "surveillance" needs of individuals by improving their knowledge about the environment (Diddi and LaRose 2006; Shoemaker 1996).

state (i.e., $t = \tau$). Their utility when they believe in a wrong state (e.g., $t = L$ but $\tau = R$) is assumed to be zero. When $\mu = \frac{1}{2}$, believing in either state will lead to v with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$, therefore consumers' (expected) utility is $\frac{v}{2}$.[16] Therefore, Equation (3.2) can also be written as $u(\mu, t) = v \cdot \mathbf{I}(t = \tau|\mu)$.

Although we specify consumers' utility based on their belief and the true state, in reality, the true state may not be observable to consumers (at least not within a certain period of time). For example, to fully understand the impact of a water pollution crisis or a lockdown policy, a long-term evaluation of its effect on people's health and life quality will be needed. Those evaluations and evidence can help consumers update their posterior beliefs, in a way that could potentiallyreverses the more likely state in their belief. However, the true state is never known with certainty unless α_L or α_R is 1. As a result, at the end of the game, consumers may not know the utility $u(\mu, t)$ but only know the expected utility, as Equation (3.3). Note that μ can be either μ_0 or μ_s in Equation (3.3). Hereinafter, we refer to $E_t[u(\mu_s, t)]$ as consumers' *conditional expected utility* after observing news s (i.e., the expected utility conditional on observing s).

$$E_t[u(\mu, t)] = u(\mu, L)(1 - \mu) + u(\mu, R)\mu = v \cdot \max\{1 - \mu, \mu\} \equiv v \cdot \Pr(t = \tau|\mu) \quad (3.3)$$

The last part of Equation (3.3) shows that consumers expected utility depends on their certainty about the true state. Specifically, consumers' expected utility

[16] This utility framework is fairly flexible. For instance, if a priori consumers prefer one state over the other, v_L and v_R can be different. In our setting, consumers only care about knowing the truth, so without loss of generality we assume $v_L = v_R = v$.

increases when they believe $t = \tau$ with a higher probability. Figure 3.1 below illustrates a consumer's expected utility given her belief. A consumer has a higher expected utility when she is more certain about the true state (i.e., μ gets closer to 0 or 1), but a lower expected utility when she feels both states may happen with a similar likelihood (i.e., μ gets closer to $\frac{1}{2}$).

Insert Figure 3.1 about here

To maximize their expected utility at the end of the game, consumers decide whether to purchase the news report from the firm. They know the firm's reporting strategy (α) and price (p), but they do not observe the content of the news report before paying the price. Only after purchase can they observe the news content (s) and update their belief (from μ_0 to μ_s). If consumers do not purchase the report, they will maintain their prior belief till the end of the game.

3.3.3 Timing of the Game

The sequence of the media firm's and consumers' decisions is summarized below.

1. *Media accuracy*: based on the prior belief μ_0, the cost coefficient c, and consumer preference for believing in the true state (v), the firm announces its reporting strategy (α_L, α_R);

2. *Media pricing*: the firm announces the price (p) for its news report;

3. *Nature chooses the true state*: the true state t is realized;

4. *News report generation*: the firm incurs the cost C to obtain a signal s based on the true state (t) and the reporting strategy (α); then it reports s in the news;

5. *Consumers' purchase and consumption of news*: consumers decide whether or not to purchase the news based on the firm's strategy (α, p) and their preference (v); if they purchase the news report, they observe s and update their belief to μ_s.

3.3.4 Model Analysis

We analyze the game by backward induction. We first discuss consumers' belief updating process and several constraints on reporting accuracy. Then we analyze consumers' purchase decisions, followed by the firm's profit maximization problem.

Consumers update their belief following the Bayes rule when they receive a news report. Though Blackwell's (1953) theorem implies that a news report always leads to an uncertainty reduction regardless of the firm's reporting strategy α (or at least remain at the original level of uncertainty), an additional constraint is needed for the posterior and prior beliefs to ensure meaningful information transmission (alternatively, informative communication): $\mu_l \leq \mu_0 \leq \mu_r$. This constraint means that the belief of R becomes weaker when l is observed and becomes stronger when r is observed.[17] It ensures that the news report is either *informative* ($\mu_l < \mu_0 < \mu_r$)

[17] Jerath and Ren (2021) analyze this constraint in a context where consumers allocate limited attention to positive and negative product information before making a purchase. They derive the same constraint, which is required to make an "information structure" (report) informative.

or *uninformative* ($\mu_l = \mu_0 = \mu_r$), but never nonsensible ($\mu_l > \mu_0 > \mu_r$). Solving these inequalities leads to $\alpha_L + \alpha_R \geq 1$. Note that $\mu_l = \mu_0 = \mu_r$ if and only if $\alpha_L + \alpha_R = 1$, which is the condition for uninformative reporting.

Consumer's ex ante expected utility from purchasing the news depends on the extent of their belief change upon consuming the news. One particularly important change is consumers believing L to be more likely under the prior ($\tau_0 = L$) shift to believing R to be more likely after consuming news ($\tau_s = R$), or vice versa. Formally, we define consumers' "belief change" based on the relationship between their prior and posterior beliefs as follows.

Definition 1. *A belief change over which state is more likely is defined as (1) $\mu_s > \frac{1}{2}$ when $\mu_0 < \frac{1}{2}$; (2) $\mu_s < \frac{1}{2}$ when $\mu_0 > \frac{1}{2}$; or (3) $\mu_s \neq \frac{1}{2}$ when $\mu_0 = \frac{1}{2}$. Mathematically, this is equivalent to $\alpha_L + \alpha_R > 1$ and $\frac{1-\alpha_L}{1-\alpha_L+\alpha_R} < \mu_0 < \frac{\alpha_L}{\alpha_L+1-\alpha_R}$.*[18]

Note that any change that leads to $\mu_s = \frac{1}{2}$, which means that consumers' posterior belief is that L and R are equally likely, does not qualify a belief change in our context. With this definition, we can analyze consumers' *ex ante expected utility* when they purchase news, given by Equation (3.4).

[18]When the news is informative (as defined in this subsection), the constraint for a belief change is equivalent to the obedient constraint in the Bayesian games of communication (Myerson 1982; Bergemann et al. 2018). Specifically, a belief change happens if and only if the more likely state in consumers' posterior beliefs (τ_s) is consistent with the news content (s), i.e., $\tau_l = L$ and $\tau_r = R$.

$$E_s[E_t[u(\mu_s, t)]] = \Pr(l) \cdot E_t[u(\mu_l, t)] + \Pr(r) \cdot E_t[u(\mu_r, t)]$$
$$= \begin{cases} v(\alpha_L(1-\mu_0) + \alpha_R\mu_0) & \text{if news may lead to a belief change,} \\ v \cdot \max\{1-\mu_0, \mu_0\} & \text{if news never lead to a belief change.} \end{cases} \tag{3.4}$$

By comparing a consumer's ex ante expected utility with and without consuming news, we can see that a news report leads to a strictly higher ex ante expected utility if and only if it may lead to a belief change. This relates to a key insight of Blackwell, that information is valuable only if it changes the recipient's optimal actions. Bergermann and Morris (2019, Proposition 1) also show this point in a general model of information design. In our model, consumers don't explicitly take action after receiving the news report and their expected utility depends on their posterior belief. The value of information (i.e., the net gain in expected utility from information consumption) depends on whether it leads to a belief change. From the firm's viewpoint, it should choose α to make the news valuable to consumers, i.e., the new report needs to satisfy the conditions presented in Definition 1.

From the viewpoint of consumers, given α, the news report is more valuable when their prior belief is closer to $\frac{1}{2}$. As their prior gets closer to $\frac{1}{2}$, the prior uncertainty is higher and then consumers have a higher WTP for news. If consumers' prior belief is so extreme that there is no belief change from news consumption, then their WTP becomes zero. In the rest of the paper, we assume that consumers will consume news only when consumption leads to a positive ex ante surplus (i.e., their WTP is greater than the price, $E_s[E_t[u(\mu_s,t)]] - E_t[u(\mu_0,t)] - p =$WTP $-p > 0$). This implies that news consumption will happen only when consumers expect a potential belief change from news, even if the price is zero. As a result, the firm needs to increase its reporting accuracy to induce consumers to purchase its news report. Lemma 1 presents the relationship between consumers' WTP and the firm's reporting accuracy.[19]

Lemma 1. *Consumers' WTP for news weakly increases as* $\mu_0 \to \frac{1}{2}$. $WTP = v(\alpha_L - 1 + (1-\alpha_L+\alpha_R)\mu_0)$ *when* $\frac{1-\alpha_L}{1-\alpha_L+\alpha_R} < \mu_0 < \frac{1}{2}$, $WTP = v(\alpha_L - (\alpha_L+1-\alpha_R)\mu_0)$ *when* $\frac{1}{2} \leq \mu_0 < \frac{\alpha_L}{\alpha_L+1-\alpha_R}$, *and* $WTP = 0$ *when* $\mu_0 \leq \frac{1-\alpha_L}{1-\alpha_L+\alpha_R}$ *or* $\mu_0 \geq \frac{\alpha_L}{\alpha_L+1-\alpha_R}$.

As the monopolist in this market, the media firm can always charge (slightly below) consumers' WTP. To increase its price, the firm needs to raise the accuracy of reporting. However, the cost of news provision also plays an important role and needs to be accounted for. In the next section, we analyze the firm's optimal strategies under different cost structures and discuss the impact of the prior belief.

3.4 Results

In this section, we examine the impact of costs on the media firm's reporting strategy. We start with the condition for news provision, then discuss the optimal accuracy levels chosen by the firm. We also construct a measure of the overall reporting inaccuracy and study how it changes with market circumstances. Our

[19]Bergemann et al. (2018) obtain similar results in Section III.A of their paper.

first finding is that consumers' extreme prior beliefs can discourage the provision of news when reporting costs are sufficiently high. The proposition below summarizes the results.

Proposition 1. *Under low reporting costs ($c \leq v$), the media firm provides news as long as consumers have prior uncertainty* ($0 < \mu_0 < 1$)*. Under high reporting costs ($c > v$), it provides news if and only if consumers have sufficiently high prior uncertainty* ($\frac{c-v}{2c} < \mu_0 < \frac{v+c}{2c}$)*.*

In our model, we assume that consumers do not know the truth with certainty under their prior belief. If they do, i.e., $\mu_0 = 0$ *or* $\mu_0 = 1$, the firm would not provide news because news does not create value for any consumer. Proposition 1 suggests that as long as uncertainty exists under consumers' prior belief, i.e., $0 < \mu_0 < 1$, the firm will always provide news when reporting is not too costly. This corresponds to situations where information on the focal issue can be easily collected, such as the price changes of a specific stock in a day. When reporting becomes sufficiently costly, consumers' prior belief plays a more important role in the firm's news provision decision. If consumers' prior belief is sufficiently extreme ($\mu_0 < \frac{c-v}{2c}$ or $\mu_0 > \frac{v+c}{2c}$), their ex ante expected utility without news consumption is already high under the prior. As a result, their WTP will be too low to cover the firm's cost, and news provision becomes unprofitable. With priors $\mu_0 = \frac{c-v}{2c}$ or $\mu_0 = \frac{v+c}{2c}$, consumers' WTP becomes just high enough to cover the cost, so news provision is still not strictly profitable for the firm. However, when the prior belief is relatively moderate ($\frac{c-v}{2c} < \mu_0 < \frac{v+c}{2c}$), i.e., consumers are fairly uncertain about the true state of the world. their WTP for a news report becomes high enough and the firm finds it profitable to produce news.

We now elaborate on the equilibrium level of accuracy chosen by the firm when news is provided. The results are presented in Proposition 2.

Proposition 2. *When the media firm provides news, it chooses a full-accuracy strategy under low reporting costs ($c \leq v$):*

$$\alpha_L = \alpha_R = 1,\ p = v \cdot min\{\mu_0, 1 - \mu_0\},$$

and a partial-accuracy strategy under high reporting costs ($c > v$):

$$\alpha_L^* = \frac{(v+c)(v - c(2\mu_0 - 1))}{4vc(1-\mu_0)},\ \alpha_R^* = \frac{(v+c)(v + c(2\mu_0 - 1))}{4vc\mu_0},\ p = v\left(\frac{c+v}{2c} - max\{\mu_0, 1-\mu_0\}\right).$$

As a result, under high costs, the news report exaggerates the presence of the ex ante more likely state.

When the cost of news provision is low, the firm will choose full accuracy. After purchasing and consuming the news, consumers will know the true state with certainty, and their posterior belief is either $\mu_l = 0$ (if the news is l, the probability that the true state is R is zero) or $\mu_r = 1$ (if the news is r, the probability that the true state is R is one). Because news is more valuable to consumers when they have higher prior uncertainty, the firm's price increases in prior uncertainty and reaches the highest level under the prior $\mu_0 = \frac{1}{2}$.

Compared with the low-cost case, the prior belief over the true state has a larger impact on the accuracy of reporting under high costs. Figure 3.2(a) shows an example of the firm's optimal reporting strategy captured by the y axis as a function of the prior belief μ_0 captured by the x axis. The regions to the left of $\frac{c-v}{2c}$ and to the right of $\frac{v+c}{2c}$ represent for the range where no news is provided, while an intermediate μ_0 allows for the existence of the news market.

Insert Figure 3.2 about here

Figure 3.2(a) illustrates how the optimal reporting accuracy changes with the prior belief when news is provided. When L is believed to be more likely than R under the prior, the media reports L with higher accuracy than R. When the prior belief favors R rather than L, reporting accuracy for R is higher than that for L. In the example of the water pollution crisis, if the prior is that the pollution has a large impact on local residents' life, the media firm will put more resources on accurately reporting a large impact with high accuracy, but only limited resources on collecting evidence to accurately reflect a small impact. In other words, the media firm will allocate more resources toward accurately reporting the more likely state based on the prior belief. Upon observing this, consumers expect to see an accurate report under the majority of circumstances and are willing to pay for the news.

Panel (b) of Figure 3.2 illustrates the news content resulting from the optimal accuracy. Interestingly, as the firm tries to build an image of being accurate under most circumstances, its news content, measured by the probability of reporting l or r, always exaggerates the "default" state (i.e., the more likely state) in consumers' prior. For example, in Panel (b), when state R happens with a probability of

70%, the firm reports r in 90% of the time (i.e., $\Pr(r) = 0.9$ when $\mu_0 = 0.7$). Such an exaggeration makes the less likely state even harder to be found in news. From consumers' perspective, this is similar to the effect of an "echo chamber," because they are largely exposed to conforming opinions in equilibrium. What's striking about this result is that it does not arise due to the media firm's intentional misreport such that media bias emerges (i.e., the firm does not truthfully report the signal it receives during investigation). Instead, the reporting inaccuracy completely stems from the firm's strategic trade-off between its production costs and consumers' willingness to pay for information to reduce their uncertainty.

After consumption, consumers' posterior belief is $\frac{c-v}{2c}$ if they receive l, or $\frac{v+c}{2c}$ if they receive r. Note that the posterior belief is independent of the prior belief. They are exactly the belief thresholds under which the firm would not provide news. The intuition is as follows. Suppose the firm chooses $(\alpha'_L, \alpha'_R) \neq (\alpha^*_L, \alpha^*_R)$ that leads to different posteriors. On the one hand, if those posteriors are between $\frac{c-v}{2c}$ and $\frac{v+c}{2c}$, the firm could have earned more profit by increasing its reporting accuracy. On the other hand, if those posteriors don't fall between $\frac{c-v}{2c}$ and $\frac{v+c}{2c}$, then the firm's marginal investment in reporting accuracy is above consumers' WTP for it. Consequentially, choosing the accuracy level (α^*_L, α^*_R) that leads to the threshold (for news provision) posterior beliefs is optimal for the firm.

Compared with the low-cost scenario ($c \leq v$), consumers' posterior belief is less extreme under high costs ($c > v$). There is remaining uncertainty after the report's realization because the high cost deters the firm from choosing full reporting accuracy. As the unit cost of news provision (c) increases, the remaining uncertainty under the posterior belief increases, and the range of μ_0 under which news is provided shrinks.

Insert Figure 3.3 about here

Figure 3.3 illustrates the relationship between equilibrium outcomes and the prior belief, μ_0. In Panel (a), the vertical axis represents the ex ante expected utility. The dashed line shows that consumers' ex ante expected utility without news consumption is lower under moderate priors (i.e., when μ_0 is closer to $\frac{1}{2}$). However, by consuming news, their ex ante expected utility can increase to a constant level. The firm earns revenue by charging the difference between the ex ante expected utility with and without news. Its revenue, as shown in Panel (b), is higher under higher prior uncertainty, so are the cost and the profit (i.e., the difference between the revenue and the cost). Panel (c) shows the levels of prior and posterior uncertainty, H (multiplied by c). Prior uncertainty is higher under moderate μ_0, so the firm needs to incur more cost to reduce uncertainty. The expected remaining uncertainty after news consumption is constant. If the uncertainty is above this level, the firm would provide more accurate news to earn a profit until the uncertainty achieves this constant level. Once the news can help consumers reach this particular level of uncertainty, the firm has no incentives to provide any further improvement over its accuracy.

Our model also allows us to study the level of reporting inaccuracy. In the previous literature such as MS and XS, media bias is defined as the expected difference between the raw information and the news, where the collected raw information is assumed to accurately present the truth in expectation. By contrast, the concept of media bias does not apply in our context. Instead, we focus on reporting inaccuracy and define it as the expected difference between the true state and the reported news. As discussed in Section 3.2, media bias can be one source of reporting inaccuracy. In our paper, we assume there is no systematic distortion from the raw material in the news report, and we focus on another source of inaccuracy: the inaccuracy in the process of investigation and information collection. We formally define reporting inaccuracy (RI) as the following.

Definition 2. $RI = E[|\mathbf{I}(t=R) - \mathbf{I}(s=r)|]$ $=\Pr(l,R) + \Pr(r,L)=$ $(1-\alpha_R)\mu_0 + (1-\alpha_L)(1-\mu_0)$.

Specifically, reporting inaccuracy is the sum of the probabilities of two types of misreporting. One is the probability of reporting l when the true state is R, and the other is the probability of reporting r when the true state is L. Intuitively, RI=0 when $\alpha_L = \alpha_R = 1$, and RI=1 when $\alpha_L = \alpha_R = 0$. RI can also be viewed as a weighted sum of the probabilities of type 1 and type 2 statistical errors. Suppose the null hypothesis is $t = R$. If the true state is indeed R, then the firm reports l (type 1 error) with probability $1 - \alpha_R$; if $t = L$, the firm reports r (type 2 error) with probability $1 - \alpha_L$. The weights of these two probabilities are the underlying probabilities that the corresponding state is the truth.

We measure the amount of reporting inaccuracy in equilibrium based on the definition above. Next, we describe the equilibrium reporting inaccuracy under different cost structures in the following proposition.

Proposition 3. *When the media firm provides news, the reporting inaccuracy* (RI) *increases when the cost of information provision* (c) *increases, but it remains unchanged when the prior uncertainty level* (H) *increases.*

Proposition 3 states that when the firm provides news, the cost of news provision leads to higher reporting inaccuracy. High costs prohibit the firm from choosing high accuracy, which leads to an increased probability of misreporting, and consumers are more likely to end up in a wrong posterior belief. Surprisingly, the amount of prior uncertainty does not influence reporting inaccuracy. One may expect there to be more inaccurate reporting in a murkier environment because evidence collection will be more difficult. However, consumers' higher valuation for accurate reporting in such an environment also incentivizes the firm to invest more. As a result, those two forces exactly cancel each other out so that reporting inaccuracy is invariant in prior belief.

Note that the results in this section show that consumers' prior beliefs play an important role in the media firm's optimal reporting strategy. To further understand the impact of prior beliefs, we consider heterogeneity in consumers' prior beliefs in the next section.

3.5 Extensions

Over a wide range of topics, consumers have heterogeneous prior beliefs about the true state of the world. For example, consumers who have had negative experiences with tap water tend to believe in the existence of water pollution. Furthermore, the degree of heterogeneity among consumers' beliefs can change over time. Surveys show that during the past two decades, Americans have become more divided in their attitudes towards political values and policy issues, such as immigrants and government regulation of business (Pew Research Center 2017). Polarization in individuals' opinions directly influence their consumption of news, and thus the media firms' provision of news. To understand the impact of polarization on news accuracy, we analyze an extension of the main model that incorporates consumers' heterogeneity in their priors.

In addition to the heterogeneity among consumers, the firm's objective behind news provision also drives its strategy. For media firms, tension always exists between profits and the public interest (Croteau and Hoynes 2006). In a second extension, we assume that the firm is a social planner that aims at improving consumers' probability of being correct about the true state. We analyze the firm's response to polarization and compare it with that of a pure profit maximizer. Finally, in a third extension, we analyze a firm with a dual objective (of obtaining profits and improving consumers' probability of being correct), which illustrates how the trade-off between a firm's commercial interest and the public interest drives the accuracy of news.

3.5.1 Impact of Polarization on News Accuracy

To study the impact of polarization, we first introduce an approach to capture polarization, then analyze its impact on the firm's choice of accuracy. Suppose consumer i has a prior belief that $\Pr(R) = \mu_0^i \in [0,1]$. We use a linear probability density function to represent a family of distributions of consumers' beliefs, as given in Equation (3.5).

$$f(\mu;\theta) = \begin{cases} 4(1-\theta)\mu + \theta & \text{if } \mu \le \frac{1}{2}; \\ 4(1-\theta)(1-\mu) + \theta & \text{if } \mu > \frac{1}{2}. \end{cases} \tag{3.5}$$

where $0 \le \theta \le 2$.

Insert Figure 3.4 about here

This density function is flexible and can describe a wide variety of distributions. Figure 3.4 displays three examples with different values of θ. When the parameter θ is small, the majority of consumers believe L and R are similarly likely (i.e., their μ are close to $\frac{1}{2}$). An increase in θ will move consumers towards both ends of the spectrum, which means their beliefs become more polarized. As θ approaches to 2, the majority of consumers believe that one state is much more likely than the other. Note that previous research has suggested using variance as a measure of polarization (DiMaggio et al. 1996). In our context, the variance of μ, $\mathrm{Var}(\mu;\theta) = \frac{1+\theta}{24}$, is a linear function of θ, which also shows that θ captures the degree of polarization in beliefs.

