Abstract

The growing influence of internet platforms acting as content aggregators is one of the most important challenges facing the media industry. We develop a simple model to understand the impact of third-party content bundling by a social platform that has a monopoly on showing user-generated content to consumers. In our model consumers can access news either directly through a newspaper’s website, or indirectly through a platform, which also offers social content. We show that content bundling, when unilaterally implemented by the platform, tends to harm publishers and to increase the dispersion of quality across outlets. News quality is more likely to increase with content bundling when the cost of providing quality is low, and when competition among publishers is strong. When content bundling follows an agreement between the platform and publisher, its effects are reversed, as publishers’ profits go up while quality dispersion goes down. In a setup with heterogeneous consumers, we also show that the platform’s ability to personalize the mix of content it shows to users induces publishers to invest more in the quality of their content.

1 Introduction

With hundreds of millions of daily active users, a few large social networks have become the dominant online media outlets for most people. The largest among these, Facebook
has reached over two billion active members across the globe who, on average, spend about an hour each day on the platform. In line with its significant consumer attention share, Facebook captured almost $70 billion worth of advertising revenues in 2019, corresponding to 21% of total worldwide digital advertising. Other successful social platforms include Tencent’s WeChat, Snap and ByteDance’s Tik Tok, among others. If, in their early days, social networks were mostly used as a way for users to share personal stories and pictures (which we refer to as user-generated content, or UGC, throughout the paper), their role has progressively evolved into one of content aggregation: an important share of the content displayed on their websites is produced by third-party publishers, who use the platforms as an alternative to their own website to reach consumers.

For many third-party content providers, large social networks represent an important distribution outlet. The news industry, in particular, has been strongly affected by such ‘content bundling’ and given the industry’s public calling this impact has been relatively well documented. In the U.S., 55% of adults get their news from social media “often” or “sometimes” in 2019, up from 44% in 2016 (Pew Research Center, 2019). Facebook alone accounts for a substantial part of traffic to publishers: a report by the U.K. Competition and Markets Authority puts the figure at 13%. Social media is the second main gateway to news, behind direct traffic to publishers, with shares of 26% and 28% respectively (Reuters, 2020). For consumers aged 18-24, social media is the main gateway (38% against 16% for direct traffic, with search coming in second at 25%).

The central question that we ask in this paper is: how does content bundling by a platform affect publishers’ incentives to invest in the quality of their content? From an economic standpoint, the situation is a double-edged sword for publishers: while social platforms provide the opportunity to reach a wider audience, newspapers regularly complain that their business is hurt by social media, citing for instance the fact that Facebook’s market share (including Instagram) on the online display advertising market is around 40% while newspapers’ advertising revenues have plummeted. Another set of concerns is that platforms have “commoditized” news: in one survey, nearly half of U.S. respondents were unable to identify the source of a news story they had accessed through social media (Reuters, 2017), leading to fears of a phenomenon of brand dilution. Moreover, publishers also complain that platforms only give them access to aggregated data generated by consumers interacting with their content through the platforms, preventing them from offering better targeting services (Competition and Markets Authority, 2020, Appendix S).

For the above reasons, the general view is that content bundling has a negative effect

---

1For mobile the figure is 17% with some heterogeneity among publishers, (see Competition and Markets Authority, 2020, Appendix S), but this number understates Facebook’s role, as a large number of consumers (47% in the U.K. in 2016) browse and read news on social media without clicking on links to access whole articles. For some, so-called ‘social native’ news providers, explicitly designed for a Facebook audience, content bundling is even more critical. For example, Buzzfeed, a leading online publisher derives 75% of its traffic from Facebook.
on newspapers’ incentive to invest in news quality. An often mentioned manifestation of lower news quality is the documented decline of newspapers’ newsroom sizes and of the overall number of journalists (see discussions by Cagé, 2016; Fan, 2013). In fact, such concerns have led lawmakers and regulators to take actions aimed at ensuring that news publishers get a larger share of the industry revenue: In 2019 the European Parliament adopted a Directive on Copyright in the Digital Single Market whose article 15, referred to as the “link tax” article, grants publishers direct copyright over “online use of their press publications by information society service providers” (following similar - unsuccessful - attempts in Spain and Germany). Similarly, Australian regulators have launched a consultation regarding a mandatory news media bargaining code, which would force Facebook and Google to bargain with news publishers (potentially as a collective) on matters such as revenue sharing, data sharing and algorithmic curation of news.

The latter point is indeed an important feature of social media platforms: thanks to the rich trove of data they collect, platforms have the ability to offer a personalized mix of content to their users, by identifying what kind of content keeps them engaged. This then leads consumers to spend more time on platforms, and to rely on them as “personalized curators” of content. It is not clear a priori whether the trend towards more personalization will lead to a softened or intensified competition between platforms and publishers (and thus to lower or higher news quality).

To study these issues, we develop a model in which a social platform and a newspaper (or publisher), both advertising-supported, compete for consumers’ attention. The newspaper produces news stories and maintains a website, which only offers news content. The social platform relies on its users to produce UGC, such as personal stories or pictures. On its website, the platform can bundle UGC with content produced by the newspaper, in which case the platform and the newspaper share advertising revenues. The platform can choose how much news to display along its own content, even though it needs to show a minimal share of UGC.

In our baseline framework, consumers are homogeneous in their demand for news, which depends on the news content’s quality. Quality is the result of an investment by the newspaper (we model quality as a demand-enhancing investment). Consumers can freely allocate their attention across the two websites, but, when on the platform, have to consume the mix of content that is offered to them (the platform plays a content curator role).

We start by comparing a benchmark where the platform cannot show news to a set-up where it can unilaterally adopt a content bundling strategy (i.e. the publisher cannot prevent content bundling). In this simple model with homogenous consumers, the publisher is always worse-off under unilateral content bundling. However, the fact that content

\[2\] Applying this directive, in 2020 the French Competition Authority enjoined Google to negotiate payments with publishers “in good faith”.

3
bundling makes the newspaper less profitable does not lead to an unambiguous decline in content quality. In fact, we find that content bundling has a negative effect on quality if and only if the cost of quality is large. In other words, our model predicts that a high-quality newspaper (with a lower cost of producing quality) should invest more under content bundling, while a low-quality one should invest less. Intuitively, when the cost of quality is high, and thus the quality of news is low, consumers have a low demand for the publisher’s content, and this demand can be satisfied by a few links on the platform. In this case, consumers only visit the platform’s website (they single-home). The returns to investment in quality are relatively small, because a higher quality only increases the share of news content shown on the platform rather than attracting consumers to the publisher’s website. When the cost of quality is low and quality is therefore high, consumers’ demand for the publisher’s content cannot be entirely satisfied by the platform (who cannot merely replicate the whole newspaper’s content). Consumers then multi-home between the platform and the publisher’s website, and direct traffic is more responsive to quality improvements than without content bundling. Indeed, because the platform shows some news content, increasing one’s consumption of news by one unit of time requires reallocating more than one unit of time from the platform’s website to the publisher’s. This core finding is important in a context of general concern for the quality of news in the age of social media. Our results indicate that the effect of social networks on news providers is not uniform, and could result in an increase in the dispersion of quality.

The results extend to a framework with subscription pricing by the publisher, and to one with competing publishers with heterogeneous costs: low-cost newspapers’ qualities increase under content bundling while the opposite is true for high-cost ones. When competition among newspapers is high, content bundling is more likely to lead to more investments in quality.

Motivated by the recent concerns with respect to the news industry we also assess the likely effect of two policy initiatives: that of a link tax and putting pressure on platforms to negotiate a compensation for the publishers whose content they use (negotiated content bundling). We find that both policies lead to results that are opposite to unilateral content bundling: while the publisher is better-off under both a link tax and negotiated content bundling, quality dispersion declines compared to the benchmark, as low-quality publishers invest more than under the benchmark, while high-quality ones invest less.

Next, we enrich the baseline model by introducing consumer heterogeneity and (imperfect) personalization of the newsfeed by the platform. Under this more realistic setup, unilateral content bundling still increases the dispersion of quality. Additionally, it may now benefit the publisher, as it increases news consumption (a market expansion effect). Moreover, we also show that an increase in the platform’s personalization capabilities induces the publisher to invest more in the quality of its content. A strategic platform may therefore commit not to fully personalize consumers’ newsfeed so as to limit the intensity
of competition with the publisher.

The next section summarizes the relevant literature. This is followed by the description and detailed analysis of a base model with homogeneous consumers. Section 5 extends the analysis to multiple publishers with heterogeneous costs while Section 6 introduces consumer heterogeneity and imperfect personalization by the platform. The paper ends with a general discussion of the results and concluding remarks.

2 Relevant literature

Internet and the news  The Internet has had a dramatic effect on the news industry, with a joint decrease of circulation and advertising revenue for the printed press (see Peitz and Reisinger, 2015, for an overview). In their online transition, newspapers have also experienced various challenges, among others self-cannibalization (Gentzkow, 2007), consumer switching behavior (Athey, Calvano, and Gans, 2018),\(^3\) or copyright violation (Cagé, Hervé, and Viaud, 2017).

