



Victor Aliende da Matta

The Market for Internet News Distribution

Dissertação de Mestrado

Thesis presented to the Programa de Pós-graduação em Economia, do Departamento de Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Timo Hiller

Co-advisor: Prof. Leonardo Rezende

Rio de Janeiro
March 2022



Victor Aliende da Matta

The Market for Internet News Distribution

Thesis presented to the Programa de Pós-graduação em Economia da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the Examination Committee:

Prof. Timo Hiller

Advisor

Departamento de Economia – PUC-Rio

Prof. Leonardo Rezende

Co-advisor

Departamento de Economia – PUC-Rio

Prof. Vinicius do Nascimento Carrasco

Departamento de Economia – PUC-Rio

Prof. Lucas Jóver Maestri

FGV/EPGE

Rio de Janeiro, March the 4th, 2022

All rights reserved.

Victor Aliende da Matta

B.Sc., Computer Science, Universidade de São Paulo, 2019

Bibliographic data

Aliende da Matta, Victor

The Market for Internet News Distribution / Victor Aliende da Matta; advisor: Timo Hiller; co-advisor: Leonardo Rezende. – 2022.

39 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2022.

Inclui bibliografia

1. Economia – Teses. 2. Teoria Econômica – Teses. 3. Organização Industrial. 4. Mercado de Notícias. 5. Notícias na Internet. 6. Agregadores de Notícias. I. Hiller, Timo. II. Rezende, Leonardo. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 620.11

To my parents, Carlos and Natalia.

Acknowledgments

I would like to first thank my advisors Prof. Timo Hiller and Prof. Leonardo Rezende for their guidance and thoughtful comments.

Then I wish to thank my family and friends, for their support and motivation throughout the years.

I would also like to thank Rafael Lincoln, who above all others influenced me to pursue this Master's program.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Financial support from FAPERJ is gratefully acknowledged.

Abstract

Aliende da Matta, Victor; Hiller, Timo (Advisor); Rezende, Leonardo (Co-Advisor). **The Market for Internet News Distribution**. Rio de Janeiro, 2022. 39p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This paper studies the role of aggregators as intermediaries in the online news industry. I propose a model where firms must appeal to consumers with differentiated tastes, trading off between a vertical dimension of quality and a horizontal dimension of relevance. In this context, depending on the relative strength of these forces, the presence of the news aggregator may either increase quality and welfare or decrease quality with an ambiguous effect on welfare. I argue that while the first scenario is more in line with the existing theoretical literature on news aggregators, the second seems to be more strongly supported by the empirical evidence. The impact of aggregators in this second scenario may substantiate concerns over the quality of news provision on the internet.

Keywords

Industrial Organization; News Market; Internet News; News Aggregator.

Resumo

Aliende da Matta, Victor; Hiller, Timo; Rezende, Leonardo. **O Mercado de Distribuição de Notícias na Internet**. Rio de Janeiro, 2022. 39p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Este artigo estuda o papel de agregadores como intermediários no mercado de notícias online. Propomos um modelo onde firmas buscam consumidores com preferências heterogêneas, balaceando entre uma dimensão vertical de qualidade e uma dimensão horizontal de relevância. Nesse contexto, dependendo da intensidade relativa dessas forças, a presença de um agregador de notícias pode aumentar a qualidade e bem estar dos consumidores ou diminuir a qualidade com efeitos ambíguos no bem estar. Argumentamos que apesar do primeiro cenário se aproximar mais da literatura teórica existente sobre agregadores de notícias, o segundo cenário parece melhor sustentado pela evidência empírica. O impacto de agregadores neste segundo cenário pode sustentar preocupações sobre a qualidade da provisão de notícias na internet.

Palavras-chave

Organização Industrial; Mercado de Notícias; Notícias na Internet; Agregadores de Notícia.

Table of contents

1	Introduction	9
2	Related Literature	12
3	Model	14
3.1	Topics	14
3.2	Firms	15
3.3	Consumers	17
3.4	Revenue	17
3.5	Aggregator	18
3.6	Timing	19
4	Equilibrium	20
4.1	Traditional Market	20
4.2	Market with an Aggregator	22
5	Effect of the Aggregator	24
5.1	Quality and Welfare	24
5.2	Discussion	25
5.3	Favoritism	26
6	Extensions	28
6.1	Advertisements Market	28
6.2	Increasing Market Size	30
7	Conclusions	31
	Bibliography	32
A	Mathematical Appendix	34
A.1	Proof of Proposition 1	34
A.2	Proof of Proposition 2	36
A.3	Proof of Proposition 3	37
A.4	Proof of Proposition 4	38
A.5	Equilibrium with increasing market size	39

1 Introduction

The internet has brought many transformations to the news media industry. These transformations have been accompanied by a growing concern over the health of the industry and the quality of news provision (Martens et al. (2018)). One key difference of the online news market is the presence of intermediaries in the form of dedicated news aggregators (such as Google News) or social media firms (such as Facebook), which attract a significant fraction of the traffic in the online news market. Indeed, Newman et al. (2021) show that only 25% of people report that their main way to get news online is by directly accessing a newspaper, while 34% report it to be either accessing social media or a news aggregator.

The role of these aggregators in the news market has attracted the attention of policy makers and economists (see Jeon (2018) for a review). In this paper, we propose a new model of the online news industry to analyse their effect. An important feature of our model is that firms and consumers have a limited capacity to produce and consume news, so that not everything can be reported on¹. This gives rise to strategic topic selection by news firms, who must produce news taking into account the trade-off between strongly appealing to a smaller audience or weakly appealing to a larger one. Modelling this trade-off is a key motivation of our model.

These constraints in turn give rise to a new role for an aggregator in the market: personalizing news. The importance of this role can indeed be seen by the fact that some of the most successful news aggregators are technology firms which use extensive data and complex algorithms in order to build personalized news feeds for consumers. Indeed, by aggregator, we mean any platform which is capable of providing people with a customized news bundle. This definition then encompasses specialized news aggregators such as Google News, but also social media websites and even messaging applications, although we do not go into the particular social incentives through which these latter websites create their customized content bundles.

In our model, when the aggregator fulfils the role of personalizing news

¹Eisensee & Strömberg (2007) provides empirical evidence of this phenomenon, Gentzkow et al. (2015) provides a discussion of news filtering by media firms

in the market, it has the effect of fragmenting the market and it reduces the quality of news provision. Its impact on quality is made through two different channels: the business stealing effect (shifting revenue from news firms to the aggregator itself) and the composition effect (shifting consumers towards lower quality news). While previous studies have focused on the first channel, the composition effect present in our model provides a new mechanism through which aggregators impact the quality of news provision, as well as the welfare of consumers and the industry structure of the news market.

Previous theoretical literature on news aggregators (most prominently Jeon & Nasr (2016)) has described how aggregators can increase the quality of news provision by concentrating a larger audience to the highest quality news articles. Under our model, this effect may also be present, but only under the scenario where consumers are initially dispersed, with the aggregator leading to a more concentrated news market and more uniform news consumption. However, empirical evidence on the effect of news aggregators, such as Athey et al. (2021), supports the opposite effect, with the aggregator disproportionately benefiting smaller news firms and mixing up the composition of news topics that people are informed about.

Our model describes how, depending upon parameters of the utility functions of consumers and of the cost function of investing in quality, the market for news under traditional technology (as opposed to aggregator technology) may be highly concentrated, even with differentiated tastes. This is because consumers prefer access to high quality mainstream news provision than a more personalized but lower quality coverage. When the aggregator enters the market, it is able to satisfy consumers in both aspects, directing them to the best news coverage at each moment. While it has been hypothesized that the reduction in the fixed costs of setting up a news media company through the advent of the internet could have eroded reputation incentives that kept firms producing high quality content (Zhuravskaya et al. (2020)), in our model, small firms may not be able to find an audience due to their incapacity to provide sufficient quality. The entry of the aggregator connects these firms to consumers by introducing customized news bundles which mix high quality mainstream content with personalized and occasionally relevant niche news.