Further, the true underlying distribution of t is assumed to be $\Pr(R) = \frac{1}{2}$. Recall that in the main model, consumers' prior belief, μ_0, is assumed to be consistent with the underlying distribution. On the contrary, there are discrepancies between consumers' priors and the underlying distribution when we incorporate heterogeneous priors into the model. The assumption $\Pr(R) = \frac{1}{2}$ means that the median consumer's prior belief is consistent with the true underlying distribution. We also assume that the prior belief of the medium, μ_0^m, is consistent with the underlying distribution of t (i.e., $\mu_0^m = \frac{1}{2}$), as in the main model.

We now consider the impact of consumer heterogeneity on the consumption and provision of news. To simplify the analysis, we assume that the firm needs to maintain $\alpha_L = \alpha_R$. To facilitate exposition, we omit the subscripts and use α to represent the firm's accuracy in the rest of this section. We first analyze consumers' purchase decisions to derive the total demand for news, then discuss the firm's trade-offs when deciding on its accuracy. Next, we study the firm's equilibrium strategy, compare it with that in the main model, and finally discuss the implication of polarization for news provision.

Recall that news consumption happens only when it leads to a strict increase in a consumer's ex ante expected utility. In other words, consumer i purchases news when $WTP^i > p$. where WTP^i can be derived based on Lemma 1 in Section (3.3.4). Corollary 1 below formally summarizes the relationship between consumers' consumption of news, their prior beliefs, and the reporting accuracy in the presence of heterogeneous prior beliefs. Though consumers with more extreme prior beliefs see relatively little value in news and generally tend not to consume news, they would still use it when the news is sufficiently accurate.

Corollary 1. *Extreme prior beliefs ($\mu_0^i \leq 1-\alpha+\frac{p}{v}$ or $\mu_0^i \geq \alpha-\frac{p}{v}$) discourage news consumption. However, as the media becomes more accurate, more consumers start consuming news.*

Aggregating all consumers' decisions leads to the total demand as a function of price, $D(\alpha, p) = \int_{1-\alpha+\frac{p}{v}}^{\alpha-\frac{p}{v}} f(\mu;\theta)d\mu$, with which we can obtain the optimal price, $p^*_{het}(\alpha)$, where the subscript "het" denotes "heterogeneous beliefs" among consumers. The next step is to find the optimal accuracy that maximizes the firm's profit. When the firm decides its accuracy, it needs to account for several effects of accuracy on its profit, summarized in Equation (3.6) below.

$$\left.\frac{d(D\cdot p - C)}{d\alpha}\right|_{p=p^*_{het}(\alpha)} = \underbrace{\frac{\partial p^*_{het}}{\partial \alpha}}_{\text{price effect}} D(p^*_{het}) + p^*_{het}\underbrace{\left(\frac{\partial D}{\partial \alpha} + \frac{\partial D}{\partial p}\frac{\partial p^*_{het}}{\partial \alpha}\right)}_{\text{demand expansion effect}} - \underbrace{\frac{\partial C}{\partial \alpha}}_{\text{cost effct}} \tag{3.6}$$

Equation (3.6) shows the three effects of accuracy on the firm's profit. First, for each consumer of the media firm, news of higher accuracy is ex ante more valuable, so the firm can charge a higher price. This is the *price effect* of accuracy. Second, as shown in Corollary 1, higher accuracy will attract consumers with relatively extreme beliefs to consume news, which leads to a higher demand (captured by $\frac{\partial D}{\partial \alpha}$). A higher demand gives the firm an incentive to raise the price, so the demand increase induced by the higher accuracy is attenuated (by the size of $\frac{\partial D}{\partial p}\frac{\partial p^*_{het}}{\partial \alpha}$). The total impact of accuracy on demand is positive, which can be viewed as a *demand expansion effect.*

Finally, increasing accuracy also has a *cost effect*, because the firm needs to expend more resources in collecting information and reducing uncertainty. To summarize, the price and demand expansion effects incentivize the firm to increase accuracy, but the cost effect drives the firm to decrease accuracy.[20]

Recall that in the main model, only the price and cost effects play a role in the firm's decision of accuracy, because the demand function $D(\alpha, p)$ is a step function of α and p under a homogeneous prior belief. In other words, as long as accuracy (price) is sufficiently high (low), all consumers would consume news, and the demand is not sensitive to a change in accuracy (price). This property allows the firm to capture a large amount of consumer surplus. We will further elaborate on this point when comparing the results of the case of heterogeneous priors with those in the main model.

Accounting for the heterogeneity in consumers' prior beliefs, we solve the profit maximization problem with Equation (3.6) and summarize the results in Lemma 2. We assume that the firm favors providing news at its breakeven point.

Lemma 2. *A profit-maximizing media firm chooses a full-accuracy strategy* ($\alpha = 1$, $p = p^*_{het}(\alpha = 1)$) *when* $c \leq \bar{c}$*; it chooses a partial-accuracy strategy* ($\alpha = \alpha^*_{het} < 1$, $p = p^*_{het}(\alpha = \alpha^*_{het})$) *when* $\theta < 1$ *and* $\bar{c} < c < \frac{v(2-\theta)}{2}$*, and does not provide news otherwise.*[21]

Figure 3.5 illustrates the results in Lemma 2. Similar to the case with homo-

[20]The indirect effects in Equation (3.6) sum up to zero: $\frac{\partial p^*_{het}}{\partial \alpha} D(p^*_{het}) + p^*_{het} \frac{\partial D}{\partial p} \frac{\partial p^*_{het}}{\partial \alpha} = \left(D(p^*_{het}) + p^*_{het} \frac{\partial D}{\partial p} \right) \frac{\partial p^*_{het}}{\partial \alpha} = \frac{\partial (D \cdot p - C)}{\partial p} \frac{\partial p^*_{het}}{\partial \alpha} = 0$, because $\frac{\partial (D \cdot p - C)}{\partial p} = 0$ when $p = p^*_{het}(\alpha)$.

[21]$\bar{c}$ is $\frac{v(\theta^2 - 6\theta + 6 - \theta\sqrt{\theta^2 - 3\theta + 3})}{9(1-\theta)}$ when $\theta < 1$, is $\frac{v}{2}$ when $\theta = 1$, and is $\frac{v(4\sqrt{(\theta^2 - 3\theta + 3)^3} - 4\theta^3 + 18\theta^2 - 18\theta)}{27(\theta - 1)^2}$ when $\theta > 1$. α^*_{het} and $p^*_{het}(\alpha)$ are given in Appendix B.

geneous priors, the firm provides full accuracy under low costs but no news under high costs. When the unit cost of providing information is relatively low compared to consumers' valuation, the price and demand expansion effects dominate. As a result, the firm chooses the highest accuracy to create a strong demand and then charges a high price. However, under high costs, consumers' WTP can no longer cover the cost. The cost-saving motivation drives the firm to choose lower accuracy despite the resulting decrease in price and demand. Consequently, news becomes uninformative and no one will purchase it, leading to a failure of news provision.

When the degree of polarization is low and the cost level is intermediate, partial accuracy is offered by the firm. Interestingly, partial accuracy is never chosen by the firm when the degree of polarization is high. We explain the intuition by considering the firm's response to a cost change. When the cost level (c) increases, the marginal cost of accuracy becomes higher. In response to this change, the firm tends to reduce the accuracy and focus on the high-valuation consumers (i.e., those with higher prior uncertainty). If the polarization level is low (i.e., $\theta < 1$), reducing accuracy can effectively cut the cost while only drives a limited number of consumers away. The firm is able to find an accuracy level to balance the three effects (defined in Equation (3.6)). However, when consumers become sufficiently polarized (i.e., $\theta \geq 1$), saving cost by reducing accuracy can end up like a downward spiral: a small decrease in the accuracy can lead to a significant drop in demand, which forces the firm to cut the price to retain consumers. However, a lower price means an even lower return to accuracy, so that the firm would further reduce its accuracy to save costs. As a result, providing no news is optimal for the media firm when the cost is above a threshold.[22]

Insert Figure 3.5 about here

By comparing the results of the main model with those of Lemma 2, we identify the impact of consumers' heterogeneous prior beliefs on news provision. In the main model, the firm pursues a full-accuracy strategy when $c < v$, and a partial-accuracy strategy otherwise. In Lemma 2, the range of costs for a full-accuracy strategy is narrower, and the range for no news expands. This is because the firm loses part of its ability to extract consumer surplus when consumers have heterogeneous prior beliefs. Heterogeneous prior beliefs imply heterogeneous WTP for news, which means that the demand becomes more sensitive to a price change. Therefore, for the same reporting accuracy, the optimal price under a heterogeneous prior is lower than that under a homogeneous prior. Furthermore, this price decreases in the degree of polarization. As a result, the return to investing in accuracy improvement decreases, which incentivizes the firm to choose lower accuracy.

While polarization can drive the firm to choose lower accuracy levels, Lemma 2 also reveals that its impact critically depends on the level of costs. We summarize the finding formally in the following proposition.

Proposition 4. *When consumers' prior beliefs become more polarized, a profit-maximizing media firm's accuracy does not change under low costs* ($c \leq \frac{8v}{27}$) *or high costs* ($c \geq v$)*, and weakly decreases under moderate costs* ($\frac{8v}{27} < c < v$)*.*

[22]The threshold for cost is $\frac{v(2-\theta)}{2}$ when $\theta < 1$, and $\bar{c}$ when $\theta \geq 1$.

Proposition 4 shows that an increase in polarization influences media accuracy only when the cost is at a moderate level. First, under low costs ($c \leq \frac{8v}{27}$), polarization does not influence the firm's accuracy. Under this condition, accurate information about the true state is readily accessible and can be collected with low expenditure, such as reporting today's temperature in a particular zip code or describing the functions of a newly-released smartphone, etc. The firm can easily obtain an accurate signal and report it in its news. Its decision is driven by the price and demand expansion effects of accuracy. In this case, the firm provides full accuracy.

Second, when the cost is sufficiently high ($c \geq v$), the firm chooses the minimum accuracy ($\alpha = \frac{1}{2}$). No one would consume news when it is of the minimum accuracy, so the market for news fails. This corresponds to the issues where accurate reporting is extremely challenging, such as examining the overall impact of a government policy or explaining the debate around a thousand-page bill in Congress. Under this condition, the cost effect is the dominant factor for choosing accuracy regardless of the degree of polarization.

Finally and interestingly, under moderate costs ($\frac{8v}{27} < c < v$), stronger polarization drives the media to choose lower accuracy. At these levels of costs, providing full accuracy is more costly than that under low costs, but not large enough to prevent news provision. In other words, none of the three effects in Equation (3.6) dominates so that the degree of polarization drives the firm's decision. When the degree of polarization increases, consumers become more certain about the state under their priors and thus have lower WTP for news. Lower WTP weakens the price effect of raising accuracy, because it not only leads to lower demand for news (in Equation (3.6), $D(p^*_{het})$ decreases), but also implies the firm cannot raise its price as much as it could for the same increase in accuracy (in Equation (3.6), $\frac{\partial p^*_{het}}{\partial \alpha}$ decreases). Under this condition, the firm needs to decrease price (p^*_{het} in Equation (3.6)) to attenuate the decline in demand, which results in a lower return on investment in improving accuracy.[23] All in all, under higher degrees of polarization, the firm is less likely to provide news, and is less accurate upon news provision. From a societal perspective, this result is significant because polarization should call for more accuracy so that consumers with different opinions ex ante can collectively move closer to the truth upon news consumption. However, Proposition 4 demonstrates that in equilibrium, the profit-maximizing media firm either decreases or makes no change to its accuracy as polarization increases, which may further discourage the consumption of news. This outcome suggests that there will be scope for third-party intervention, especially when the underlying issues are important political or economic ones. In the next subsection, we furthur extend the model to study how the firm's objective influences its response to polarization.

3.5.2 Media Firm As An Objective-Welfare Maximizer

In addition to generating revenues for media firms, news also has educational and public service functions for society (Croteau and Hoynes 2006). In this subsection, we first discuss how to measure social welfare from an objective perspective

[23]Greater polarization may increase or decrease the demand expansion effect of accuracy in Equation (3.6), depending on the level of reporting cost. Compared with the changes in other parts of Equation (3.6), this change has less impact on the firm's decision on accuracy.

(i.e., objective welfare) and then analyze the game assuming that the media firm intends to maximize objective welfare. Finally, we compare the results with those in Section 3.5.1 to understand how the firm's objective (i.e., maximizing profits vs. maximizing objective welfare) influences the market outcome in the presence of polarization.

To ensure a fair comparison between the results in this subsection and those in Section 3.5.1, we retain three assumptions made in Section 3.5.1. First, the distribution of consumers' prior beliefs is described by $f(\mu, \theta)$, as defined in Equation (3.5). Second, the true underlying distribution of t is $\Pr(t = R) = \frac{1}{2}$, and the media firm's prior belief is consistent with the underlying distribution (i.e., $\mu_0^m = \frac{1}{2}$). Third, the media firm maintains $\alpha_L = \alpha_R = \alpha > \frac{1}{2}$ when news is provided.

When consumers and the firm have different beliefs about the true state of the world, they have different evaluations of consumers' probability of being correct about the truth, and thus different evaluations of welfare. From consumer i's perspective, her probability of being correct is $\Pr(t = \tau^i|\mu^i)$, which is evaluated based on her belief μ^i. We refer to $\Pr(t = \tau^i|\mu^i)$ as consumer i's *subjective* probability of being correct, which is her expected utility when $v = 1$ (see Equation (3.3)). By contrast, from the firm's perspective, consumer i's probability of being correct is $k^i = \Pr(t = \tau^i|\mu^m)$, which is evaluated based on the media firm's belief μ^m. For a media firm that cares about consumers' objective welfare, k^i is a more reasonable measure to rely on than $\Pr(t = \tau^i|\mu^i)$. k^i is consumer i's *objective* probability of being correct, because the firm's prior is consistent with the underlying distribution of the truth. To best serve the public interest, a media firm should (in expectation) use news to lead consumers objectively closer to the truth, rather than merely let consumers subjectively believe that they are closer to the truth. In the rest of the paper, we simply use "consumer i's probability of being correct" to refer to k^i, unless otherwise specified.

To aid the discussion, we use Figure 3.6 to illustrate the relationship between consumers' subjective and objective probabilities of being correct, and how they relate to subjective and objective welfare. We normalize $v = 1$ so that each consumer's expected utility equals her subjective probability of being correct. In Figure 3.6(a), the V-shape line is each consumer's expected utility (and subjective probability of being correct) without news consumption, which is a function of one's prior belief, μ_0^i. For example, if consumer i believes R happens with a probability of 0.6 under the prior (i.e., $\mu_0^i = 0.6$), then the more likely state in her mind is $\tau_0^i = R$, and she subjectively believes that she is correct 60% of the time (i.e., $\Pr(t = R|\mu_0^i) = 0.6$). The area below the V-shape line represents the baseline level of subjective welfare enjoyed by all consumers without consuming news.

If the firm provides news with accuracy α and price p, consumers with prior beliefs $1-\alpha+\frac{p}{v} < \mu_0^i < \alpha-\frac{p}{v}$ will consume news (see Corollary 1). These consumers pay p to access news, and their ex ante expected utility is $v\alpha = \alpha$ (based on Equation (3.4)). Consumer i's ex ante expectation of her (subjective) probability of being correct becomes $E_s[\Pr(t = \tau_s^i|\mu_s^i)] = \alpha$ if she consumes news.[24] After paying p, those consumers expect a surplus from news consumption (blue triangle), while the

[24]When news may lead to consumer i's belief change, based on Equation (3.3), $E_s[\Pr(t = \tau_s^i|\mu_s^i)] = E_s\left[\frac{E_t[u(\mu_s^i,t)]}{v}\right] = \frac{E_s[E_t[u(\mu_s^i,t)]]}{v}$. Based on Equation (3.4), $E_s[E_t[u(\mu_s^i,t)]] = v(\alpha(1-\mu_0^i) + \alpha\mu_0^i) = v\alpha$, so $E_s[\Pr(t = \tau_s^i|\mu_s^i)] = \frac{v\alpha}{v} = \alpha$.

firm obtains revenues (red rectangle), and the sum of these two parts is the ex ante expectation of the increase in gross *subjective welfare* due to news consumption. Specifically, consumer i's ex ante surplus from news consumption is the difference between her WTP and the price, given in the following equation.

$$\begin{aligned} CS^i &= E_s[E_t[u(\mu_s^i, t)]] - E_t[u(\mu_0^i, t)] - p \\ &= v(E_s[\Pr(t = \tau_s^i | \mu_s^i)] - \Pr(t = \tau_0^i | \mu_0^i)) - p \\ &= \begin{cases} v(\alpha - (1 - \mu_0^i)) - p & \text{if } \mu_0^i < \frac{1}{2}, \\ v(\alpha - \mu_0^i) - p & \text{if } \mu_0^i \geq \frac{1}{2}. \end{cases} \end{aligned} \tag{3.7}$$

In Equation (3.7), consumer i's ex ante surplus depends on the increase in her expected (subjective) probability of being correct because of news consumption, $E_s[\Pr(t = \tau_s^i | \mu_s^i)] - \Pr(t = \tau_0^i | \mu_0^i)$. For example, in Figure 3.6(a), if a consumer with prior $\mu_0^i = 0.6$ consumes the news of accuracy $\alpha = 0.9$, her expected (subjective) probability of being correct will increase by $0.3 (= 0.9 - 0.6)$. Her ex ante consumer surplus will be $v(\alpha - \mu_0^i) - p = 1(0.9 - 0.6) - 0.2 = 0.1$. In our analysis in Section 3.5.1, a profit-maximizing firm optimizes its producer surplus, which equals to its revenue (red rectangle) less its cost (which is not plotted in Figure 3.6). By contrast, a subjective-welfare-maximizing firm optimizes the total surplus, i.e., the sum of the producer surplus (red rectangle less the cost) and the (ex ante) consumer surplus (blue triangle). Importnatly, in our model, the firm focuses on improving objective welfare, which is different from the expected total surplus because the latter depends on consumers' subjective probability of being correct. We will further elaborate on this difference in the subsequent discussion about Figure 3.6.

Insert Figure 3.6 about here

Figure 3.6(b) shows consumers' (ex ante) gross *objective welfare*, which is the multiplication of consumers' value of being correct (v, which is 1 in this figure) and their ex ante objective probability of being correct ($E[k^i]$). Without news, the expected probability of being correct ($E[k^i]$) is 0.5 for all consumers. Given that the truth is either state with equal likelihood, all consumers are expected to be correct half of the time. The area below the horizontal line at 0.5 represents the baseline level of objective welfare. Comparing these two figures, we see that without news, a consumer subjective probability of being correct (the V-shape line in Figure 3.6(a)) is weakly higher than the objective probability (0.5, as in Figure 3.6(b)). This is because the prior biases in consumers' beliefs give them a false sense of certainty and make them overestimate their probability of being correct.

When consumers with prior beliefs $1-\alpha+\frac{p}{v}<\mu_0^i<\alpha-\frac{p}{v}$ consume news, their ex ante objective probability of being correct ($E_s[k_s^i]=E_s[\Pr(t=\tau_s^i|\mu_s^m)]$) is α.[25] Note that this equals to their ex ante expectation of the (subjective) probability of being correct in Panel (a). The intuition is as follows. As discussed in Section 3.3.4, consumers would consume news only when news may lead to a belief change. Under this condition, the state which is more likely under consumer i's posterior

[25] When news may lead to consumer i's belief change, there are $\tau_l^i=L$ and $\tau_r^i=R$. Under this condition, consumer i's ex ante probability of being correct is $E_s[k_s^i]=E_s[\Pr(t=\tau_s^i|\mu_0^m)]=\Pr(s=l,t=\tau_l^i|\mu_0^m)+\Pr(s=r,t=\tau_r^i|\mu_0^m)=\Pr(s=l,t=L|\mu_0^m)+\Pr(s=r,t=R|\mu_0^m)=\alpha$.

(τ_s^i) always depends on the news content (s) that she observes (i.e., $\tau_l^i = L$ and $\tau_r^i = R$). Furthermore, the news content (s) is consistent with the true state (t) with a probability of α (i.e., $\Pr(s = l|t = L)= \Pr(s = r|t = R) = \alpha$). As a result, τ_s^i is expected to be consistent with t with a probability of α, regardless of being evaluated based on the consumer's belief or the firm's belief.

By comparing Figures 3.6(a) and 3.6(b), we observe that news consumption has a greater impact on consumers' objective probability of being correct than on their expected subjective probability. Recall that in the example, when $\alpha = 0.9$, a consumer with $\mu_0^i = 0.6$ will get an increase of 0.3 in her ex ante subjective probability of being correct. From the firm's perspective, consumer i's objective probability of being correct increases by $0.4(= 0.9 - 0.5)$ because of news consumption. In Figure 3.6(b), the gray rectangle represents the increase in consumers' ex ante gross objective welfare. This increase in objective welfare can be interpreted as the (objective) value of moving consumers closer to the truth. It is greater than the increase in subjective welfare, i.e., the sum of the ex ante consumer surplus and the firm's revenue. The is because when consumers overestimate their probability of being correct, they will underestimate the value of information in news.

Next, we analyze the game assuming that the firm's objective is to maximize objective welfare. Consumers consume news only when it strictly increases their expected utility, which requires $p < WTP^i$ for consumer i. Accounting for consumers' consumption decisions, the firm solves $\max_{\alpha,p} v \int_i \left(E[k^i] - \frac{1}{2}\right) - C$.[26] Lemma 3 below

[26] Note that solving $\max_{\alpha,p} v \int_i \left(E[k^i] - \frac{1}{2}\right) - C$ and solving $\max_{\alpha,p} v \int_i E[k^i] - C$ will generate the same solution. To be consistent with Figure 3.6(b), we use the former one.

describes the firm's optimal strategies.