An issue which is particularly relevant to this paper is the emergence of news aggregators, such as Yahoo News or Google News (see Jeon, 2018, for a survey). As in our paper, the central question is how these intermediaries impact the consumption of news as well as the quality of content produced. On the theory side, Jeon and Nasr (2016) and Dellarocas, Katona, and Rand (2013) model aggregators as enabling consumers to find high-quality news more easily. They find that the entry of an aggregator tends to increase competition among websites, leading to higher quality. The impact on newspapers’ profit depends on which effect is stronger: business stealing or market expansion. Rutt (2011) studies how the presence of an aggregator affects newspapers’ choice of business model, and shows that it has different effects on the quality provided by free versus paying outlets. In George and Hogendorn (2012), the aggregator reduces the cost of multi-homing for consumers. In contrast to the way we model the platform’s behaviour, in these papers, aggregators are non-strategic and do not produce their own content, but merely replicate the experience of a newspaper.

A recent series of empirical papers examine the impact of aggregators on the news industry. Using disputes between Google News and Spanish publishers (Athey, Mobius, and Pal, 2017; Calzada and Gil, 2016) or the Associated Press (Chiou and Tucker, 2017), empirical research finds that Google News increases overall news consumption. In particular, Athey, Mobius, and Pal (2017) document that this effect is mostly present for small publishers, who cannot rely on brand recognition to attract users and therefore benefit most from the aggregator. In relation to the theoretical work on aggregators,\(^3\) see also Ambrus, Calvano, and Reisinger (2016) and Anderson, Foros, and Kind (2018) for the study of consumer multi-homing.
these papers suggest that the demand-expansion effect of aggregators dominates. George and Hogendorn (2013) study the consequences of a redesign of Google News, and find that news aggregators can potentially also change the composition of news consumption. Sismeiro and Mahmood (2018) study the impact of social media on news consumption using a global outage of Facebook in 2013, and show that the outage reduces traffic to newspapers’ websites and changes the types of articles that consumers read.

In our model, the platform allocates consumers’ attention by choosing the mix of content that it displays. In this respect it is similar to a search engine, which allocates traffic through its ranking and design (see de Cornière and Taylor, 2014; Burguet, Caminal, and Ellman, 2015). However, in these papers the intermediary enjoys an exogenous bottleneck position: consumers have to use the search engine to find content. In contrast, our mechanism is one where the allocation of attention while on the platform (i.e. content bundling) determines how consumers allocate their attention between the platform and the newspaper. The gatekeeping role of the platform thus emerges endogenously.

A few recent papers explicitly examine the relationship between social media and the news industry. Social platforms have been accused of fostering echo chambers and spreading fake news. Some critics argue that platforms should be held responsible for the content displayed on their websites (see Rolnik et al., 2019, for a recent overview of the issues). Abreu and Jeon (2020) and Campbell, Leister, and Zenou (2019) study how the distribution of news through a social network can lead to more polarized content being produced. Berman and Katona (2020) look at the effect of various curation algorithms on content quality and link formation in a social network. Our work is related as it focuses on the fundamental effect of content bundling on publishers’ incentive to invest in content quality.

As Alaoui and Germano (2020), we also assume that consumers are time constrained in their consumption of media and our results resonate with theirs in that competition between content suppliers (including the social network) may distort consumers’ media consumption (in our case, when there is imperfect personalization). However, we focus on consumers’ time allocation across qualitatively different content providers and we abstract away from the editorial process of publishers when multiple topics are present.

**News quality** In practice, what constitutes news quality is a complex issue (Lacy and Rosenstiel, 2015), with many relevant dimensions.\(^4\) Empirical researchers have used the number of pages, size of the newsroom and the number of Pulitzer prize winners per staff member as measures of quality (see, e.g. Berry and Waldfogel (2010) and Angelucci and Cagé (2016)). In this paper, we define quality as a demand-enhancing investment, which applies to the content itself (e.g. better investigative journalism). Notice that

\(^4\)Bogart, 2004 cites characteristics such as “integrity, fairness, balance, accuracy, comprehensiveness, diligence of discovery, authority, breadth of coverage, variety of content, reflection of the entire home community, vivid writing, attractive makeup, packaging or appearance, and easy navigability”. 

6
this approach does not encompass situations where consumers have heterogeneous views on what constitutes high quality as, for example, in the literature on media bias (see Gentzkow, Shapiro, and Stone, 2015; Puglisi and Snyder Jr, 2015, for recent surveys of media bias).

3 Baseline model

**Content and preferences** We consider a model where consumers can consume two kinds of content: news and UGC. UGC is produced by a monopolist social platform (indexed by 0), at no cost, and its quality is exogenous. News are produced by a monopolist publisher, indexed by 1.

Consumers are homogeneous in their relative preferences between news and UGC (an assumption we relax in Section 6). If we denote by $x_U$ the consumption of UGC and $x_N$ that of news, we assume that consumers’ utility function is

$$U(x_U, x_N, q) = x_U + (1 + q)x_N - \frac{(x_N)^2}{2}.$$ (1)

Here $q$ denotes the quality of the publisher’s content, which is produced at cost $C(q) = \frac{cq^2}{2}$. Equation 1 captures the ideas that consumers enjoy both kinds of content, that their relative preference for news is increasing in its quality, and that news offers diminishing marginal utility. Consumers are time constrained and we assume that they have one unit of time to allocate between consumption of news and UGC. Maximizing $U$ under the constraint that $x_U + x_N \leq 1$, we obtain the demand for news of quality $q$: $D_N(q) = q$. Throughout the paper we assume that $c$ is large enough that we always have interior solutions (that is $q \in (0, 1)$).

**Content bundling** Consumers divide their time between the platform and the publisher’s website. While the publisher’s website only shows news, the platform has a content curating role and may offer a mix of UGC and news, a strategy we refer to as content bundling. More precisely, for each unit of time spent on the platform’s website, consumers are exposed to a share $\lambda$ of news content, and a share $1 - \lambda$ of UGC, where $\lambda$ is chosen by the platform. We assume that the platform is restricted to choosing $\lambda < \Lambda$, where $\Lambda \in (0, 1]$ is an exogenous bound on the amount of news it can show. The parameter $\Lambda$ captures in a reduced-form way the idea that the platform needs to have some original content to be successful, and cannot only show content produced by the publisher. A limit

---

5See also Angelucci and Cagé (2016) for a model where newspapers can invest more in “hard news”, which is a horizontal feature.

6As we show in Section 5, our results also hold with several heterogeneous newspapers. Having one newspaper corresponds to the case of a highly differentiated news outlet, while also allows us to obtain closed-form solutions.
on Λ also captures the idea that the platform has market power as long as it can claim to be the sole provider of UGC.\textsuperscript{7} Therefore, if a consumer spends $T_0$ units of time on the platform and $T_1 = 1 - T_0$ units of time on the publisher’s website, his consumption of news equals $x_N = T_1 + \lambda T_0$. We call $\lambda T_0$ the indirect consumption of news.

We start by considering the case of what we call “unilateral content bundling”, in the sense that the publisher has no control on how much of its content the platform bundles in its newsfeed. Later, we examine the case when the publisher can deny its content to the platform and the two parties engage in bargaining about content sharing (“negotiated content bundling”).

**Revenues** We assume that both the platform and the publisher are purely advertising supported.\textsuperscript{8} Moreover, we assume the following: (i) the publisher is no more efficient than the platform at monetizing attention, because for instance the former sells ad space through an intermediary while the latter is vertically integrated (“owned and operated advertising supply”); (ii) the publisher generates more revenues from direct consumption than from indirect consumption of news (for instance because of “brand dilution”, or because indirect news consumption does not allow publishers to obtain data about consumers); (iii) the platform generates more advertising revenues (per unit of time) from showing UGC than from showing news, either because of revenue sharing or because of the risk that a consumer who clicks a link to the publisher does not come back to the platform.

Formally, we normalize to 1 the platform’s revenue when it shows UGC (all revenues are expressed per unit of time). Its revenue when it shows news is $\phi_0 \leq 1$, as per assumption (iii) above. We denote the publisher’s revenue from direct traffic as $\beta \leq 1$, as per assumption (i) and its revenue from indirect news consumption is denoted $\phi_1 < \beta$, as per assumption (ii).

Overall, for given $T_0$, $T_1$ and $\lambda$, the platform’s revenue equals $R_0 = (1 - \lambda + \lambda \phi_0)T_0$, while the publisher’s revenue is $R_1 = \beta T_1 + \phi_1 \lambda T_0$. The platform’s profit equals its revenue ($\pi_0 = R_0$), and the publisher’s profit is $\pi_1 = R_1 - \frac{\phi_1^2}{2}$.

For now, we interpret $\phi_0$ and $\phi_1$ as exogenous parameters that depend on the characteristics of the advertising technologies used by the firms, and not as the outcome of a bargaining game between the publisher and the platform (see Section 4.4).

**Timing and equilibrium** In the first stage of the game, the publisher chooses the quality of its content $q$. In the second stage, the platform chooses how much news to incorporate to its own content, $\lambda$. In the third stage, consumers choose how to allocate their attention across the two websites. We look for subgame perfect equilibria.