More precisely, we build a model of long-run equilibrium in the news market, where firms make zero profit, and find that if in the traditional market consumers favor high relevance niche content bundles over high quality mainstream ones, entry of an aggregator increases quality and welfare. This happens because firms providing higher quality content can now capture an audience previously fragmented among horizontal lines. This is consistent with

previous findings from the theoretical literature on news aggregators.

But if consumers instead favor high quality mainstream content bundles over high relevance niche ones in the traditional market, entry of an aggregator has the opposite effect: fragmenting the market and decreasing news quality. Consumers now have their news bundles personalized and quality provision cannot be sustained at the same level. The welfare effect is then ambiguous: quality is lost but relevance is gained.

In this latter scenario, the aggregator changes the composition of news topics people are informed about and disproportionately benefits smaller news firms, consistent with the empirical evidence of Athey et al. (2021) regarding Google News. The decrease in quality predicted by the model in this scenario is reminiscent of growing concerns about the online news market.

Our model is meant to capture the mechanisms laid out above in a stylized manner, with a number of simplifying assumptions. For instance, we assume that consumers get their news from a single source and that all news is financed through advertisements. These assumptions, which we discuss in more detail in the third chapter, are common in the literature (Bisceglia (2020), Jeon & Nasr (2016)). Furthermore, they provide a good illustration of the mechanisms proposed and serve as a suitable benchmark. As mentioned before, we also introduce assumptions that lead us to a zero profit equilibrium condition, so that we only explore the behaviour of firms acting in a competitive market.

Our framework potentially has a number of policy consequences. While aggregators may reduce quality by limiting the extent through which preference externalities can finance investment in news production, beyond the well characterized business stealing effect, they may also increase competition and provide consumers with information on a wider variety of topics. Regulation of news aggregators must consider all these effects in order to be truly effective.

We end this paper with possible extensions to the model and a direction for future research. Most prominently, we sketch the possibility of deriving endogenous prices for advertisements. We find that, under our model, aggregators compete with news firms in the advertising market in a different way than news firms compete amongst themselves, with a stronger impact on advertisement prices. This represents yet another mechanism through which aggregators may impact the quality of news provision.

2

Related Literature

The empirical literature robustly presents that news aggregators increase news consumption (Jeon (2018) reviews the nascent literature around news aggregators, both theoretical and empirical). Athey et al. (2021) and Calzada & Gil (2020) utilize natural experiments arising from judicial disputes in Spain and Germany to identify this effect. Chiou & Tucker (2017) utilizes a contract dispute between Google News and the Associated Press to study the effect of the exit of a major news provider from the platform and finds that it decreases news consumption from other content providers. Sismeiro & Mahmood (2018) analyze the effect of a short term Facebook outage and finds that it also decreases news consumption.

Most of the theoretical literature works on the assumption that aggregators direct news consumers to the highest quality option available, where quality is defined as a broad spectrum of positive attributes of news articles which represent a fixed cost for the firm and increase consumer utility.

Jeon & Nasr (2016) utilizes this assumption, together with the empirical findings that aggregators increase news consumption, to conclude that aggregators increase news quality and consumer welfare. In contrast, our model strategically separates topic selection from other measures of news quality. Ultimately, under this perspective, it becomes an empirical question whether aggregator technology selects for quality or increased topic affinity. In the latter case, we show that the aggregator decreases quality and may lead to a decrease in welfare in the long run.

Bisceglia (2020), while not directly discussing news aggregators, is closely related to our discussion. In his model of an ideologically differentiated duopolistic news market, a shift from traditional to internet news consumption changes the incentives between investment in the intensive margin of news quality and extensive margin of news coverage, leading to a decrease in quality. While his effects on news quality is driven by lower competition in exclusive articles, our model has free entry and lower quality provision results from preference externalities and market size, that is, a pecuniary externality resulting from increasing returns to scale and differentiated tastes. Our results on quality provision are also related to a shift in the structure of the news

industry, from more to less concentration.

Our framework also connects to a different literature looking at the impact of social media on news consumption (Zhuravskaya et al. (2020)). While the underlying technology of social media platforms works quite differently from that of standard news aggregators, they ultimately also produce personalised streams of news articles. We believe that the description of our model whereby consumers forego higher quality news outlets towards low quality niche content describes an important mechanism to explain why consumers seek information in social media.

3 Model

3.1 Topics

By topic, we mean a category of news articles, which follows an exogenous separation of information on all possible states of the world into independent variables. As such, each topic is subject to an exogenous stochastic relevance shock, which makes reading about that topic more interesting to consumers. This shock reflects the events going on in the real world that consumers may want to read about.

This definition of a topic is meant to capture the wide variety of things consumers may want to read about while providing a stylized model of the heterogeneity in their tastes. A prominent example, which is of particular interest to policy makers and economists because of its impact in politics, is that of local news. In our model, each location would represent a particular topic (e.g. “New York news”). Our definition of topic however is also meant to be applicable to other dimensions beyond local/non-local news. Newspapers themselves commonly separate between news categories such as “sports” and “finance”. This separation however is still constrained by having to produce a single product that appeals to a large audience and therefore limits itself, in general, to broad distinctions.

Among topics, our main distinction is that between those that are broad and niche. A broad topic interests all consumers while niche topics interest only a small audience. This distinction has a number of economic consequences. While consumers benefit from firms catering to their niche interests, they may be reluctant to forego the higher quality that is provided by a firm with mainstream appeal. Again, one prominent example of such a trade-off is the local vs national news market. While local news firms should be able to provide more relevant content to their audiences, George & Waldfogel (2006) shows evidence that consumers do switch from local news outlets to higher quality national ones.

For simplicity, we will assume that there is one broad topic (generally referred to by index b) and n niche topics, with each consumer being interested

by a single niche topic and each niche topic being interesting to the same number of consumers, that is, niche topics equally partition the sets of consumers. Let $T = \{b, 1, 2, \dots, n\}$ be the set of all topics.

The results we will establish are all stated in terms of a sufficiently large value of n . This assumption, by guaranteeing that niche topics are small relative to the broad topic, is what makes precise the distinction between broad and niche topics and gives it sharp economic consequences.

Each topic $t \in T$ is *relevant* independently with exogenous probability π , reflecting a state of the world that is more interesting regarding that particular topic. For example, if t represents politics, then the day of the 2021 United States Capitol attack would be a day where reading about t is more attractive. Let $r_t \sim \text{Bernoulli}(\pi)$ be the random variable that indicates whether topic t is relevant.

3.2

Firms

News firms produce a single “unit” of news, competing for the attention of consumers by choosing their coverage of topics, broad or niche, and investing in quality.

Quality is meant to reflect features of articles which are generally appreciated such as depth, thoroughness and timeliness. Investment in quality is modelled as a choice of a vector $\mathbf{q} \in R_+^T$, assigning a non-negative real number q_t for each topic t . This choice incurs a cost of $\sum_{t \in T} C(q_t)$, where C is a strictly increasing differentiable cost function. Therefore, since quality is a fixed cost of production, we have increasing returns to scale regarding quality investment, which has been recognized as an important component of the media industry at least since Shaked & Sutton (1987). Importantly, firms’ choice of quality is independent from realizations of relevance, which reflect longer term decisions the firm needs to make, such as how many journalists to hire.

We assume that firms only pay the quality cost when they are active, that is, when they have at least one consumer, which is meant to reflect a long-run equilibrium in the market, being contested by potential entrants. This hypothesis allows us to explore strategic interactions in content provision while leading firms to invest at a competitive level, earning zero profit (as will be seen in later sections). By doing so, we simplify away from the effect of imperfect competition on the equilibrium supply of quality.

