Competition for Attention in the News Media Market*

HENG CHEN
University of Hong Kong
WING SUEN
University of Hong Kong

April 23, 2019

Abstract. We provide a framework for analyzing monopolistic competition in the media market, where a large number of news outlets provide similar but differentiated news content and consumers allocate their attention among these news outlets. Entry of new firms raises the overall informativeness of the industry unambiguously, but reduces the quality of news reports produced by existing firms: owners invest less in news accuracy and editors take less effort in news reporting. Further, there exists an endogenous limit to media proliferation: the number of firms that can be supported in equilibrium is bounded even when entry cost is zero.

Keywords. monopolistic competition, news quality, attention allocation, sender-receiver game

JEL Classification. D83, D84, L15, L82

*We thank Jimmy Chan, Alessandro Lizzeri, Andrea Prat, Antonio Russo, and seminar participants at Academia Sinica, Hitoshibashi University, Hong Kong Baptist University, Peking University, Seoul National University, Waseda University, University of Copenhagen, Australasian Economic Theory Workshop, Media Economics Workshop, Public Economic Theory Conference, International Economic Association Congress, Conference on Advances in Information Economics, Economics of Media Bias Workshop, Asian Meeting of the Econometric Society, and Royal Economic Society Annual Conference for useful comments and discussions. This research project is partly funded by the General Research Fund of the Research Grants Council of Hong Kong (Project No. 17501516) and by the Seed Fund for Basic Research of Hong Kong University (Project No. 201811159002).

1. Introduction

The news media industry has experienced a profound transformation. Gone are the days when a consumer got his news from just one or two local newspapers and a few radio and television stations. The modern news consumer faces a menu of options much richer than before: free dailies and online newspapers, news programs on cable, and an almost unlimited amount of information on the internet available at no cost. Thanks to the lowering cost of producing and distributing news, the last few decades have witnessed a dramatic increase in the number of news content providers.[1] The current news industry can therefore best be described by monopolistic competition, given a large number of media firms compete and produce similar but differentiated news content.

Monopolistic competition in the context of the news media market poses new and interesting questions for economic analysis. First, the advertising-supported free content business model (i.e., media firms sell content to news consumers for free and sell their "eyeballs" to advertisers for revenue) is not new, but has become almost ubiquitous in the news media industry.[2] Even the most traditional print media has been shaken: in the past two decades, free newspapers emerged to become "a solitary beacon of hope" in a declining newspaper industry (Anderson 2009). Given that media firms do not charge news consumers directly, their attention paid to media has become "the new currency of business," a term coined by Davenport and Beck (2013). When news can be provided for free, the real constraint facing consumers is the limited attention available for processing information. Instead of using a model in which consumers allocate their budget across different goods, modeling how consumers rationally allocate their attention across information providers is crucial for understanding the nature of competition among news firms.

Second, in response to an enriched media environment, typical news consumers who wish to take an informed action, engage in multi-homing to obtain and aggregate information from multiple news sources.[3] Because different media firms gather

[1] Many of these new outlets have gained substantial popularity among news consumers. According to the 2010 wave of the Biennial Media Consumption Survey conducted by Pew Research Center, 26 online news sites were named by at least one percent of the respondents in the United States as their source for accessing news in 2010; but the comparable figure was only 17 sites in 2006.

[2] This business model is now so prevalent that media measurement technology has to evolve from measuring the size of viewership (e.g., Nielson ratings) and readership (e.g., ABC audit reports) to measuring the detailed aspects of information consumption, especially the amount of attention that consumers pay to media outlets (Napoli 2013; Webster 2014).

[3] The Pew Personal News Cycle Survey in 2014 finds that an average American adult uses four types of media (e.g., print media, radio and television, and the internet) every week for getting news, and the Pew Online News Survey in 2010 finds that 68 percent of respondents access online news from more than two websites. Gentzkow and Shapiro (2011) also demonstrate that a significant share of consumers acquire news from multiple outlets either online or offline. A recent theoretical literature analyzes

their news from diverse but related sources, the interaction between news firms and news consumers resembles a sender-receiver game with multiple senders who possess non-identical information. Competition among news firms is shaped by the outcome of this game, and the welfare implications are therefore different from that obtained from conventional welfare analysis in which goods produced by different firms enter directly into consumers' utility functions.

In this paper, we build a model of monopolistic competition in the news industry that incorporates the features described above. Our model has three sets of actors. News consumers take an action about an uncertain state and they may receive news about the state from multiple news outlets. They decide how much attention to give to each outlet and their attention, in turn, generates revenue for the news firms. Owners of news firms are profit oriented. They invest in an infrastructure to conduct news gathering and investigative research—the more they invest, the more accurate are the facts obtained. Editors of news firms are unbiased. They intend to inform the public, but incur a cost when they take effort to facilitate the information transmission. Based on the facts obtained from news gathering, they craft news stories optimally. There are a large number of such firms in the news market, each producing a differentiated product (because the facts obtained are not identical and because editors adopt different reporting strategies). The news quality, consumers' attention allocation and the influence of different media outlets on consumers' actions are determined jointly.

Our model exhibits strategic complementarity between news consumers and news firms, as well as strategic substitution among firms. Strategic complementarity arises because the more attention news consumers give to a news outlet, the more incentive this outlet has to improve its quality; and of the higher quality are news reports, the more willing are consumers to pay attention to them. This feature of strategic complementarity is consistent with empirical evidence from online media outlets (Sun and Zhu 2013).[4] Strategic substitution among firms arises because of the "attention diversion" mechanism: an improvement in the quality of news from other outlets shifts attention away from an individual firm, which reduces the incentive for its owner to invest in news accuracy. This feature of strategic substitution is broadly consistent with the evidence provided by Gentzkow (2007).

We use our model to explore emerging issues in the media market. As the number of media outlets seems to explode in the past decades, will the trend of media proliferation continue forever? In a standard model of monopolistic competition in the goods market (Dixit and Stiglitz 1977), the number of firms grows to infinity as fixed cost

the implications and consequences of muti-homing behavior of consumers in attention economy, e.g., White and Jain (2018).