Lemma 3. *An objective-welfare-maximizing media firm chooses a full-accuracy strategy* ($\alpha = 1, p = 0$) *when* $\theta < 1$ *and* $c \leq v(1+\theta)$*, or* $\theta \geq 1$ *and* $c \leq 2v$*; it chooses a partial-accuracy strategy* ($\alpha = \alpha^*_{het,k} < 1, p = 0$) *when* $\theta < 1$ *and* $v(1+\theta) < c < 2v(2-\theta)$*, and does not provide news otherwise.*[27]

Figure 3.7 illustrates the results in Lemma 3. The overall pattern is similar to that of a profit-maximizing firm: the firm offers full accuracy under low costs, offers partial accuracy under moderate costs if polarization is low, and offers no news under high costs. Notable differences from comparing Figures 3.5 and 3.7 are that an objective-welfare maximizer is more likely to provide news than a profit maximizer, and its accuracy is weakly higher than that of a profit maximizer. This is because the payoff gain from raising accuracy for an objective-welfare maximizer is larger than the profit gain from raising accuracy for a profit maximizer. First, when a firm raises its accuracy, the gain in objective welfare (viewed from the firm's perspective) is larger than the increase in subjective welfare (viewed from consumers' perspectives), because consumers' incorrect prior beliefs make them underestimate the benefit of news. Second, only part of the increase in subjective welfare translates into the firm's profit, because the firm cannot fully capture consumer surplus when charging a uniform price from heterogeneous consumers. Consequently, an objective-welfare maximizer tends to choose higher accuracy levels than a profit maximizer. Importantly, this result illustrates that financial incentives alone may

[27] $\alpha^*_{het,k} = \frac{v(7-5\theta)-c}{6v(1-\theta)}$.

not be sufficient to generate news reports that maximize objective welfare under relatively high costs of reporting.

Insert Figure 3.7 about here

Lemma 3 also shows how an objective-welfare maximizer responds to polarization and how costs influence the equilibrium strategy. We summarize those findings in the following proposition.

Proposition 5. *When consumers' prior beliefs become more polarized, an objective-welfare-maximizing media firm's accuracy does not change under low costs* ($c \leq v$) *or high costs* ($c \geq 4v$)*. Its accuracy weakly increases under moderately low costs* ($v < c < 2v$)*, and weakly decreases under moderately high costs* ($2v \leq c < 4v$)*.*

Fundamentally different from that of a profit-maximizing firm, the social planer media firm may increase its accuracy when consumers become more polarized. Specifically, this happens when $v < c < 2v$. The intuition is as follows. An objective-welfare maximizer firm offers the news for free to encourage news consumption, and consumers with relatively high prior uncertainty (i.e., those with $1 - \alpha < \mu_0^i < \alpha$) consume the news. When polarization increases, consumers become more extreme, and some of them would stop consuming news, leading to a decline in objective welfare. In response to such a decline, the firm would raise its reporting accuracy when the cost of reporting is not prohibitively high. Two factors strengthen the firm's incentives to choose higher accuracy. On the one hand, under higher degrees of polarization, the firm can achieve a larger increase in demand from the same increase in accuracy (i.e., $\frac{\partial^2 D(\alpha,0)}{\partial\theta\partial\alpha} > 0$) as long as the level of accuracy is sufficiently high.[28] On the other hand, when the firm raises accuracy to convert non-consumers, each conversion generates a payoff gain of $v(\alpha - \frac{1}{2})$, which is invariant in how extreme the converted consumer's prior belief is. Overall, the objective-welfare maximizing firm has stronger motivations to convert extreme consumers in the presence of polarization, so it would raise accuracy levels when reporting costs are not too high ($c < 2v$).

For a profit-maximizing firm, the benefit from raising accuracy is different, which depends on its financial return. Polarization decreases consumers' WTP for news. As we pointed out in the discussion following Proposition 4, consumers' extreme prior beliefs make them see less value in news. Holding everything else constant, when polarization increases, the demand decreases so that the firm is only able to charge a lower price for the same accuracy level. Moreover, for the same increase in accuracy, the additional price the firm can charge is smaller under greater polarization. In summary, polarization weakens the profit-maximizing firm's incentives to improve accuracy, so the firm reduces its accuracy and saves costs. Again, consumers' underestimation of the benefit of news (because of their incorrect prior beliefs) is the key driver behind the difference between a profit maximizer and an objective-welfare maximizer. It limits the profit maximizer's ability to extract consumer surplus, which in turn hurts consumers' probability of being correct.

[28] Please see Appendix B for details.

3.5.3 Media Firm with Dual Objectives

In the two previous extensions, we assume that the media firm either maximizes profits or objective welfare. In practice, many media firms may need to balance between the two distinctive objectives instead of completely focusing on one. To capture this feature, in this subsection, we consider the case where the media firm has a dual objective of earning profits and improving objective welfare, solving $\max_{\alpha,p} \Omega = \delta \cdot D \cdot p + (1-\delta)v \int_i \left(E[k^i] - \frac{1}{2}\right) - C$. In the objective function, $\delta \in [0,1]$ is the weight of the profit-seeking objective, and $1-\delta$ is the weight of the objective-welfare-improving objective. The two previous extensions are special cases of this function: $\delta = 1$ corresponds to a profit maximizer and $\delta = 0$ corresponds to an objective-welfare maximizer. To provide more generalized insights, in this extension, we focus on the case where the firm accounts for both objectives, i.e., $0 < \delta < 1$.

We first analyze consumers' decisions on news consumption. They consume news when their WTP is above the price of news. Then we solve the firm's optimal price and optimal accuracy level. We summarize the firm's equilibrium strategies in the lemma below.

Lemma 4. *When the media firm has a dual objective of earning profits and improving objective welfare, its equilibrium strategies are summarized in Table 3.1, where $p^*_{het,d}(\alpha)$ is uniquely determined by $\frac{\partial \Omega(\alpha,p)}{\partial p} = 0$, $\alpha^*_{het,d}$ is uniquely determined by $\frac{\partial \Omega(\alpha, p^*_{het,d}(\alpha))}{\partial \alpha} = 0$ on $\alpha \in (\frac{1}{2}, 1)$, and $\alpha^{**}_{het,d} = \frac{5}{6} + \frac{2v(1-\delta)-c}{6v(1-\delta)(1-\theta)}$. The expressions of $\underline{c}_d$ and $\overline{c}_d$ are given in Appendix B.*

Insert Table 3.1 about here

Insert Figure 3.8 about here

The results in Lemma 4 are illustrated by Figure 3.8(a). Intuitively, the firm chooses full accuracy under low costs and weakly reduces its accuracy as costs increase. The level of accuracy also decreases in the strength of the firm's profit-seeking motivation (δ). As in our discussion about Lemma 3, an objective-welfare maximizer experiences a larger payoff gain from raising its accuracy than a profit maximizer. Therefore, a media firm's accuracy decreases when it weighs more heavily towards earning profits. By contrast, a less profit-seeking firm tends to encourage the consumption of news. When δ is sufficiently low, it can even offer news for free. We summarize a media firm's equilibrium pricing strategy in the proposition below.

Proposition 6. *In the equilibrium where a media firm provides news, it always provides news with a positive price if its profit-seeking motivation is sufficiently strong* ($\frac{2}{3} < \delta \leq 1$). *If its profit-seeking motivation is weak* ($0 \leq \delta \leq \frac{2}{3}$), *it provides news for free under high costs* ($c \geq \overline{c}_d$) *or high polarization* ($\theta \geq \frac{\delta}{1-\delta}$), *but with a positive price otherwise.*

In Figure 3.8(b), we provide an example with $\delta = 0.25$, which allows us to visually compare a dual-objective firm's equilibrium strategy with that in the two previous extensions. A profit maximizer always sets a positive price for news, while an objective-welfare maximizer always offers news for free. By contrast, a dual-objective firm's pricing strategy is more nuanced: it only provides news for free when its profit-seeking motivation is weak. Under this condition, it provides news for free under high costs or high polarization, but with a positive price otherwise.

We first explain why the firm charges a positive price under low costs and low polarization. On the one hand, low costs mean that it is relatively easy for the firm to offer high accuracy. On the other hand, under low polarization, most consumers are highly uncertain about the truth and thus they have a high demand for accurate news. If the firm charges a small price for news, most consumers will pay to get news, except for a small group of extreme consumers. Compared with providing news for free, charging a small price leads to a gain in revenue and a loss in objective welfare, the latter of which is due to the extreme consumers who do not consume news. Under low polarization, the gain outweighs the loss so the firm provides news with a positive price.

Now, we discuss what happens if the degree of polarization or the cost changes. When the degree of polarization increases, the firm gradually decreases its price until it becomes zero. First, in terms of revenue, greater polarization means consumers become more certain and less willing to pay for news, so the firm needs to cut the price. Second, in terms of objective welfare, polarization raises the benefit of converting non-consumers–the same price cut can attract more consumers under higher θ–as long as the level of accuracy is sufficiently high (see the discussion following Proposition 5). Both factors drive the firm to reduce its price. Likewise, an increase in the cost also leads to a drop in price. Under higher costs, the firm saves production costs by reducing its accuracy, which decreases consumers' demand. As a result, the firm lowers its price.

In addition to the pricing strategy, we are also interested in the firm's choice of accuracy. Recall that in the presence of polarization, a profit-maximizing firm always weakly decreases its accuracy, while an objective-welfare-maximizing firm may increase or decrease its accuracy, depending on its cost level. The following proposition presents our finding on a dual-objective firm's accuracy.

Proposition 7. *In the equilibrium where offering partial accuracy with a positive price is optimal, greater polarization (a higher θ) leads to higher accuracy when the dual-objective firm sufficiently values ($\delta < \frac{3}{4}$) objective welfare and the degree of polarization is sufficiently high ($\theta > \tilde{\theta}$). Otherwise, greater polarization leads to weakly lower accuracy.*[29]

Proposition 7 suggests that when a dual-objective firm offers partial accuracy with a positive price, it may increase its accuracy in response to stronger polarization. This happens when the firm sufficiently values objective welfare and the degree of polarization is sufficiently high. As we noted in the discussion following Proposition 5, greater polarization strengthens an objective-welfare maximizer's incentives to raise accuracy but weakens a profit maximizer's incentives to do so. When the firm accounts for both incentives, its response to polarization depends on the weight it puts on each objective, δ. If it sufficiently values making more

[29] $\tilde{\theta}$ is uniquely determined by $\frac{\partial \alpha^*_{het,d}}{\partial \theta} = 0$.

consumers better-informed, it will react more like an objective-welfare maximizer, raising its accuracy.

For a given firm (i.e., fixing the value of δ), greater polarization leads to higher accuracy when the degree of polarization is already sufficiently high. The intuition is as follows. Under low degrees of polarization, most consumers have high uncertainty and are willing to pay for news. In this case, the decrease in the profit-seeking incentives caused by greater polarization (i.e., not able to raise the price as much for the same increase in accuracy, lower demand, and lower price, as mentioned in the discussion following Proposition 4) is larger than the increase in the objective-welfare-improving incentive caused by it (i.e., demand become more sensitive to accuracy). Consequently, the equilibrium level of accuracy decreases. As the degree of polarization (θ) increases from low to high, the demand's sensitivity to accuracy not only increases in θ (i.e., $\frac{\partial}{\partial\theta}\left(\frac{\partial D(\alpha,p^*_{het,d}(\alpha))}{\partial\alpha}\right) > 0$), but also increases at a faster rate (i.e., $\frac{\partial^2}{\partial\theta^2}\left(\frac{\partial D(\alpha,p^*_{het,d}(\alpha))}{\partial\alpha}\right) > 0$). When θ is above a threshold, the decrease in the profit-seeking incentives caused by greater polarization is dominated by the increase in the objective-welfare-improving incentive, so the accuracy level rises. To summarize, when the degree of polarization in the environment becomes sufficiently high, the objective-welfare-improving objective dominates the firm's decision-making. The firm, therefore, raises its accuracy to counter the greater division among consumers' prior beliefs.

3.6 Concluding Remarks

In this paper, we study a media firm's news provision strategy. Specifically, we conceptualize the news provision strategy as the firm choosing reporting accuracy in a world with a binary state. Consumers want to learn about the true state, which cannot be directly observed. By consuming news from the firm, they can update their beliefs and possibly obtain a higher utility from a better understanding of the true state. We analyze the firm's optimal news provision strategy under different cost levels. First, news provision depends on both the cost of reporting and the prior beliefs of consumers. When the cost of reporting is sufficiently high, consumers' extreme prior beliefs can discourage news provision. Second, conditional on the firm providing news, it allocates more resources to accurately reporting the state that is ex ante more likely so that consumers expect an accurate report under the majority of circumstances. As a result, the relatively more likely state is exaggerated in the news report.

Interestingly, as one may expect more reporting inaccuracy in an environment with higher prior uncertainty, we find that reporting inaccuracy does not depend on prior beliefs. This is because consumers' valuation for news is higher under greater prior uncertainty, which then drives the firm to invest more in news provision. In equilibrium, the overall impact of the prior beliefs on reporting inaccuracy nets out to be zero.

Our analysis also sheds novel light on the impact of polarization in consumers' prior beliefs on news provision. We find that the firm becomes less likely to provide accurate news in a more polarized environment. Importantly, the impact of polarization is most pronounced when the cost of news provision is at an intermediate level. This is because the firm tends to respond to polarization by cutting the price under low costs, and by not providing news under high costs. On the contrary, if the firm aims at moving consumers closer to the truth rather than earning higher profits, it may increase the accuracy of reporting when polarization increases. This is because the firm evaluates social welfare from an objective viewpoint, which is not influenced by consumers' subjective beliefs over the state of the world. If the firm has a dual objective of earning profits and improving objective welfare, its strategy depends on its value for each objective and how polarized consumers are. Under high polarization, the incentive to make consumers better-informed drives the firm's decision, so the firm tends to provide news for free and to raise accuracy in response to greater polarization.

Our research provides important insights for several groups of stakeholders. For information providers, especially news media firms and platforms, our analysis shows that their brand positioning and objective have important implications for market outcomes. A media firm that puts more emphasis on moving consumers closer to the truth is more likely to supply accurate information, especially in the presence of consumer polarization. Our findings also highlight the undesirable consequences of polarization, which should put consumers on notice and be cautious about the influence of extreme opinions and beliefs. For policymakers, our study shows that the cost of providing information influences the equilibrium outcome, such that polarization may not lead to a drop in reporting accuracy if reporting cost is sufficiently low. As one of the important sources of information, if policymakers can improve the consistency and transparency in their communication to media, they can reduce the cost of reporting and thus help improve the accuracy of news. In addition, the government can support organizations and campaigns that advocate media literacy. Improving consumers' media literacy will raise their valuation for learning about the true state, which encourages the production of accurate information. Finally, for advertisers that advertise in media markets, our analysis shows that a more accurate media outlet attracts a more diverse audience, which also has implications for targeted advertising.

Our model builds on the existing framework of information design and offers new insights into the relationship between news provision strategy and market conditions. There are several paths to extend our research. First, it would be interesting to consider the impact of competition on reporting inaccuracy. In previous studies, when the provision of accurate news is costless, competition can decrease reporting inaccuracy when consumers want accurate information (GS), but increase reporting inaccuracy when consumers prefer like-minded news (MS; XS). In our model, consumers seek the truth, but the accuracy of news is costly. The introduction of cost makes the impact of competition on reporting inaccuracy less clear. Suppose two firms–a low-cost incumbent and a high-cost entrant–are competing for consumers with heterogeneous prior beliefs. When both firms are in the market, it is likely that the low-cost firm targets consumers with high uncertainty under the prior, while the high-cost firm serves news to those with lower uncertainty. On the one hand, competition can drive the incumbent to offer higher accuracy, which leads to lower inaccuracy. On the other hand, consumers served by the high-cost firm may encounter more inaccuracy than before. As a result, a priori, the impact of competition on reporting inaccuracy under costly news provision is ambiguous and merits further studies.

Second, there is a growing discussion on echo chamber and news polarization (Boxell et al. 2017; Flaxman et al. 2016). Echo chamber describes an information environment in which individuals are largely exposed to conforming opinions. This can result from selective exposure to like-minded news (Iyengar and Hahn 2009), limited attention (Che and Mierendorff 2019), and a platform's curation algorithm (Berman and Katona 2020). We find an exaggeration effect that is similar to echo chamber in our current model, and some further exploration can deepen our understanding. For example, does competition increase or decrease this effect? When consumers have low heterogeneity in prior beliefs, we expect that competition will increase accuracy and thus decrease the level of exaggeration. When consumer heterogeneity is high, we expect the opposite to happen, because firms will have enough differentiation and don't need to compete on accuracy.

Parameter Range	Strategy	Accuracy (α)	Price (p)
$\delta \leq \frac{\theta}{1+\theta}$ and $c \leq \min\{v(1-\delta)(1+\theta), 2v(1-\delta)\}$	Full accuracy for free	1	0
$\delta > \frac{\theta}{1+\theta}$ and $c \leq \underline{c}_d$	Full accuracy with a positive price	1	$p^*_{het,d}(\alpha = 1)$
$\theta < 1$, $\delta \leq \frac{\theta}{1+\theta}$ and $v(1-\delta)(1+\theta) < c < 2v(1-\delta)(2-\theta)$, or $\theta < 1$, $\frac{\theta}{1+\theta} < \delta < \frac{1}{2}$ and $\overline{c}_d \leq c < 2v(1-\delta)(2-\theta)$	Partial accuracy for free	$\alpha^{**}_{het,d}$	0
$\theta < 1$, $\delta > \frac{\theta}{1+\theta}$ and $\underline{c}_d < c < \overline{c}_d$	Partial accuracy with a positive price	$\alpha^*_{het,d}$	$p^*_{het,d}(\alpha^*_{het,d})$
Otherwise	No news	–	–

Table 3.1: Equilibrium Strategy When $0 < \delta < 1$

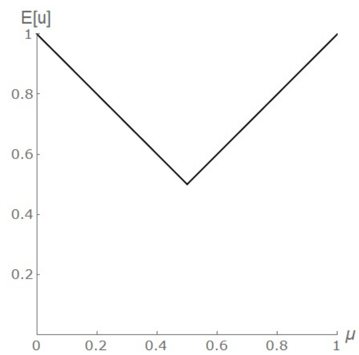
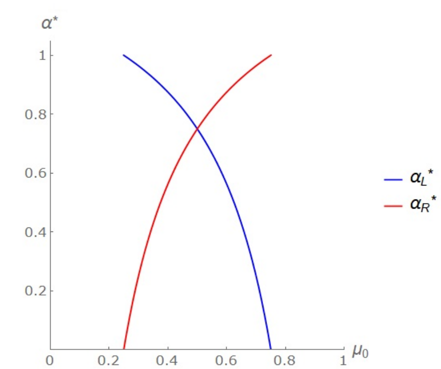
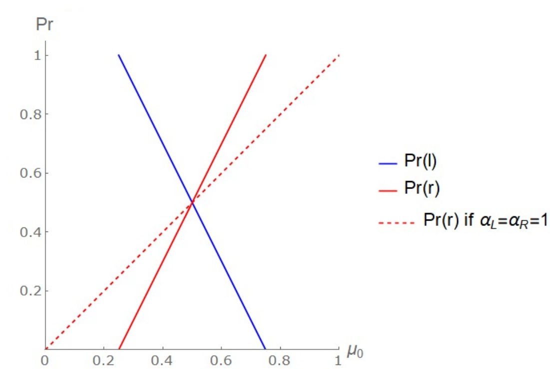


Figure 3.1: Consumer's Expected Utility Under A Given Belief ($v = 1$)



(a) Reporting Accuracy

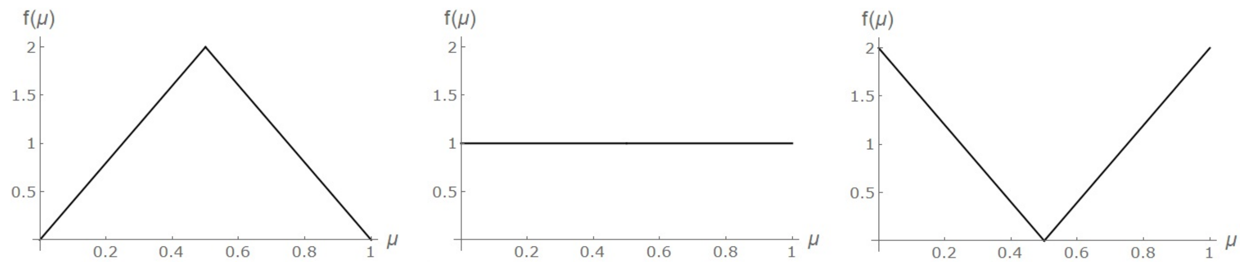


(b) Reported Content

Figure 3.2: Reporting Accuracy and Reported Content in Equilirbium ($v = 1$, $c = 2$)

(a) Ex Ante Expected Utility (b) Revenue and Cost (c) Uncertainty

Figure 3.3: Equilibrium Outcomes ($v = 1$, $c = 2$)



(a) $\theta = 0$ (b) $\theta = 1$ (c) $\theta = 2$

Figure 3.4: Examples for $f(\mu; \theta)$

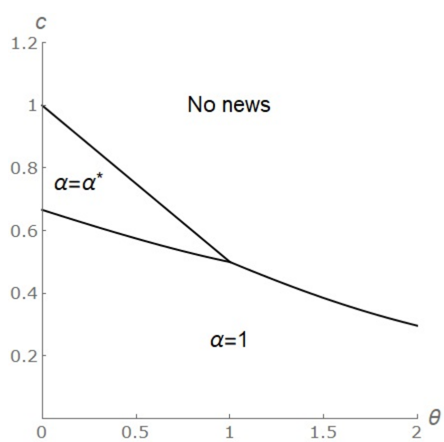
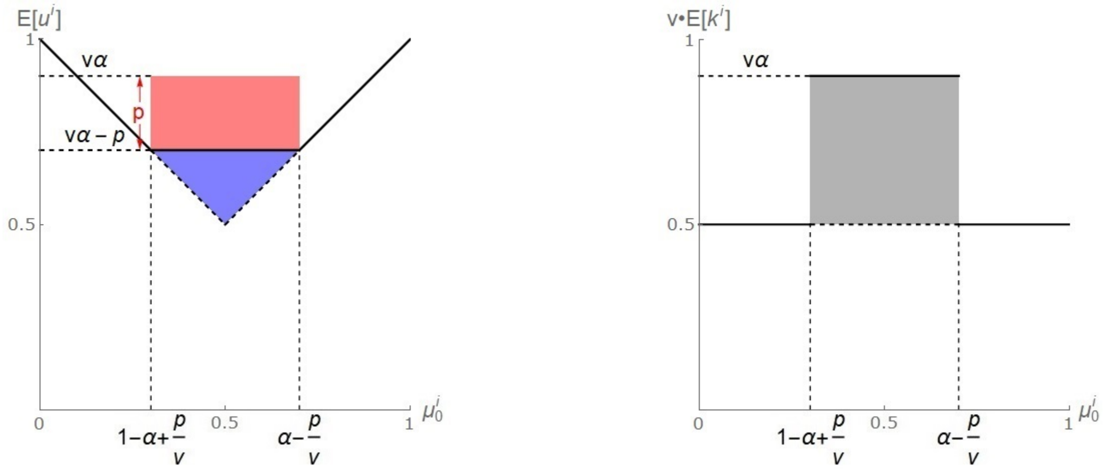


Figure 3.5: The Firm's Equilibrium Strategy ($v = 1$)



(a) Ex Ante Consumer Surplus (Blue) and Revenue (Red)

(b) Change in Gross Objective Welfare

Note*:* These two figures show the gross welfare because the costs of news provision are not plotted. In Panel (a), the solid line is consumers' ex ante expected utility if they do not consume news, or consumers' ex ante expected utility less the price if they consume news. In Panel (b), the solid line is consumers' ex ante probability of being correct.