Firms strategy on coverage, which defines the topic of an article that will be presented to consumers, on the other hand can depend on realizations of

relevance. We model it as a function $\sigma : 2^T \rightarrow \Delta(T)$, mapping the state of the world, given by a binary vector indicating whether each topic is relevant, to a probability distribution between topics.

The strategy of firms is then a tuple (q, σ) of a quality vector q and a topic selection function σ .

As an example, suppose $T = \{b, x, y\}$ where b refers to national news and x, y refers to local news of two regions X, Y with equal populations. A firm may enter the market setting $q = (2, 1, 1)$ and implementing a symmetric topic selection function where:

1. $\sigma((1, r_x, r_y)) = (1, 0, 0)$ for all r_x, r_y .
2. $\sigma((0, 1, 1)) = (0, 1/2, 1/2)$.
3. $\sigma((0, 1, 0)) = (0, 1, 0)$ and $\sigma((0, 0, 1)) = (0, 0, 1)$.
4. $\sigma((0, 0, 0)) = (1, 0, 0)$.

The cost of operating this firm is $C(2) + 2C(1)$.

The intuition behind this strategy space is of a news firm who must make rigid quality and news gathering investments, for example, by hiring a certain number of journalists and allocating them to certain sections of the newspapers, which then go on to engage in particular news gathering activities, reflecting the fixed cost quality component of a firm's strategy. Journalists may then pitch article ideas and drafts to an editorial body, who make the final decision regarding what will be show to consumers. This editorial side is reflected by the topic selection function.

The “unit” of news that a firm produces is then a probability distribution between articles, as a function of quality investment, topic selection strategy and stochastic state of the world, from which consumers will have access to a single realization. This assumption embodies the limitation news firms face in deciding what information to show to their consumers. This “unit” of news may best be interpreted as the homepage of a newspaper, as that is the uniform product that is presented to consumers. Real news firms may have more than one homepage, allowing consumers to self select into groups of different interests and consume a different news bundle. The underlying assumption of this paper is that such discrimination between consumers is relatively trivial compared to what is implemented by news aggregators and, in particular, technology driven news aggregators.

We assume that the market features an infinite number of news firms, reflecting free entry. This is meant to reflect the long-term nature of the game

we are modelling and the fact that the modern online news media market has very few barriers to entry.

3.3 Consumers

There is a continuum of mass N of consumers with unit demand for news. As described before, each consumer is interested in the broad topic b and some niche topic k . Consumers single-home, that is, consume content from a single news firm, and receive utility from reading news only on topics they are interested in. The utility consumers may get from a news article depends on its quality and relevance. Importantly, consumers choose between firms without knowing the realization of relevance, maximizing expected utility, since they are not aware of the state of the world before reading the news.

The fact that the consumers can only choose one firm is meant to reflect the limited time they have to dedicate to news. This assumption goes together with the fact that consumers are only interested in two news topics. What is essential to the model is that consumers do not have enough time to go over news content until they exhaust all their topics of interest, as this would make different solutions to the compilation of news bundles irrelevant, something which does not seem realistic.

The utility of consumer c , who is interested in niche topic k , of the firm implementing strategy (q, σ) is then given by:

$$E_r[\sigma_b(r)(q_b + r_b R_H + (1 - r_b)R_L) + \sigma_k(r)(q_k + r_k R_H + (1 - r_k)R_L)]$$

Where R_H and R_L are constants which reflect the utility consumers get from high and low topic relevance. We assume that $R_H > R_L > 0$.

3.4 Revenue

Firms' revenue comes from a competitive advertising market with a fixed amount of advertisements that can be displayed per article and no nuisance costs¹. We capture in reduced form the additional revenue that companies can generate by making articles more interesting to consumers with three tiers of ad exposure: firms receive no revenue by presenting a consumer with an article on a topic she is not interested in, they receive p_l revenue by presenting an article

¹We therefore abstract from deeper strategic elements related to advertisement-financed media companies in general. We further explore the role of the advertising market in chapter 6.

on an interesting topic but with low relevance realization and they receive $p_l + p_h$ by presenting an interesting and relevant article. These three tiers are intuitively related to the amount of engagement consumers may have with a news website depending on what she finds there and therefore the amount of exposure to advertising that the consumer can be submitted to. Revenue per consumer (of type k) is then given by:

$$E_r[R_k(\sigma)] = E_r[\sigma_b(r)(r_b p_h + p_l) + \sigma_k(r)(r_k p_h + p_l)]$$

The assumption of advertisement financed media companies is in line with the previous theoretical literature on internet news and aggregators (Jeon & Nasr (2016), Bisceglia (2020)). Most importantly, in the context of this model, it establishes a strong connection between audience size and revenue, which in turn impacts quality. Many news firms on the internet have introduced *paywalls* to their content, which could lead to important strategic considerations: if niche firms could charge their consumers enough to be able to provide quality at a mainstream level, we should expect them to have a larger presence in the market. However, this does not seem to be the case in practice. Indeed, it seems more common for large news firms to introduce paywalls while small firms rely on advertising. If this is indeed so, the conclusions of this model should be expected to hold. It would of course be interesting to extend the current model to a flexible financing model encompassing both price and advertising competition, but we leave this project for future research.

3.5 Aggregator

We then introduce a news aggregator into the market. The news aggregator is a monopolist firm with a special technology that directs consumers to news firms, based on realizations of relevance and the personal preferences of the consumer. Effectively, by using the aggregator, instead of choosing between firms without knowledge of the realization of relevance, consumers can pick the firm that provides the highest utility under a particular realization. As such, consumer utility under the aggregator is:

$$E_r[\max_f \{\sigma_b^f(r)(q_b + r_b R_H + (1 - r_b)R_L) + \sigma_k^f(r)(q_k + r_k R_H + (1 - r_k)R_L)\}]$$

Where f stands for any news firm. This technology of the aggregator is the analogue in our model of the assumption that aggregators choose between firms to maximize quality, but here quality is also composed of the exogenous

realization of relevance. In reality, aggregators often achieve this by collecting information on user behaviour upon being exposed to different news articles. Since users display more engagement with articles they find more interesting, the aggregator is able to identify the most promising news articles and direct more consumers to them.

But the aggregator also serves as a substitute to news firms insofar as it diverts users away from their homepages and sometimes provide snippets as well². In our model, this is flexibly reflected by the aggregator keeping to itself a fraction $0 < \tau < 1$ of the revenue. In this manner, τ captures the business stealing effect of the aggregator.

3.6 Timing

We will analyse the pure-strategy subgame-perfect Nash equilibria of a game with three phases. First, firms enter the market simultaneously, choosing quality and their topic selection strategy. Second, consumers decide between news firms or between news firms and the aggregator. Third, relevance shocks are revealed, consumers are presented with their articles and all players receive their payoffs.

²We assume that users do not derive additional utility from this. This is mainly because news firms also provide such snippets in their own homepages. Furthermore, aggregators which function through links to news sites provide snippets that are written by the news firms themselves.

4 Equilibrium

Let $Q := C^{-1}([p_l + \pi p_h]N)$. This quantity summarizes key parameters of the model: cost structure, revenue structure and market size. This quantity arises from what we will call “the mainstream strategy”. A firm may enter the market with strategy $((Q, 0, 0, \dots, 0), \sigma^*)$ where $\sigma_b^*(r) = 1 \forall r$ (and therefore $\sigma_t^*(r) = 0 \forall t \neq b$ since $\sigma^*(r)$ is a probability distribution) and be viable as long as it can attract all consumers. We will generally state assumptions in terms of Q , giving them clear economic interpretation.