[4] They find that online media outlets, such as blogs, take more effort to raise the quality of their products when the attention of news consumers exogenously increases.

of entry goes to zero, and consumers spend an infinitesimal fraction of their expenditure on purchasing the goods from each firm. In the attention economy described by our model, a similar outcome does not obtain. The endogeneity of news quality is the key to understanding this result. If news quality is assumed to be homogenous and fixed among competing firms, it is indeed optimal for consumers to spread their attention evenly and thinly among all firms as the number of news firms grows. However, a firm that gets only a tiny fraction of consumers' attention does not have the incentive to provide quality news. The strategic complementarity between attention allocation by consumers and quality news provision by firms produces a downward spiral, which gives rise to a discontinuous demand function. The implication is that firms cease operating if the attention they can attract is lower than a threshold, which puts an upper bound on the number of firms that can be supported in equilibrium.

We also use the model to study the classical issue of the effect of entry on media competition. We show that there is a quantity-quality trade-off. With more news firms in the market, consumers get to read more stories on the same topic, but the quality of these stories decreases because of the attention diversion effect. In product markets, the relationship between competition and product quality is theoretically ambiguous (Schmalensee 1979; Spence 1975).[5] However, even if news quality is to fall, it is not obvious a priori whether consumers will become more or less informed as new firms enter the market. One contribution of our paper is to provide a framework that leads to an unambiguous answer in evaluating this quantity-quality trade-off. We derive a representation which reduces the sender-receiver game with multiple heterogeneous senders to an aggregative game (Dubey et al. 2006; Acemoglu and Jensen 2013), in which each sender's payoff can be reduced to a function of his own action and an aggregate variable H which reflects the overall informativeness of the media industry as a whole.[6] The equilibrium value of this aggregate variable H provides a summary measure of how effective the media industry as a whole is in informing the public. It can also serve as a useful measure of welfare for this industry.[7] Armed with this welfare measure, we show that overall informativeness of the media industry must

[5] The news media is a platform industry, selling news consumers' attention to advertisers. A recent study of another platform industry—supermarkets—finds that more intense competition raises the quality of supermarket service; in particular, it reduces the probability of product stockouts (Matsa 2010).

[6] The Cournot model with homogeneous goods is an aggregative game (Novshek 1985), but in general, quantity competition in product markets with heterogeneous goods cannot be reduced to an aggregative game.

[7] This welfare measure does not take into account the cost of attention paid by consumers. However, because attention is also valued by advertisers, it is not obvious whether it should be treated as a net cost unless one takes a specific stance on what drives the value of advertising. In any case, the media market is special compared to other product markets because the information produced by the media generates significant externalities through its influence on public opinions. For public policy purposes, focusing on the informativeness of the media appears is much more important than focusing on the cost of attention incurred by consumers or the profits earned by media firms.

increase when new firms enter, despite attention diversion that reduces the quality of news provided each individual firm.

2. Literature Review

Our model of attention allocation builds on Sims (2003), in which the assimilation of available information is plagued by receiver noise.[8] We borrow from Myatt and Wallace (2012), who distinguish information accuracy from clarity, the two distinct aspects of news quality; however, in our model, both information accuracy and clarity are endogenously determined in equilibrium. Our approach therefore also differs from existing works on endogenous information acquisition, where the information structure of the underlying signals is taken as exogenous (e.g., Hellwig and Veldkamp 2009; Myatt and Wallace 2012; and Chen et al. 2015).[9]

Our work contributes to a recent strand of the media economics literature that focuses on the news provision of media outlets that are *not partisan*. Perego and Yuksel (2018) model competition for news consumers at the extensive margin and show that greater competition leads to a smaller but more homogenous reader group for each newspaper; as a consequence, media outlets tilt their resources toward topics closer to the preferences of readers of their own segment and away from topics of general interest. Nimark and Pitschner (2018) study news selection in a setting where readers extract information from both the content of the news and the topic choice made by editors. In our paper, because news consumers spread their attention across multiple firms, competition occurs on the intensive margin. Firms have only one issue to cover but may choose how to present their stories and how much to invest to improve the quality of news sources. Galperti and Trevino (2018) study an environment where an arbitrarily large number of firms engage in perfect competition and emphasize the role of coordination motive among news consumers. Thus, their analysis is complementary to our work, in which we characterize the news industry with monopolistic competition and news readers consume news only to take an informed action.

In the literature on partisan media bias, many studies also focus on how news content can be manipulated by profit-driven news firms. Anderson and McLaren (2012) show that biased media owners may choose to withhold information from the public, and such a strategy can be profitable when rational consumers do not know when the news is being withheld. Bernhardt et al. (2008) show that news firms may try to

[8] Hellwig et al. (2012) review the subsequent research modeling inattention with alternative approaches.

[9] One of the few exceptions is Dewan and Myatt (2008). In a political economy setting, they study the communication strategy of leaders in a beauty contest game and allow leaders to choose the clarity of their messages to seek attention. In our model, clarity is chosen by editors who enjoy wielding influence on the action of news consumers and expressing their own beliefs.

to grab partisan voters' attention by suppressing information that the audience may not like to hear. In our model, firm owners and editors attract readers' attention by increasing the news quality. Sobbrio (2014) studies endogenous news accuracy in a model with partisan bias: the news firms can choose editors based on their ideological preferences, who in turn choose news supply, and consumers turn to like-minded editors for news. In our model, consumers, owners and editors are unbiased; and owners and editors make independent decisions on news accuracy and reporting strategy.

Our paper is related to the media competition literature in general (see Prat and Strömberg 2013 and Gentzkow et al. 2016 for a review). Gentzkow and Shapiro (2006) demonstrate that when competition is not sufficiently strong, media firms may bias their reports towards the prior belief of Bayesian consumers, because readers would discount the quality of the news firms when they obtain news far different from their prior belief. Mullainathan and Shleifer (2005) show that the confirmation bias of consumers may lead to an equilibrium where the media presents biased news to cater to consumers, even when competitors are present. Chan and Suen (2008) study how news firms use coarsened information to influence voters.