Figure 3.6: Alternative (Gross) Objective Functions of the Firm ($v = 1$, $\alpha = 0.9$, $p = 0.2$)

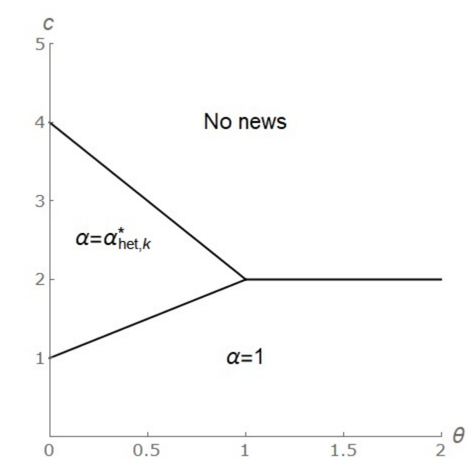
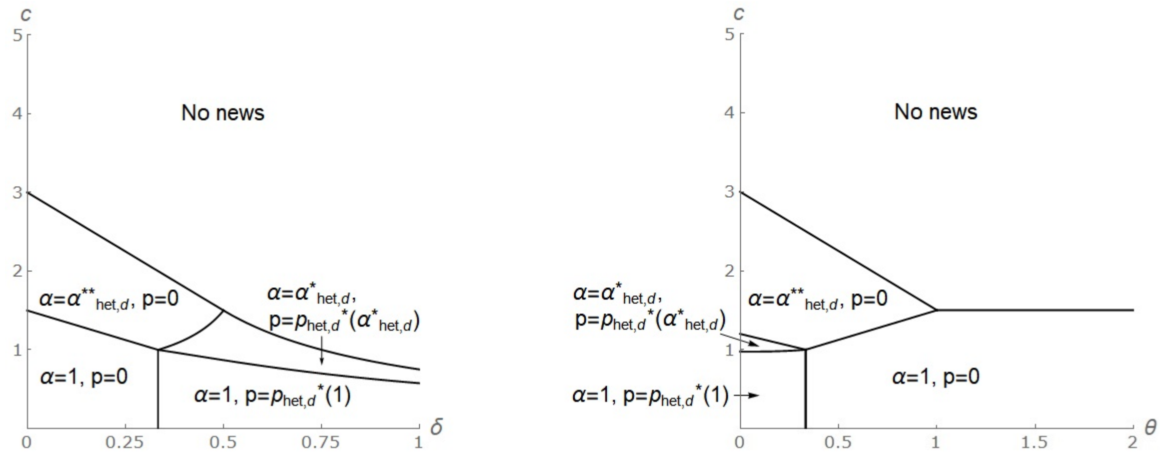


Figure 3.7: An Objective-Welfare-Maximizing Media Firm's Equilibrium Strategy ($v = 1$)



(a) When $\theta = 0.5$

(b) When $\delta = 0.25$

Figure 3.8: A Dual-Objective Media Firm's Equilibrium Strategy ($v = 1$)

Chapter 4: Conclusion

In this dissertation, we study media platforms' digital content provision and consumers' consumption decisions. We analyze platforms' pricing strategy in the presence of consumers' endogenous content consumption in the first essay. Our analysis contributes to the literature on paywall strategies. First, our paper is the first to endogenize content consumption in the presence of paywalls. Second, we incorporate consumers' heterogeneity in consumption costs, which provides a novel explanation for the existence of metered paywalls and generates new insights into the optimal paywall design.

We investigate media firms' product design strategy when consumers are truth-seeking in the second essay. Our research on the accuracy of news adds to the literature on the news consumption and information design. We show that reporting inaccuracy can be a result of media firms' unbalanced resource allocation in the presence of costly news provision, and it will lead to the exaggeration of the default state. Moreover, while polarization among consumers calls for more accurate reporting, the profit-seeking motivation may drive a firm to reduce its investment on the accuracy, leading to more inaccuracies.

There are several future directions to extend the research in this dissertation.

First, the impact of competition on media firms' strategies is an important issue. Under competition, it is more difficult for media firms to charge a price for their content. It is not clear whether media firms should respond to competition by cutting price or providing more content for free. The implications of competition for the accuracy of news also merit further study. As consumers become more polarized, competing media firms may differentiate on their reporting strategies, especially in the presence of reporting costs.

Second, the firm's content production and consumers' preference for content in the first essay can be further explored, for example, allowing the media firm to endogenously decide the quantity and quality of its content. Moreover, all the content provided by the firm is assumed to be of the same category and thus has the same value for consumers. Incorporating multiple categories of content and consumers' heterogeneous taste will allow us to study richer designs of paywalls (e.g., freemium paywalls).

Finally, the phenomenon of echo chamber can be incorporated in the model of the second essay. We find an effect similar to echo chamber in our current framework, and further exploration can deepen our understanding of this phenomenon. Depending on the extent of consumer heterogeneity, competition may increase or decrease this effect. It will also be interesting to incorporate some behavioral factors of consumers (e.g., confirmation bias) and see how this effect will be affected.

Appendix A: Proof of Lemmas and Propositions in Chapter 2

A.1 Proof of Lemmas and Propositions

A.1.1 Proof of Lemma 1

When $\alpha = 0$, all consumers are heavy readers. We first analyze consumers' utility maximization problem. Consumers decide whether to read for free or pay to subscribe. The heavy reader who is indifferent between the free version and the paid version has a valuation of $\overline{v}_H = \frac{p}{1-f}$. Next, we analyze the firm's profit maximization problem. We consider the two following cases:

1. To make some readers subscribe ($\overline{v}_H \in (0, V)$), the firm needs to set the price $p \in (0, (1-f)V)$. In this case, $\frac{\partial \pi}{\partial f} < 0$, which leads to a corner solution, $f = 0$. This is a hard paywall. Solving the first-order condition for price (FOC_p) with $f = 0$, we can obtain $p_{h,H} = \frac{V-A}{2}$. This price falls in $(0, (1-f)V)$ when $A < V$.

2. If all readers read for free, the firm only earns advertising revenue. This happens when $p \in \{0\} \cup [(1-f)V, \infty)$. In this case, $\pi = Af$, which always increases in f. The firm would choose $f = 1$, which is a no-paywall strategy.

Comparing the profits from these two strategies, we can find: the hard paywall with $p = p_{h,H}$ is optimal when $A \in [0, V)$; and no-paywall is optimal when $A \in [V, \infty)$.

A.1.2 Proof of Lemma 2

When $0 < \alpha < 1$, if the firm earn both subscription and advertising revenues, it can obtain subscriptions from both heavy and light readers or only from heavy readers. Besides, it can also choose to only earn advertising revenue. In the following, we analyze each case separately.

A.1.2.1 If the Firm Obtains Subscriptions From Both H and L Segments

We now consider a case where some heavy readers subscribe ($\overline{v}_H \in (0, V)$) and some light readers subscribe ($\overline{v}_L \in (0, V)$). While it is easy to derive $\overline{v}_H = \frac{p}{1-f}$, we need to solve $u(f; v, c) = u(\min\{\frac{v}{c}, 1\}; v, c) - p$ to derive $\overline{v}_L$. To determine the value of $\min\{\frac{v}{c}, 1\}$, we should consider whether the marginal subscriber (whose $v = \overline{v}_L$) finishes all the content (i.e., reads 1).

1. If she does not finish reading all the content ($\frac{\overline{v}_L}{c} < 1$ so that $\min\{\frac{\overline{v}_L}{c}, 1\} = \frac{\overline{v}_L}{c}$), then $\overline{v}_L = fc + \sqrt{2pc}$. The demand and impressions are Equations (2.3) and (2.4) respectively. In this case, $\frac{\partial \pi}{\partial f} < 0$, so the corner solution $f = 0$ is optimal (i.e., a hard paywall). Solving FOC_p with $f = 0$, we can obtain $p_{h,HL} = \frac{V-A}{2(1-\alpha)} + \frac{9c\alpha^2 - 3\alpha\sqrt{16c(1-\alpha)(V-A)+9c^2\alpha^2}}{16(1-\alpha)^2}$. This price satisfies $\overline{v}_H \in (0, V)$, $\overline{v}_L \in (0, V)$, and $\frac{\overline{v}_L}{c} < 1$ in the following parameter ranges: $\alpha \in (0, \frac{2}{3}]$,$A \in$

$\left[0, \frac{(2-3\alpha)V}{2}\right)$, and $c \in \left(\frac{2(V-A)}{2+\alpha}, \frac{2(1-\alpha)V^2}{(2-3\alpha)V-2A}\right)$; or $\alpha \in (0, \frac{2}{3}]$,$A \in \left[\frac{(2-3\alpha)V}{2}, V\right)$, and $c \in \left(\frac{2(V-A)}{2+\alpha}, \infty\right)$; or $\alpha \in \left(\frac{2}{3}, 1\right)$,$A \in [0, V)$, and $c \in \left(\frac{2(V-A)}{2+\alpha}, \infty\right)$.

2. If she finishes reading the whole newspaper ($\frac{\overline{v}_L}{c} \geq 1$ so that $\min\{\frac{\overline{v}_L}{c}, 1\} = 1$), then $\overline{v}_L = \frac{2p+(1-f^2)c}{2(1-f)}$. The demand is Equation (2.3) and the impressions are Equation (2.4). In this case, we also find that $\frac{\partial \pi}{\partial f} < 0$, so a hard paywall ($f = 0$) is chosen. Solving FOC_p with $f = 0$, we can obtain $p_{h,HL2} = \frac{2V-2A-c\alpha}{4}$. This price satisfies $\overline{v}_H \in (0, V)$, $\overline{v}_L \in (0, V)$, and $\frac{\overline{v}_L}{c} \geq 1$ in the following parameter ranges: $A \in [0, V)$ and $c \in \left(0, \frac{2(V-A)}{2+\alpha}\right]$. In this case, all subscribers finish reading the whole newspaper (i.e., reads 1 unit of content). By contrast, in the case of a hard-HL paywall, the marginal subscriber only reads $\frac{\overline{v}_L}{c}$ unit of content.

A.1.2.2 If the Firm Obtains Subscriptions Only From the H segment

Similar to previous, heavy readers subscribing means that $\overline{v}_H \in (0, V)$. Light readers don't subscribe, which means that $\overline{v}_L \in [V, \infty]$. In this case, $D = D_H$. The calculation for impressions is more nuanced, because no light readers would hit the paywall if the meter limit is sufficiently high (i.e., $f > \frac{V}{c}$). Therefore, we separate the analysis into two subcases: $f \in \left[0, \frac{V}{c}\right)$ and $f \in \left[\frac{V}{c}, 1\right)$.

1. When $f \in \left[0, \frac{V}{c}\right)$, some light readers hit the paywall. In this subcase, $I_L = \alpha\left(\int_0^{fc} \frac{v}{cV} dv + \int_{fc}^{V} \frac{f}{V} dv\right)$. Solve the first-order conditions, we obtain the optimal paywall design, $f_m = \frac{V}{c} - \frac{(1-\alpha)(V-A)^2}{4c\alpha A}$ and $p_m = \frac{V-A}{2}\left(1 - \frac{V}{c} + \frac{(1-\alpha)(V-A)^2}{4c\alpha A}\right)$. Denote $\underline{A} = \frac{(1-\sqrt{\alpha})^2 V}{1-\alpha}$. The meter limit falls in $\left(0, \frac{V}{c}\right)$ when $A \in (\underline{A}, V)$. Be-

sides, the price satisfies $\overline{v}_H \in (0, V)$ and $\overline{v}_L \in [V, \infty]$ in the following parameter ranges: $A \in \left(\underline{A}, \frac{(1-\alpha)V}{1+3\alpha}\right]$ and $c \in \left[\frac{(V-A)^3-2\alpha(V+A)(V-A)^2+\alpha^2(V^3+AV^2+11A^2V+3A^3)}{16\alpha^2A^2}, \infty\right)$; or $A \in \left(\frac{(1-\alpha)V}{1+3\alpha}, \frac{(1-\alpha)V}{1+\alpha}\right)$ and $c \in \left[\frac{(V-A)^2-\alpha(V+A)(V-3A)}{4\alpha A}, \infty\right)$; or $A = \frac{(1-\alpha)V}{1+\alpha}$ and $c \in \left(\frac{(V-A)^2-\alpha(V+A)(V-3A)}{4\alpha A}, \infty\right)$; or $A \in \left(\frac{(1-\alpha)V}{1+\alpha}, V\right)$ and $c \in \left[\frac{-(V-A)^2+\alpha(V+A)^2}{4\alpha A}, \infty\right)$. When $A \in [0, \underline{A}]$, $f_m \leq 0$ so that the firm would choose a hard paywall ($f = 0$). Solving FOC_p with $f = 0$, we can obtain $p_{h,H} = \frac{V-A}{2}$. This is a hard paywall targeting heavy readers. The price satisfies $\overline{v}_H \in (0, V)$ and $\overline{v}_L \in [V, \infty]$ when $c \in \left[\frac{V^2}{V-A}, \infty\right)$.

2. When $f \in \left[\frac{V}{c}, 1\right)$, no light reader hits the paywall. Thus, $I_L = \alpha \int_0^V \frac{v}{cV} dv$. We find that $\frac{\partial \pi}{\partial f} < 0$, so the corner solution $f = \frac{V}{c}$ is optimal. Solving FOC_p with $f = \frac{V}{c}$, we obtain $p = \frac{(c-V)(V-A)}{2c}$. This price satisfies $\overline{v}_H \in (0, V)$ and $\overline{v}_L \in [V, \infty]$ when $A \in [0, V)$ and $c \in (V, \infty)$.

A.1.2.3 If the Firm Only Earns Advertising Revenue

If no consumer subscribes ($\overline{v}_H$,$\overline{v}_L \in [V, \infty)$), the firm only earns advertising revenue (i.e., $D = 0$). Similar to the case where the firm obtains subscriptions only from the H segment, we separately discuss the subcases of $f \in \left[0, \frac{V}{c}\right)$ and $f \in \left[\frac{V}{c}, 1\right)$. We also discuss $f = 1$ because it is also a subcase where all revenues are from advertising.

1. When $f \in \left[0, \frac{V}{c}\right)$, some light readers hit the paywall and $I_L = \alpha \left(\int_0^{fc} \frac{v}{cV} dv + \int_{fc}^V \frac{f}{V} dv\right)$. In this subcase, $\frac{\partial \pi}{\partial f} > 0$, so that any $f \in \left[0, \frac{V}{c}\right)$ is strictly dominated by $f \in \left[\frac{V}{c}, 1\right)$.

2. When $f \in \left[\frac{V}{c}, 1\right)$, no light reader hits the paywall and $I_L = \alpha \int_0^V \frac{v}{cV} dv$. In this subcase, we also find that $\frac{\partial \pi}{\partial f} > 0$, so the firm would deviate to $f = 1$.

3. When $f = 1$, the firm chooses a no-paywall strategy $p_n = 0$.

Given any combination of parameters (V, α, A and c), we can compare the profits earned by the strategies in all the cases above. The one that earns the highest profit is the equilibrium strategy. The results of the comparison are:

1. A hard paywall with $p = p_{h,HL2}$ is optimal when $A \in [0, V)$ and $c \in \left(0, \frac{2(V-A)}{2+\alpha}\right]$;

2. A hard paywall with $p = p_{h,HL}$ is optimal when $A \in [0, V)$ and $c \in \left(\frac{2(V-A)}{2+\alpha}, \underline{c}\right]$;[1]

3. A hard paywall with $p = p_{h,H}$ is optimal when $A \in [0, \underline{A}]$ and $c \in (\underline{c}, \infty)$;

4. A metered paywall with $f = f_m$ and $p = p_m$ is optimal when $A \in (\underline{A}, V)$ and $c \in (\underline{c}, \infty)$;

5. No-paywall is optimal when $A \in [V, \infty)$.

6. A metered paywall with $f = \frac{V}{c}$ and $p = \frac{(c-V)(V-A)}{2c}$ is never optimal.

A.1.3 Proof of Lemma 3

When $\alpha = 1$, all consumers are light readers. The firm may choose to earn revenues from both subscription and advertising or only earn advertising revenue. Similar to the proof of Lemma 2, we also need to consider whether the marginal subscriber (whose $v = \overline{v}_L$) finishes all the content.

[1] $\underline{c}(V, \alpha, A)$ is the value of c such that $\pi_{h,HL}(V, \alpha, A, c) = \pi_m(V, \alpha, A, c)$ when $A \in (\underline{A}, V)$, and the value of c such that $\pi_{h,HL}(V, \alpha, A, c) = \pi_{h,H}(V, \alpha, A, c)$ when $A \in [0, \underline{A}]$.

1. If some light readers subscribe ($\overline{v}_L \in (0, V)$) and the marginal subscriber does not finish all the content ($\frac{\overline{v}_L}{c} < 1$), then $\overline{v}_L = fc + \sqrt{2pc}$. In this subcase, $\frac{\partial \pi}{\partial f} < 0$, so the firm chooses a hard paywall ($f = 0$). Solving FOC_p with $f = 0$, we can obtain $p_{h,L} = \frac{2(V-A)^2}{9c}$. This price satisfies $\overline{v}_L \in (0, V)$ when $A \in [0, V)$ and $c \in \left[\frac{2(V-A)}{3}, \infty\right)$.

2. If some light readers subscribe and the marginal subscriber finishes all the content ($\frac{\overline{v}_L}{c} \geq 1$), then $\overline{v}_L = \frac{2p+(1-f^2)c}{2(1-f)}$. In this subcase, $\frac{\partial \pi}{\partial f} < 0$ so that the firm would choose a hard paywall ($f = 0$). Solving FOC_p with $f = 0$, we can get $p = p_{h,HL2}$. This price satisfies $\overline{v}_L \in (0, V)$ when $A \in [0, V)$ and $c \in \left(0, \frac{2(V-A)}{3}\right]$.

3. In the subcase where no light reader subscribes ($\overline{v}_L \in [V, \infty)$), the firm only earns advertising revenue. The profit always increases in f when $c \in (0, V)$ and $f \in [0, 1]$; or when $c \in [V, \infty)$ and $f \in [0, \frac{V}{c}]$. As a result, a no-paywall strategy is chosen.

Comparing these solutions, we can find: a hard paywall with $p = p_{h,HL2}$ is optimal when $A \in [0, V)$ and $c \in \left(0, \frac{2(V-A)}{3}\right]$; a hard paywall with $p = p_{h,L}$ is optimal when $A \in [0, V)$ and $c \in \left[\frac{2(V-A)}{3}, \infty\right)$; and no-paywall is optimal when $A \in [V, \infty)$.

A.1.4 Proof of Proposition 1

As we find in Lemma 2, the boundary between a hard-HL paywall equilibrium and a metered paywall equilibrium is $c = \underline{c}$ (when $A \in (\underline{A}, V)$). An increase in A could make the firm shift from a metered paywall to a hard-HL paywall under the following condition: $V \in (0, \infty)$, $\alpha \in (0, 1)$, and $c \in \left(\underline{c}(\underline{A}), \max_{A \in (\underline{A}, V)} \underline{c}(A)\right)$, where $\underline{c}$ is uniquely determined by $\pi_{h,HL} = \pi_m$. We illustrate the existence of such a strategy shift with the following numerical example: $V = 1$, $\alpha = 0.25$. In this case, $\underline{c}(\underline{A}) = 1.85$ and $\max_{A \in (\underline{A}, V)} \underline{c}(A) = 2.08$. When $c = 2$, $\pi_m > \pi_{h,HL}$ when $A = 0.38$, but $\pi_m < \pi_{h,HL}$ when $A = 0.40$. Hence, the firm launches a metered paywall when $A = 0.38$ but a hard-HL paywall when $A = 0.40$. Figure 2.2 in the paper illustrates the equilibrium strategies.

A.1.5 Proof of Proposition 2

Based on Lemma 2, $f_m = \frac{V}{c} - \frac{(1-\alpha)(V-A)^2}{4c\alpha A}$, so $\frac{\partial f_m}{\partial V} = \frac{-(1-\alpha)V+(1+\alpha)A}{2c\alpha A}$. When $A \in \left(\underline{A}, \frac{(1-\alpha)V}{1+\alpha}\right)$, $\frac{\partial f_m}{\partial V} < 0$; when $A \in \left(\frac{(1-\alpha)V}{1+\alpha}, V\right)$, $\frac{\partial f_m}{\partial V} > 0$. Under a metered paywall, when there is an increase in V, less content will be provided for free under low ad rates, while the opposite is true under high ad rates. The intuition behind this result is discussed in Section 5.1.

A.1.6 Proof of Lemma 4

Based on Lemma 2, $f_m = \frac{V}{c} - \frac{(1-\alpha)(V-A)^2}{4c\alpha A}$, so $\frac{\partial f_m}{\partial \alpha} = \frac{(V-A)^2}{4c\alpha^2 A} > 0$. This implies that under a metered paywall, the amount of free content increases in α. The intuition behind this result is discussed in Section 5.1.