\textsuperscript{7}The unconstrained case where $\Lambda = 1$ is a special case of our analysis. Allowing $\Lambda < 1$ enriches the analysis, by allowing us to consider shifts in the overall strategy of a platform as, for example, when Facebook announced in 2019 that it would show less news content going forward.

\textsuperscript{8}Our results continue to hold if the publisher can charge a subscription price.
4 Equilibrium

4.1 Benchmark without content bundling

We start by providing the analysis of the case where the platform cannot display content produced by the publisher on its own website. Formally, this corresponds to $\Lambda = 0$. We denote by $\hat{q}$ the equilibrium quality.

In the last stage of the game, given a quality $q$, consumers allocate their attention so as to satisfy their demand for each type of content, $D_N = q$ and $D_U = 1 - q$, choosing $T_1 = q$ and $T_0 = 1 - q$. The publisher’s profit is $\beta q - \frac{c q^2}{2}$. Maximizing this profit over $q$, we obtain:

Lemma 1. In the benchmark case without content bundling, the equilibrium quality of news and the publisher’s profit are\[ \hat{q} = \frac{\beta}{c}, \quad \text{and} \quad \hat{\pi}_1 = \frac{\beta^2}{2c}. \]

4.2 Equilibrium with unilateral content bundling

Let us now solve for the equilibrium of the game with content bundling through backward induction.

Stage 3: Consumers Suppose that quality of news is $q$ and that the platform has chosen $\lambda$. Then, as we have seen, consumers’ desired share of news consumption is also $q$. We need to distinguish two cases: (i) If $\lambda \geq q$, the optimal allocation of attention is $T_0 = 1$ because by spending all their time on the platform, consumers already obtain (weakly) more news than they would like. (ii) If $q > \lambda$, consumers must spend a positive amount of time on each website to consume their optimal mix of content. The optimal allocation of attention $(T_0, T_1) = (1 - T_1, T_1)$ solves\[ q = T_1 + \lambda T_0 = \lambda + T_1(1 - \lambda). \]

Lemma 2. Allocation of consumers’ attention Given a quality $q$ and a share of news shown on the platform $\lambda$, the share of time that consumers spend on the publisher’s website is $T_1(q, \lambda) = \max\{\frac{q - \lambda}{1-\lambda}, 0\}$.

Stage 2: Platform The platform pursues a dual objective, which is to maximize the time consumers spend on its website while minimizing the amount of external content it shows them. Here again we need to distinguish two cases:

If $q < \Lambda$, notice first that it is never optimal to set $\lambda > q$: indeed, such a strategy is dominated by setting $\lambda = q$, as this would still maximize $T_0$ while reducing the quantity
of news showed on the platform. Second, setting $\lambda < q$ is also suboptimal: in that case, consumers would adjust their allocation of attention so as to consume a share $q$ of news and $1 - q$ of UGC, but some of the news consumption would occur on the publisher’s website. Setting $\lambda = q$ leads to consumers receiving a share $q$ of news, exclusively through the platform.

If $q > \Lambda$, the platform cannot offer the optimal mix on its website. The best it can do is to set $\lambda = \Lambda$. Any lower $\lambda$ would lead to the same mix of content being consumed, with more time spent on the publisher’s site.

**Lemma 3. Platform’s strategy** The optimal strategy for the platform is to set $\lambda = \min\{q, \Lambda\}$.

**Stage 1: Publisher** Given the above analysis, the publisher’s profit can be written as follows:

$$\pi_1(q) = \begin{cases} 
\phi_1 q - \frac{c q^2}{2} & \text{if } q \leq \Lambda, \\
\phi_1 \Lambda \frac{1}{1 - \Lambda} + \frac{\beta}{1 - \Lambda} - \frac{c q^2}{2} & \text{otherwise.}
\end{cases}$$

(B2)

Broadly speaking, the publisher has two strategies. It can either choose a relatively low quality ($q \leq \Lambda$), in which case its content will be consumed exclusively through the platform. Such a strategy therefore induces consumers to single-home on the platform and it resonates to the emergence of so-called “social native” publishers whose primary distribution is via social networks. Alternatively, it can choose a relatively high quality ($q > \Lambda$) in order to attract some direct traffic (equal to $\frac{q - \Lambda}{1 - \Lambda}$), thereby leading to consumers multihoming, which is the typical strategy of established news publishers. The relative profitability of these strategies depends on the cost of producing quality, captured by the parameter $c$.

Maximizing (2) leads to the following result:

**Lemma 4. Equilibrium quality** Let $q_L \equiv \frac{\phi_1}{c} < q_H \equiv \frac{\beta - \Lambda \phi_1}{c(1 - \Lambda)}$. Let $c^* \equiv \frac{\beta + \phi_1(1 - 2\Lambda)}{2\Lambda(1 - \Lambda)}$.

The equilibrium quality is $q^* = q_H$ if $c \leq c^*$, and $q^* = q_L$ otherwise.

**Proof.** The two possible values of $q^*$ are the maximizers of the two expressions in (2).

If $\Lambda < q_L$ (i.e. $c < \frac{\phi_1}{\Lambda}$), profit is continuous and increasing for $q \leq \Lambda$ (given by $\phi_1 q - \frac{c q^2}{2}$) and again for $q \in [\Lambda, q_H]$ (given by $\phi_1 \Lambda \frac{1}{1 - \Lambda} + \frac{\beta}{1 - \Lambda} - \frac{c q^2}{2}$), and decreasing for $q > q_H$. The optimal quality is thus $q_H$.

When $\Lambda > q_H$ (i.e. $c > \frac{\beta - \phi_1 \Lambda}{\Lambda(1 - \Lambda)}$), profit is increasing over $[0, q_L]$ and decreasing for $q > q_L$, which implies that the optimal quality is $q_L$.

When $\Lambda \in (q_L, q_H)$ (i.e. $c \in (\frac{\phi_1}{\Lambda}, \frac{\beta - \phi_1 \Lambda}{\Lambda(1 - \Lambda)}))$, profit has two local maximizers, at $q_L$ and $q_H$. Comparison of the value of profit at these two points reveal that $q_H$ dominates $q_L$ for $c < c^*$.
4.3 Effects of content bundling and comparative statics

Now, we are ready to evaluate the effect of content bundling on newspapers’ content quality. Comparing Lemmas 1 and 4, we can state the following result:

**Proposition 1.** Compared to a benchmark without content bundling:

(i) Content bundling lowers the publisher’s profit.

(ii) The quality of the publisher’s content and consumer surplus increase if and only if $c < c^*$.

**Proof.** (i) With or without content bundling, news consumption is equal to $q$. Because under content bundling part of the consumption occurs through the platform, revenue is always lower under content bundling for a given $q$. Let $\hat{R}_1(q)$ and $R_1^*(q)$ be the respective revenue functions in the benchmark and with content bundling for a given $q$. We then have, for $q = q^*$, $\hat{R}_1(q^*) > R_1^*(q^*)$. Subtracting $C(q^*)$, we obtain $\hat{R}_1(q^*) - C(q^*) > \pi_1^*$. By revealed preferences, $\hat{\pi}_1 \geq \hat{R}_1(q^*) - C(q^*)$, which implies that $\hat{\pi}_1 > \pi_1^*$.

(ii) Because $\beta > \phi_1$, it is straightforward to check that $q_L = \frac{\phi_1}{c} < \hat{q} = \frac{\beta}{c}$ and that $q_H = \beta - \Lambda \phi_1 > \hat{q}$. Because consumers always get their optimal content mix, consumer surplus moves in the same direction as quality of news.

The first part of Proposition 1 is fairly intuitive: because content bundling does not increase news consumption but diverts some consumption onto the platform, it cannot benefit the publisher. Note that this result does not always hold with heterogeneous consumers, as we will see in Section 6.

The intuition for the second part of the Proposition is as follows: when $c$ is large, content bundling leads the publisher to adopt a low-quality strategy leading consumers to single-home on the platform, where the marginal revenue is $\phi_1$, which is less than the marginal revenue from direct traffic ($\beta$). Publishers who rely exclusively on the platform (i.e. social native publishers) therefore invest less than they would absent content bundling.

The story is different for publishers with a low cost of quality. Under content bundling, they compete with the platform to attract direct traffic. It turns out that the amount of direct traffic under content bundling is more sensitive to quality than under the benchmark. Indeed, when a consumer wishes to increase his news consumption by 1 unit of time under content bundling, he needs to increase the time he spends on the publisher’s site by $1/(1 - \Lambda) > 1$, because the time he spends on the platform already allows him to consume news. While the value of direct traffic is also lower because of indirect revenues ($\beta - \Lambda \phi_1$ instead of $\beta$), which we refer to as a softening effect, the overall marginal revenue is $\frac{\beta - \Lambda \phi_1}{1 - \Lambda} > \beta$, giving the publisher incentives to increase its quality.

Using a broader interpretation of our model, namely that there are multiple publishers that differ in their cost of producing quality (see Section 5 for a formal treatment of that case), Proposition 1 suggests that content bundling has a heterogeneous effect on
newspapers: for high-quality newspapers (with low $c$), content bundling increases the incentive to invest in quality, while for low-quality newspapers it reduces their incentive. A testable prediction of the model is then that content bundling should increase the variance of the distribution of quality.