Our work also enriches the literature on sender-receiver games by introducing a number of distinctive features that have not been thoroughly researched. First, much of this literature focus on strategic information revelation with two competing senders who intend to influence the receiver's decision (e.g., Ambrus and Takahashi 2008; Battaglini 2002; Bhattacharya and Mukherjee 2013; Chan and Suen 2009; Krishna and Morgan 2001; but see Gentzkow and Kamenica 2015 for a multi-sender model of persuasion), where senders are assumed to be fully informed about the true state. In our model, a large number of senders compete with one another to both influence receivers' actions and attract valuable attention from them. Our assumption that senders acquire noisy and non-identical signals about the state is both realistic and crucial for modeling media quality.[10] Second, in our model, both senders and receivers make decisions on information acquisition, which are endogenously complementary. This feature is different from that in Dewatripont and Tirole (2005), in which one sender and one receiver can both exert effort to improve the amount of information understood by the receiver, but the complementarity in effort is exogenously assumed in the communication technology.

[10] See also Kartik et al. (2017) for a persuasion game model in which multiple senders have noisy and non-identical signals.

3. Media Influence and Attention Allocation

3.1. Editors and Readers

There is a continuum of ex ante homogenous news consumers indexed by $i \in [0,1]$, who acquire information from the media about an uncertain state θ and take an action q_i. In the news market, there is a large but finite number of media firms indexed by $j \in \{1, \ldots, J\}$. Firms and consumers share a common prior belief that θ is normally distributed, with mean μ and variance σ_θ^2. Each media firm is endowed with some evidence about the issue, i.e., a noisy signal about the true state. Let $x_j = \theta + \epsilon_j$ represent such a signal, where ϵ_j is normally distributed with mean 0 and variance $\sigma_{\epsilon j}^2$ (and is independent of the state and independent across different media firms).[11] Let $\gamma_j \equiv \sigma_\theta^2 / (\sigma_\theta^2 + \sigma_{\epsilon j}^2)$ represent the *accuracy* of news outlet j. The accuracy of media firms is summarized by the vector $\gamma = (\gamma_1, \ldots, \gamma_J)$. In this section, γ is assumed to be exogenous, while it will be endogenously determined in Section 4.

Each media firm j has an editor who does not observe the state θ, but only the noisy signal x_j. Editor j writes a news story y_j about θ to inform the news consumers. The editor prefers that the aggregate action taken by news consumers, $Q = \int_0^1 q_i \, \mathrm{d}i$, is close to the true θ, which represents the journalistic goal of informing the public.[12] The news story y_j conveys "new" information beyond the commonly known information on this issue, i.e., the prior knowledge (mean) μ. The editor j needs to take effort to integrate, represent, and illustrate the differences between the substance of the news story from what was previously know about the issue. Therefore, it is natural to assume that the cost that the editor incurs is in proportion to the distance $(y_j - \mu)^2$: it takes more effort for editors to write and deliver a story which is further away from the prior belief of readers (or to include more new content besides the previously known information). Editor j chooses y_j to maximize his payoff,

$$U_j = -\mathrm{E}\left[(Q - \theta)^2 \,\middle|\, x_j \right] - \phi_j (y_j - \mu)^2, \tag{1}$$

where ϕ_j is interpreted as the marginal cost of the editor's effort. A higher ϕ_j means that the editor incurs a higher cost to convey new information in his story.

In this paper, we assume that the editors are not biased in a partisan or behavioral fashion. Similar assumptions have been adopted in the recent literature, such as Perego and Yuksel (2018) and Galperti and Trevino (2018). It also resembles that in

[11] It is realistic to expect that the noise term may be correlated across media firms conditional on the state. We consider such a scenario in Section 4.5.

[12] As defined by American Press Association, "[T]he central purpose of journalism is to provide citizens with accurate and reliable information they need to function in a free society." (See `https://americanpressassociation.com/principles-of-journalism/`.)

Nimark and Pitschner (2018), where news providers are benevolent and select news to maximize the utility of their readers.[13]

In general, the reporting strategy is a function of the signal, $y_j(x_j)$. We focus on linear equilibria in which the reporting strategy takes a linear form:

$$y_j = \alpha_j x_j + \alpha_{j0}. \tag{2}$$

A high α_j means that the story closely reflects the evidence (or the underlying signal) obtained from the news investigation, while a low α_j represents a "cookie-cutter" style of reporting that produces standardized stories which fail to reflect all the nuances of the evidence. Two polar cases help illustrate. When ϕ_j goes to infinity, the editor j only reports $y_j = \mu$, i.e., $\alpha_j = 0$ and $\alpha_{j0} = \mu$, because it is extremely costly for him to integrate any new information. When ϕ_j is zero, the editor j can describe and report whatever he knows with ease, that is, $y_j = x_j$, i.e., $\alpha_j = 1$ and $\alpha_{j0} = 0$. We use α_j to denote the strategy of the editor j.[14]

The strategy of news editors is summarized by the vectors $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_J)$ and $\boldsymbol{\alpha_0} = (\alpha_{10}, \ldots, \alpha_{J0})$. The assumption of rational Bayesian news consumers implies that the intercepts $\boldsymbol{\alpha_0}$ of the linear reporting strategies do not systematically distort consumers' actions. Our analysis focuses on the vector of slopes $\boldsymbol{\alpha}$.

Each news consumer chooses to acquire information from the media about θ, including which news reports he wants to pay attention to and how much attention he pays to each report. If consumer i picks up the news report y_j, he reads the news content with a reader noise η_{ji} attached to the actual report. That is, he observes

$$\hat{y}_{ji} = y_j + \eta_{ji}, \tag{3}$$

where $\eta_{ji} \sim N(0, \sigma^2_{\eta ji})$ is independent of y_j and independent across news consumers. This specification captures the idea that an individual has limited capacity to process all the information contained in a story; he reads the content of a news story with actual or interpretive errors. The variance of interpretive errors or reader noise is not exogenous, and it depends on the attention or capacity spent on the news story. News consumer i can read a news story with greater precision by paying more attention to it.

[13] Political and behavioral bias are not the major theme of this line of works, which instead focuses on the determination of new quality.

[14] The strategy α_j may correspond to the notion of "clarity," in the information acquisition literature, as conceptualized in Myatt and Wallace (2012). For a news article, its clarity refers to how easy readers can understand the news content, in contrast to the definition of news accuracy, which is the amount of information contained. Whether readers can understand the news content with ease is related to how the news content is presented, which is captured by the reporting strategy in this model. This interpretation will become clearer once the information extracted from outlet j by the news consumer is fully described in equation (5) below.