A.1.7 Proof of Proposition 3

In Lemma 2, we obtain f_m and p_m, based on which we can calculate the size of information flow under a metered paywall (I_m). Since Equation (2.8) is helpful with the explanation, we copy it here again:

$$\frac{\partial I_m}{\partial \alpha} = \underbrace{-\left(1-\frac{p_m}{V}\right)}_{\text{smaller H segment size (-)}} + \underbrace{\frac{f_m(2V-f_m c)}{2V}}_{\text{larger L segment size (+)}} + \underbrace{\left(\frac{(1-\alpha)(V-A)}{2V}+\frac{\alpha\,(V-f_m c)}{V}\right)\frac{\partial f_m}{\partial \alpha}}_{\text{paywall-hitting readers get more free content (+)}}$$

.

Under a metered paywall, the exact condition for $\frac{\partial I_m}{\partial \alpha} > 0$ can be obtained by plugging in p_m, f_m, and $\frac{\partial f_m}{\partial \alpha}$. After simplifying the inequality, we can find that $\frac{\partial I_m}{\partial \alpha} > 0$ when $\alpha < \widetilde{\alpha}$ *and* $c < \widetilde{c}_\alpha$, and $\frac{\partial I_m}{\partial \alpha} \leq 0$ otherwise. $\widetilde{\alpha}(V, A, c)$ is the minimum between the two following values: the value of α such that $\frac{\partial I_m}{\partial \alpha} = 0$, and $\alpha(\underline{c})$, which is the inverse function of $\underline{c}(\alpha)$. When $\alpha \geq \alpha(\underline{c})$, the firm no longer chooses a metered paywall but adopts a hard-HL paywall. $\widetilde{c}_\alpha(V, A) = \frac{V(V+A)^2}{2(V-A)A}$.

A.1.8 Proof of Proposition 4

Based on Lemma 1 to Lemma 3, $p_{h,HL} = \frac{V-A}{2(1-\alpha)} + \frac{9c\alpha^2 - 3\alpha\sqrt{16c(1-\alpha)(V-A)+9c^2\alpha^2}}{16(1-\alpha)^2}$, $p_{h,HL2} = \frac{2V-2A-c\alpha}{4}$, $p_{h,H} = \frac{V-A}{2}$, and $p_m = \frac{V-A}{2}\left(1-\frac{V}{c}+\frac{(1-\alpha)(V-A)^2}{4c\alpha A}\right)$. We can obtain that:

$\frac{\partial p_{h,HL}}{\partial c} = \frac{3\alpha}{16(1-\alpha)^2}\left(3\alpha - \frac{8(1-\alpha)(V-A)+9c\alpha^2}{\sqrt{16c(1-\alpha)(V-A)+9c^2\alpha^2}}\right) < 0$, because $\left(3\alpha\sqrt{16c(1-\alpha)(V-A)+9c^2\alpha^2}\right)^2 - \left(8(1-\alpha)(V-A)+9c\alpha^2\right)^2 = -64(1-\alpha)^2(V-A)^2 < 0$;

$\frac{\partial p_{h,HL2}}{\partial c} = -\frac{\alpha}{4} < 0$; $\frac{\partial p_{h,H}}{\partial c} = 0$; $\frac{\partial p_m}{\partial c} = \frac{(V-A)}{2}\frac{f_m}{c} > 0$.

A.1.9 Proof of Lemma 5

When the price for each unit of content is r, a light reader's utility maximizes at $n = \frac{v-r}{c}$, so she would consume $\min\{\frac{v-r}{c}, 1\}$. The media firm's profit function is $\pi(r) = (r + A)(I_H + I_L)$, where I_i is the impressions from segment $i = H, L$.

1. If $r \leq 0$, then $I_H = \alpha$ and $I_L = (1-\alpha)\int_0^V \min\{\frac{v-r}{c}, 1\}\frac{1}{V}dv$;

2. If $0 < r \leq V$, then $I_H = \alpha(1 - \frac{r}{V})$ and $I_L = (1-\alpha)\int_r^V \min\{\frac{v-r}{c}, 1\}\frac{1}{V}dv$;

3. If $r > V$, $I_H = I_L = 0$.

The next step is solving the profit maximization problem and identifying the parameter range where the optimal price holds. We first focus on the case where $0 < \alpha \leq 1$. We can obtain that $r = \frac{\sqrt{(A-c)^2+6Vc/\alpha}-A-2c}{3} < 0$ is optimal when $A > \max\{\frac{2V-c\alpha}{2\alpha}, \frac{2c(1+\alpha)-3V\alpha}{2\alpha}\}$, and $r = \frac{V-2A-2c(1-1/\alpha)}{4} < 0$ is optimal when $A \in (\frac{2c(1-\alpha)+V\alpha}{2\alpha}, \frac{2c(1+\alpha)-3V\alpha}{2\alpha}]$. Denote $\overline{A}_{PPU} = \max\{\frac{2V-c\alpha}{2\alpha}, \frac{V\alpha+2c(1-\alpha)}{2\alpha}\}$, we can rewrite the firm's negative PPU strategy as $r = \max\{\frac{\sqrt{(A-c)^2+6Vc/\alpha}-A-2c}{3}, \frac{V-2A-2c(1-1/\alpha)}{4}\}$ when $A > \overline{A}_{PPU}$.

We can also obtain that $r = \frac{2V-2A-c\alpha}{4} > 0$ is optimal when $c \leq \frac{2V}{4-\alpha}$ and $A < \frac{2V-c\alpha}{2}$, or $\frac{2V}{4-\alpha} < c < V$ and $\frac{(4-\alpha)c-2V}{2} < A < \frac{2V-c\alpha}{2}$, and $r = \tilde{r} = \frac{(2V-A)\alpha+2c(1-\alpha)}{3\alpha} - \frac{\sqrt{((V+A)\alpha+2c(1-\alpha)^2)-2c(V+A)\alpha(1-\alpha)}}{3\alpha} > 0$ is optimal when $\frac{2V}{4-\alpha} < c < V$ and $A \leq \frac{(4-\alpha)c-2V}{2}$, or $c \geq V$ and $A < \frac{V(V\alpha+2c(1-\alpha))}{2(V\alpha+c(1-\alpha))}$. Denote $\underline{A}_{PPU} = \max\{\frac{2V-c\alpha}{2}, \frac{V(V\alpha+2c(1-\alpha))}{2(V\alpha+c(1-\alpha))}\}$, we can rewrite the firm's positive PPU strategy as $r = \max\{\frac{2V-2A-c\alpha}{4}, \tilde{r}\} > 0$ when $A < \underline{A}_{PPU}$.

When $\underline{A}_{PPU} \leq A \leq \overline{A}_{PPU}$, $r = 0$ is the most profitable and thus the equilibrium strategy.

When $\alpha = 0$, all consumers are heavy readers. By solving the profit maximization problem, we can obtain $r = \frac{V-A}{2}$ when $A < V$, and $r = 0$ when $A \geq V$. A negative PPU strategy is never used.

Proof of Proposition 5

To find the equilibrium strategy, we compare the profit obtained from the optimal paywall strategy (as described in Lemma 1-3) and that obtained from the optimal PPU strategy (as described in Lemma 5). As mentioned in Section 2.6.1, a PPU strategy with $r = 0$ is equivalent to a no-paywall strategy, while a PPU strategy with $r = \frac{2V-2A-c\alpha}{4}$ earns the same profit as a hard-HL2 paywall. Next, we compare a hard-H paywall with a PPU strategy with $r = \tilde{r}$. Under the condition where a hard-H paywall is chosen (as in Lemma 2), $\pi(f = 0, p_{h,H})$ is always dominated by $\pi(r = \tilde{r})$. Define $\tilde{A}_{PPU}$ as the value of A such that $\pi(f = 0, p = p_{h,HL}) = \pi(r = \tilde{r})$ when $c \leq \underline{c}$, and the value of A such that $\pi(f_m, p_m) = \pi(r = \tilde{r})$ when $c > \underline{c}$. Due to the complexity of equations, we haven't analytically solved $\tilde{A}_{PPU}$, and only numerically show the results in Proposition 5.

A.2 Other Results and Proof

A.2.1 Free Content Provision

Lemma 5. *Under a metered paywall, the amount of free content increases in the ad rate* (A)*, but decreases in light readers' consumption cost* (c)*.*

Proof: Based on Lemma 2, $f_m = \frac{V}{c} - \frac{(1-\alpha)(V-A)^2}{4c\alpha A}$. We can obtain that:

$\frac{\partial f_m}{\partial A} = \frac{(1-\alpha)(V^2-A^2)}{4c\alpha A^2} > 0$ when $A < V$;

$\frac{\partial f_m}{\partial c} = -\frac{f_m}{c} < 0$.

Discussion: First, if an increase in the ad rate does not shift the type of paywall (as in Proposition 1), then a firm that operates a metered paywall would increase the meter limit. Because it wants to increase the consumption of content by providing more free content. Second, when light readers' consumption costs increase, they would consume less and generate fewer ad impressions. Hence, the firm has a smaller incentive to offer free content and would lower the meter limit.

A.2.2 Information Flow

Lemma 6. *Under a metered paywall, the size of information flow increases in the ad rate* (A)*, but decreases in light readers' consumption cost* (c)*.*

Proof: Based on Lemma 2, we know f_m and p_m, so we can calculate the size of information flow, I_m. We can obtain that:

$\frac{\partial I_m}{\partial A} = -(1-\alpha)\frac{1}{V}\frac{\partial p_m}{\partial A} + \alpha\frac{V-f_m c}{V}\frac{\partial f_m}{\partial A} > 0$, where $\frac{\partial p_m}{\partial A} = \frac{1}{2}\left(-(1-f_m) - (V-A)\frac{\partial f_m}{\partial A}\right) < 0$, and $V - f_m c > 0$.

$\frac{\partial I_m}{\partial c} = -(1-\alpha)\frac{(V-A)}{2V}\frac{f_m}{c} - \alpha\left(\frac{f_m^2}{2V} + \frac{f_m(V-f_m c)}{cV}\right) < 0.$

when $A < V$; $\frac{\partial f_m}{\partial c} = -\frac{f_m}{c} < 0.$

Discussion: Under a metered paywall, when the ad rate increases, the firm gives more content for free (Lemma 5) and reduces its subscription price (Lemma 8). A lower price can convert more readers to subscribers, which leads to a larger subscription demand. After becoming subscribers, those consumers read more than they would with a free version, resulting in a larger information flow.

When light readers' consumption costs increase, they have a weaker appetite for content. In response to it, the firm decreases the amount of free content (Lemma 5), and raises the price accordingly (Proposition 4). As a result, the consumption by non-subscribers decreases while that by subscribers remains the same, so the total consumption declines.

Lemma 7. *Under a metered paywall, the size of information flow may decrease in the upper-bound of consumer valuation for content (V). This happens under a low ad rate and a large difference in consumption costs.*

Proof:

$$\frac{\partial I_m}{\partial V} = (1-\alpha)(\underbrace{-\frac{(1-f_m)A}{2V^2}}_{\text{H,cancel subscription (-)}} + \underbrace{\frac{V-A}{2V}\frac{\partial f_m}{\partial V}}_{\text{H,hitting}}) + \alpha(\underbrace{\frac{V-f_m c}{V}\frac{\partial f_m}{\partial V}}_{\text{L,hitting}} + \underbrace{\frac{f_m^2 c}{2V^2}}_{\text{L,not hitting (+)}})$$

Based on Proposition 2, whether the second and the third terms in the above equation are positive depends on the ad rate, as the conditions for $\frac{\partial f_m}{\partial V} < 0$ and $\frac{\partial f_m}{\partial V} \geq 0$. After plugging in $\frac{\partial f_m}{\partial V}$, we can find that $\frac{\partial I_m}{\partial V} < 0$ when $A \in (\underline{A}, V)$ and

$c \in (\max\{\underline{c}, \widetilde{c}_V\}, \infty)$, where $\widetilde{c}_V(V, \alpha, A)$ is the value of c such that $\frac{\partial I^*}{\partial V} = 0$. Denote $\widetilde{A}_V(V, \alpha)$ as the value of A such that $\widetilde{c}_V = \underline{c}$.

Insert Figure A.1 about here

Discussion: I_m decreases in V when the difference in consumption costs is high, as shown in Figure B.1 above. Based on the expression of $\frac{\partial I_m}{\partial V}$ above, an increase in V leads to changes in I_m to four different groups of consumers. The first term is the decline in the information flow due to the heavy readers that cancel the subscription and only read the free version because the subscription gets more expensive when V increases (Lemma 8). The changes in consumption by paywall-hitting heavy readers (of size $\frac{V-A}{2V}$, captured by the second term) and paywall-hitting light readers (of size $\frac{V-f_m c}{V}$, captured by the third term) depend on the change in the meter limit. The last term is the additional consumption from light readers who did not hit the paywall.

When the ad rate is sufficiently low, i.e., $A < \min\{\frac{(1-\alpha)V}{1+\alpha}, \widetilde{A}_V\}$, the information flow decreases in V.[2] As V increases, the firm would decrease the amount of free content, which results in a decline in consumption from all paywall-hitting readers ($\frac{\partial f_m}{\partial V} < 0$, so the second and third terms in $\frac{\partial I_m}{\partial V}$ are negative). Under low ad rates, these two losses are significant: The size of paywall-hitting readers in the H segment

[2]For the case where $A \in \left(\frac{(1-\alpha)V}{1+\alpha}, \widetilde{A}_V\right)$, as V increases, the information flow decreases even though the firm offers more free content. This happens when the light reader segment is small. In this case, the amount of free content is small and the first term in the expression of $\frac{\partial I_m}{\partial V}$ is large. It dominates all the other three effects so that overall the information flow decreases.

($\frac{V-A}{2V}$ in the second term) is large under low ad rates, and so is the size of paywall-hitting light readers ($\frac{V-f_m c}{V}$ in the third term), because the meter limit is low under low ad rates. A low meter limit also implies that the only gain in consumption ($\frac{f_m^2 c}{2V^2}$, the last term) from light readers who do not hit the paywall, is small. As a result, the three losses (the first, the second, and the third terms) in $\frac{\partial I_m}{\partial V}$ dominate the gain (the last term), and the information flow drops.

When the ad rate is high, the change in I_m from a change in V depends on the difference in consumption costs. First, I_m increases in V under a high ad rate and a low difference in consumption costs ("high A, low c"), the latter of which means that even light readers want to read a lot. Under a metered paywall, these conditions imply a large amount of free content. To take advantage of the strong ad market, as V increases, the firm would offer even more content for free ($\frac{\partial f_m}{\partial V} > 0$), so paywall-hitting readers have more to read and contribute more ad revenues. From the expression of $\frac{\partial I_m}{\partial V}$, all terms except the first are gains and the total gain in I_m outweighs the loss from subscribing heavy readers (the first term). Second, I_m decreases in V under a high ad rate and a high difference in consumption costs ("high A, high c"). Compared with the case of "high A, low c," the amount of free content becomes smaller, so the loss from subscribing heavy readers' cancellation ($-\frac{(1-f_m)A}{2V^2}$ in the first term) becomes more prominent, while the gain from light readers who do not hit paywall (captured by the last term, $\frac{f_m^2 c}{2V^2}$) decreases. Further, the gain in ad impressions from all paywall-hitting readers (the second and the third terms) also become smaller under a high difference in consumption costs. This is because light readers have a weaker desire for content under this condition, and the amount of free content is small and becomes less responsive to the upper-bound change ($\frac{\partial f_m}{\partial V}$ becomes smaller when c is high). In summary, once the difference in consumption costs is above a certain threshold, the loss from one group (subscribing heavy readers) will dominate the gains from the other three groups of readers, which leads to a smaller amount of information flow.

A.2.3 Subscription Price

Lemma 8. *Under a metered paywall, the subscription price decreases in the ad rate* (A) *and the size of the light reader segment* (α)*, but increases in the upper-bound of consumer valuation for content* (V)*.*

Proof: Based on Lemma 2, $p_m = \frac{V-A}{2}\left(1 - \frac{V}{c} + \frac{(1-\alpha)(V-A)^2}{4c\alpha A}\right)$. We can obtain that:

$\frac{\partial p_m}{\partial A} = \frac{1}{2}\left(-(1-f_m) - (V-A)\frac{\partial f_m}{\partial A}\right) < 0$ when $A < V$, because $\frac{\partial f_m}{\partial A} > 0$ (Lemma 5);

$\frac{\partial p_m}{\partial \alpha} = -\frac{(V-A)^3}{8c\alpha^2 A} < 0$ when $A < V$;

$\frac{\partial p_m}{\partial V} = \frac{1}{2}\left((1-f_m) - (V-A)\frac{\partial f_m}{\partial V}\right) = \frac{(3+\alpha)(V-A)^2+4\alpha(cA-V^2)}{8c\alpha A} > 0$ when $A \in (\underline{A}, V)$ and $c \in (\underline{c}, \infty)$.

Discussion: Following the discussion of Lemma 5, when the ad rate increases, the firm offers more free content under a metered paywall. The availability of free content weakens the demand for subscriptions, so the firm has to cut the price. Similarly, based on Lemma 4, the amount of free content increases when the proportion of light readers increases. That would also result in a lower subscription price. Finally, when consumers' valuation for content increases, they have a higher WTP for subscription, so the firm increases the price.

A.2.4 Correlation between c and v

A.2.4.1 Positive Correlation between c and v

When $c = kv$, all light readers want to consume $\frac{1}{k}$ unit of content. When $k > 1$ and $f < \frac{1}{k}$, they would subscribe if $u(\frac{1}{k}; v, kv) - p > u(f; v, kv)$. When $k \leq 1$, they would subscribe if $u(1; v, kv) - p > u(f; v, kv)$. Heavy readers' decisions are similar to those in the main model. The firm chooses (p, f) to maximize its profit. However, the model does not generate closed-form solution when being solved with general values of α. In the following, we present the results when α take two special values: $\alpha = \frac{1}{2}$ and $\alpha = 1$.

When $\alpha = \frac{1}{2}$, the firm's equilibrium strategy is summarized in Table B.1, where $\underline{k} = \frac{((3k-2)\sqrt{1+2k}-2k^2+k-2)V}{2(4-k)k}$ when $A \leq \underline{A}$, $= \frac{(\sqrt{2k^3-k^2-k}+2k^2-7k-1)V}{2k^2-10k-1}$ when $\underline{A} < A < \frac{2V}{3}$, and $= 1$ when $\frac{2V}{3} \leq A < V$. The equilibria are very similar to those in the main model.

Insert Table A.1 about here

When $\alpha = 1$, c and v are perfectly positively correlated, and either a hard paywall or no paywall will be chosen. Specifically, the firm launches a hard paywall with $p = \frac{(2-k)V-2A}{4}$ when $k \leq 1$ and $A < \frac{(2-k)V}{2}$, or a hard paywall with $p = \frac{V-2A}{4k}$

when $k > 1$ and $A < \frac{V}{2}$, and no paywall otherwise. When the consumption cost is relatively low ($k < 1$), the area for a no-paywall strategy expands with k. This is because consumers want more content than the total amount of available content (i.e., $\frac{1}{k} > 1$) under this condition. Consequently, an increase in k has no impact on the amount of content consumed (and thus the ad revenue from a no-paywall strategy). On the other hand, an increase in k leads to a drop in consumers' WTP. As the ad revenue remains invariant and the WTP decreases, the firm shifts from a hard paywall to no paywall.

A.2.4.2 Negative Correlation between c and v

When $c = k(V-v)$, light readers want to consume $\frac{k(V-v)}{v}$. The consumer who is indifferent between the free and paid versions has $u(\min\{\frac{v}{k(V-v)}, 1\}; v, k(V-v)) - p = u(f; v, k(V-v))$. Denote the valuation of this consumer as $\hat{v}_L$. Among light readers, those with $v \in [0, \frac{fkV}{1+fk})$ would read for free, those with $v \in [\frac{fkV}{1+fk}, \hat{v}_L]$ would read f, while those with $v \in (\hat{v}_L, 1]$ would pay and read $\min\{\frac{v}{k(V-v)}, 1\}$. The next step is to solve the firm's profit maximization problem. Numerical analysis suggests metered paywall can still emerge in equilibrium (as in Figure B.1(b)).

When $\alpha = 1$, c and v are perfectly negatively correlated, and a metered paywall does not earn more profit than a hard paywall or no paywall. Specifically, when $A \leq \frac{V}{1+k}$, a hard-H paywall (with $p = \frac{V-A}{2}$) is the most profitable. When $\frac{V}{1+k} < A < V$, a hard-H paywall with $p = \frac{(V-A)^2}{2kA}$ and a metered paywall with $f \in \left(0, \frac{A(1+k)-V}{kV}\right]$ and $p = \frac{(1+fk)(V-A)^2}{2kV}$ are equally profitable. If we assume that operating a metered paywall is slightly more costly than a hard paywall (e.g., a very small cost for tracking how much content each consumer already consumed), then a hard paywall is preferred by the firm. When $A \geq V$, the firm chooses no paywall in equilibrium.

A.2.5 Heterogeneity in v

In this subsection, we discuss the role of heterogeneity in v. First, we show the existence of a metered paywall assuming consumers have heterogeneous valuation and homogeneous consumption costs. Second, we prove that the heterogeneity in v alone cannot support a metered paywall in equilibrium.

In the main model, we assume that consumers have heterogeneous $v \sim U[0, V]$ and heterogeneous c, where α proportion of the consumers with $c > 0$ and $1 - \alpha$ proportion of them with $c = 0$. Here, we consider an alternative distribution for valuation: $v = v_L$ for $\delta \in (0, 1)$ proportion of the consumers and $v = v_H > v_L$ for $1 - \delta$ proportion of the consumers. All consumers have the same consumption cost $c \in (v_L, v_H)$. In this setting, consumers with v_L would consume $\min\{1, \frac{v_L}{c}\}$ if they subscribe. They would pay to subscribe when $u(\frac{v_L}{c}; v_L, c) - p \geq u(f; v_L, c)$, where $f \in \left[0, \frac{v_L}{c}\right)$. Consumers with v_H would consume 1 if they subscribe, which happens when $u(1; v_H, c) - p \geq u(f; v_H, c)$, where $f \in [0, 1)$. Solving the firm's profit maximization problem using the approach in Lemma 2, we can obtain the following table.