We will assume throughout that $R_H - R_L > Q$. If this were not the case, for any state of the world, consumers would prefer to consume news of a single topic, given its high provision of quality. In that case, there would be no incentive for firms to differentiate and no room for the aggregator to play in the market in the context of our model, giving rise to a single firm dominating the market in all scenarios.

4.1 Traditional Market

First, we characterize the equilibrium in the market without an aggregator. In this scenario, firms compete to create the bundle of news that provides highest expected utility to consumers. The revenue of firm f is given by:

$$R^f = \sum_{k \in T \setminus \{b\}} \frac{N}{n} D_f^k E_r [R_k(\sigma^f)]$$

$$U_k(q, \sigma, r) = \sigma_b(r)(q_b + r_b R_H + (1 - r_b) R_L) + \sigma_k(r)(q_k + r_k R_H + (1 - r_k) R_L)$$

$$R_k(\sigma) = \sigma_b(r)(r_b p_h + p_l) + \sigma_k(r)(r_k p_h + p_l)$$

Where D_f^k is the demand for firm f from consumers of type k . If firm f provides the maximum expected utility for these consumers, D_f^k will be equal to 1. When there are ties, consumers are indifferent between options. In equilibrium, however, consumers will all consume from a single firm. To see this, notice that if consumers were split between firms, then one of the firms

could increase their quality and attract the other firm's consumers. The firm then makes expected revenue equal to $E_r[R_k(\sigma)]$ from each consumer of type k . This revenue depends on whether the firm is present interesting and relevant content to those consumers. With this, we establish the following proposition:

Proposition 1. For sufficiently high n :

- If $(1 - \pi)\pi(R_H - R_L) < Q$, then the market features a single active firm in any equilibrium. There exists an equilibrium where this firm plays the mainstream strategy $((Q, 0, \dots, 0), \sigma^*)$. Furthermore, if (q, σ) is an equilibrium strategy for the firm, then $(q, \sigma) \rightarrow ((Q, 0, \dots, 0), \sigma^*)$ as $n \rightarrow \infty$.
- If $(1 - \pi)\pi(R_H - R_L) \geq Q$, then there are n active firms in equilibrium. Each active firm $k \in [1, 2, \dots, n]$ features only the broad topic b or the niche topic k , always favoring the topic $t \in \{b, k\}$ which has the highest relevance.

Proof: Provided in the Mathematical Appendix.

This result establishes the two possibilities that are supported by the market in our model, which depend on the relative strength of two forces: the relevance that can be gained by specialization and the quality that can be gained by reaching a wider audience.

In the first scenario, increasing returns to scale for quality provision are strong enough that consumers prefer a mainstream firm, giving rise to significant preference externalities (Anderson & Waldfogel (2015)). This externality is also behind multiple equilibria: though one firm dominates the market, it can choose to favor certain groups of consumers. No firm can profitably enter as long as it cannot attract these favored groups and other consumers prefer the quality they gain from these groups' contribution to the firm's scale than what another firm can offer them. However, for sufficiently high n , the firm's capacity to display favoritism vanishes.

In the second scenario, we have that consumers care more about specialized news than the higher quality that could be provided. Therefore, consumers are divided into specialized firms. The left hand side term in the inequality $(1 - \pi)\pi(R_H - R_L)$ expresses the marginal provision of relevance that can be gained by adopting a specialized strategy as compared to the mainstream one. In particular, the term $(1 - \pi)\pi$ (where π is the probability that a topic is relevant) serves as a crucial weighting factor when comparing both strategies.

This scenario is most similar to what is usually assumed by the literature. In the case of Jeon & Nasr (2016) and Bisceglia (2020), consumers are assumed

to be divided between firms that are exogenously differentiated in an ideological spatial dimension.

Which scenario best describes the real world is an empirical question, though difficult to properly characterize by itself. The news media is in general quite concentrated while also displaying significant horizontal differentiation among firms that are present in the market. But as we will see in the following sections, the impact of an aggregator differs dramatically in each scenario. It is this impact which will suggest to us which scenario is more relevant for explaining the current online news media market.

Before that, however, it is worthwhile to further describe the effects of the assumption that n is sufficiently high. First, this effect imposes vanishingly small quality provision that can be attained by appealing to niche audiences. Though this is naturally a simplification, it allows us to neatly compare the trade-off between broad and niche appeal: just as it is the marginal provision of relevance in the niche strategy that matters for our results, it is also only the marginal provision of quality in the mainstream strategy that is significant, which, given our assumptions, approaches Q . Second, it strengthens the constraint that the firm cannot provide coverage of all topics. Since catering to one group of consumers alienates all others, firms can only effectively display favoritism towards a fixed quantity of groups, which become insignificant as n grows. Again, this simplification allows us to neatly characterize the relevance gains of a niche strategy.

4.2 Market with an Aggregator

We now consider the market with an aggregator. If consumers choose to use the aggregator, the revenue of firm f is given by:

$$R_A^f = (1 - \tau) \sum_{k \in T \setminus b} \frac{N}{n} E_r \left[D_f^a(k) R_k(\sigma^f) \right]$$

$$U_k(q, \sigma, r) = \sigma_b(r)(q_b + r_b R_H + (1 - r_b) R_L) + \sigma_k(r)(q_k + r_k R_H + (1 - r_k) R_L)$$

$$R_k(\sigma) = \sigma_b(r)(r_b p_h + p_l) + \sigma_k(r)(r_k p_h + p_l)$$

Where $D_f^a(k)$ is how many consumers the aggregator sends to the firm, which is chosen according to which firm would provide the highest utility to consumers. This function is very similar to what we found in the traditional market, but with important differences. Most important is that now, instead of

aiming to maximize the expected utility for each group, firms must maximize a posteriori utility in order to gain an audience, since the aggregator leads consumers to the most relevant firm after the shock takes place. This leads us to our next proposition:

Proposition 2. For sufficiently high n , then in all equilibria consumers have the same utility and the market features $n + 1$ active firms, with a single active firm actively covering only the broad topic with quality Q^A and n active firms actively covering only a single niche topic with quality $Q^n < Q^A$, where to actively cover a topic is defined by covering the topic while having at least one consumer.

Proof: Provided in the Mathematical Appendix.

When the aggregator enters the market, it essentially limits competition to within a certain topic. Consumers always prefer to receive relevant news and, conditional on that, the highest quality option available. Higher quality provision can only be attained by a bigger audience, which gives rise to equilibria where each firm dominates its own topic. All consumers prefer the aggregator since only through the aggregator can they access all the relevant news at the highest level of quality. The equilibrium then becomes essentially unique, besides from the fact that firms have to take decisions regarding what to do even when it does not attract any consumers.

The provision of quality for the broad topic under the aggregator is given by $Q^A = C^{-1}((1-\tau)[(1-\pi(1-\pi))p_l + \pi p_h]N)$. It is straightforward to establish that $Q^A < Q = C^{-1}([p_l + \pi p_h]N)$. This difference is determined by the share τ of revenue appropriated by the aggregator as well as the diverted attention $\pi(1-\pi)$ from the broad topic to niche ones. We discuss the differences in overall quality provision with and without the aggregator in the next section.

5 Effect of the Aggregator

5.1 Quality and Welfare

As we saw in Proposition 1, there are two different scenarios which can describe equilibria in the traditional market. In contrast to that, the market with an aggregator always has the same equilibria. We therefore derive the effect of the entry of the aggregator for each scenario in the following proposition:

Proposition 3. For sufficiently high n :

- If $(1 - \pi)\pi(R_H - R_L) < Q$, then the quality of every article is lower with the aggregator than in the traditional market. Welfare is increasing if and only if:

$$(1 - \pi)\pi(R_H - R_L) > [1 - \pi(1 - \pi)](Q - Q^A) + \pi(1 - \pi)Q$$

- If $(1 - \pi)\pi(R_H - R_L) \geq Q$, then quality of the broad topic is higher with the aggregator than in the traditional market and welfare is increasing.