**Comparative statics** Let us now examine the effect of the parameters of the model on the equilibrium under content bundling.

First, let us consider the effects of $\phi_1$, the indirect monetization parameter. An increase in $\phi_1$ can be viewed as a way to model a “link tax”, forcing the platform to pay the publisher as is considered in recently discussed policy proposals.\(^9\) If the publisher has a low cost ($c < c^\dagger$) and therefore a higher quality, its incentive to invest in quality goes down with $\phi_1$ (i.e. $\frac{\partial q_H}{\partial \phi_1} < 0$) because of what we call the “softening effect”: as indirect monetization improves, attracting direct traffic becomes relatively less profitable. On the other hand, the incentives to produce higher quality increase with $\phi_1$ if the publisher has a high cost ($c > c^\dagger$) because in this case all news is consumed on the platform and a higher quality increases indirect traffic only. In summary, a link tax does not lead to a general increase in quality. Rather, it leads to a reduction in newspapers’ quality dispersion. We expand on the idea of policy interventions aimed at compensating publishers in the next subsection and in Appendix A.2 (where we use the model with heterogeneous consumers of Section 6). Results are qualitatively similar there.

Next, an increase in the advertising revenue from direct traffic $\beta$ makes it more profitable for the publisher to adopt a “high quality” strategy ($\frac{\partial q_H}{\partial \beta} > 0$), and the corresponding quality also increases ($\frac{\partial q_H}{\partial \beta} > 0$). If the publisher adopts the “low quality” strategy, its quality choice is unaffected by $\beta$.

Finally, one may look at the effects of a change in $\Lambda$, which one could interpret as a strategic choice by the platform to give relatively more or less prominence to externally produced content, as for instance when Facebook announced that it would reduce the share of news in the newsfeed in 2019. We have seen that the platform is not entirely free to set $\Lambda$ as it has to show some UGC to retain its identity and/or market power. Nevertheless, in the model, a reduction in $\Lambda$ would lead to a lower $q_H$, because the reduced sensitivity of direct traffic would more than offset the increased relative value of such traffic.\(^{10}\) Notice that, starting from $\Lambda = 0$, a small increase in $\Lambda$ increases quality for all types of publishers (because $c^\dagger \rightarrow +\infty$ when $\Lambda \rightarrow 0$).

These results are summarized below:

**Proposition 2.** Assume that $\Lambda < 1/2$. Under unilateral content bundling, comparative statics results are given by the following table:

---


10The logic is similar to the one explaining that quality increases under content bundling for $c < c^\dagger$. 
Subscription pricing by the publisher  Our framework assumes an advertising revenue model for both the publisher and the platform. It is important to realize however, that assuming a mixed revenue model by the publisher, which also includes a subscription fee does not change our results. Suppose indeed, that the publisher can set a price \( p \) that consumers must pay in order to access its content. A consumer who does not pay \( p \) only consumes UGC and gets a surplus of 1 (by Equation 1). Because content bundling does not distort consumers’ content mix, for a given quality \( q \) the gross consumer surplus is the same with and without content bundling, equal to \( 1 + \frac{q^2}{2} \). Therefore consumers’ willingness to pay to access a publisher of quality \( q \) is not affected by content bundling, and is equal to \( v(q) = \frac{q^2}{2} \). Because the publisher can set \( p = v(q) \), it maximizes \( v(q) + R_1(q) - c(q) \), where \( R_1 \) is the only thing changing with content bundling. It follows that Proposition 1 continues to hold with subscription pricing.

4.4 Negotiated content bundling

The analysis so far shows that the publisher is always worse-off under unilateral content bundling (i.e. when the platform is free to display the publisher’s content) compared to a benchmark without content bundling. What happens however, when the publisher can deny its content to the platform and the latter is forced to bargain with the publisher? Policy initiatives pushing social platforms to negotiate with news publishers have been proposed by regulators multiple times and more recently, by regulators in France and Australia, in particular. Moreover, Facebook has explicitly explored this avenue, notably when it introduced Instant Articles in 2015 and Facebook News in 2019.

To understand the economics of such a context, suppose that the timing is the following: in the first stage, the publisher chooses its quality \( q \). In the second stage, the publisher and the platform engage in Nash bargaining (with equal bargaining power) to determine a lump-sum payment to the publisher. In the third stage, the platform chooses the quantity of news it shows to its users subject to the constraint \( \lambda \leq \Lambda \), and gets all the advertising revenue when consumers use its website \( (\phi_0 = 1, \phi_1 = 0) \). In stage 4, consumers allocate their attention among the two websites so as to consume their optimal mix. Direct traffic to the publisher’s website generates revenues of \( \beta \) per unit of time.

Our modelling of bargaining can be interpreted as a formal integration of the publisher’s content into the platform’s feed, where the platform gets all the advertising revenue. The constraint \( \lambda \leq \Lambda \) captures the idea that, even with a formal agreement, the platform
must show enough original content to be viable. We assume that when the platform is indifferent between several values of $\lambda$ it chooses the smallest one. In case bargaining fails, the outcome is the one without content bundling, described in Lemma 1.

**Equilibrium** We solve the game by backward induction. In Stages 3 and 4, consumers face the same decision as in the baseline analysis, and their optimal behavior is given by Lemma 2: If $\lambda \geq q$, $T_1(q, \lambda) = 0$. If $q > \lambda$, $T_1(q, \lambda) = \frac{q - \lambda}{1 - \lambda}$.

Given that $\phi_0 = 1$, the platform simply seeks to maximize the amount of time users spend on its website. If $q \leq \Lambda$, this entails setting $\lambda = q$, which leads to $T_0(q, q) = 1$. If $q > \Lambda$, the platform sets $\lambda = \Lambda$ and $T_0(q, \Lambda) = \frac{1 - q}{1 - \Lambda}$.

At stage 2 (the bargaining stage), the cost $C(q)$ is sunk and does not affect the bargaining outcome. Given a quality $q$, the outside option payoffs are, respectively, $\beta q$ for the publisher and $1 - q$ for the platform.

Let $F$ be the lump sum payment to the publisher in case of an agreement. If the publisher has chosen $q \leq \Lambda$, we know that the platform will choose $\lambda = q$ and obtain advertising revenues equal to 1, while the publisher will get zero revenues. By the logic of Nash bargaining, the negotiated payment $F$ must then be the solution to

$$\max_F (F - \beta q) (1 - F - (1 - q)),$$

that is

$$F = \frac{q(1 + \beta)}{2} \equiv F_1(q).$$

If the publisher has chosen $q > \Lambda$, the parameters of the negotiation are different. Indeed, even with an agreement, the publisher will receive a share of traffic $T_1(q, \Lambda) = \frac{q - \Lambda}{1 - \Lambda}$, and a revenue $\beta T_1$. The negotiated payment $F$ must then be the solution to

$$\max_F \left( \beta \frac{q - \Lambda}{1 - \Lambda} + F - \beta q \right) \left( \frac{1 - q}{1 - \Lambda} - F - (1 - q) \right),$$

that is

$$F = \frac{(1 - q)(1 + \beta)\Lambda}{2(1 - \Lambda)} \equiv F_2(q).$$

At stage 1, anticipating the outcome of the bargaining stage, the publisher maximizes

$$\pi_1(q) = \begin{cases} 
F_1(q) - \frac{c q^2}{2} & \text{if } q \leq \Lambda \\
F_2(q) + \beta \frac{q - \Lambda}{1 - \Lambda} - \frac{c q^2}{2} & \text{otherwise.}
\end{cases}$$

Let us denote $q_1(c) \equiv \arg\max_q F_1(q) - \frac{c q^2}{2} = \frac{1 + \beta}{2c}$ and $q_2(c) \equiv \arg\max_q F_2(q) + \beta \frac{q - \Lambda}{1 - \Lambda} - \frac{c q^2}{2} = \frac{\beta (2 - \Lambda) - \Lambda}{2(1 - \Lambda)c}$. Maximizing (7) leads to the following result.
Lemma 5. Under negotiated content bundling, the equilibrium quality is:

\[ \tilde{q} = \begin{cases}  
q_2(c) & \text{if } c < \frac{\beta(2-\Lambda)-\Lambda}{2(1-\Lambda)\Lambda} \\
\Lambda & \text{if } c \in \left[\frac{\beta(2-\Lambda)-\Lambda}{2(1-\Lambda)\Lambda}, \frac{1+\beta}{2\Lambda}\right] \\
q_1(c) & \text{if } c > \frac{1+\beta}{2\Lambda}. 
\end{cases} \]

Proof. First, one can check that \( q_1(c) > q_2(c) \) and that profit is continuous. If \( \Lambda < q_2(c) \), profit is increasing over \([0, q_2(c)]\) and decreasing afterwards. The condition \( \Lambda < q_2(c) \) can be rewritten as \( c < \frac{\beta(2-\Lambda)-\Lambda}{2(1-\Lambda)\Lambda} \). If \( \Lambda > q_1(c) \) (i.e. \( c > \frac{1+\beta}{2\Lambda} \)), profit is increasing over \([0, q_1(c)]\) and decreasing afterwards. Finally, for \( \Lambda \in [q_2(c), q_1(c)] \), profit is increasing over \([0, \Lambda]\) and decreasing afterwards.