Insert Table A.2 about here

Figure B.2 illustrates the results with a numerical example. We find that a metered paywall can be the equilibrium strategy when consumers have sufficiently large heterogeneity in valuation, the size of the low valuation segment is small, and the ad rate is sufficiently high.

Insert Figure A.2 about here

Next, we assume that consumers have no cost for consuming content and examine whether heterogeneity in valuation leads to metered paywalls.

Theorem 1 below presents the result. We find that a metered paywall strategy is never more profitable than a hard or a no-paywall strategy. The intuition is the following: when consumers have heterogeneity in valuation and zero consumption cost (only heavy readers exist in the market), they display no heterogeneity in the amount of content desired. In this case, when the firm gives some content for free, it can earn more profits either by reducing the amount of free content and raising the price when ad rates are low, or by giving more content for free when ad rates are high. When the ad rate is at a threshold level, a metered paywall earns the same amount of profit as a hard paywall. In summary, heterogeneity in v alone does not make a metered paywall more profitable than a hard paywall or no paywall.

Theorem 1. *When consumers have heterogeneous valuation for content and no cost of consumption, a metered paywall strategy is never more profitable than a hard paywall or a no-paywall strategy.*

Proof: We separately prove this theorem for two cases: v discretely distributed and continuously distributed.

(i) Suppose $v \in \mathbf{V} = \{v_1, ..., v_n | v_i < v_{i+1}\}$. Assume that $\Pr(v = v_i) = \alpha_i \geq 0$ $(i = 1, 2, ..., n)$ and $\Sigma_{i=1}^{n} \alpha_i = 1$. For each given pair of (f, p) that satisfies $0 \leq f < 1$ and $0 < p \leq v_n(1-f)$, there exists $j \in \{1, 2, ..., n\}$ such that $v_{j-1}(1-f) < p \leq v_j(1-f)$ (define $v_0 = 0$). All consumers with $v_i \geq v_j$ would pay the price p and consume all content, while all consumers with $v_i \leq v_{j-1}$ do not pay and consume f. Let $D_j = \Sigma_{i=j}^{n} \alpha_i$, then the profit function can be written as $\pi(f, p) = pD_j + A(f(1-D_j)+D_j)$ for $p \in (v_{j-1}(1-f), v_j(1-f)]$. This is a increasing function of p on $p \in (v_{j-1}(1-f), v_j(1-f)]$, therefore $\pi(f, v_{j-1}(1-f) < p \leq v_j(1-f)) \leq \pi(f, v_j(1-f))$. We rewrite $\pi(f, v_j(1-f))$ as the following:

$$\begin{aligned} \pi(f, v_j(1-f)) &= v_j(1-f)D_j + A(f(1-D_j) + D_j) \qquad \text{(A.2.1)} \\ &= (v_j + A)D_j + (A(1-D_j) - v_j D_j)f \end{aligned}$$

$\pi(f, v_j(1-f))$ is function of v_j and f. Based on the equation above, it is a linear function of f. If $A > \frac{v_j D_j}{1-D_j}$, then $\pi(f, v_j(1-f))$ increases as $f \to 1$, so it is less profitable than a no-paywall strategy (which earns a profit of $\pi(1, 0) = A$). If $A < \frac{v_j D_j}{1-D_j}$, then $\pi(f, v_j(1-f))$ optimize at $f = 0$, which corresponds to a hard paywall strategy which earns a profit of $\pi(0, v_j) = (v_j + A)D_j$. If $A = \frac{v_j D_j}{1-D_j}$, then f

has no impact on the profit, which means a metered paywall is not more profitable than a hard paywall. To summarize, $\pi(f, v_j(1-f)) \leq \max\{A, (v_j + A)D_j\}$. Note that at $j = 1$, there is $D_j = 1$ so that $\pi(0, v_1) = v_1 + A > A = \pi(1, 0)$. Therefore, a no-paywall strategy is always dominated by a hard paywall with $p = v_1$.

Based on the analysis above, the firm would always choose a hard paywall for given $i \in \{1, 2, ...n\}$. In equilibrium, it chooses a hard paywall with $p^* = v_i$, where $i = \arg\max_{i \in \{1,2,...,n\}}(v_i + A)D_i$.

(ii) Suppose $v \sim G(v)$, which is defined on $v \in [0, \infty)$. The firm's profit is $\pi(f, p) = pD + A(f(1-D) + D)$, where $D = 1 - G(\frac{p}{1-f})$. Denote the valuation of the indifferent consumer as $\bar{v}$, then there is $p = \bar{v}(1-f)$. We can then rewrite the profit function as a function of $\bar{v}$ and f.

$$\begin{aligned}\pi(f, \bar{v}(1-f)) &= \bar{v}(1-f)D(\bar{v}) + A(f(1-D(\bar{v})) + D(\bar{v})) \\ &= (\bar{v} + A)D(\bar{v}) + (A(1-D(\bar{v})) - \bar{v}D(\bar{v}))f\end{aligned} \tag{A.2.2}$$

In the equation above, $D(\bar{v}) = 1 - G(\bar{v})$. Now, we can consider the firm's problem as choosing $\bar{v}$ and f to maximize π. Equation (A.2.2) is a linear function of f. Therefore, for a given $\bar{v}$, it either optimizes at $f = 0$ (if $A < \frac{\bar{v}D(\bar{v})}{1-D(\bar{v})}$), or is dominated by $f = 1$ (if $A > \frac{\bar{v}D(\bar{v})}{1-D(\bar{v})}$). f has no impact on π if $A = \frac{\bar{v}D(\bar{v})}{1-D(\bar{v})}$. In other words, a metered paywall does not earn more profits than a hard paywall or no paywall. Similar to the proof in (i), in equilibrium, the firm chooses a paywall strategy with $p^* = \arg\max_{p \in [0,\infty)}(p + A)D(p)$, where $p > 0$ corresponds to a hard paywall and $p = 0$ corresponds to no paywall.

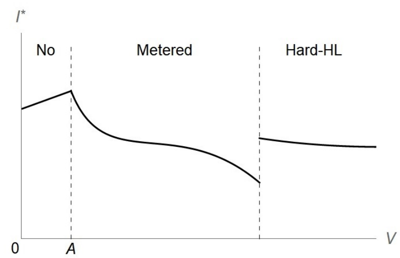
Parameter Range	Paywall Strategy	Source of Subscribers	Readers' Consumption Behavior	Meter Limit (f)	Subscription Price (p)
$A < \frac{2V}{3}$ and $k \leq 1$, or $\frac{2V}{3} \leq A < V$ and $k < \frac{4(V-A)}{2V-A}$	Hard-HL2	H, L	All subscribers read the whole newspaper	0	$\frac{(2-k)V}{4-k} - \frac{A}{2}$
$A < \frac{2V}{3}$ and $1 < k \leq \underline{k}$	Hard-HL	H, L	Some subscribers only read part of the newspaper	0	$\frac{2V-3A}{4k+2}$
$A \leq \underline{A}$ and $k > \underline{k}$	Hard-H	H		0	$\frac{V-A}{2}$
$\underline{A} < A < V$ and $k > \underline{k}$	Metered	H		$\frac{1}{k}$	$\frac{(V-A)(1-1/k)}{2}$
$\frac{2V}{3} \leq A < V$ and $\frac{4(V-A)}{2V-A} \leq k \leq 1$, or $A \geq V$	No-paywall	-		1	0

Note: In this table, we highlight the difference in readers' consumption behavior under hard-HL2 and hard-HL paywalls. The consumption behaviors under other types of paywall are similar to those discussed in the proof of Lemma 1-3.

Table B.1: Equilibrium Strategy When $c = kv$ among Light Readers ($\alpha = \frac{1}{2}$)

Parameter Range	Paywall Strategy	Source of subscribers	Meter Limit (f)	Subscription Price (p)
$\max\{1 - \frac{v_L^2}{(c-v_L)(2v_H-c-v_L)}, 0\} < \delta \leq 1$ and $A > \max\{\frac{c(2v_H-c)(1-\delta)-v_L^2}{2v_L\delta}, 0\}$	Hard	Both	0	$\frac{v_L^2}{2c}$
$\delta \leq 1 - \frac{v_L^2}{c(2v_H-c)}$, and $A \leq \min\{\frac{(2v_H-v_L)(1-\delta)}{2\delta}, \frac{c(2v_H-c)(1-\delta)-v_L^2}{2v_L\delta}\}$	Hard	High value	0	$v_H - \frac{c}{2}$
$v_L < \frac{c(2v_H-c)}{2v_H}$, $\delta \leq 1 - \frac{v_L^2}{(c-v_L)(2v_H-c-v_L)}$, and $A > \frac{(2v_H-v_L)(1-\delta)}{2\delta}$	Metered	High value	$\frac{v_L}{c}$	$(v_H - \frac{c+v_L}{2})(1 - \frac{v_L}{c})$

Table B.2: Equilibrium Strategy Under Heterogeneous v and Homogeneous c ($v_L < c < v_H$)



(a) Information Flow

Figure B.1: Impact of the Upper-bound of Valuation (V) When $c > \max\{\underline{c}, \widetilde{c}_V\}$

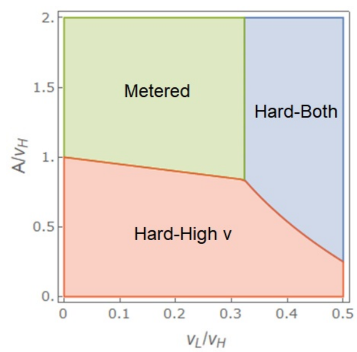


Figure B.2: Equilibrium Strategy Under Heterogeneous v and Homogeneous c ($\delta = 0.5$,$c/v_H = 0.5$)

Appendix B: Proof of Lemmas and Propositions in Chapter 3

B.1 Proof of Lemmas and Propositions

B.1.1 Proof of Lemma 1

To prove Lemma 1, we need to derive Equation (3.2) first. $E_t[u(\mu,t)] = u(\mu,L)(1-\mu) + u(\mu,R)\mu$. When $\mu < \frac{1}{2}$, $E_t[u(\mu,t)] = \mathbf{I}(\mu < \frac{1}{2}, t = L) \cdot v(1-\mu) = v(1-\mu)$; when $\mu > \frac{1}{2}$, $E[u(\mu,t)] = \mathbf{I}(\mu > \frac{1}{2}, t = R) \cdot v\mu = v\mu$; when $\mu = \frac{1}{2}$, $E[u(\mu,t)] = \frac{v}{2}$. Thus we can rewrite $E_t[u(\mu,t)] = v \cdot \max\{1-\mu, \mu\}$. We can also prove $E_t[u(\mu,t)] = v \cdot \Pr(\tau = t|\mu)$. When $\mu < \frac{1}{2}$, there is $\tau = L$, and $E_t[u(\mu,t)] = \mathbf{I}(\mu < \frac{1}{2}, t = L) \cdot v(1-\mu) = v \cdot \mathbf{I}(\tau = L, t = L) \Pr(t = L|\mu) = v \cdot \Pr(t = \tau|\mu)$. Similarly, $E_t[u(\mu,t)] = v \cdot \Pr(t = \tau|\mu)$ when $\mu > \frac{1}{2}$ and $\mu = \frac{1}{2}$, so $E_t[u(\mu,t)] \equiv v \cdot \Pr(t = \tau|\mu)$.

Now consider consumers' WTP for news, which depends on the difference between the expected (gross) utility without and with news consumption (Equation (3.4)). When consumers do not use news, their expected utility is $E_t[u(\mu_0,t)] = v \cdot \max\{1-\mu_0, \mu_0\}$.

When consumers use news, $E_s[E_t[u(\mu_s,t)]] = \Pr(l) \cdot E_t[u(\mu_l,t)] + \Pr(r) \cdot E_t[u(\mu_r,t)]$.

(i) When $\mu_0 < \frac{1}{2}$, consumers believe L to be more likely, a belief change implies that $\mu_r > \frac{1}{2}$. The expected utility from news consumption becomes $\Pr(l) \cdot v(1 - \mu_l) + \Pr(r) \cdot v\mu_r = v(\alpha_L(1-\mu_0) + \alpha_R\mu_0)$. The expected utility from no purchase is $v(1-\mu_0)$. Compare these two expected utility:

$$
\begin{aligned}
v(\alpha_L(1-\mu_0) + \alpha_R\mu_0) - v(1-\mu_0) &= v \Pr(r) \frac{\alpha_R\mu_0 - (1-\alpha_L)(1-\mu_0)}{\Pr(r)} \\
&= v \Pr(r) \left(\mu_r - (1-\mu_r)\right) \qquad \text{(B.1)} \\
&= v \Pr(r) \left(2\mu_r - 1\right) > 0 \text{ when } \mu_r > \frac{1}{2}
\end{aligned}
$$

If no belief change happens after consuming news, consumers expected utility is $\Pr(l) \cdot v(1-\mu_l) + \Pr(r) \cdot v(1-\mu_r) = v(1-\mu_0)$.

(ii) when $\mu_0 > \frac{1}{2}$, a belief change implies that $\mu_l < \frac{1}{2}$. The expected utility from news consumption is also $v(\alpha_L(1-\mu_0)+\alpha_R\mu_0)$, and the expected utility from no purchase is $v\mu_0$. Similarly, $v(\alpha_L(1-\mu_0) + \alpha_R\mu_0) - v\mu_0 = v \Pr(l) (1 - 2\mu_l) > 0$ when $\mu_l < \frac{1}{2}$. If no belief change happens, the expected utility is $\Pr(l) \cdot v\mu_l + \Pr(r) \cdot v\mu_r = v\mu_0$.

(iii) when $\mu_0 = \frac{1}{2}$, a belief change means that $\mu_s \neq \frac{1}{2}$ and the informative constraint implies that $\pi_L + \pi_R > 1$. The expected utility from news consumption is $\frac{v(\alpha_L+\alpha_R)}{2}$. The expected utility from no purchase is $\frac{v}{2}$. The informative constraint suggests that the signal is valuable.

By combining (i)-(iii), we obtain Equation (3.4). Now we can compare the difference between consumers' expected utility with and without consuming news.

(i) when $\mu_0 < \frac{1}{2}$, a belief change implies that $\mu_r > \frac{1}{2}$, which is equivalent to $\mu_0 > \frac{1-\alpha_L}{1-\alpha_L+\alpha_R}$. $WTP = v(\alpha_L - 1 + (1-\alpha_L+\alpha_R)\mu_0)$;

(ii) when $\mu_0 > \frac{1}{2}$, a belief change implies that $\mu_l < \frac{1}{2}$, which leads to $\mu_0 < \frac{\alpha_L}{\alpha_L+1-\alpha_R}$. $WTP = v(\alpha_L - (\alpha_L + 1 - \alpha_R)\mu_0)$;

(iii) when $\mu_0 = \frac{1}{2}$, a belief change implies that $\mu_s \neq \frac{1}{2}$. $WTP = v(\alpha_L - (\alpha_L + 1 - \alpha_R)\mu_0)$.

Combining (i)-(iii) leads to Lemma 1.

B.1.2 Proof of Proposition 1

The cost function $C(\alpha; \mu_0) = c(H(\mu_0) - E_s[H(\mu_s)])$, where $E_s[H(\mu_s)] = \Pr(l)H(\mu_l) + \Pr(r)H(\mu_r)$.

Properties of the cost function: (i) convexity in α_t $(t = L, R)$:

$\frac{\partial^2 C}{\partial \alpha_R^2} = 2c\mu_0^2(1-\mu_0)^2 \left(\frac{\alpha_R^2}{\Pr(r)^3} + \frac{(1-\alpha_R)^2}{\Pr(l)^3} \right) > 0$;

$\frac{\partial^2 C}{\partial \alpha_L^2} = 2c\mu_0^2(1-\mu_0)^2 \left(\frac{\alpha_L^2}{\Pr(l)^3} + \frac{(1-\alpha_L)^2}{\Pr(r)^3} \right) > 0$;

where $\Pr(r) = (1-\alpha_L)(1-\mu_0) + \alpha_R\mu_0$, $\Pr(l) = (1-\alpha_R)\mu_0 + \alpha_L(1-\mu_0)$.

(ii) increase in α_t when $\alpha_L + \alpha_R > 1$:

$\frac{\partial C}{\partial \alpha_R} = c\mu_0^2(1-\mu_0)\gamma$, where $\gamma = \frac{(\alpha_R \Pr(l) + (1-\alpha_R)\Pr(r))(1-\mu_0)(\alpha_L+\alpha_R-1)}{\Pr(r)^2 \Pr(l)^2} > 0$.

$\frac{\partial C}{\partial \alpha_L} = c\mu_0(1-\mu_0)^2\delta$, where $\delta = \frac{(\alpha_L \Pr(r) + (1-\alpha_L)\Pr(l))\mu_0(\alpha_L+\alpha_R-1)}{\Pr(r)^2 \Pr(l)^2} > 0$.

The firm's problem is $\max_\alpha WTP(\alpha, \mu_0) - C(\alpha, \mu_0)$. The first order condition for α_L is: $\frac{\partial WTP-C}{\partial \alpha_L} = c(1-\mu_0)(\frac{v}{c} - \mu_0^2\gamma)$. We can prove that $\mu_0^2\gamma < 1 \Leftrightarrow \mu_0^2 \Pr(l)^2\alpha_R^2 - \mu_0^2 \Pr(r)^2(1-\alpha_R)^2 < \Pr(l)^2 \Pr(r)^2$

$\Leftrightarrow -\mu_0^2 \Pr(r)^2(1-\alpha_R)^2 < \Pr(l)^2(\Pr(r)^2 - \mu_0^2\alpha_R^2)$

$\Leftrightarrow -\mu_0^2 \Pr(r)^2(1-\alpha_R)^2 < \Pr(l)^2(\Pr(r) + \mu_0\alpha_R)(1-\mu_0)(1-\alpha_L)$.

When $c \leq v$, we can obtain that $\frac{\partial(WTP-C)}{\partial \alpha_L} > 0$ because $\mu_0^2\gamma < 1$. Similarly,

$\frac{\partial(WTP-C)}{\partial \alpha_R} > 0$. In this case, the firm would choose $\alpha_L = \alpha_R = 1$. Consumers' posterior belief after news consumption is $\mu_l = 0$ or $\mu_r = 1$. Therefore, as long as consumers' prior belief $\mu_0 \neq 0$ or 1, the firm is able to charge a positive price for news. So news is provided when $0 < \mu_0 < 1$.

When $c > v$, there exist interior solutions to the first order conditions (FOC): $\frac{\partial(WTP-C)}{\partial \alpha_L} = 0$ and $\frac{\partial(WTP-C)}{\partial \alpha_R} = 0$. Note that the firm's strategy needs to ensure a positive profit (i.e., $p = WTP(\alpha, \mu_0) - C(\alpha, \mu_0) > 0$) and the provision of informative news (i.e., $\alpha_L + \alpha_R > 1$). With these constraints, the firm can earn a positive profit only when $\frac{c-v}{2c} < \mu_0 < \frac{v+c}{2c}$. No news is provided when $\mu_0 \leq \frac{c-v}{2c}$ or $\mu_0 \geq \frac{v+c}{2c}$.

B.1.3 Proof of Proposition 2

In the proof of Proposition 1, we show that the firm chooses $\alpha_L = \alpha_R = 1$ when $c \leq v$ and $0 < \mu_0 < 1$. Consumers' expected (gross) utility from news consumption is v, so the firm can charge $p = WTP - v \cdot \max\{1-\mu_0, \mu_0\} = v \cdot \min\{1-\mu_0, \mu_0\}$.

When $c > v$, solving the firm's problem leads to $\alpha_L = \frac{(v+c)(v-c(2\mu_0-1))}{4vc(1-\mu_0)}$ and $\alpha_R = \frac{(v+c)(v+c(2\mu_0-1))}{4vc\mu_0}$. The posterior beliefs caused by the news is $\mu_l = \frac{c-v}{2c}$ or $\mu_r = \frac{v+c}{2c}$. When news is provided ($\frac{c-v}{2c} < \mu_0 < \frac{v+c}{2c}$), it charges $p = WTP - v \cdot \max\{1-\mu_0, \mu_0\} = v(\frac{c+v}{2c} - \max\{\mu_0, 1-\mu_0\})$. For news consumers, the probability of receiving news l is $\Pr(r) = \frac{v+c(2\mu_0-1)}{2v}$; and $\Pr(l) = 1 - \Pr(r)$. When $\mu_0 > \frac{1}{2}$, the degree of exaggeration is $\Pr(r) - \mu_0 = \frac{(c-v)(2\mu_0-1)}{2v}$.

B.1.4 Proof of Proposition 3

$RI = (1-\alpha_L)(1-\mu_0) + (1-\alpha_R)\mu_0 = \frac{c-v}{2c}$.

We can find that $\frac{\partial RI}{\partial c} = \frac{v}{2c^2} > 0$ and $\frac{\partial RI}{\partial \mu_0} = 0$.

B.1.5 Proof of Corollary 1

Consumers only consume news when WTP>0, so the consumers who are indifferent between consuming news or not satisfies $1-\alpha+\frac{p}{v} < \mu_0^i < \alpha - \frac{p}{v}$. We can find that consumers with extreme prior beliefs do not consume news. Define the demand in the presence of heterogeneous beliefs as the following:

$$\begin{aligned} D(\alpha,p) &= \int_{1-\alpha+\frac{p}{v}}^{\alpha-\frac{p}{v}} f(\mu;\theta)d\mu \\ &= \int_{1-\alpha+\frac{p}{v}}^{\frac{1}{2}} (4(1-\theta)\mu_0+\theta)d\mu_0 + \int_{\frac{1}{2}}^{\alpha-\frac{p}{v}} (4(1-\theta)(1-\mu_0)+\theta)d\mu_0 \qquad \text{(B.2)} \\ &= \frac{(v(2\alpha-1)-2p)(v(2\alpha(\theta-1)-2\theta+3)-2p(\theta-1))}{v^2} \end{aligned}$$

If $p=0$, the demand is $D(\alpha,0) = (2\alpha-1)(3-2\theta+2\alpha(\theta-1))$, which increases in α: $\frac{\partial D(\alpha,0)}{\partial \alpha} = 8-6\theta+8\alpha(\theta-1) > 0$ when $\frac{1}{2} < \alpha < 1$ and $0 \le \theta \le 2$.