Proof: Provided in the Mathematical Appendix.

Let $\Delta R = (1 - \pi)\pi(R_H - R_L)$ and $\partial Q := [1 - \pi(1 - \pi)](Q - Q^A) + \pi(1 - \pi)Q$.

What happens in the first scenario ($\Delta R < Q$) is that, as the traditional market is completely concentrated on a single firm, the entry of the aggregator causes the fragmentation of the market. Consumers divert their attention from high quality low relevance news to low quality high relevance news, which leads to an unambiguous decrease in quality.

This in turn causes an ambiguous change in welfare. The condition for welfare to be increasing can be interpreted by comparing, on the left hand side, the gains from increased news personalization ΔR to, on the right hand side, the quality decrease ∂Q . This decrease has three channels: there is a mechanical decrease which comes from the change in the composition of content from mainstream to niche news, there is the decrease in quality of broad content

which is caused by the well known *business stealing effect* of the aggregator, since it retains a portion of the industry's revenue and finally there is a loss in quality which comes from a scaling down of broad content as consumers no longer dedicate as much time as they did to viewing it.

Intuitively, there is a pecuniary externality since consumers do not internalize the effect of their readership in the provision of high quality content. The aggregator, even if it was free ($\tau = 0$), can decrease welfare in the market since it gives people an incentive to act in a way that is prejudicial to the bottom line of high quality news firms, limiting their capacity to collect revenue from consumers. It acts as if to decrease the property rights of news firms over their content, on the one hand empowering consumers, but on the other harming them, as quality decreases in equilibrium. Under the context of our model, this can take place even as there is increased consumption of news, with the compositional effect generally driving consumers to read more articles, even as the utility they get from each article falls, reminiscent as well of current concerns about information abundance.

But in the second scenario ($\Delta R \geq Q$), the aggregator instead allows for one firm to specialize in broad content, attracting an audience from all groups of consumers, concentrating the market and increasing quality and welfare.

The situation can be easily summarised in the following picture, identifying the consequences of the location of ΔR in the parameter space:

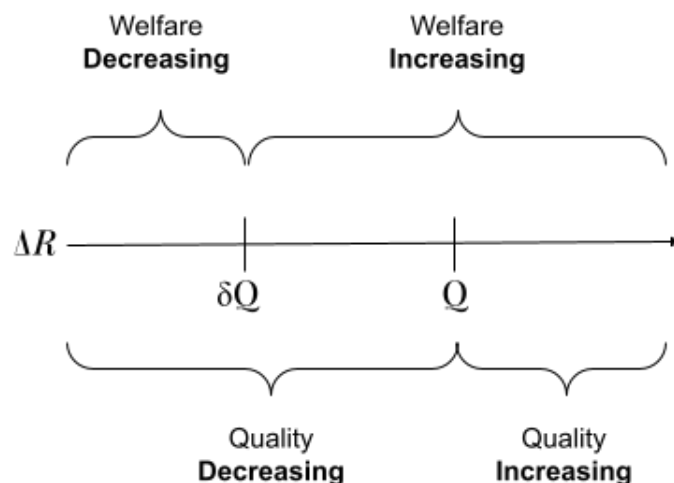


Figure 5.1: Parameter Space and Consequences

5.2 Discussion

While the second scenario, where the aggregator promotes quality investment, is reminiscent of the investigation in Jeon & Nasr (2016) since the

aggregator increases competition on quality within a particular topic, it is the first scenario which seems to be more strongly corroborated by the empirical literature.

Athey et al. (2021) utilizes a natural experiment introduced by a copyright reform enacted on December 2014 in Spain that caused Google News to shut down in the country. Utilizing data on individual news consumption, they find that the aggregator changed the profile of news topics that users were exposed to and caused increased news consumption but this increase was entirely towards smaller news outlets.

Though difficult to ascertain empirically, concerns have been raised over the quality of news provision on the internet, as discussed by Bisceglia (2020) and Martens et al. (2018). Consumers report less trust on online media sources and by some measures, like the length of news articles or originality of reporting, quality seems to be diminished. We are not aware of any studies that specifically try to measure the quality of articles consumed through aggregators.

Our model predicts that decreasing quality and shifting towards smaller news firms are part of one underlying mechanism: personalization of news. While increased competition on the news market may have many positive effects on consumers, not only from the perspective of market power but also media power (the political power of media companies, see Prat (2015)), it may come with negative consequences.

We have seen that our theoretical model cannot guarantee that aggregators increase welfare. There is however suggestive evidence that they do: as Jeon (2018) reports, legal disputes in Germany led Google News to establish an opt-in system for news media companies. However, after initially staying out of the platform, many publishers eventually decided to opt-in. Under the lens of our model, this would suggest that the news firms could not offer a more attractive product to consumers than the aggregator, suggesting indeed an increase in welfare.

5.3 Favoritism

Concentrating still on the first scenario, we can elaborate on another effect of the aggregator.

As we saw in Proposition 1, in the traditional market, the dominating firm may not play the mainstream strategy and instead display favoritism towards some group or groups of consumers, while still retaining the attention of the others, because of the higher quality level that can be sustained by

maintaining the market concentrated. Such deviations are necessarily welfare decreasing for all consumers not among those favored, which tend to be the large majority, since these consumers are pushed content of no interest to them, while gaining nothing in return.

The multiple possible strategies that the dominating firm may play while retaining its position in the market may be interpreted as editorial leeway that news firms have to push content under a different motivation than maximizing revenue. Therefore, in this model, the news firm has the possibility of exercising editorial power in a costless or, more precisely, payoff equivalent, manner. This possibility is here despite of the assumption of perfect competition, again, because of pecuniary externalities introduced by the nature of the technology, which displays increasing returns to scale.

This mechanism is reminiscent of the literature on media bias (see Gentzkow et al. (2015) for a theoretical survey). Among this literature, our model presents an interesting possibility. Though the “bias” in content displayed by the dominating news firm is supported by consumer demand, this demand comes from a minority of consumers and yet it still affects news provision over the entire market.

This possibility however vanishes when the aggregator enters the market. In that case, if a firm chooses to display favoritism towards some fraction of consumers, the aggregator will simply direct the remaining consumers to other news sites for that fraction of content. This will then cause the deviating firm to lose market share on its mainstream content and will leave room for another firm to replace it. The aggregator thus takes away editorial power from news firms by removing the possibility that a portion of their content sustains the existence of another portion, making their bundle still attractive to consumers.

This however may come at a cost: although the aggregator eliminates the possibility of mismatch between what consumers want and what they presented to read, this can have negative effects. Ideological bubbles and confirmation bias have become an important concern regarding consumers’ access to information. In this respect, the aggregator will always allow consumers to become more extreme.

6 Extensions

6.1 Advertisements Market

So far, we have assumed that the market for advertisements is perfectly competitive and that prices are set exogenously, we now relax this assumption. This would not only serve a robustness check, adding realism to the model, but also allow us to tackle interesting questions regarding firm behaviour on the market for advertisements. Among these, we can explore the consequences of policy proposals which aim to increase the revenue of news firms by establishing compensatory schemes between them and the most prominent aggregators. What we describe below is still an early attempt at tackling this problem, only suggesting a promising avenue for future research.

There is a extensive literature on advertising in media markets (see Anderson & Jullien (2015) for a review). For now, we abstract from common concerns of the literature over the two sided nature of the market by maintaining that advertisements do not cause nuisance to consumers. Instead, we focus only on the impact in the provision of news quality of having oligopolistic competition in the market for advertisements.