Intuitively, when the cost of producing quality is small, the publisher operates in the parameter region such that \( q > \Lambda \) and therefore chooses \( q_2(c) \). Conversely, when \( c \) is large, it is more profitable for the publisher to choose a strategy of exclusive distribution through the platform, and then \( \tilde{q} = q_1(c) \).

Comparison with benchmark and unilateral content bundling Straightforward algebra allow us to compare the equilibrium under bargaining to the equilibrium under the benchmark (Lemma 1) and under unilateral content bundling (Lemma 4). (The proof is omitted for brevity).

Proposition 3. There exist \( c_1 \) and \( c_2 \) such that:

1. Quality is lower under negotiated content bundling compared to the benchmark if and only if \( c < c_1 \).
2. Quality is lower under negotiated content bundling compared to unilateral content bundling if and only if \( c < c_2 \).

Figure 1 illustrates the result.

The intuition for the result is as follows. Suppose \( c \) is large, so that the content is exclusively distributed through the platform under negotiated content bundling (low-quality content with single-homing on the platform). The point of investing in quality is not to affect the joint surplus before the negotiation (because this joint surplus is constant and equal to 1 in the region where \( q < \Lambda \)), but only to shift the parties’ outside options. A \( \Delta \) quality increase improves the publisher’s outside option by \( \beta\Delta \) and reduces the platform’s one by \( \Delta \). Under Nash bargaining, this results in an increase of \( \frac{1+\beta}{2\Delta} \) of the publisher’s payoff, which is larger than the corresponding increase without content bundling (\( \beta\Delta \)).

When \( c \) is small and we are in the region such that \( q > \Lambda \), an increase in \( q \) still affects the parties’ outside options, but it also entails an inefficiency from the parties’ joint profit.

15
perspective, as it attracts more traffic to the publisher’s website, where advertising revenues are lower than on the platform. The joint profit is indeed \( 1 - \frac{(1-\beta)(q-\Lambda)}{1-\Lambda} \), decreasing in \( q \). This inefficiency dampens the publisher’s incentives to invest.

If one worries about the negative effect of social platforms’ content bundling on newspapers’ profitability, putting their survival at risk, a solution based on bargaining appears promising as it ensures that publishers are better-off than under unilateral content bundling. However such a regime is not neutral from a quality perspective, and leads to a decrease in the variance of quality, as high-quality newspapers invest less while low-quality ones invest more than under alternative regimes. In sum, similar to the effect of a link tax, forcing negotiations between the platform and newspapers may only lead to a reduction in the variation across newspapers’ qualities and not to a general increase of quality.

5 Competing publishers

In this section we extend the model by introducing competition among publishers, and show that our main insights regarding the effects of content bundling continue to hold.

Suppose that there are \( N \) publishers, and that they are heterogeneous with respect to their cost of providing quality, \( c_i q_i^2 / 2 \). We order firms so that \( c_1 \leq \ldots \leq c_N \). These cost parameters are known to everyone.

We denote \( Q \equiv (q_1, \ldots, q_N) \). Consumers are homogenous, and the demand for content by publisher \( i \) is \( D_i(Q) = q_i + \gamma \sum_{j \neq i} (q_i - q_j) \). The parameter \( \gamma \geq 0 \) measures the intensity of competition between the newspapers, which we assume is symmetric for notational
simplicity.

The platform can unilaterally choose to host some of the publishers’ content as in the baseline model. We assume that it is constrained not to show more than a share $\Lambda$ of any single publisher’s content: $\lambda_i \leq \Lambda$ for all $i \in \{1, ..., N\}$. We also assume that $NA \leq 1$ (the platform has to show some original content to consumers).

Publishers’ (per-unit of time) revenue from direct traffic is $\beta$, their revenue from indirect traffic is $\phi_1$. The platform’s revenue when it shows UGC is 1, and $\phi_0$ when it shows publishers’ content.

The timing is the following: First, publishers choose their quality $q_i$. Second, the platform decides how much of each publisher’s content to host, $\lambda_i$. Third, consumers choose how to allocate their attention across websites, and payoffs are realized.

**Benchmark: no content bundling** Suppose that the platform does not have the ability to display publishers’ content. Then publisher $i$ maximizes $\beta D_i(Q) - c_i q_i^2 / 2$. The first-order condition defining the equilibrium is thus

$$\hat{q}_i = \frac{\beta (1 + \gamma)}{c_i}.$$  

Equilibrium quality decreases with a firm’s cost to produce quality. Moreover, as competition intensifies, firms invest more because demand is more responsive to quality increases.

**Unilateral content bundling** As some of our examples below illustrate, there can be a multiplicity of equilibria in the game with content bundling. A class of equilibria of particular interest is that of *Monotone Pure Strategy Equilibria* (MPSE), which satisfy the property that $q_i^* > q_j^*$ only if $c_i < c_j$. For this class of equilibria we can show the following.

**Proposition 4.** In any MPSE of the game with unilateral content bundling by the platform, there exists an $i^\dagger \in \{0, ..., N\}$ such that $i$ invests more under unilateral content bundling than under the benchmark ($q_i^* \geq \hat{q}_i$) if and only if $i \leq i^\dagger$.

The intuition for Proposition 4 is closely related to the case with one publisher. In an MPSE, firms with a relatively large cost decide to adopt a strategy of low quality content with consumers single-homing on the platform, for which the marginal return to investment is lower than under the benchmark. Firms with a lower cost decide to adopt a strategy of high quality with multi-homing, which leads them to invest more because direct traffic is more sensitive to quality than under the benchmark.

Proposition 4 also implies that content bundling tends to increase the variance of quality, by making high-cost publishers invest less and low-cost publishers invest more.

\footnote{The case where $i^\dagger = 0$ corresponds to the case where all firms reduce their quality.}
Existence of MPSE  Notice that Proposition 4 does not guarantee the existence of an MPSE. Restricting attention to duopoly settings, we can find many combinations of parameters such that an MPSE exists. For instance, suppose that $\gamma = 1/10, \beta = 9/10, \phi_1 = 2/3, \Lambda = 2/10, c_1 = 3, c_2 = 6$. Without content bundling, equilibrium qualities are $\hat{q}_1 = 0.33$ and $\hat{q}_2 \approx 0.16$. With content bundling, there exists a unique MPSE, with $q_1^* \approx 0.35$ and $q_2^* \approx 0.12$.

Symmetric duopoly  Suppose that $c_1 = c_2 = c$. One can show that there are two kinds of pure strategy equilibria: one in which both firms choose $q_1^* = q_L < \hat{q}$ such that $D_i(Q_L^*) < \Lambda$, and all content is accessed through the platform (what we call single-homing equilibrium), and one in which they choose $q_1^* = q_H > \hat{q}$ such that $D_i(Q_H^*) > \Lambda$ and both publishers receive some direct traffic (multi-homing equilibrium). There also exists a mixed strategy equilibrium, where firms randomize between a high quality $(q_h \in (\hat{q}, q_H))$ and a low quality $q_L$. The mixed strategy equilibrium converges to the pure strategy ones as we approach their region. The details of the formal analysis are relegated to the Appendix.

Figure 2 illustrates the equilibrium configurations for some combinations of parameters. Notice that there is an area where both types of pure-strategy equilibrium exist, when $\gamma$ is small enough. To understand the intuition, let us fix $\gamma = 0$. Even though publishers do not compete, they exert an externality on one another: when $q_j$ is large, it is easier for firm $i$ to attract direct traffic than when $q_j$ is small. In the former case, because consumers spend some time $T_j$ on publisher j’s website, they spend less time on the platform, and it is thus harder for the platform to satisfy the demand for news from publisher $i$. Formally, publisher $i$ only needs to set $q_i \geq (1-T_j)\Lambda$ to attract direct traffic, as opposed to $q_i \geq \Lambda$ when $T_j = 0$. This leads to a form of strategic complementarity: the return on quality is larger when the other publisher invests more.

Just like in the monopoly case, we find that content bundling increases qualities when $c$ is small, and reduces it when $c$ is large. The figure also indicates that as the intensity of competition increases, we are more likely to be in a multi-homing equilibrium, where quality is larger than without content bundling. Even though we have not been able to formally prove this result, the intuition is as follows: when $\gamma$ is large enough, competition leads firms to choose quality above $\Lambda$, i.e. to operate in a region where they attract direct traffic even in the presence of content bundling by the platform. As we saw above, direct traffic is more responsive to quality under content bundling, which explains why firms have an extra incentive to invest.

6  Heterogeneous consumers and platform analytics

So far, we have assumed homogeneous consumers and as a result, the social network’s content strategy was reflected by a single number $\lambda$ and each consumer’s newsfeed contained
Figure 2: Equilibrium configurations for the symmetric case (with $\beta = 0.9$, $\Lambda = 0.2$, and $\phi_1 = 2/3$). (The grey area is excluded because we would have $q_1 + q_2 > 1$).

the same proportions of news and UGC. In reality however, there is a large variation of preferences with respect to content across members on a social network. More importantly, a fundamental appeal of social platforms is that by observing their members’ behavior the platform can personalize each member’s newsfeed. In this section we extend our basic model to capture these important characteristics.