Let z_{ji} represent the amount of attention devoted to news outlet j. The noise reduction technology is specified as:

$$\sigma^2_{\eta ji} = \frac{\chi^2}{z_{ji}},$$

where χ is a constant capturing the technological aspect of the information assimilation process. If consumer i pays no attention to the news story j, i.e., $z_{ji} = 0$, the variance of the reader noise is infinity and the news content is totally uninformative. If consumer i pays an infinite amount of attention to the news story j, the variance of the reader noise is zero and consumer i obtains the story y_j precisely. This noise reduction technology is commonly adopted in the attention allocation literature; it specifies that the precision of the noise is linearly related to the attention devoted to the information source (Myatt and Wallace 2012; Mondria and Quintana-Domeque 2013).[15]

The information set available to consumer i is an array of his perceived reports, $(\hat{y}_{1i}, \ldots, \hat{y}_{Ji})$. Given his information set, consumer i chooses action q_i to maximize $-\mathrm{E}[(q_i - \theta)^2]$. The optimal action strategy $q_i(\hat{y}_{1i}, \ldots, \hat{y}_{Ji})$ is in general a function of the J perceived news reports. In a linear equilibrium with Gaussian signals, the optimal action $q_i = \mathrm{E}[\theta \mid \hat{y}_{1i}, \ldots, \hat{y}_{Ji}]$ is also linear:

$$q_i = \beta_{0i} + \sum_{j=1}^{J} \beta_{ji} \hat{y}_{ji}. \tag{4}$$

Because news consumers are ex ante identical, we focus on equilibria in which their strategies are identical (but their actions may be different since each consumer perceives a different report $\hat{y}_{ji}$ based on the same story y_j). From here on, for $j = 0, 1, \ldots, J$, we write $\beta_{ji} = \beta_j$ for all news consumer i. The common action strategy of news consumers is represented by the constant β_0 and the vector of *action weights* $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)$ that they attach to the perceived stories of the news outlets. The constant does not play a role in the subsequent analysis and we focus on the action weights $\boldsymbol{\beta}$.

Note that the reporting strategy of the editors' plays a role in both supply and demand sides of the information transmission process. On the one hand, given the linear reporting strategy, the information content in $\hat{y}_{ji}$ can be written as:

$$\frac{\hat{y}_{ji} - \alpha_{j0}}{\alpha_j} = \theta + \left(\epsilon_j + \frac{1}{\alpha_j} \eta_{ji} \right). \tag{5}$$

[15] In spirit, this specification is the same as the seminal idea of rational inattention proposed by Sims (2003) that the amount of information conveyed is increased when the receiver devotes more information-processing capacity to the underlying signal. The subtle difference is the noise reduction technology. Because entropy-based technology is neither linear nor separable, we prefer the linear technology for its tractability.

The effective reader's noise, $\epsilon_j + \eta_{ji}/\alpha_j$, is smaller when α_j is higher for a fixed amount of attention z_{ji}. That is, for the same amount of attention paid, clearer reports (with higher α_j) give rise to smaller reader's noise. On the other hand, news stories are less difficult to digest, when they are presented in a clearer fashion. Such a cognitive aspect is captured by assuming that the marginal cost of attention devoted to clearer stories is lower. In particular, we let $p_j = p/\alpha_j^2$ represent the marginal cost of giving attention to media outlet j, where p is the common component and α_j^2 is the outlet specific component.[16]

To sum up, the objective of the news consumer i is to choose q_i and z_{ji} to maximize his net payoff:

$$\max_{z_{ji}} \left\{ \max_{q_i} \left\{ -\mathrm{E}\left[(q_i - \theta)^2 \,\middle|\, \hat{y}_{1i}, \ldots, \hat{y}_{Ji} \right] \right\} - \sum_{j=1}^{J} \frac{p}{\alpha_j^2} z_{ji} \right\}. \tag{6}$$

Because news consumers are homogenous ex ante, they make the same information choice. In what follows, we suppress the subscript i and write z_j for z_{ji} unless it causes confusion. The attention allocation of news consumers is summarized by the vector $\boldsymbol{z} = (z_1, \ldots, z_J)$.

The timing of the game is as follows. Editors simultaneously choose the stories y_j to publish based on the news sources x_j endowed. News consumers choose their attention allocation $\boldsymbol{z}$ and their actions q_i based on the perceived stories $\hat{y}_{ji}$ they read. The editors and news consumers play a sender-receiver game with multiple senders and multiple receivers. In equilibrium, taking accuracy γ as given, the reporting strategies of editors (summarized by $\boldsymbol{\alpha}$) and the attention allocation and action strategies of news consumers (summarized by $\boldsymbol{z}$ and $\boldsymbol{\beta}$) are best responses to one another.

We take two steps to analyze this model. In Section 3.2, we first fix the attention allocation $\boldsymbol{z}$ chosen by readers and study how editors' reporting strategy $\boldsymbol{\alpha}$ respond to action weights $\boldsymbol{\beta}$ chosen by readers; and vice versa. The solution to this sender-receiver game allows us to characterize the influence of individual news outlets, and to derive an aggregate variable that summarizes the influence of the media industry as a whole. In Section 3.3, we study the attention allocation decision $\boldsymbol{z}$ of news consumers. Then, we fully characterize the equilibrium with $(\boldsymbol{\alpha}^*, \boldsymbol{\beta}^*, \boldsymbol{z}^*)$.

[16] This assumption of allowing the marginal cost to vary in α_j is not crucial for our results. In online Appendix B, we explain how it improves the tractability of our analysis, and why our results are robust when the marginal cost does not depend on α_j.