We extend the model outlined above in the scenario where $p_h = 0$ by first explicitly introducing that consumers spend a certain amount of time reading news. We assume that consumers spend 0 units of time with an article about a topic that is not of interest to them and 1 unit of time reading an article that is interesting. Furthermore, we assume that users of the aggregator spend a fraction τ of their time on it. Firms who acquire consumers' time may then sell it to advertisers while choosing a certain ad intensity a . In that way, a firm who can attract a consumer for t_i time and chooses ad-intensity a_k has $t_i a_k$ ad slots from that consumer to sell to advertisers.

In the market, there is a continuum of advertisers that derive some benefit from reaching consumers.¹ When multiple firms sell ad slots from a single

¹We assume that advertisers do not value reaching a certain consumer for more than one ad slot and that firms cannot discriminate between advertisers. Advertisers are ordered so that $\omega(a)$ is the willingness-to-pay of the a -th highest advertiser. The revenue of a monopolist firm k that sells advertisements is then $R(t_i a_k) = \omega(t_i a_k) t_i a_k$ from consumer i .

consumer, they engage in Cournot competition, with the market bearing an equilibrium price for all available ad slots.²

As before, we analyse pure-strategy subgame-perfect Nash equilibria of the game. However, the timing now consists of 4 phases: firms simultaneously enter the market with their product decisions, followed by consumers deciding between firms. Then, each firm simultaneously chooses their ad intensity, bringing to market a certain amount of ad slots per consumer. In the final phase, relevance shocks are revealed and players receive their payoffs.

In the traditional market, as expected, news firms act as monopolists, since consumers single-home³. Even when the aggregator enters the market, however, news firms still do not compete. This is because the time consumers spend between multiple news firms is conditioned on different realizations of the relevance shock. Therefore, news firms sell ad slots which have independent value from one another. The aggregator however does provide effective competition, as it always retains some of the time budget of consumers and also acts on the advertising market.

The result is then that ad prices fall, and with decreased revenue, news firms invest less in quality. This extension to the model therefore presents yet another way the aggregator threatens the bottom line of news firms and perhaps leads to a decline in news quality on the internet.

Concerns over the impact of aggregators in news firms' revenues have led to legislation in many countries such as Germany, Spain and Australia. The most common form of such legislation is to enforce that platforms and news firms reach an agreement on compensation over the use of news content. We sketch out the effect of such legislation in the context of our model.

Because of perfect competition in our model of the news market, the situation would not change if news firms were to individually bargain with the platform. They may however attempt to collectively bargain, as indeed happens in reality. For simplicity, we will consider what happens if all news firms would bargain collectively. Using the Nash bargaining solution, the aggregator's profits are split equally between it and the news firms.

We can therefore establish:

Naturally, we assume that $R'(0) > 0$, $R'' < 0$ such that R has an unique interior maximum.

²We allow firms to sell ad slots conditional on stochastic realizations of the relevance variable. This seems to most accurately reflect the real online advertisements market, where advertisements are often sold in terms of click-through rates, and therefore need to be observably presented to consumers in order to generate revenue.

³Ad prices are then set at the monopoly level such that $R' = 0$.

Proposition 4. In the traditional market, advertisement prices are set at the monopoly level. With the aggregator, advertisement prices are smaller, because of increased competition. Lower prices lead to lower news quality. If collective bargaining is introduced, there is decreased competition, with ad prices rising but still below the monopoly level.

Proof: Provided in the Mathematical Appendix.

6.2

Increasing Market Size

So far, we have assumed demand for news to be fixed. In this section, we relax this assumption, discussing some of the new properties of this generalized model. Assume now that consumers have an outside option with value $g \sim G$ with support on $[0, \infty[$. Indeed, assume that every type of consumer is distributed in the same manner. We describe in the appendix the assumptions necessary for the game to have an equilibrium where analogous versions of Proposition 1 and 2 hold.

Fully exploring this version of the model is an area for future research. However, from what we have so far, two interesting features seem to come about.

The first one is the possibility of an inefficient equilibrium. This is because of a positive externality in the market, whereby investment in quality by one firm increases the demand for news and increases revenues for another firm. Thus would come about if the marginal cost of quality for a firm would be lower than the marginal revenue for all firms, which is possible, but not necessary, in the equilibrium characterized above.

The second feature is a new mechanism whereby the aggregator impacts the provision of quality in the news market. The increase in relevance that is necessarily caused by the aggregator increases the demand for news at any level of quality. Because of increasing returns to scale, the increase in demand itself leads to an increase in quality. This has then the opposite effect than the other mechanisms operating through the aggregator in the second scenario, where the aggregator fragments the market. Which effect is stronger depends on the shape of the demand curve.

7 Conclusions

In this paper, we propose a new model to interpret the effect of aggregators on the online news industry, which strongly features the personalization of news. In a world that produces vast amounts of information, platforms which simplify consumers' search for news articles of interest may have a first order role in shaping the industry and therefore the information that is available to consumers.

This model adds previously neglected dimensions to the understanding of the impact of aggregators on the online news market such as the compositional effect, which drives people to consume niche news that were not viable under previous technology. This effect is naturally accompanied by the proliferation of small news companies, serving this novel demand.

It is the distinction between broad and niche topics that most characterizes the domain of applicability of our model and suggests its proper interpretation. We expect to find that the conclusions of our model should hold whenever the broad topics being considered present an overwhelming larger audience, and with it a decisively larger potential level of quality, than the niche topics.

As we have already mentioned on a couple of occasions, this model suggests a number of directions for future research. Perhaps the most promising avenue would be to enrich our modelling of how firms make their revenue, either by further developing the market for advertisements, considering ad-nuisance for instance, or allowing for the implementation of paywalls. We could also explore the consequences of restricting entry, leading to a market with imperfect competition.

Another interesting area for future research is to the role of aggregators as we have modelled them in the spread of misinformation, as well as other pathological features of aggregated content such as *clickbait*. In this paper, we assume that aggregators truly know the utility that each article would provide to each consumer. In the real world, aggregators have important technological constraints on how they select news articles, which create a trade-off between more versatile technological options and the manual curation still very much in place in traditional news firms.

Bibliography

- TARSKI, A.. **A lattice-theoretical fixpoint theorem and its applications.** Pacific Journal of Mathematics, 5(2):285 – 309, 1955.
- ANDERSON, S. P.; WALDFOGEL, J.. **Preference Externalities in Media Markets.** In: HANDBOOK OF Media Economics, volumen 1, p. 3–40. Elsevier, 2015.
- PRAT, A.. **Chapter 16 - media capture and media power.** In: Anderson, S. P.; Waldfogel, J. ; Strömberg, D., editors, HANDBOOK OF MEDIA ECONOMICS, volumen 1 de **Handbook of Media Economics**, p. 669–686. North-Holland, 2015.
- ATHEY, S.; MOBIUS, M. ; PAL, J.. **The impact of aggregators on internet news consumption.** Technical report, National Bureau of Economic Research, 2021.
- SHAKED, A.; SUTTON, J.. **Product differentiation and industrial structure.** The Journal of Industrial Economics, 36(2):131–146, 1987.
- BISCEGLIA, M.. **The Unbundling of Journalism.** 2020.
- CALZADA, J.; GIL, R.. **What do news aggregators do? evidence from google news in spain and germany.** Marketing Science, 39(1):134–167, 2020.
- CHIOU, L.; TUCKER, C.. **Content aggregation by platforms: The case of the news media.** Journal of Economics & Management Strategy, 26(4):782–805, 2017.
- EISENSEE, T.; STRÖMBERG, D.. **News droughts, news floods, and us disaster relief.** The Quarterly Journal of Economics, 122(2):693–728, 2007.
- GENTZKOW, M.; SHAPIRO, J. M. ; STONE, D. F.. **Chapter 14 - Media Bias in the Marketplace: Theory.** In: Anderson, S. P.; Waldfogel, J. ; Strömberg, D., editors, HANDBOOK OF Media Economics, volumen 1 de **Handbook of Media Economics**, p. 623–645. North-Holland, 2015. ISSN: 2213-6630.