To introduce consumer heterogeneity in a tractable way, we modify consumers’ utility function in (1) by assuming that a consumer of type $\theta$ who consumes a quantity $x_N$ of news (of quality $q$) and $x_U$ of UGC derives utility:

$$U(x_U, x_N, q, \theta) = x_U + (1 + \theta + q)x_N - \frac{(x_N)^2}{2},$$

where $\theta$ is uniformly distributed on $[0, 1]$. Given the constraint that $x_U + x_N \leq 1$, the demand for news of a consumer of type $\theta$ is $\min\{\theta + q, 1\}$. As in the baseline model, consumers allocate their unit of attention across two websites: one operated by the publisher, and one by the platform.

We assume that the platform is imperfectly informed about consumers’ types. More specifically, the platform observes a consumer’s type $\theta$ with probability $\alpha$, and observes nothing with probability $1 - \alpha$. We interpret the parameter $\alpha$ as the exogenous quality of the platform’s analytics technology, enabling it to infer consumers’ preferences.\(^{12}\)

The platform can adapt the content it shows to a consumer depending on his type,
choosing a higher share of news (λ) to higher types. Consumers whose type the platform does not observe all face the same λ. We assume that the maximal amount of news that the platform can show, Λ, is arbitrarily close but strictly below one, meaning that the platform can show as much news as it wants to high types (when it identifies them), but consumers who only desire to read news visit the publisher directly. Note that here, the platform does not have to worry about its identity as a social network as it will always show UGC to low type consumers.

For notational simplicity, we also assume that the platform shares with the newspaper the total advertising revenue when advertising on behalf of the newspaper, i.e. φ₁ + φ₀ = 1. In this case, we simply denote by φ the share of ad revenues that the platform keeps for itself. Given our previous assumptions on the φ-s and β, this means that 1 − φ < β < 1.

The timing is as follows: at τ = 1, the newspaper chooses a quality q, publicly observed, and incurs cost c(q). At τ = 2, the platform observes the type of a share α of consumers, and chooses the share of news λ(θ, q) it shows to its users (for consumers whose type it does not observe, the platform’s content bundling strategy is λ(q)). At τ = 3, consumers observe the λ they face, and choose t₀(θ, q, λ), the time they spend on the platform as a function of their type, of the quality of news and of the platform’s content mix. The resulting news consumption is xₕ = λt₀ + t₁. We look for subgame-perfect equilibria.

6.1 Benchmark: UGC-only newsfeed

When the platform cannot bundle news content alongside UGC (i.e. Λ = 0) consumers can consume their desired mix of content: a consumer of type θ then spends t₁(θ, q, 0) = min{θ + q, 1} units of time on the newspaper’s website, and t₁(θ, q, 0) = 1 − t₁(θ, q, 0) on the platform. Thus, the total time spent on the newspaper is

\[ T₁(q, 0) = \int_0^1 \min{\theta + q, 1} d\theta = \int_0^{1-q} (\theta + q) d\theta + \int_{1-q}^1 d\theta = \frac{1 + 2q - q^2}{2}. \]

The profit of the newspaper is βT₁(q, 0) − cq²/2, which leads to the following equilibrium quality and profit in the benchmark:

\[ \hat{q} = \frac{\beta}{\beta + c} \quad \text{and} \quad \hat{π}_1 = \frac{\beta(2\beta + c)}{2(\beta + c)}. \]  

(10)

As in the base model, quality increases in β and decreases with c.

6.2 Equilibrium with content bundling

Under content bundling consumers’ time allocation across platforms depends on θ. A consumer of type θ to whom the platform shows a share λ of news will allocate his attention so as to come as close as possible to his desired consumption mix (see Anderson
and Neven, 1989, for a related analysis of combinable goods). The optimal allocation of attention mirrors that of Lemma 2, where we replace \( q \) by \( \theta + q \). Following the tie-breaking rule we also observe that consumers such that \( \theta + q \geq 1 \) do not visit the platform (i.e. \( t_0(\theta, q, \lambda) = 0 \)).

The platform’s problem consists in choosing a share of news \( \lambda(\theta, q) \) to show to consumers of type \( \theta \), and a share \( \overline{\lambda}(q) \) to show to consumers whose type it does not observe. First, the platform’s decision regarding consumers whose type it can observe is straightforward: \( \lambda(\theta, q) = \min\{\theta + q, 1\} \). By showing each consumer his ideal content mix, the platform ensures that consumers with a positive demand for UGC (such that \( \theta < 1 - q \)) allocate all their attention to its website: \( t_0(\theta, q, \lambda(\theta, q)) = 1 \) for all \( \theta < 1 - q \) (consumers such that \( \theta \geq 1 - q \) only visit the newspaper).

For the \( 1 - \alpha \) unidentified consumers, whose \( \theta \) is unknown to the platform, the problem is more subtle. If the platform displays a share \( \overline{\lambda} \geq q \) of news content, the total amount of attention that it receives from this segment is

\[
T_0(\overline{\lambda}, q) = \int_0^{\overline{\lambda}-q} 1d\theta + \int_{\overline{\lambda}-q}^{1-q} \frac{1 - (\theta + q)}{1 - \overline{\lambda}} d\theta = \frac{1 + \overline{\lambda} - 2q}{2}. \tag{11}
\]

The platform’s profit is then \((1 - \overline{\lambda} + \overline{\lambda}\phi_0)T_0(\overline{\lambda}, q)\), which is maximized for \( \overline{\lambda}(q) = q + \frac{\phi}{2(1 - \phi)} \).

The first term, \( q \), corresponds to the demand for news of the lowest type (\( \theta = 0 \)). Beyond this level, the platform’s optimal strategy depends on the share \( \phi \) of revenues it captures when it shows news: for large values of \( \phi \) the platform has an incentive to show a lot of news content to its users.

At \( \tau = 1 \), the publisher can anticipate the platforms strategy and consumers’ time allocation across sites and chooses its quality so as to maximize its profit. Adding revenues from consumers across sites, the publisher’s optimal choice of quality is as follows (see the proof in the Appendix):

**Lemma 6.** With heterogeneous consumers, under content bundling by the platform, the profit-maximizing quality level for the newspaper is

\[
q^* = \frac{\beta + 1 - \phi + \alpha(\beta - (1 - \phi))}{2(1 - \phi + c)}. \tag{12}
\]

The proof of the Lemma is in the Appendix. As in the homogeneous case, quality is increasing in \( \beta \) and decreasing in \( c \) and \( \phi \). Interestingly, it is also increasing in \( \alpha \) (see below for a discussion).
6.3 Effects of content bundling

Now, we can compare the equilibrium outcomes with and without content bundling.\(^\text{13}\) We start with the publisher’s profit:

**Proposition 5.** There exists \( \beta \in [1 - \phi, 1) \) such that the newspaper is better-off under content bundling if and only if \( \beta < \beta^* \).

Content bundling has two effects on the newspaper: first, as in the baseline model, content bundling diverts news consumption onto the platform, thereby reducing the newspaper’s revenue per-unit of attention. This effect is negative.

Second, unlike the baseline model with homogeneous consumers, content bundling increases news consumption (as long as \( \alpha < 1 \)): by setting \( \lambda(q) > 0 \), the platform induces some low-types to read more news than they would otherwise, while higher types still consume their optimal mix. This is a market expansion effect. If the newspaper’s revenue from direct traffic, \( \beta \) is close to its revenue from indirect traffic \((1 - \phi)\), the first effect dominates, while the second effect dominates for \( \beta \) close to 1.

Of central interest is the newspaper’s incentive to invest in quality under content bundling compared to the benchmark. Simple algebra yields the following result.

**Proposition 6.** The equilibrium quality of news increases under content bundling compared to the benchmark if and only if \( c \leq \frac{1 + \alpha}{1 - \alpha} \beta \).

This result mirrors Proposition 1 in the case with heterogeneous consumers: the effects of content bundling on quality depend on the publisher’s cost of quality, as a result of two opposite effects: (i) increasing demand for news of low-types is less valuable, as this consumption takes place on the platform; (ii) direct traffic by relatively higher types is more sensitive to quality under content bundling.

Lastly, we examine the impact of the platform’s analytic capabilities by asking: suppose that the platform were to use the optimal content bundling strategy, given its analytic capability \( \alpha \), how would an increase in said ability affect the equilibrium? From (12) we directly see the following:

**Proposition 7.** The equilibrium quality of news is an increasing function of the platform’s analytics capability, \( \alpha \).

With personalization, competition between the platform and the publisher is more intense. In particular, turning a consumer into an exclusive news reader \((\theta + q \geq 1)\) is much more valuable to the publisher when the platform offers personalized content.