3.2. The Sender-Receiver Game

Each individual editor j chooses a story y_j to maximize his payoff U_j described in equation (1), given the strategies of news consumers and of other editors. In a linear equilibrium where other editors follow the reporting strategy (2) and news consumers follow the identical action strategy (4), the aggregate action is

$$Q = \beta_0 + \beta_j y_j + \sum_{k \neq j} \beta_k (\alpha_k x_k + \alpha_{k0}).$$

Substitute this expression into the objective function (1), the first-order condition for y_j is

$$\mathrm{E}\left[\beta_j \left(\beta_0 + \beta_j y_j + \sum_{k \neq j} \beta_k (\alpha_k x_k + \alpha_{k0}) - \theta \right) + \phi_j \left(y_j - \xi_j \right) \middle| \, x_j \right] = 0.$$

For $k \neq j$, we have $\mathrm{E}[x_k \mid x_j] = \gamma_j x_j + (1 - \gamma_j)\mu$. Therefore, the solution to the first-order condition gives

$$y_j = \frac{\gamma_j \beta_j}{\beta_j^2 + \phi_j} \left(1 - \sum_{k \neq j} \alpha_k \beta_k \right) x_j + \text{constant}. \tag{7}$$

Thus, when news consumers and other editors adopt linear strategies, the best response for editor j indeed takes the linear form (2), with

$$\alpha_j = \frac{\gamma_j \beta_j}{\beta_j^2 + \phi_j} \left(1 - \sum_{k \neq j} \alpha_k \beta_k \right), \tag{8}$$

and with α_{j0} equal to the constant in equation (7). The intercept α_{j0} plays no role in the subsequent analysis because its equilibrium value is commonly known.

An important feature of this multi-sender game is that the reporting strategy of news reports exhibits strategic substitution: equation (8) shows that a higher α_k ($k \neq j$) lowers α_j. This feature arises because all editors prefer that the public be informed. If other editors are producing clearer news stories, then an individual editor can free ride on their efforts.

Fix news consumers' action strategy $\boldsymbol{\beta}$ and their attention allocation $\boldsymbol{z}$, equation (8) holds for every $j = 1, \ldots, J$. The solution to this equation system gives $\boldsymbol{\alpha} = \boldsymbol{a}(\boldsymbol{\beta}; \boldsymbol{z})$.

Lemma 1. *The reporting strategy $\alpha_j = a_j(\boldsymbol{\beta}; \boldsymbol{z})$ of news outlet j is given by:*

$$\alpha_j = \frac{1}{\beta_j} \frac{\frac{\gamma_j \beta_j^2}{(1-\gamma_j)\beta_j^2 + \phi_j}}{1 + \sum_k \frac{\gamma_k \beta_k^2}{(1-\gamma_k)\beta_k^2 + \phi_k}}. \tag{9}$$

It increases in news accuracy γ_j and decreases in cost ϕ_j. It decreases in the accuracy and increases in cost of other news outlets.

The reporting strategy α_j chosen by editor j increases with accuracy γ_j. When the underlying news signal x_j is more informative, the incentive to inform the public about the signal is higher. It is also intuitive that α_j decreases in the editor's cost ϕ_j. Finally, because of the strategic substitution effect, the response with respect to other news outlets' accuracy and cost is opposite to that with respect to own accuracy and cost.

Turning to news consumers' action strategy, the quadratic loss function implies that $q_i = \mathrm{E}[\theta \mid \hat{y}_{1i}, \ldots, \hat{y}_{Ji}]$. To derive this conditional expectation, we recall equation (5) and let τ_j represent the precision of the combined noise term (relative to the precision of the prior belief). We have

$$\tau_j = \frac{1}{\frac{1-\gamma_j}{\gamma_j} + \frac{\chi^2}{z_j \alpha_j^2 \sigma_\theta^2}}. \tag{10}$$

Fix editors' reporting strategy $\boldsymbol{\alpha}$ and news consumers' attention allocation $\boldsymbol{z}$, the optimal action rule $\boldsymbol{\beta} = \boldsymbol{b}(\boldsymbol{\alpha}; \boldsymbol{z})$ can be obtained from the linear Bayesian updating formula.

Lemma 2. *The action weight $\beta_j = b_j(\boldsymbol{\alpha}; \boldsymbol{z})$ of news outlet j's report is given by:*

$$\beta_j = \frac{1}{\alpha_j} \frac{\tau_j}{1 + \sum_k \tau_k}. \tag{11}$$

It increases in the accuracy γ_j of news outlet j and in the attention z_j paid to it. It decreases in the accuracy of and the attention given to other news outlets.

News consumers increase their reliance on the news from outlet j when they pay more attention to it and when its underlying news source is more accurate. Similarly, news consumers decrease their reliance on report j when they pay more attention to other reports and when the other news sources are more accurate.

Given $\boldsymbol{z}$, the equilibrium $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ of the sender-receiver game can be obtained by solving (9) and (11) jointly. To state our formal result, define for $j = 1, \ldots, J$ the quantity:

$$h_j \equiv 1 - \frac{\chi}{\sigma_\theta}\sqrt{\frac{\phi_j}{z_j \gamma_j}}; \tag{12}$$

and for any subset of media outlets $G \subseteq \{1, \ldots, J\}$, define

$$H_G \equiv \frac{\sum_{j \in G} \frac{\gamma_j}{1-\gamma_j} h_j}{1 + \sum_{j \in G} \frac{\gamma_j}{1-\gamma_j}} \tag{13}$$

Proposition 1. *Given attention allocation $\mathbf{z}$ and accuracy profile $\boldsymbol{\gamma}$, there exists an equilibrium $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ of the sender-receiver game with a set of media outlets $G \subseteq \{1, \ldots, J\}$ such that: (a) if $j \in G$, then*

$$\hat{\alpha}_j \hat{\beta}_j = \frac{\gamma_j}{1-\gamma_j}(h_j - H_G) > 0; \tag{14}$$

and (b) if $j \notin G$, then $\hat{\alpha}_j \hat{\beta}_j = 0$.

In the proof of Proposition 1, we provide separate formulas for $\hat{\alpha}_j$ and $\hat{\beta}_j$, but equation (14) gives the product of the two. Note that $\hat{\alpha}_j \hat{\beta}_j$ measures the *influence* of news sources from media outlet j on the aggregate action of consumers. That is because, with a linear action rule and a linear reporting strategy, we have

$$\hat{\alpha}_j \hat{\beta}_j = \frac{\mathrm{Cov}[Q, x_j]}{\sigma_\theta^2}.$$

The influence of news outlet j depends on the magnitude of h_j relative to H_G. From equation (12), we see that h_j increases in accuracy and attention but decreases in ϕ_j.