- GEORGE, L. M.; WALDFOGEL, J.. **The new york times and the market for local newspapers**. *American Economic Review*, 96(1):435–447, 2006.
- JEON, D.-S.. **Economics of News Aggregators**. TSE Working Papers 18-912, Toulouse School of Economics (TSE), Apr. 2018.
- JEON, D.-S.; NASR, N.. **News Aggregators and Competition among Newspapers on the Internet**. *American Economic Journal: Microeconomics*, 8(4):91–114, Nov. 2016.
- MARTENS, B.; AGUIAR, L.; GGMEZ, E. ; MUELLER-LANGER, F.. **The Digital Transformation of News Media and the Rise of Disinformation and Fake News**. *SSRN Electronic Journal*, 2018.
- NEWMAN, N.; FLETCHER, R.; SCHULZ, A.; ANDI, S.; ROBERTSON, C. T. ; NIELSEN, R. K.. **Reuters institute digital news report 2021**. Reuters Institute for the Study of Journalism, 2021.
- ANDERSON, S. P.; JULLIEN, B.. **Chapter 2 - the advertising-financed business model in two-sided media markets**. In: Anderson, S. P.; Waldfogel, J. ; Strömberg, D., editors, **HANDBOOK OF MEDIA ECONOMICS**, volumen 1 de **Handbook of Media Economics**, p. 41–90. North-Holland, 2015.
- SISMEIRO, C.; MAHMOOD, A.. **Competitive vs. Complementary Effects in Online Social Networks and News Consumption: A Natural Experiment**. *Management Science*, 64(11):5014–5037, Nov. 2018.
- ZHURAVSKAYA, E.; PETROVA, M. ; ENIKOLOPOV, R.. **Political Effects of the Internet and Social Media**. *Annual Review of Economics*, 12(1):415–438, Aug. 2020.

A

Mathematical Appendix

A.1

Proof of Proposition 1

Solving the game by backwards induction, consumers will choose the firm that provides maximum expected utility and, in equilibrium, will always choose a single firm in the case of ties, since if they didn't, either one firm would earn negative revenue or one firm could increase its profit by marginally increasing quality and attracting the other firms' consumers.

We now analyse the first phase of the game where firms enter.

Let (Q, σ^*) be the strategy of the described mainstream firm that features only the broad topic and invests the revenue that comes from all consumers in the market into quality for this topic, setting $Q = C^{-1}([p_l + \pi p_h]N)$. It is easy to see that, if all consumers prefer this firm, it makes 0 profit and therefore it would not be viable to increase quality further for any topic.

Let σ be any other topic selection function and (q, σ) be some firm's strategy. A consumer who is interested in niche topic t would derive utility:

$$E_r[q_b\sigma(b; r) + q_t\sigma(t; r) + (r_b\sigma(b; r) + r_t\sigma(t; r))(R_H - R_L) + (\sigma(b; r) + \sigma(t; r))R_L]$$

Where b represents the broad topic. This expression must be greater than $Q + \pi(R_H - R_L) + R_L$ for the consumer to prefer this alternative firm over the mainstream one. It is immediate that $E_r[r_b\sigma(b; r)(R_H - R_L)] \leq \pi(R_H - R_L)$.

Notice however that $E_r[q_b\sigma(b; r) + q_t\sigma(t; r)] \leq E_r[\sigma(b; r) + \sigma(t; r)] \max\{q_b, q_t\}$ and $\max\{q_b, q_t\}$ can be bounded above by:

$$C^{-1} \left(\left[E_r[\sigma(b; r)]p_l + E_r[r_b\sigma(b; r)]p_h + \sum_{k \neq b} (E_r[\sigma(k; r)]p_l + E_r[r_k\sigma(k; r)]p_h) / n \right] N \right)$$

Which then yields that, $\forall t' \neq \{b, t\}$, $E_r[q_b\sigma(b; r) + q_t\sigma(t; r)]$ is bounded above by:

$$E_r[1 - \sigma(t'; r)]C^{-1} \left(\left[p_l + \pi p_h + \sum_{k \neq b} (E_r[\sigma(k; r)]p_l + E_r[r_k \sigma(k; r)]p_h) / n \right] N \right)$$

This last expression yields Q when evaluated at σ^* . If we take the derivative of this expression with respect to $\sigma(t'; r)$ and let $n \rightarrow \infty$ we get $-P(r)C^{-1}([p_l + E_r[r_b \sigma(b; r)]p_h]N)$, which is always negative. In words, for sufficiently large n , the maximum possible marginal provision of quality generated by additional revenue from any alternative strategy cannot compensate consumers for the loss of utility associated with being pushed content that is not interesting to them. The positive effect on quality of only increasing niche content that is interesting to the consumer vanishes toward 0.

Let $\kappa = R_H/R_L$. In order for the consumer to prefer the alternative as regards the provision of relevance we must have that:

$$E_r[r_t \sigma(t; r)(R_H - R_L) + \sigma(t; r)R_L] \geq (1 - E_r[\sigma(b; r)])R_L$$

But that means $E_r[\sigma(t; r)]R_H > E_r[r_t \sigma(t; r)R_H + (1 - r_t)\sigma(t; r)R_L] \geq (1 - E_r[\sigma(b; r)])R_L$ and therefore $E_r[\sigma(t; r)] \geq (1 - E_r[\sigma(b; r)])/\kappa$. This implies that at most κ groups of consumers can prefer an alternative strategy on relevance.

Suppose $(1 - \pi)\pi(R_H - R_L) \geq Q$. Then consumers prefer specialized firms over the mainstream firm, since they would rather have a quality of 0 than miss out on coverage of the niche topic when it is relevant. If a firm tries to appeal to more than one group of consumers, since topics can be relevant simultaneously, it would need to provide partial coverage of at least one niche topic. But by the arguments above this could at most be preferable to κ groups of consumers. The maximum extra quality one group could receive from this larger firm is bounded by $C^{-1}([p_l + p_h]N\kappa/n)$ which naturally goes to 0 as $n \rightarrow \infty$. This extra quality cannot then compensate the group that receives partial coverage of his preferred topic. The only equilibrium is then for each group of consumers to have its own specialized firm, investing all its revenues in quality. This is so because if the firm makes profits, then another firm with no consumers would rather enter the market with a higher quality. Equilibrium is found when multiple firms offer the highest possible quality, with one of them being chosen by consumers but all of them making 0 profit.

Now, suppose $(1 - \pi)\pi(R_H - R_L) < Q$. Again, the arguments above imply that any alternative (not mainstream) strategy on topic selection can only be preferred by κ groups of consumers and quality on any topic is bounded by $C^{-1}([p_l + p_h]N\kappa/n)$ which again goes to 0 as $n \rightarrow \infty$. Under these conditions,

the maximum utility a consumer would get from an alternative strategy goes to $0 + (\pi + (1 - \pi)\pi)(R_H - R_L) + R_L$ which is smaller than $Q + \pi(R_H - R_L) + R_L$ by assumption. We therefore conclude that no alternative viable strategy can attract consumers away from the mainstream firm and the situation where multiple firms play the mainstream strategy where only one is chosen by consumers, by arguments similar to the previous paragraph, is an equilibrium.