\(^{13}\)All proofs are in the Appendix. Furthermore, there, we show additional analysis, where we explore and discuss the impact of a “link tax”, as proposed by European legislators.
because non-exclusive news readers spend all their time on the platform, even if the share of news they read is close to one.\footnote{A similar logic would work if there was a limit to how much news the platform can show.}

The result that personalization leads the publisher to invest more in the quality of its content has some interesting implications. In particular, the platform’s profit is not always increasing in $\alpha$: under some parameters, even if the platform had the ability to offer a fully personalized content mix to all its users, it would not be optimal to do so, in order to deter the publisher from investing too much in quality. We can show that the platform tends to optimally limit the extent of personalization when $\beta/(1 - \phi)$ is large, i.e. when the publisher’s incentive to compete fiercely to attract attention on its own website is large.

Finally, with respect to consumer surplus, our findings are ambiguous. When we compare content bundling to the benchmark, two effects must be taken into account, namely the impact of bundling on quality, and its impact on the equilibrium mix of content. While the first effect can go either way, as discussed previously, the second one can only be negative, with low-type consumers being offered too much news compared to their optimal level. Importantly, note that the effect of more personalization ($\alpha$) is unambiguously positive: we saw that quality goes up with $\alpha$ (Proposition 7), and personalization reduces the content mix distortion.

7 Discussion and concluding remarks

Social networks have gained tremendous importance in the last decade, claiming a significant share of consumer attention. They have achieved such prominence by leveraging network effects and, more recently, by successful content bundling, whereby third party content is presented in their users’ “newsfeed”. This strategy, in turn, has started to fundamentally transform media production and consumption. For some content domains with so-called public calling, such as news for example, concerns have been raised about the overall health of the industry. In particular, given the market power of social platforms, it is important to understand whether news providers’ maintain adequate incentives to invest in content quality. Accordingly, policy makers have pushed for ways to encourage transfers between social platforms and news providers.

In this paper we develop a simple model of competition for attention between a social platform and third-party content providers, allowing us to shed light both on the strategic motives for content bundling and on its implications for publishers and their incentives to invest in quality. By bundling news content with UGC, the platform increases the share of attention it receives from consumers: part of their demand for news is satisfied by the platform, which induces them to spend less time on the newspaper’s website. We show that the practice of unilateral content bundling tends to harm publishers, but has a
heterogenous impact on their incentives to invest in the quality of their content: initially low quality outlets invest less under content bundling, while initially high quality ones invest more. While the result is most clearly exposed in a set-up with one publisher and homogenous consumers, it holds with heterogenous publishers and consumers.

When the platform and the publisher negotiate for the former to use the latter’s content, the results are reversed: the publisher is better-off than under the benchmark, but the dispersion of quality goes down, as low quality outlets invest more and high quality ones invest less. This highlights a potential trade-off associated with current proposals (e.g. in Australia and the European Union) to regulate relations between publishers and platforms by forcing platforms to negotiate with publishers.

In order to keep the model parsimonious, we have abstracted away from several interesting considerations. In particular, the structure of the social network is notably absent from the analysis. An interesting avenue for future research would be to study environments where consumers have heterogeneous preferences and where the social network exhibits homophily. We have also mostly abstracted away from users’ behavior on the platform, regarding, for instance, their decision to share third-party content or to produce UGC. Finally, we have abstracted away from an in-depth interpretation of content quality. Recent research in the context of news (see e.g. Garz et al. (2018)) shows that the consumption of news is driven by a complex web of psychological factors. These considerations may shed different light on our results w.r.t news quality in particular.

Our analysis focuses on the impact of a social network on publishers. We have illustrated our results for a particular content domain, news, because of its special importance for public life. However, our model readily applies to third-party publishers in other content domains (e.g. games, videos) who also seek to be present in consumers’ ‘newsfeed’ on social media. Beyond social networks narrowly defined, the modeling framework also seems to be applicable to a broader set of interactions between multi-sided platforms and third-party ‘content’ providers.

For example, video distribution platforms such as Netflix, Hulu or Amazon Prime Video all offer third-party content alongside shows they produce themselves. Here, the role of newspapers is played by movie studios or TV networks who can monetize their content independently but are attracted by the platforms’ customer base. In the e-commerce sector, Amazon.com also offers consumers the possibility to buy some products from third-party merchants or from Amazon itself. In the search engine market, Google often displays a result box that contains information produced by a website (e.g. a definition, the rating of a merchant or of a movie), with a small link to that website. While some consumers may click that link, many will simply stop there, having obtained the relevant information. This practice has been denounced by some websites (e.g. Yelp) who argue that it makes it difficult for them to attract visitors. This case raises issues of copyright protection, and also of efficiency gains.
These examples share some of the characteristics of our framework, and some of our insights might apply there as well. However, they also have some specific features, which are not captured by our model. This calls for future research in this area.\footnote{In quite different contexts, Moorthy, Chen, and Tehrani (2018) and Hagiu, Jullien, Wright, et al. (forthcoming) also investigate hybrid forms of competition, where a firm distributes the product of its rivals.}

References


Angelucci, Charles and Julia Cagé (2016). “Newspapers in times of low advertising revenues”.


Competition and Markets Authority (2020). “Online platforms and digital advertising market study”.


Pew Research Center (2019). “Americans are Wary of the Role Social Media Sites Play in Delivering the News”.


27
A Appendix

A.1 Omitted proofs

Proof of Proposition 4. Suppose that, in equilibrium, consumers spend $T_0$ units of time on the platform. Any firm such that $D_i(Q^*) \leq \Lambda T_0$ does receive any direct traffic, as the platform can set $\lambda_i T_0 = D_i(Q^*)$ and satisfy consumers’ demand for publisher $i$’s content.

Firm with $D_i(Q^*) > \Lambda T_0$ receives some direct traffic in equilibrium. Indeed, the platform optimally sets $\lambda_i = \Lambda$ and consumers complete their demand for publisher $i$’s content by spending $t_i = D_i(Q^*) - \Lambda T_0$ units of time on its website.

Let $\mathcal{H}$ be the set of firms such that $D_i(Q^*) > \Lambda T_0$, of cardinal $N_H$. Let $\mathcal{L}$ be the complementary set of firms. By the assumption that the equilibrium is monotone, when $N_H \geq 1$ we can write that $\mathcal{H} = \{1, \ldots, i^\dagger\}$ and $\mathcal{L} = \{i^\dagger + 1, \ldots, N\}$.

We must have

$$T_0 = 1 - \sum_{i \in \mathcal{H}} (D_i(Q^*) - \Lambda T_0) \iff T_0 = \frac{1 - \sum_{i \in \mathcal{H}} D_i(Q^*)}{1 - \Lambda N_H}$$

Take a firm $i \in \mathcal{H}$. We have $\frac{\partial T_0}{\partial q_i} = \frac{1 - (N - N_H)\gamma}{1 - \Lambda N_H} < 0$.

The profit of a firm $i \in \mathcal{H}$ is

$$\pi_i^\mathcal{H}(Q) = \beta (D_i(Q) - \Lambda T_0) + \phi_1 \Lambda T_0 - c_i q_i^2$$

The equilibrium first-order condition is then

$$\frac{\partial \pi_i^\mathcal{H}(Q)}{\partial q_i} = 0 \iff \beta \frac{\partial D_i(Q^*)}{\partial q_i} - (\beta - \phi_1) \Lambda \frac{\partial T_0}{\partial q_i} = c_i q_i^* \quad (13)$$

If we compare with the first-order condition without content bundling, $\beta \frac{\partial D_i(Q)}{\partial q_i} = c_i q_i$, we see that $q_i^* > q_i$ because $\beta > \phi_1$ and $\frac{\partial T_0}{\partial q_i} < 0$.

For a firm $i \in \mathcal{L}$,

$$\pi_i^\mathcal{L}(Q) = \phi_1 D_i(Q^*) - c_i q_i^2$$

The first-order condition is then

$$\frac{\partial \pi_i^\mathcal{L}(Q)}{\partial q_i} = 0 \iff \phi_1 \frac{\partial D_i(Q^*)}{\partial q_i} = c_i q_i \quad (14)$$

which, compared to the benchmark, indicates a lower quality.