B.1.6 Proof of Lemma 2

As we calculate in the proof of Corollary 1, the demand is

$$D(\alpha,p) = \frac{(v(2\alpha-1)-2p)(v(2\alpha(\theta-1)-2\theta+3)-2p(\theta-1))}{v^2}$$

. Solving the FOC for p leads to $p^*_{het} = \frac{v(4\alpha(\theta-1)-3\theta+4)-v\sqrt{4\alpha^2(\theta-1)^2-2\alpha(\theta-1)(3\theta-4)+3\theta(\theta-3)+7}}{6(\theta-1)}$.

The next step is finding the optimal accuracy. The news needs to be informative (i.e., $\alpha > \frac{1}{2}$). Solving the FOC for α leads to $\alpha^*_{het} = \frac{v^2(2-\theta)^2+4vc(4-3\theta)-9c^2+(v(2-\theta)-3c)\sqrt{v^2(2-\theta)^2-2vc(2-\theta)+9c^2}}{16vc(1-\theta)}$. This is an interior solution when $\theta < 1$ and $\frac{v(\theta^2-6\theta+6-\theta\sqrt{\theta^2-3\theta+3})}{9(1-\theta)} < c < \frac{v(2-\theta)}{2}$. When $\theta < 1$ and $c \leq \frac{v(\theta^2-6\theta+6-\theta\sqrt{\theta^2-3\theta+3})}{9(1-\theta)}$, the firm makes the highest profit from $\alpha = 1$. When $\theta < 1$ and $c \geq \frac{v(2-\theta)}{2}$, the firm would not make a positive profit from news, so no news is provided.

When $\theta \geq 1$, the second order condition for α, $\frac{\partial^2(D(p^*_{het})p^*_{het}-C)}{\partial\alpha^2} > 0$ always hold when $D(p^*_{het})p^*_{het} - C > 0$, so the firm chooses either $\alpha = 1$ or $\alpha = \frac{1}{2}$. In the latter case ($\alpha = \frac{1}{2}$), the firm would rather not provide news because no one would use news. $\alpha = 1$ is chosen when $c < \frac{v(4\sqrt{(\theta^2-3\theta+3)^3}-4\theta^3+18\theta^2-18\theta)}{27(\theta-1)^2}$. When $c \geq \frac{v(4\sqrt{(\theta^2-3\theta+3)^3}-4\theta^3+18\theta^2-18\theta)}{27(\theta-1)^2}$, no news is provided. We define $\bar{c}= \frac{v(\theta^2-6\theta+6-\theta\sqrt{\theta^2-3\theta+3})}{9(1-\theta)}$ when $\theta < 1$, $\bar{c} = \frac{v}{2}$ when $\theta = 1$, and $\bar{c} = \frac{v(4\sqrt{(\theta^2-3\theta+3)^3}-4\theta^3+18\theta^2-18\theta)}{27(\theta-1)^2}$ when $\theta > 1$.

B.1.7 Proof of Proposition 4

Based on the definition of $\bar{c}$ (in the Proof of Lemma 2), we can find that $\frac{\partial\bar{c}}{\partial\theta} < 0$. Since the firm chooses $\alpha = 1$ when $\theta < 1$ and $c \leq \bar{c}$, or $\theta \geq 1$ and $c < \bar{c}$, we can obtain that $\alpha = 1$ is optimal when $c \leq \min_\theta \bar{c} = \bar{c}(\theta = 2) = \frac{8}{27v}$. Similarly, the firm would not provide news when $\theta < 1$ and $c \geq \frac{v(2-\theta)}{2}$, or $\theta \geq 1$ and $c \geq \bar{c}$. We can also obtain that no provision is optimal when $c \geq \frac{v(2-0)}{2} = v$. To summarize, the optimal level of accuracy is not influenced when $c \leq \frac{8}{27v}$ or $c \geq v$.

When $\frac{8}{27v} < c < v$, as θ increases, the firm may shift from $\alpha = 1$ to not providing news, or first shift from $\alpha = 1$ to $\alpha = \alpha^*$and then to not providing news. From consumers' perspective, not providing news and providing uninformative news (i.e., $\alpha = \frac{1}{2}$) are effectively equivalent, because no one would use news in the latter case. When the firm choose $\alpha = \alpha^*$, the accuracy decreases in θ ($\frac{\partial \alpha^*}{\partial \theta} < 0$). In summary, in this range of costs, the accuracy of the firm weakly decreases in θ.

B.1.8 Proof of Lemma 3

Compared with Lemma 2, the only difference here is the firm's objective function. Therefore, consumers' demand is still

$D = \frac{(v(2\alpha-1)-2p)(v(2\alpha(\theta-1)-2\theta+3)-2p(\theta-1))}{v^2}$. The firm's problem is $\max_{\alpha,p} v \int_i \left(E[k^i] - \frac{1}{2}\right) - C$, where the aggregate level of consumers' probability of being correct, $\int_i \left(E[k^i] - \frac{1}{2}\right) = \left(\alpha - \frac{1}{2}\right) D$, always increases in D. This is because $\alpha > \frac{1}{2}$ when news is provided. When the firm chooses a price to maximize the aggregate level of consumers' probability of being correct, it will chooses the price that maximizes D, which is $p = 0$.

The next step is finding the optimal accuracy. The news needs to be informative (i.e., $\alpha > \frac{1}{2}$). Solving the FOC for α leads to $\alpha^*_{het,k} = \frac{v(7-5\theta)-c}{6v(1-\theta)}$. This is an interior solution when $\theta < 1$ and $v(1+\theta) < c < 2v(2-\theta)$. When $\theta < 1$ and $c \leq v(1+\theta)$, the net expected welfare is maximized at $\alpha = 1$. When $\theta < 1$ and $c \geq 2v(2-\theta)$, providing news always lead to a decrease in the net expected welfare, so no news is provided.

When $\theta \geq 1$, the second order condition for α, $\frac{\partial^2 (v \int_i \left(E[k^i] - \frac{1}{2}\right) - C)}{\partial \alpha^2} > 0$ always hold when $v \int_i \left(E[k^i] - \frac{1}{2}\right) - C > 0$. Under this condition, the firm would choose either $\alpha = 1$ or $\alpha = \frac{1}{2}$. $\alpha = 1$ is chosen when $c < 2v$. When $c \geq 2v$, $\alpha = \frac{1}{2}$ is chosen, which is equivalent to no provision of news.

B.1.9 Proof of Proposition 5

Based on the results of Lemma 3, the firm always chooses $\alpha = 1$ when $c \leq \min(\min_{\theta\in[0,2]}\{v(1+\theta), 2v\}) = v$. Similarly, the firm would not provide news when $c \geq \max(\max_{\theta\in[0,2]}\{2v(2-\theta), 2v\}) = 4v$. To summarize, the optimal level of accuracy is not influenced when $c \leq v$ or $c \geq 4v$.

When $v < c < 2v$, as θ increases, the firm chooses $\alpha = \alpha^*_{het,k}$ when $\theta < \frac{c}{v} - 1$, and $\alpha = 1$ when $\theta \geq \frac{c}{v} - 1$. Under this condition, $\frac{\partial \alpha^*_{het,k}}{\partial \theta} > 0$, so the accuracy weakly increases in θ. To understand the intuition behind the increase of accuracy, we can analyze the firm's incentive for increasing accuracy using the following equation:

$$\frac{d(v\int_i \left(E[k^i] - \frac{1}{2}\right) - C)}{d\alpha} = \underbrace{v}_{\text{incremental accuracy effect}} D(\alpha,0) + v\left(\alpha - \frac{1}{2}\right) \underbrace{\frac{\partial D(\alpha,0)}{\partial \alpha}}_{\text{demand expansion effect}} - \underbrace{\frac{\partial C}{\partial \alpha}}_{\text{cost effct}} \tag{B.3}$$

The interior optimal accuracy, $\alpha^*_{het,k}$, makes Equation (B.3) equal to zero, which means the first and second terms (incremental accuracy and demand expansion effects) cancel out the third term (cost effect). When θ increases, the first term decreases (i.e., $\frac{\partial D(\alpha,0)}{\partial \theta} < 0$), the second term increases when α is sufficiently large ($\frac{\partial^2 D(\alpha,0)}{\partial\theta\partial\alpha} > 0$ when $\alpha > \frac{3}{4}$), while the third term remains constant. Under the condition of $v < c < 2v$, there is $\alpha^*_{het,k} > \frac{3}{4}$. Under such a high accuracy level, an increase in θ will enlarge the demand expansion effect, which dominates the decrease in $D(\alpha, 0)$. As a result, the firm is incentivized to raise its accuracy.

When $2v \leq c < 4v$, as θ increases, the firm shifts from $\alpha = \alpha^*_{het,k}$ to not providing news. Under this condition, $\frac{\partial \alpha^*_{het,k}}{\partial \theta} < 0$ when , so the accuracy weakly decreases in θ.

B.1.10 Proof of Lemma 4

A dual objective firm solves $\max_{\alpha,p} \Omega = \delta \cdot D \cdot p + (1-\delta)v \int_i \left(E[k^i] - \frac{1}{2}\right) - C$. Based on the proof of Lemma 3, $\int_i \left(E[k^i] - \frac{1}{2}\right) = \left(\alpha - \frac{1}{2}\right) D$. The first step is solving for the $p^*(\alpha)$ that maximizes Ω. Solving the FOC for price, we can obtain

$$p^*_{het,d}(\alpha) = \frac{1}{6\delta(\theta-1)}(v((\theta - 1 + \delta(5-4\theta) + 2\alpha(3\delta-1)(\theta-1))$$
$$- \sqrt{\delta^2(\theta-2)^2 + \delta(\theta-2)(\theta-1) + (1+4\alpha^2)(\theta-1)^2 - 2\alpha(\theta-1)((\delta+2)\theta - 2(\delta+1))}).$$

Note that this price needs to satisfy $p^*_{het,d}(\alpha) \geq 0$ and $D(p^*_{het,d}(\alpha)) \geq 0$.

Next, we solve for the optimal level of α. It is worth noting that both $p \geq 0$ and $\alpha \leq 1$ can be binding. We separately analyze the following cases: [i] Neither constraint binds; [ii] $p > 0$ and $\alpha = 1$; [iii] $p = 0$ and $\alpha < 1$; [iv] $p = 0$ and $\alpha = 1$.

1. When neither constraint binds, we solve the FOC $\frac{\partial \Omega(p = p^*_{het,d}(\alpha), \alpha)}{\partial \alpha} = 0$ and obtain $\alpha = \alpha^*_{het,d} = \frac{9c^2\delta^2 - v^2(2-\theta)^2 + 4cv((\delta+2)\theta - 2(\delta+1))}{16cv(\theta-1)} - \frac{(3c\delta + v(\theta-2))\sqrt{9c^2\delta^2 + 2cv\delta(\theta-2) + v^2(\theta-2)^2}}{16cv(1-\theta)}$. The corresponding optimal price is $p^*(\alpha^*_{het,d})$. The solution $(\alpha^*_{het,d}, p^*_{het,d}(\alpha^*_{het,d}))$ is an interior solution when $\theta < 1$, $\delta > \frac{\theta}{1+\theta}$ and $\underline{c}_d < c < \bar{c}_d$, where $\underline{c}_d$ is the value of c such that $\alpha^*_{het,d} = 1$ if $\theta \in [0, 1)$; $\bar{c}_d = \frac{v(1-\delta)(2-\theta)}{2-3\delta}$ if $\theta \in [0, 1)$ and

$\delta \in (\frac{\theta}{1+\theta}, \frac{1}{2})$, and $= \frac{v(2-\theta)}{2\delta}$ if $\theta \in [0, 1)$ and $\delta \in [\frac{1}{2}, 1]$.

2. When $p > 0$ and $\alpha = 1$: in this case, the optimal price is $p^*_{het,d}(1) = \frac{v(1-\delta-\theta+2\delta\theta-\sqrt{\delta^2(\theta-2)^2+(\theta-1)^2-\delta(2-3\theta+\theta^2)})}{6\delta(\theta-1)}$. This is a postive price when $\delta > \frac{\theta}{1+\theta}$. To make the constraint bind, we also need $\frac{\partial\Omega(p^*_{het,d}(\alpha),\alpha)}{\alpha}\Big|_{\alpha=1} > 0$. It is also necessary that providing news leads to a (weak) improvement in Ω: $\Omega(p^*_{het,d}(1), 1) \geq \Omega(0, \frac{1}{2})$. These constraints are satisfied when $c \leq \underline{c}_d$, where $\underline{c}_d = \frac{v}{2\delta}$ if $\theta = 1$, and is the value of c such that $\Omega(p^*_{het,d}(1), 1) = \Omega(0, \frac{1}{2})$ if $\theta \in (1, 2]$.

3. When $p = 0$ and $\alpha < 1$: in this case, we solve the FOC $\frac{\partial\Omega(0,\alpha)}{\partial\alpha} = 0$ and obtain $\alpha = \alpha^{**}_{het,d} = \frac{5}{6} + \frac{2v(1-\delta)-c}{6v(1-\delta)(1-\theta)}$. This is an interior solution (i.e., $\alpha^{**}_{het,d} \in (\frac{1}{2}, 1)$) when $\theta < 1$, $\delta \leq \frac{\theta}{1+\theta}$ and $v(1-\delta)(1+\theta) < c < 2v(1-\delta)(2-\theta)$, or $\theta < 1$, $\frac{\theta}{1+\theta} < \delta < \frac{1}{2}$ and $\bar{c}_d \leq c < 2v(1-\delta)(2-\theta)$.

4. When $p = 0$ and $\alpha = 1$: in this case, we need $\frac{\partial\Omega(0,\alpha)}{\alpha}\Big|_{\alpha=1} > 0$ and $\Omega(0, 1) \geq \Omega(0, \frac{1}{2})$, which are satisfied when $\delta \leq \frac{\theta}{1+\theta}$ and $c \leq \min\{v(1-\delta)(1+\theta), 2v(1-\delta)\}$.

For the parameter range that is not covered by any of the cases above, no news is provided.

B.1.11 Proof of Proposition 6

Consider the conditions for each case in Lemma 4 and focus on the case where $p = 0$, we can find that $p = 0$ is never chosen when $\delta > \frac{2}{3}$. When $0 \leq \delta \leq \frac{2}{3}$, the firm chooses $p = 0$ only when $c > \bar{c}_d$ (i.e., choosing $(p = 0, \alpha = \alpha^{**}_{het,d})$) or when

$\theta \geq \frac{\delta}{1-\delta}$ (i.e., choosing $(p = 0, \alpha = 1)$).

B.1.12 Proof of Proposition 7

Consider the area where $(p^*_{het,d}(\alpha^*_{het,d}), \alpha^*_{het,d})$ is the equilibrium strategy. We can find $\frac{\partial \alpha^*_{het,d}}{\partial \theta} < 0$ always hold when $\delta \geq \frac{3}{4}$. When $\delta < \frac{3}{4}$, there is $\frac{\partial \alpha^*_{het,d}}{\partial \theta} < 0$ when $\theta < \tilde{\theta}$, and $\frac{\partial \alpha^*_{het,d}}{\partial \theta} > 0$ when $\theta > \tilde{\theta}$.

B.2 Other Results and Proof

B.2.1 Asymmetric Value for States

In many cases, consumers may put more value on being correct about one state than the other. For example, having the correct understanding about the existence of a high pollution (state R) is more important than that of a low pollution (state L). In other words, being correct about the existence of a high pollution is more valuable ($v_R > v_L$). In the main model, we assume $v_L = v_R = v$. In this subsection, we relax this assumption and analyze the game. The following lemma summarizes the conditions for news provision and equilibrium accuracy levels under this situation.

Lemma 5. *The conditions for news provision and the corresponding levels of accuracy are summarized in Table B.1, where* $\tilde{c} = \max\{\frac{(v_L+v_R)^2}{4v_L}, \frac{(v_L+v_R)^2}{4v_R}\}$, $\phi = \frac{v_L}{v_L+v_R}$ *and* $\gamma = \frac{c}{v_L}$. *The price is always* $p = v_L\alpha_L(1-\mu_0) + v_R\alpha_R\mu_0 -$

$$\begin{cases} v_L(1-\mu_0) & \textit{if } \mu_0 \leq \frac{1}{2} \\ v_R\mu_0 & \textit{if } \mu_0 > \frac{1}{2} \end{cases}.$$

Insert Table B.1 about here

The first two columns of Table B.1 shows the condition for news provision. The results under low ($c \leq \min\{v_L, v_R\}$) or high ($c > \tilde{c}$) costs are similar to those in Proposition 1. The result under moderate costs ($\min\{v_L, v_R\} < c \leq \tilde{c}$) is more interesting. Under this condition, news provision happens when consumers do not have sufficiently strong belief on the more valuable state. In the main model, consumers see little value in news if they have extreme beliefs. When knowing the two states have different values, an extreme belief about the more valuable state will discourage learning more than an extreme belief about the less valuable state. Consumers an extreme belief about the more valuable state expect a smaller loss from being wrong about the true state, so they have less incentive to learn. As a result, news provision is less likely when consumers have an extreme belief about the more valuable state.

The third column of Table B.1 presents the equilibrium levels of accuracy, which depend on both the cost and the prior belief. To compare these results with those in the main model, we illustrate the equilibrium accuracy under $c > \tilde{c}$ using a numerical example, as in Figure B.1. Figure B.1(a) shows that news is provided when consumers believe the less valuable state (L) to be more likely, or believe the more valuable state (R) to be a little more likely. Compared with the case of $v_L = v_R = 1$, the increase in v_R makes news provision possible when consumers have relatively extreme beliefs in L. As learning about R becomes more valuable, consumers have stronger demands for information if they have relatively extreme beliefs in L, but weaker demand for it if they have relatively extreme beliefs in R. In response, the firm increases the accuracy of reporting R and decreases that of reporting L. Consequently, the news report will overproportionally presents the more valuable state. However, if consumers have sufficiently extreme beliefs about the less valuable state (e.g., low pollution), the news report may downplay the likelihood of the more valuable state (e.g., high pollution). In other words, consumers' extreme beliefs make the media firm underreport the relatively more important state, which can potentially result in serious consequences.

Insert Figure B.1 about here

B.2.2 Ex Post Polarization and the Impact of News

In Section 3.5.1, we use θ to measure the level of polarization in consumers' prior beliefs. If we generate a measure for the ex post (expected) level of polarization, we can quantify the impact of news on the level of polarization. In this subsection, we extend the analysis in Section 3.5.1 to find the impact of news on polarization. Since the variance of consumers' posterior beliefs is mathematically complex, we use another measure to reflect the level of polarization among consumers' beliefs. Under the prior, the shape of consumers' belief distribution is symmetric around $\frac{1}{2}$. By measuring the average belief of consumers who believe L to be more likely (i.e., consumers whose prior belief $\mu_0^i < \frac{1}{2}$) and examining how it changes with news consumption, we can also find the impact of news on the level of polarization.

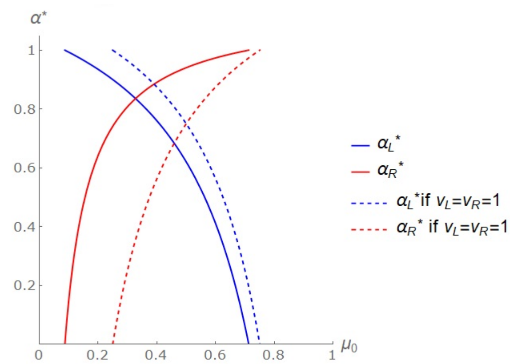
Under the prior, the average prior belief of consumers whose prior belief $\mu_0^i < \frac{1}{2}$ is $2\int_0^{1/2} \mu(\theta + (4 - 4\theta)\mu)d\mu = \frac{1}{3} - \frac{\theta}{12}$. The integral is multiplied by 2 because the probability density function needs an adjustment when calculating the average of half of the market. To calculate the average posterior belief, we need to consider consumers' news consumption behaviors as well as the content of news. Based on Corollary 1, only the consumers with $1 - \alpha + \frac{p}{v} < \mu_0^i < \alpha - \frac{p}{v}$ consume news. If $s = l$, the average posterior belief of consumers whose prior belief $\mu_0^i < \frac{1}{2}$ is $\Delta_l = 2\int_0^{1-\alpha_{het}^* + \frac{p_{het}^*(\alpha_{het}^*)}{v}} \mu(\theta + (4 - 4\theta)\mu)d\mu + 2\int_{1-\alpha_{het}^* + \frac{p_{het}^*(\alpha_{het}^*)}{v}}^{1/2} \mu_l(\theta + (4 - 4\theta)\mu)d\mu$. The counterpart under $s = r$, Δ_r, can be calculated in a similar way, with μ_l being replaced by μ_r. We use the expected value, $\Delta = \Pr(s = l)\Delta_l + \Pr(s = r)\Delta_r$, to measure the ex post level of polarization. The following figure illustrates the results using three numerical examples.

Insert Figure B.2 about here

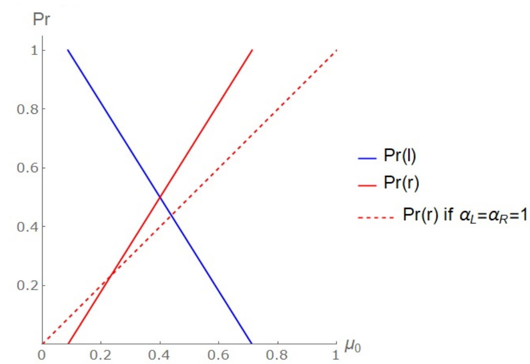
In the figure above, the horizontal axis is θ, which measures the level of polarization under the prior. The black line denoted by "$c > 1$" represents the average posterior belief of consumers whose prior belief $\mu_0^i < \frac{1}{2}$ when no news is provided, which is the equilibrium outcome under high costs (i.e., $c > 1$). In this case, their average posterior belief is the same as their average prior belief. As θ increases, the average belief becomes closer to 0, which corresponds to a relatively high level of polarization. The blue line represents the case where $c = 0.7$, where news is provided under low polarization (i.e., $\theta < 0.6$). The provision and consumption of news make the average posterior belief closer to $\frac{1}{2}$. The blue dashed line represents the case where $c = 0.3$, where news is always provided with full accuracy. Compared this case with $c = 0.7$, we can find that the average posterior belief is closer to $\frac{1}{2}$ with when news is more accurate. Moreover, as θ increases, the average posterior belief under fully accurate news also gets closer to 0, but the distance between the line with fully accurate news and that without news becomes larger under high values of θ. In other words, as θ increases, fully accurate news has more impact on polarization when the level of polarization under the prior is sufficiently high. This is because under higher levels of polarization, the firm lowers down the price to attract consumers with more extreme prior beliefs, which resulting in having more impact on the average posterior belief.