We now go on to characterize all equilibria. Define $Q' = Q - (1 - \pi)\pi(R_H - R_L)$. Q' is the minimum average quality a firm must provide to a consumer in order for him to possibly prefer that firm over the mainstream firm. The firm therefore must generate a revenue of at least $C(Q')$ in order to be able to pay for that level of quality, which means that it must attract some minimum number $m > C(Q')/(p_l + p_h)$ of consumers. Let $\nu = N/m$ and notice that, since each firm can only provide higher relevance than the mainstream firm for at most κ groups of consumers, there can be at most $\nu\kappa N/n$ consumers that receive higher relevance under their current firm. As $n \rightarrow \infty$, this number goes to 0 and $\mu = (N - \nu\kappa N/n) \rightarrow N$.

For the situation to sustain an equilibrium then these consumers must receive strictly higher average quality than $Q'' = [p_l + \pi p_h]\mu$. Since $2C(Q'') \rightarrow 2C(Q) = 2[p_l + \pi p_h]N > [p_l + \pi p_h]N + (1 - \pi)p_h N/n$, the market cannot sustain more than one active firm.

Now, similarly to above, since this firm revenue is equal to $E_r[\sigma(b; r)(p_l + r_b p_h)N + \sum_{k \neq b} \sigma(k; r)(p_l + r_k p_h)N/n] = E_r[(1 - \sum_{k \neq b} \sigma(k; r))(p_l + r_b p_h)N + \sum_{k \neq b} \sigma(k; r)(p_l + r_k p_h)N/n]$, we have that the derivative of revenue is always negative away from strategy σ^* . This means that a firm playing any other strategy would have strictly lower revenue and therefore strictly lower quality than Q . In order to maintain equilibrium as $n \rightarrow \infty$, we must have that this firm's strategy approaches (Q, σ^*) . \square

A.2

Proof of Proposition 2

First, notice that consumers always prefer a relevant topic over a not relevant one, since $R_H > Q + R_L$ and Q is the maximum quality that can be offered with non-negative profits. From that, it follows that in equilibrium all consumers will only be shown a low relevance topic if no relevant topic is available, because otherwise a firm could enter offering 0 quality on the relevant topic and make a profit attracting consumers through the aggregator.

However, because of the constraints on firm technology, a single firm can only offer maximal relevance to consumers of a single type (firms need to decide which topic to feature when both are relevant). We must then have that there

are at least n active firms and consumers must either use the aggregator or a firm specializing in their type. Since maximal quality is only be achieved with one active firm per topic, we have that the number of active firms that actively cover niche topics is exactly n .

Now consider the provision of the broad topic. Since a firm must either attract consumers through the aggregator or consumers of a single type and no consumer is shown the broad topic when it has a relevance disadvantage over the niche topic, total revenue from the broad topic is bounded by $\rho = (1 - \tau)[(1 - \pi(1 - \pi))p_l + \pi p_h]N$, for sufficiently large n . We must then have that a firm enters the market supplying the broad topic for every realization with quality $Q^A = C^{-1}(\rho)$, for either providing the broad topic in less realizations or with higher quality would lead to negative profits, while providing the broad topic with lower quality could not be sustained in equilibrium since another firm could take the market and make positive profits.

From that, we conclude that all consumers use the aggregator (in order to be shown the broad topic at high quality and the niche topic at high relevance) and niche topics are supplied at a quality of $Q^n = (1 - \tau)[\pi(1 - \pi)(p_l + p_h)]N/n$.

We therefore concluded that all consumers receive utility equal to:

$$\pi(Q^A + R_H) + \pi(1 - \pi)(Q^n + R_H) + (1 - \pi)^2(Q^A + R_L)$$

□

A.3

Proof of Proposition 3

By what we have established above, when $(1 - \pi)\pi(R_H - R_L) \geq Q$, all consumers receive maximal relevance. In the traditional market, they are served by a specialized firm offering a level of quality which goes to 0 as n grows. When consuming the aggregator, consumers can instead consume the broad topic at quality level $Q^A = C^{-1}((1 - \tau)[(1 - \pi(1 - \pi))p_l + \pi p_h]N)$ which is strictly greater than 0. Therefore, quality and welfare increases under the aggregator.

However, when $(1 - \pi)\pi(R_H - R_L) < Q$, then consumer welfare in the traditional market approaches $Q + \pi(R_H - R_L) + R_L$ while welfare under the aggregator approaches $[1 - \pi(1 - \pi)]Q^A + (\pi + \pi(1 - \pi))(R_H - R_L) + R_L$, which immediately yields the inequality. It is easy to see as well that $Q > Q^A$ and therefore, similarly to the case above, quality is decreasing.

A.4

Proof of Proposition 4

As before, in the traditional market, consumers of a same type cannot be split between multiple firms since, if they were, another firm could enter offering a higher quality and attracting all consumers of a certain type. This leads directly to the fact that firms act as monopolists in the advertising market and set ad quantities and prices at the monopoly level. Since ad prices are fixed at this level, equilibria are the same as described under the model with exogenous prices.

With the entry of the aggregator, the situation is also similar to the case with exogenous ads. Indeed, under any price for ads, since consumers always prefer to read relevant news over irrelevant news, again the market necessarily sustains at least n firms, each covering one of the niche topics. This is because each topic will be covered by only one firm, since if it were covered by more than one, another firm could enter at a higher quality. This of course includes the broad topic, which is covered at a high level of quality. Since in each state of the world only one firm grabs the attention of consumers, firms do not compete on the advertising markets, since advertisers maintain the same level of interest in all of their products.

The aggregator does enter as a competitor. To solve for the level of ads, we have that firms FOCs are:

$$\begin{aligned} A &= (1 - \tau)a_k + \tau a_a \\ a_k &= \frac{\omega(A)}{-\omega'(A)} \frac{1}{(1 - \tau)} \\ a_a &= \frac{\omega(A)}{-\omega'(A)} \frac{1}{\tau} \end{aligned}$$

Where a_k is the advertising intensity of the active news media firm while a_a is the advertising intensity of the aggregator. This implies that $A > a_m$, where a_m is the monopoly ad level that solves $a_m = \frac{\omega(a_m)}{-\omega'(a_m)}$, since $\omega'(A)A + 2\omega(A) = 0 \Rightarrow \omega'(A)A + \omega(A) < 0 \Rightarrow A > a_m$. The revenue each firm can extract from a single consumer is $\omega(A)A/2$ which is less than the monopolist revenue. Therefore, firms invest less in quality.

When Nash bargaining is introduced, the FOCs for equilibrium in the ads market is:

$$\begin{aligned} A' &= (1 - \tau)a'_k + \tau a'_a \\ a'_k &= \frac{\omega(A')}{-\omega'(A')} \frac{1}{(1 - \tau)} - \frac{\tau}{2(1 - \tau)} a'_a \end{aligned}$$

$$a'_a = \frac{\omega(A')}{-\omega'(A')} \frac{1}{\tau}$$

Which implies that $a_m < A' < A$, since $\omega'(A')A' + \frac{3}{2}\omega(A') = 0 \Rightarrow \omega'(A') + \omega(A') < 0$ and $\omega'(A') + 2\omega(A') > 0$.

Therefore news firms make more revenue than in the situation without bargaining, but not as much as without the aggregator.

A.5

Equilibrium with increasing market size

We assume that the distribution G and cost function C are such that the revenue and cost functions only cross once (ie G is concave and C is convex). The problem of the firms remain essentially the same, and analogous versions of Proposition 1 and 2 continue to hold. When the aggregator is in the market, equilibrium in quality is characterized by:

$$\sum_{k \in T \setminus \{b\}} \frac{N}{n} G(E_r[U_k(q, \sigma, r)]) E_r[R_k^b] = C(q_b)$$

$$\frac{N}{n} G(E_r[U_k(q, \sigma, r)]) E_r[R_k^k] = C(q_k) \quad \forall k \in T \setminus \{b\}$$

And it is easy to see that there is an unique best response by each firm, which is increasing in the strategy of others. Given this, we can establish that the game has an equilibrium by Tarski's theorem (Tarski (1955)).