Equilibria of the symmetric duopoly case. We start by looking for pure strategy equilibria.
Suppose that firms have chosen $q_i$ in the first stage. Recall that we assume that the demand for content $i$ is $D_i(Q) = q_i + \gamma(q_i - q_j)$. Several configurations are possible:

(I) If $Q$ is such that $D_i(Q) \leq \Lambda$ for both firms, the platform can attract all direct traffic by setting $\lambda_i = D_i(Q)$. The profit of each firm is

$$\pi_i^I = \phi_1 D_i(Q) - \frac{cq_i^2}{2} \quad (15)$$

(II) If $D_i > \Lambda$ for some $i$, publisher $i$ will attract some direct traffic even if the platform sets $\lambda_i = \Lambda$ (which it optimally does). Such direct traffic is equal to $D_i(Q) - \Lambda T_0$. The time spent on the platform must therefore satisfy $T_0 \leq 1 - (D_i(Q) - \Lambda T_0)$, i.e. $T_0 \leq \frac{1-D_i(Q)}{1+\Lambda}$. If $D_j \leq \Lambda \frac{1-D_i(Q)}{1+\Lambda}$, the platform can satisfy all the demand for publisher $j$'s content by setting $\lambda_j = D_j \frac{1+\Lambda}{1-D_i(Q)}$. In such a configuration, only the platform and publisher $i$ receive direct traffic. Profits are given by

$$\pi_{H}^{II} = \phi_1 \Lambda \frac{1-D_i(Q)}{1+\Lambda} + \beta \frac{D_i(Q) - \Lambda}{1+\Lambda} - \frac{cq_i^2}{2} \quad \text{and} \quad \pi_{L}^{II} = \phi_1 D_j(Q) - \frac{cq_j^2}{2} \quad (16)$$

(III) If $D_i > \Lambda \frac{1-D_i(Q)}{1+\Lambda}$ for $i \in \{1, 2\}$, both publishers receive direct traffic even if $\lambda_i = \Lambda$. The allocation of time is the solution to the system

$$D_i(Q) = T_i + \Lambda T_0, \quad \text{and} \quad T_0 + T_1 + T_2 = 1$$

i.e.

$$T_i = \frac{D_i(Q) - \Lambda (D_i(Q) + D_j(Q))}{1-2\Lambda}, \quad \text{and} \quad T_0 = \frac{1-D_i(Q) - D_j(Q)}{1-2\Lambda}$$

Profits are then given by

$$\pi_{H}^{III} = \phi_1 \Lambda \frac{1-D_i(Q) - D_j(Q)}{1-2\Lambda} + \beta \frac{D_i(Q) - \Lambda (D_i(Q) + D_j(Q))}{1-2\Lambda} - \frac{cq_i^2}{2} \quad (17)$$

On the left panel of Figure 3 we represent these configurations as a function of $D_1(Q)$ and $D_2(Q)$. On the right panel, we represent them as a function of $q_1$ and $q_2$.

There are several potential equilibria in pure strategy: (i) in the interior of region I, (ii) in the interior of region II, (iii) in the interior of region III, (iv) at the boundary between I & II, (v) at the boundary between II & III, (vi) at the boundary between all three regions (when $q_1 = q_2 = \Lambda$).

Because profit is not quasi-concave, for each potential equilibrium we need to check conditions that guarantee that there is neither a local nor a global deviation. Doing so, we can exclude most cases as we find that the only possible pure strategy equilibria are interior equilibria in region I or III. (See Figure 2).
Figure 3: Allocation of attention as a function of demand for content (left) and quality (right). I: \( T_1 = T_2 = 0 \); II: \( T_1 = 0 \) or \( T_2 = 0 \); III: \( T_1 > 0 \) and \( T_2 > 0 \).

**Equilibrium in region I** In such an equilibrium, both firms maximize (15), and neither firm find it optimal to deviate into II or III. The quality \( q^*_I = \frac{\phi_1(1+\gamma)}{c} \) is lower than the benchmark quality \( \hat{q} = \frac{\beta(1+\gamma)}{c} \).

**Equilibrium in region III** In such an equilibrium, both firms maximize (17), and neither firm find it optimal to deviate into I or II. The quality \( q^*_{III} = \frac{\phi_1(1+\gamma)}{c} \) is greater than the benchmark quality \( \hat{q} \).

**Mixed strategy equilibrium** There also exist a set of parameters such that no pure strategy exists. Such a situation is illustrated in Figure 4, where the arrows indicate profitable deviations from candidate equilibria A, B, C, D.

In such cases we have been able to construct mixed-strategy equilibria that take the following form: each firm plays some \( q_l \) with probability \( \sigma \), and some \( q_h > q_l \) with probability \( 1 - \sigma \). The qualities \( q_l \) and \( q_h \) are such that \((q_l, q_l)\) is in region I, \((q_l, q_h)\) and \((q_h, q_l)\) are in region II, and \((q_h, q_h)\) is in region III.

By (15) and (16), the firm with the lowest quality in region II has the same marginal profit as a firm in region I. Therefore, if a firm chooses a quality \( q_l \), the first-order condition implies that \( q_l = q^*_I = \frac{\phi_1(1+\gamma)}{c} \).

On the other hand, \( q_h \) must satisfy \( \sigma \frac{\partial \pi_{II}^I(q_h,q_l)}{\partial q_h} + (1 - \sigma) \frac{\partial \pi_{I}^I(q_h,q_h)}{\partial q_h} = 0 \).

The value of \( \sigma \) is then such that each firm is indifferent between \( q_l \) and \( q_h \).  

---

30
This figure explains why we may not have a pure strategy equilibrium.

Figure 4: Example of non-existence of pure-strategy equilibrium. Dashed arrows indicate profitable deviations from potential equilibria A, B, C and D.

Proof of Lemma 6. Adding visitors across sites, the publisher’s profit reads as follows:

\[
\pi_1(q) = \alpha(1 - \phi) \int_0^{1-q} \lambda(\theta, q) d\theta + (1 - \phi)(1 - \alpha) \int_0^{\lambda(q)-q} \lambda(q) d\theta \\
+ (1 - \alpha) \int_{\lambda(q)-q}^{1-q} (\lambda(q)(1 - \phi)t_0(\theta, q, \lambda(q)) + \beta t_1(\theta, q, \lambda(q))) d\theta + \beta \int_{1-q}^1 d\theta \\
- c(q). \tag{18}
\]

Term (i) in equation (18) corresponds to the consumers that the platform identifies and to whom it shows a tailored mix of content (\(\lambda(\theta, q) = \theta + q\)), inducing them to spend all their time on the platform. The newspaper’s revenue over these consumers is \((1 - \phi)\) per unit of time. Term (ii) represents the consumers that the platform does not identify and who have a relatively low type \((\theta < \lambda(q) - q)\): these consumers spend all their time on the platform, are exposed to a share \(\lambda(q)\) of news (which is more than enough to satisfy their demand for news), over which the newspaper gets a revenue \((1 - \phi)\) per unit of time. Term (iii) corresponds to the consumers who are not identified by the platform and who wish to complement the news they get from the platform by direct traffic to the newspaper. They spend \(t_0(\theta, q, \lambda(q)) = \frac{1 - \theta - q}{1 - \lambda(q)}\) units of time on the platform’s website, generating revenue \(\lambda(q)(1 - \phi)\) per unit of time for the newspaper, and \(t_1(\theta, q, \lambda(q)) = 1 - t_0(\theta, q, \lambda(q))\) units of time on the newspaper’s website, generating a revenue of \(\beta\) per unit of time. Finally, term (iv) represents the consumers who only visit the newspaper’s website.
Given that \( \bar{\lambda}(q) = q + \frac{\phi}{2(1 - \phi)} \), after some algebra we obtain the Lemma.

**Proof of Proposition 5.** When \( \beta = 1 - \phi \), \( \pi_1(q^*) - \tilde{\pi}_1 = \frac{(1 - \alpha)^2}{\phi(1 - \phi)} \geq 0 \).

When \( \beta = 1 \) the newspaper is necessarily worse-off under content bundling. Indeed, in such a case the industry revenue is constant and equal to 1. Let \( \hat{R}_0(q) \) and \( \hat{R}_1(q) \) be the platform’s and the newspaper’s revenue without content bundling if news quality is \( q \), and \( \tilde{R}_0(q) \) and \( \tilde{R}_1(q) \) their revenues under content bundling with quality \( q \). Because \( R_0 + R_1 = 1 \) and \( \tilde{R}_0(q) > \hat{R}_0(q) \) for any \( q \) (optimal content bundling necessarily increases the platform’s revenue), we have

\[
\hat{\pi}_1(q^*) \equiv \hat{R}_1(q^*) - c(q^*) < \tilde{R}_1(q^*) - c(q^*) \leq \tilde{R}_1(\tilde{q}) - c(\tilde{q})
\]

where the last inequality comes from a revealed-preferences argument.

Last, the difference between the newspaper’s profit with and without content bundling is decreasing in \( \beta \).

**A.2 Effects of a link tax in the model with heterogenous consumers**

Within the framework of the extended model, one can think of a link tax as a reduction in \( \phi \) (the share of advertising revenues accruing to the platform). It is straightforward to show the following result:

**Proposition 8.** Consider a link tax that takes the form of a small reduction in \( \phi \).

(i) For any \( \phi \), there exists \( \overline{\alpha}(\phi) \) and \( \overline{\beta}(\phi) > 1 - \phi \) such that the link tax decreases the publisher’s profit if \( \alpha < \overline{\alpha}(\phi) \) and \( \beta < \overline{\beta}(\phi) \). Otherwise, the link tax increases the publisher’s profit.

(ii) The link tax leads the publisher to invest less in the quality of its content if \( c \leq \beta \frac{1 + \alpha}{1 - \alpha} \), and to invest more otherwise.

Like in the baseline model with homogeneous consumers, the effect of the link tax on equilibrium quality is positive for high-cost publishers, and negative otherwise. The intuition for the ambiguous effect on profit is the following: a decrease in \( \phi \) increases the publisher’s revenues for a given \( \bar{\lambda} \), but it also leads the platform to show less news to consumers whose optimal mix is not observed (i.e. to decrease \( \bar{\lambda} \)). This second effect is more costly if there are many such consumers (i.e. if \( \alpha \) is small). The decrease in news consumption on the platform is partially offset by consumers spending more time on the newspaper’s website, but this is not enough if \( \beta \) is relatively small.