An important by-product of Proposition 1 is that it provides a summary measure of the informativeness of the media as a whole, given by H_G. The variable H_G is a weighted average of the influence of each outlet j, adjusted by a factor less than 1.[17] The influence of a news outlet depends on the characteristics of other news outlets only through this aggregate variable H_G. If we sum equation (14) over j, we can show that

$$H_G = \sum_j \hat{\alpha}_j \hat{\beta}_j = \frac{\mathrm{Cov}[Q, \theta]}{\sigma_\theta^2}.$$

In other words, H is the *total influence* of the news industry as a whole. The higher is H, the more effective is the industry in informing the public to choose an aggregate

[17] The weight on h_j is $\gamma_j/(1-\gamma_j) = \sigma_\theta^2/\sigma_{\epsilon j}^2$, which reflects the precision of signal x_j (relative to the prior). The adjustment factor is $\sum_j \sigma_{\epsilon j}^{-2} / \left(\sigma_\theta^{-2} + \sum_j \sigma_{\epsilon j}^{-2}\right)$, which reflects the combined precision of all news sources as a fraction of total precision (news sources plus prior).

action Q that closely tracks the true state θ. Furthermore, using Lemma 2, we have

$$H_G = \frac{\sum_j \tau_j}{1 + \sum_j \tau_j}.$$

Thus, we sometimes also refer to H_G as *total media informativeness* (relative to the prior). This variable plays a key role in our model.

We call the set G in Proposition 1 the *active media group*, because news outlets in this group have positive influence. Given an equilibrium with active media group G, Proposition 1 determines a unique strategy profile $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ corresponding to that equilibrium.[18] However, there exist multiple equilibria in this sender-receiver game, with a different active media group in each equilibrium. Coordination failure is the reason behind equilibrium multiplicity. If stories from news outlet j do not contain any information content (i.e., $\alpha_j = 0$), news consumers do not act on it; and if consumers' actions do not put any weight on stories from this outlet (i.e., $\beta_j = 0$), its editor has no incentive to present any facts obtained. Because of this coordination problem, given any equilibrium with active media group G, it remains an equilibrium if we delete an arbitrary subset of media outlets from this group.[19] Of course, the influence of the remaining outlets in the active media group will be different after the deletion. The following result shows that an equilibrium with a larger active media group (in the set inclusion order) is more informative than one with a smaller active media group.

Lemma 3. *Suppose there is an equilibrium of the sender-receiver game with active media group G. For any $G' \subset G$, there is an equilibrium with active media group G' with $H_{G'} < H_G$.*

3.3. Attention Allocation and Equilibrium

Because action q_i of news consumer i is chosen to be equal to the posterior mean of θ, the expected value of the quadratic loss function $(q_i - \theta)^2$ is simply the posterior variance of θ. Since the posterior precision of θ is equal to the prior precision plus the precisions from all the signals about θ, at the attention allocation stage, news consumers' objective (6) can be written as

$$\max_{z} \left\{ -\frac{\sigma_\theta^2}{1 + \sum_j \tau_j} - \sum_j \frac{p}{\alpha_j^2} z_j \right\},$$

[18] If $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ is an equilibrium profile corresponding to G, it remains an equilibrium profile when we replace both $\hat{\alpha}_j$ and $\hat{\beta}_j$ by their negative values. In general, these alternative profiles are payoff equivalent.

[19] A babbling equilibrium (one in which G is the empty set) is an extreme manifestation of this coordination failure. This feature of coordination failure will be extended and generalized in the monopolistic competition problem. See our discussion regarding Proposition 5.

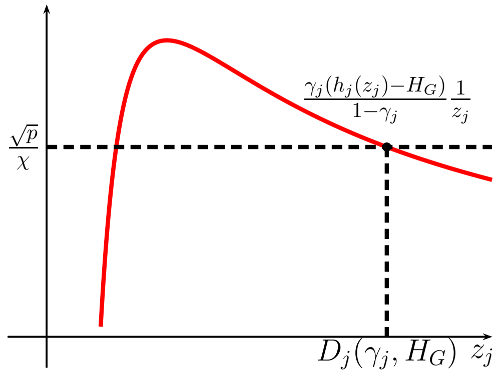


Figure 1. *The reduced-form marginal benefit of attention is increasing then decreasing in z_j. The key equation (16) either has two solutions or no solution. When there are two solutions, the demand function is defined to be the larger root.*

where τ_j is given by equation (10) and increases in z_j. The first-order conditions for z_j are:

$$\frac{\tau_j}{1+\sum_k \tau_k}\frac{1}{z_j} - \frac{\sqrt{p}}{\chi} = 0. \tag{15}$$

By Lemma 2, $\tau_j/(1+\sum_k \tau_k)$ is simply the influence $\alpha_j\beta_j$ of news outlet j. In other words, equation (15) shows that the attention given to a news outlet is proportional to its influence. To be sure, we do not ascribe a causal interpretation to this relationship, because both attention and influence are jointly determined.

Using equation (14) for $\alpha_j\beta_j$, and writing $h_j(z_j)$ to emphasize the dependence of h_j on z_j according to equation (12), we can express the first-order conditions (15) and the equilibrium conditions for media influence (14) solely in terms of z_j, given by the following key equation:

$$\frac{\gamma_j(h_j(z_j)-H_G)}{1-\gamma_j}\frac{1}{z_j} = \frac{\sqrt{p}}{\chi}. \tag{16}$$

Equation (16) can be interpreted as a reduced-form first-order condition for z_j, after taking into account the fact that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are endogenously determined in the sender-receiver game. We illustrate this point with Figure 1. The left-hand side of the equation is increasing then decreasing in z_j. For fixed $\boldsymbol{\alpha}$, we have $\partial^2 V/\partial z_j^2 < 0$. Diminishing returns to attention, reflected in the term $1/z_j$, accounts for the decreasing part of the graph. However, news consumers put more action weight β_j on news outlet j as they pay more attention to it. By Lemma 1, a higher action weight β_j induces editor j to present the evidence better in equilibrium, i.e., increase his choice of reporting strategy α_j. But as it increases, the marginal benefit from paying attention also increases (i.e., $\partial^2 V/\partial\alpha_j\partial z_j > 0$). This effect is reflected in the term $h_j(z_j)$, which increases in z_j, and it accounts for the increasing part of the graph in Figure 1.