Cost Level	Prior Belief	Accuracy
$c \leq \min\{v_L, v_R\}$	$0 < \mu_0 < 1$	$\alpha_L = \alpha_R = 1$
$\min\{v_L, v_R\} < c \leq \tilde{c}$	$0 < \mu_0 < \sqrt{v_L/c}$ if $v_L < v_R$, or $1 - \sqrt{v_R/c} < \mu_0 < 1$ if $v_L > v_R$,	$\alpha_L = \frac{1-\mu_0\sqrt{c/v_L}}{1-\mu_0}$, $\alpha_R = 1$ if $v_L < v_R$, or $\alpha_L = 1$, $\alpha_R = \frac{1-(1-\mu_0)\sqrt{c/v_R}}{\mu_0}$ if $v_L > v_R$
$c > \tilde{c}$	$\frac{v_L}{v_L+v_R} - \frac{v_L+v_R}{4c} < \mu_0 < \frac{v_L}{v_L+v_R} + \frac{v_L+v_R}{4c}$	$\alpha_L = \frac{1+4\gamma\phi(1-\phi)}{2(1-\mu_0)}\left(\frac{1}{4\gamma\phi} + \phi - \mu_0\right)$, $\alpha_R = \frac{1+4\gamma\phi^2}{2\mu_0}\left(\frac{1}{4\gamma\phi} - \phi + \mu_0\right)$

Table B.1: Conditions for News Provision and Equilibrium Accuracy



(a) Reporting Accuracy



(b) Reported Content

Figure B.1: Reporting Accuracy and Reported Content in Equilirbium ($v_L = 1$, $v_R = 1.5$, $c = 2$)

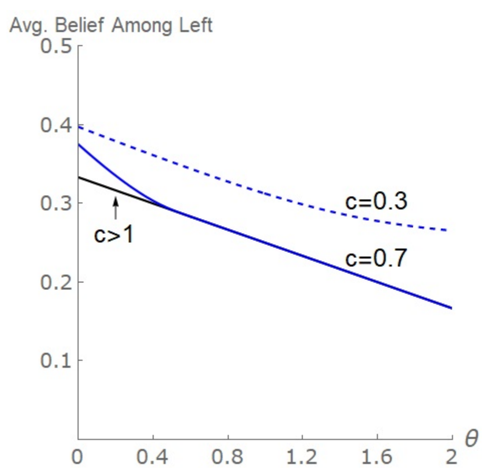


Figure B.2: Expected Ex Post Polarization ($v = 1$)

Bibliography

[1] Abramowitz AI, Saunders KL (2008) Is polarization a myth? *The Journal of Politics* 70(2):542–555.

[2] Acemoglu D, Chernozhukov V, Yildiz M (2016) Fragility of asymptotic agreement under Bayesian learning. *Theoretical Economics* 11(1):187–225.

[3] Adams R (2011) New York Times readies pay wall; paper will charge for bundled digital service, allow some free access. *The Wall Street Journal* (January 24) https://www.wsj.com/articles/SB10001424052748704213404576100033883758352

[4] Alpert LI (2019) The Atlantic, propped up by Laurene Powell Jobs, charts new course. *The Wall Street Journal* https://www.wsj.com/articles/the-atlantic-propped-up-by-steve-jobss-widow-charts-new-course-11566651600.

[5] Amaldoss W, Du J, Shin W (2021) Media platforms' content provision strategies and sources of profits. *Marketing Science* 40(3):527–547.

[6] Anderson SP, Coate S (2005) Market provision of broadcasting: A welfare analysis. *The Review of Economic Studies* 72(4):947–972.

[7] Anderson SP, Jullien B (2015) The advertising-financed business model in two-sided media markets. Anderson SP, Waldfogel J, Stromberg D, eds. *Handbook of Media Economics*. (Elsevier), 41–90.

[8] Appel G, Libai B, Muller E, Shachar R (2020) On the monetization of mobile apps. *International Journal of Research in Marketing* 37(1):93-107.

[9] Aral S, Dhillon PS (2021) Digital paywall design: Implications for content demand and subscriptions. *Management Science* 67(4):2381–2402.

[10] Armstrong M (2006a) Competition in two-sided markets. *The RAND Journal of Economics* 37(3):668–691.

[11] Armstrong M (2006b) Recent developments in the economics of price discrimination. Blundell R, Newey W, Persson T, eds. *Advances in Economics and Econometrics: Theory and Applications*. (Cambridge University Press), 97–141.

[12] Arrese Á (2016) From gratis to paywalls. *Journalism Studies* 17(8):1051–1067.

[13] Augenblick N, Rabin M (2021) Belief movement, uncertainty reduction, and rational updating. *The Quarterly Journal of Economics* 136(2):933–985.

[14] Baliga S, Hanany E, Klibanoff P (2013) Polarization and ambiguity. *American Economic Review* 103(7):3071–3083.

[15] Baron DP (2006) Persistent media bias. *Journal of Public Economics* 90(1–2):1–36.

[16] Bergemann D, Bonatti A (2019) Markets for information: An introduction. *Annual Review of Economics* 11:85–107.

[17] Bergemann D, Bonatti A, Smolin A (2018) The design and price of information. *American Economic Review* 108(1):1–48.

[18] Bergemann D, Morris S (2019) Information design: A unified perspective. *Journal of Economic Literature* 57(1):44–95.

[19] Berman R, Katona Z (2020) Curation algorithms and filter bubbles in social networks. *Marketing Science* 39(2):296–316.

[20] Blackwell D (1953) Equivalent comparisons of experiments. *The Annals of Mathematical Statistics* 24(2):265–272.

[21] Boxell L, Gentzkow M, Shapiro JM (2017) Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences* 114(40):10612–10617.

[22] Bureau of Labor Statistics (2019) *American time use survey*

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

[24] Chatterjee P, Zhou B (2021) Sponsored content advertising in a two-sided market. *Management Science* 67(12):7560–7574.

[25] Che YK, Mierendorff K (2019) Optimal dynamic allocation of attention. *American Economic Review* 109(8):2993–3029.

[26] Chellappa RK, Mehra A (2018) Cost drivers of versioning: Pricing and product line strategies for information goods. *Management Science* 64(5):2164–2180.

[27] Chen Y, Xie J (2007) Cross-market network effect with asymmetric customer loyalty: Implications for competitive advantage. *Marketing Science* 26(1):52–66.

[28] Chiou L, Tucker C (2013) Paywalls and the demand for news. *Information Economics and Policy* 25(2):61–69.

[29] City of Flint (2015) *City of Flint Water System Questions & Answers.*

[30] Cook JE, Attari SZ (2012) Paying for what was free: Lessons from the New York Times paywall. *Cyberpsychology, Behavior, and Social Networking* 15(12):682–687.

[31] Croteau D, Hoynes W (2006) *The Business of Media: Corporate Media and the Public Interest* (Pine Forge Press).

[32] Diddi A, LaRose R (2006) Getting hooked on news: Uses and gratifications and the formation of news habits among college students in an Internet environment. *Journal of Broadcasting & Electronic Media* 50(2):193–210.

[33] DiMaggio P, Evans J, Bryson B (1996) Have American's social attitudes become more polarized? *American Journal of Sociology* 102(3):690–755.

[34] Dixit AK, Weibull JW (2007) Political polarization. *Proceedings of the National Academy of Sciences* 104(18):7351–7356.

[35] Doctor K (2017) LA Times Tops 100,000 in Digital Subscriptions. https://www.thestreet.com/opinion/l-a-times-tops-100-000-in-digital-subscriptions-14315543.

[36] Ely J, Frankel A, Kamenica E (2015) Suspense and surprise. *Journal of Political Economy* 123(1):215–260.

[37] Erb R (2015) Who wants to drink Flint's water? *Detroit Free Press* https://www.freep.com/story/news/local/michigan/2015/01/22/water-woes-latest-hit-flint/22193291/.

[38] Felton R (2014) Flint residents raise concerns over discolored water. *Detroit Metro Times* https://www.metrotimes.com/detroit/flint-residents-raise-concerns-over-discolored-water/Content?oid=2231724.

[39] Fiorina MP, Abrams SJ, Pope JC (2005) *Culture War: The Myth of A Polarized America* (Longman).

[40] Flaxman S, Goel S, Rao JM (2016) Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80(S1):298–320.

[41] Fleisher R, Bond JR (2004) The shrinking middle in the US Congress. *British Journal of Political Science* 34(3):429–451.

[42] Gabszewicz JJ, Laussel D, Sonnac N (2001) Press advertising and the ascent of the 'Pensée Unique.' *European Economic Review* 45(4–6):641–651.

[43] Gal-Or E, Dukes A (2003) Minimum differentiation in commercial media markets. *Journal of Economics & Management Strategy* 12(3):291–325.

[44] Gal-Or E, Geylani T, Yildirim P (2012) The impact of advertising on media bias. *Journal of Marketing Research* 49(1):92–99.

[45] Gentzkow M, Kamenica E (2014) Costly persuasion. *American Economic Review* 104(5):457–62.

[46] Gentzkow M, Shapiro JM (2006) Media bias and reputation. *Journal of Political Economy* 114(2):280–316.

[47] Glaeser EL, Sunstein CR (2009) Extremism and social learning. *Journal of Legal Analysis* 1(1):263–324.

[48] Godes D (2020) Media bias in the presence of feedback. Working Paper.

[49] Godes D, Ofek E, Sarvary M (2009) Content vs. advertising: The impact of competition on media firm strategy. *Marketing Science* 28(1):20–35.

[50] Grieco E, Sumida N, Fedeli S (2018) About a third of large U.S. newspapers have suffered layoffs since 2017. https://www.pewresearch.org/fact-tank/2018/07/23/about-a-third-of-large-u-s-newspapers-have-suffered-layoffs-since-2017/

[51] Guo H, Zhao X, Hao L, Liu D (2019) Economic analysis of reward advertising. *Production and Operations Management*.

[52] Ha A (2017) Medium now lets anyone publish behind its paywall. https://techcrunch.com/2017/10/10/medium-opens-up-its-partner-program-allowing-anyone-to-publish-behind-its-paywall/

[53] Halbheer D, Stahl F, Koenigsberg O, Lehmann DR (2014) Choosing a digital content strategy: How much should be free? *International Journal of Research in Marketing* 31(2):192–206.

[54] Han SP, Park S, Oh W (2016) Mobile app analytics: A multiple discrete-continuous choice framework. *MIS Quarterly* 40(4):983–1008.

[55] Herrnstein RJ, Prelec D (1991) Melioration: A theory of distributed choice. *Journal of Economic Perspectives* 5(3):137–156.

[56] Holden E, Fonger R, Glenza J (2019) Revealed: water company and city officials knew about Flint poison risk. *The Guardian* (December 10) https://www.theguardian.com/us-news/2019/dec/10/water-company-city-officials-knew-flint-lead-risk-emails-michigan-tap-water.

[57] Huang E (2009) The causes of youths' low news consumption and strategies for making youths happy news consumers. *Convergence* 15(1):105–122.

[58] Iyengar S, Hahn KS (2009) Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication* 59(1):19–39.

[59] Iyengar S, Sood G, Lelkes Y (2012) Affect, not ideology, a social identity perspective on polarization. *Public Opinion Quarterly* 76(3):405–431.

[60] Iyengar S, Westwood SJ (2015) Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science* 59(3):690–707.

[61] Iyer G, Yoganarasimhan H (2021) Strategic Polarization in Group Interactions. *Journal of Marketing Research* 58(4).

[62] Jacobson GC (2000) Party polarization in national politics: The electoral connection. *Polarized Politics: Congress and the President in A Partisan Era.* 17–18.

[63] Jerath K, Ren Q (2021) Consumer rational (in) attention to favorable and unfavorable product information, and firm information design. *Journal of Marketing Research* 58(2):343–362.

[64] Jones T 'The greatest crisis in American journalism is the death of local news,' plus a bleak outlook in Reading and Seattle's tips for converting readers to subscribers.
https://www.poynter.org/newsletters/2019/the-greatest-crisis-in-american-journalism-is-the-death-of-local-news-plus-a-bleak-outlook-in-reading-and-seattles-tips-for-converting-readers-to-subscribers/

[65] Kamada Y, Öry A (2020) Contracting with word-of-mouth management. *Management Science* 66(11):5094–5107.

[66] Kamenica E (2019) Bayesian persuasion and information design. *Annual Review of Economics* 11:249–272.

[67] Kamenica E, Gentzkow M (2011) Bayesian persuasion. *American Economic Review* 101(6):2590–2615.

[68] Kennedy M (2016) Lead-laced water in Flint: A step-by-step look at the makings of a crisis.
https://www.npr.org/sections/thetwo-way/2016/04/20/465545378/lead-laced-water-in-flint-a-step-by-step-look-at-the-makings-of-a-crisis.

[69] Khazan O (2019) The trouble with America's water. The Atlantic (September 11)
https://www.theatlantic.com/health/archive/2019/09/millions-american-homes-have-lead-water/597826/.

[70] Kind HJ, Nilssen T, Sorgard L (2009) Business models for media firms: Does competition matter for how they raise revenue? *Marketing Science* 28(6):1112–1128.

[71] Kramer J (2009) Appreciating (and counting) loyal readers. *MinnPost* https://www.minnpost.com/inside-minnpost/2009/10/appreciating-and-counting-loyal-readers/.

[72] Kumar V, Anand B, Gupta S, Oberholzer-Gee F (2013) *The New York Times paywall*

[73] Lambrecht A, Goldfarb A, Bonatti A, Ghose A, Goldstein DG, Lewis R, Rao A, Sahni N, Yao S (2014) How do firms make money selling digital goods online? *Marketing Letters* 25(3):331–341.

[74] Lambrecht A, Misra K (2017) Fee or free: when should firms charge for online content? *Management Science* 63(4):1150–1165.

[75] Lee AM (2013) News audiences revisited: Theorizing the link between audience motivations and news consumption. *Journal of Broadcasting & Electronic Media* 57(3):300–317.

[76] Levendusky M (2013) Partisan media exposure and attitudes toward the opposition. *Political Communication* 30(4):565–581.

[77] Levendusky MS (2013) Why do partisan media polarize viewers? *American Journal of Political Science* 57(3):611–623.

[78] Levy R (2021) Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review* 111(3):831–70.

[79] Lewis A (2018) *Metered pay models comprise bulk of newspaper website pay models* (Missouri School of Journalism, University of Missouri).

[80] Lewis WA (1941) The two-part tariff. *Economica* 8(31):249–270.

[81] Li H, Jain S, Kannan P (2019) Optimal design of free samples for digital products and services. *Journal of Marketing Research* 56(3):419–438.

[82] Lin S (2020) Two-sided price discrimination by media platforms. *Marketing Science* 39(2):317–338.

[83] Liu Y, Putler DS, Weinberg CB (2004) Is having more channels really better? A model of competition among commercial television broadcasters. *Marketing Science* 23(1):120–133.

[84] Maskin E, Riley J (1984) Monopoly with incomplete information. *The RAND Journal of Economics* 15(2):171–196.

[85] Media Insight Project (2014) *The personal news cycle: How Americans choose to get their news* (American Press Institute and AP-NORC Center for Public Affairs Research).

[86] Media Insight Project (2017) *Paying for news: Why people subscribe and what It says about the future of journalism* (American Press Institute and AP-NORC Center for Public Affairs Research).

[87] Mele N, Skibinski M, Spector M (2019) *Digital pay-meter playbook* (Shorenstein Center and Lenfest Institute).

[88] Moorthy KS (1984) Market segmentation, self-selection, and product line design. *Marketing Science* 3(4):288–307.

[89] Mullainathan S, Shleifer A (2005) The market for news. *American Economic Review* 95(4):1031–1053.

[90] Mussa M, Rosen S (1978) Monopoly and product quality. *Journal of Economic Theory* 18(2):301–317.

[91] Myerson RB (1982) Optimal coordination mechanisms in generalized principal–agent problems. *Journal of Mathematical Economics* 10(1):67–81.

[92] Nelson LD, Meyvis T, Galak J (2009) Enhancing the television-viewing experience through commercial interruptions. *Journal of Consumer Research* 36(2):160–172.

[93] Niculescu MF, Wu DJ (2014) Economics of free under perpetual licensing: Implications for the software industry. *Information Systems Research* 25(1):173–199.

[94] Nielsen M, Stewart RT (2020) Persistent disagreement and polarization in a Bayesian setting. *The British Journal for the Philosophy of Science* 72(1).

[95] Oh H, Animesh A, Pinsonneault A (2016) Free versus for-a-fee: The impact of a paywall. *MIS Quarterly* 40(1):31–56.

[96] Papanastasiou Y (2020) Fake news propagation and detection: A sequential model. *Management Science* 66(5):1826–1846.

[97] Pattabhiramaiah A, Sriram S, Manchanda P (2019) Paywalls: Monetizing online content. *Journal of Marketing* 83(2):19–36.

[98] Pew Research Center (2011) *The state of the news media 2011: An annual report on American journalism*. (Pew Research Center, Washington, DC).

[99] Pew Research Center (2017) *The partisan divide on political values grows even wider*. (Pew Research Center, Washington, DC).

[100] Prasad A, Mahajan V, Bronnenberg B (2003) Advertising versus pay-per-view in electronic media. *International Journal of Research in Marketing* 20(1):13–30.

[101] Rabin M, Schrag JL (1999) First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics* 114(1):37–82.

[102] Ratner RK, Kahn BE, Kahneman D (1999) Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research* 26(1):1–15.

[103] Redden JP (2008) Reducing satiation: The role of categorization level. *Journal of Consumer Research* 34(5):624–634.

[104] Rochet JC, Tirole J (2003) Platform competition in two-sided markets. *Journal of the European Economic Association* 1(4):990–1029.

[105] Rogowski JC, Sutherland JL (2016) How ideology fuels affective polarization. *Political Behavior* 38(2):485–508.

[106] Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27(3):379–423.

[107] Shi Z, Zhang K, Srinivasan K (2019) Freemium as an optimal strategy for market dominant firms. *Marketing Science* 38(1):150–169.

[108] Shoemaker PJ (1996) Hardwired for news: Using biological and cultural evolution to explain the surveillance function. *Journal of Communication* 46(3):32–47.

[109] Simon FM, Graves L (2019) *Pay models for online news in the US and Europe: 2019 update* (Reuters Institute and University of Oxford).

[110] Sims CA (2003) Implications of rational inattention. *Journal of Monetary Economics* 50(3):665–690.

[111] Smith M, Bosman J, Davey M (2019) Flint's water crisis started 5 years ago. It's not over. *The New York Times* (April 25) https://www.nytimes.com/2019/04/25/us/flint-water-crisis.html.

[112] Spence M, Owen B (1977) Television programming, monopolistic competition, and welfare. *The Quarterly Journal of Economics* 91(1):103–126.

[113] Stampler L (2019) Condé Nast will put every single publication behind a paywall by year's end. https://fortune.com/2019/01/23/conde-nast-will-put-every-single-publication-behind-a-paywall-by-years-end/.

[114] Steiner J, Stewart C, Matějka F (2017) Rational Inattention Dynamics: Inertia and Delay in Decision-Making. *Econometrica* 85(2):521–553.

[115] Steiner PO (1952) Program patterns and preferences, and the workability of competition in radio broadcasting. *The Quarterly Journal of Economics* 66(2):194–223.

[116] Stulberg A (2017) Testing news paywalls: Which are leaky, and which are airtight?

[117] Tåg J (2009) Paying to remove advertisements. *Information Economics and Policy* 21(4):245–252.

[118] Veolia North America (2015) *Water Quality Report* https://www.cityofflint.com/wp-content/uploads/Veolia-REPORT-Flint-Water-Quality-201503121.pdf

[119] Walker M U.S. newsroom employment has fallen 26% since 2008. https://www.pewresearch.org/fact-tank/2021/07/13/u-s-newsroom-employment-has-fallen-26-since-2008/.

[120] Wall Street Journal (1996) The Wall Street Journal Launches Interactive Edition. https://www.wsj.com/articles/SB849581696318836500.

[121] Watson A (2019) Average time spent per day with digital media in the United States from 2011 to 2018. https://www.statista.com/statistics/262340/daily-time-spent-with-digital-media-according-to-us-consumsers/.

[122] Weyl EG (2010) A price theory of multi-sided platforms. *American Economic Review* 100(4):1642–1672.

[123] Webster SW, Abramowitz AI (2017) The ideological foundations of affective polarization in the US electorate. *American Politics Research* 45(4):621–647.

[124] Wood T, Porter E (2019) The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* 41(1):135–163.

[125] Woodford M (2009) Information-constrained state-dependent pricing. *Journal of Monetary Economics* 56:S100–S124.

[126] Xiang Y, Sarvary M (2007) News consumption and media bias. *Marketing Science* 26(5):611–628.

[127] Xiang Y, Soberman DA (2013) Consumer favorites and the design of news. *Management Science* 60(1):188–205.

[128] Xiang Y, Soberman DA (2011) Preview provision under competition. *Marketing Science* 30(1):149–169.

[129] Yildirim P, Gal-Or E, Geylani T (2013) User-generated content and bias in news media. *Management Science* 59(12):2655–2666.

[130] Zhu Y, Dukes A (2015) Selective reporting of factual content by commercial media. *Journal of Marketing Research* 52(1):56–76.

[131] Zimper A, Ludwig A (2009) On attitude polarization under Bayesian learning with non-additive beliefs. *Journal of Risk and Uncertainty* 39(2):181–212.