The hump-shaped reduced-form marginal benefit of attention implies that some media outlet j can be endogenously ignored by news consumers if its accuracy is sufficiently low. In Figure 1, a low enough γ_j will shift the marginal benefit curve below $\sqrt{p}/\chi$. The flip side of this argument is also true: if the marginal cost of attention p is sufficiently low, there exists an equilibrium in which all outlets belong to the active media group.

Proposition 2. *Fix any set of media firms $G \subseteq \{1, \ldots, J\}$, there exists $\tilde{p}_G$ such that for any $p \leq \tilde{p}_G$, there is an equilibrium in which the corresponding profile $\{z_j^*, H_G^*\}$ for $j \in G$ is determined by equations (13) and (16).*

Proposition 2 is established by showing that the system of equations represented by (13) and (16) has a solution when p is low. Further, given an equilibrium attention allocation $\boldsymbol{z}^*$ characterized by Proposition 2, equilibrium reporting strategy $\boldsymbol{\alpha}^*$ and action weights $\boldsymbol{\beta}^*$ are obtained from Proposition 1.

3.4. Information Overflow?

In this section, we analyze an emerging issue in the media market, utilizing our framework with endogenous reporting strategy choices by editors and attention allocation by news consumers. As discussed in the introduction, the most salient development of the news market in the last few decades is the proliferation of media outlets. Figure 2 shows the number of online news sites using data from the Guide to Online News Startups of *Columbia Journalism Review*, which keeps track of online news sites that satisfy the following four criteria: the outlet has to be primarily devoted to original reporting and content production; it should have full-time employees; it is independent and not the web arm of a legacy media entity; and it attracts financial support through advertising, grants, or other revenue sources to sustain its operation. On the one hand, the upward trend shown in this figure is unmistakable. On the other hand, the absolute number of independent online news content producers—298 in 2011—cannot be considered exceptionally large by historical comparison. For example, Lee (1947) documents that the number of daily newspapers in America rose from 1,610 in 1899 to 2,600 in 1909. Despite the fact that technology has drastically reduced the cost of starting an online news site, the number of news content providers has increased but has not exploded. A few natural questions arise: Will the current trend of proliferation of news media outlets continue indefinitely? How do consumers cope with the abundance of information? Will they eventually become perfectly informed as the cost of operating a news firm falls?

In much of the existing literature, the growth of media outlets is assumed exogenously and it is predicted that news consumers would get better and better informed

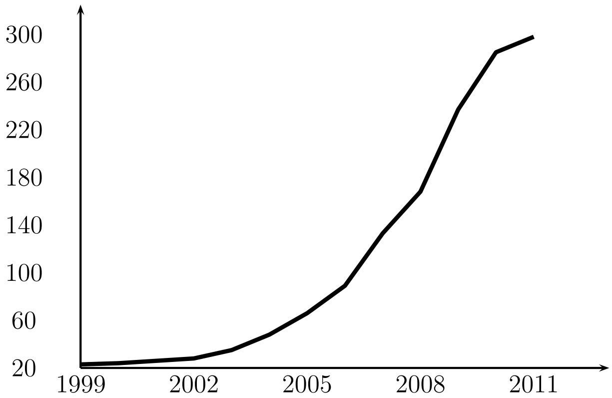


Figure 2. *Number of online news sites. Source: Compiled from the Guide to Online News Startups.*

as the number of media firms continues to grow. For example, Chan and Suen (2008) show that, in a Hotelling model in which media outlets compete for audience size, the proliferation of news firms as entry cost shrinks to zero produces the full-information outcome (see also Chan and Stone 2013). However, our model with endogenous news quality and competition for attention predicts otherwise: both the informativeness of the news industry and the number of active firms would eventually reach a limit, despite the seemingly unbounded availability of news sources.

Two factors are at play in our model that limit the benefit from having a new entry into the media industry. The first is the cost of attention. The marginal benefit from having more information falls as news consumers become better informed, while the marginal cost of attention does not. In the limit, even if the number of news outlets goes to infinity so that perfect information is feasible, news consumers rationally choose to remain partially uninformed (by paying very little attention to each news outlet). The second factor is novel in our model. We show that even if the number of news outlets goes to infinity, only a finite number of them can be active in equilibrium because news quality is endogenous in our model.

To illustrate these two factors in the simplest way, we study a special case of the model, in which media firms are identical, i.e., accuracy γ and cost ϕ are the same across firms. We first consider the effect of costly attention and shut down the mechanism of endogenous reporting strategy. To do so, we fix α exogenously at some $\overline{\alpha}$. Let the number of media firms be J, then the optimal attention given to each news outlet is:

$$z = \begin{cases} \frac{\chi}{\sqrt{p}} \frac{\gamma}{1-\gamma} \left(\frac{1 - \frac{\sqrt{p}\chi}{\overline{\alpha}^2 \sigma_\theta^2}}{1 + J\frac{\gamma}{1-\gamma}} \right) & \text{if } \overline{\alpha}^2 > \sqrt{p}\chi/\sigma_\theta^2; \\ 0 & \text{otherwise.} \end{cases}$$

When $\overline{\alpha}^2 > \sqrt{p}\chi/\sigma_\theta^2$, we have $z > 0$ regardless of J. That is, news consumers spread their attention evenly among different news outlets. As the number of firms increases, the attention z given to each news outlet falls and the total media informativeness increases. But total informativeness approaches a limit strictly less than 1:

$$\overline{H} \equiv \lim_{J\to\infty} \frac{1}{1+J\tau} = 1 - \frac{\sqrt{p}\chi}{\overline{\alpha}^2\sigma_\theta^2}.$$

The higher is the marginal cost of attention p, the lower is $\overline{H}$, which reflects the rational decision of news consumers to remain partially uninformed when attention is costly.

Next, we allow for the effect of the endogenous determination of the reporting strategy. In this case, the equilibrium value of α is determined by, among other things, editors' marginal cost ϕ. To maintain comparability with the model of exogenous reporting strategy in the previous case, we choose the parameter ϕ so that the equilibrium value of α is equal to $\overline{\alpha}$.

Proposition 3. *In any equilibrium in which reporting strategy of the active firms is equal to $\overline{\alpha}$, equilibrium informativeness satisfies*

$$H_G^* \leq 1 - \frac{3}{2}\frac{\sqrt{p}\chi}{\overline{\alpha}^2\sigma_\theta^2} < \overline{H},$$

regardless of the total number of news firms J in the industry.

Even when the number of news firms goes to infinity, total media informativeness is bounded away from $\overline{H}$. This result obtains because strategic substitution in the model with endogenous reporting strategy imposes a constraint on the number of active media firms in the industry. Figure 3 illustrates this. When there are more active media outlets, holding attention to each firm constant, the influence of each firm will fall. This follows from equation (14), in which an increase in H_G (induced by a larger number of active firms) causes $\alpha\beta$ to fall. But because attention is proportional to influence via equation (15), z also falls in response. This is shown by the decrease from z^1 to z^2 in Figure 3, as the number of active firms increases from n_1 to n_2 and total influence H increases. If the number of firms further increases to n_3, the additional benefit from writing clear factual stories to inform the public is very small because news readers are already very well-informed. Instead, a fraction of news editors (e.g., $n_3 - n_2$) will choose not to take effort to write factual stories or stop writing at all (i.e., $\alpha_j = 0$), and they do not gain any attention from readers who are interested in learning about the state. Equilibrium cannot sustain a symmetric outcome in which all n_3 news outlets are active and receive positive attention from serious news consumers.

This mechanism of endogenous quality is particularly relevant in the media environment nowadays. It has been often observed that new media sites happened to catch

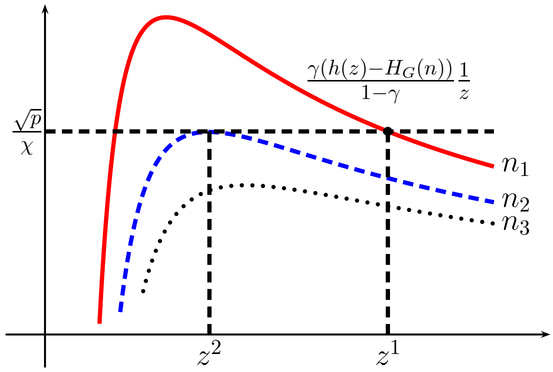


Figure 3. *When the number of firms increases from n_1 to n_2, total media informativeness increases and the solid curve shifts down accordingly. The equilibrium amount of attention devoted to each outlet drops from z^1 to z^2. In this case, n_2 is the largest number of media that consumers can pay attention to, and z^2 is the smallest possible amount of attention paid to each medium. Any equilibrium with a larger number of active firms (e.g., n_3) cannot be sustained.*

the attention of news consumers and became prominent content providers by capitalizing on it, but they declined quickly when competitors diverted consumers' attention away from them and eventually closed down as the quality of contents deteriorated. The rise of fall of news site Digg is a case in point.[20] More generally, this mechanism is consistent with the fact that the vast majority of independent online news sites fail to grab the attention of more than one percent of news consumers.[21]

Proposition 4. *For any given ϕ and γ, there exists an upper bound $\overline{n}(\phi,\gamma)$ such that in any equilibrium with active media group G, the number of active firms in G is lower than $\overline{n}$, regardless of the total number of firms J in the industry. Furthermore, $\overline{n}$ increases in γ and decreases in ϕ.*

It follows from Proposition 4 that, when the total number of firms J is large enough, news consumers' attention allocation is asymmetric: they pay attention to at most $\overline{n}$ media outlets and ignore the rest. Since the number of active firms remains finite, the attention that each active media outlet receives remains bounded away from zero even when J goes to infinity (it cannot fall below z^2 in Figure 3).

Propositions 3 and 4 offer a prediction for the future of the media industry: the seemingly ever-increasing trend in the number of news outlets, illustrated in Figure 2, will eventually cease, and news consumers will never choose to be fully informed either. The strategic complementarity between news consumers and news firms that

[20] See "Digg Failed Because 'Social Media Grew Up'," *Wall Street Journal*, July 13, 2012.

[21] The Pew Media Consumption Survey indicates that 26 online news sites (some of which are the web arm of legacy media entities) were named by at least one percent of the respondents as their source of accessing news in 2010, which implies that the majority of the 298 independent online news sites identified in the Guide to Online News Startups do not attract a lot of attention.

we analyze above implies one crucial difference between the equilibrium of the news market and that of the product markets: the "zero-profit condition" does not necessarily apply. Even if there is no entry barrier to the news industry, new firms may not be able to enter the market and grab enough attention when all the existing firms earn positive profits.

Relating to the existing literature, our analysis reveals that models with exogenous and endogenous information quality can deliver qualitatively different predictions about the news market. Our results that attention allocation is asymmetric even among symmetric media, and that there exists an upper bound for the number of media firms that can receive attention and a lower bound for the amount of attention that has to be paid to each active firm, are driven by the endogenous choice of editorial strategy. They are different from the predictions of models with exogenous reporting strategy, where news consumers diversify their attention to all existing firms, with each firm getting an amount of attention that approaches zero.[22]

4. Monopolistic Competition of News Firms

In this section, we enrich our framework and further endogenize another element of media quality—accuracy—which is taken as given in previous sections. Toward this end, we first introduce a profit-seeking owner in each media firm, assuming that the owner decides on the investment in news gathering (or investigations) and that they turn the attention received into advertisement revenues. We then characterize the equilibrium with endogenous accuracy, reporting strategy, attention allocation, and action weights.

In Section 4.4, we use this framework to study the effects of increased media competition. Our model rationalizes two seemingly contradictory facts as media competition strengthens: that news consumers perceive news quality to have declined, but that they spend more time with news nowadays. This application also illustrates the necessity of modeling media competition at the intensive margin, as a similar prediction cannot be obtained when media firms compete only for the audience.

4.1. Owners

Each media firm j consists of two players: an owner who runs the business for profit, and an editor whose objective is described in equation (1). All players share a common normal prior about the state. We choose to separate the two roles within a media firm so as to capture the norm of editorial independence in modern journalism, which pre-

[22] In Myatt and Wallace (2012), the quality of signals is exogenous and receivers rationally ignore very unclear signals. But they pay attention to all of them if they are of the same quality and clear enough, which differs from the prediction of our model with endogenous information quality.

scribes that owners do not get too involved in the decisions of the editorial office.[23] It also allows us to endogenize both accuracy (chosen by owners) and reporting strategy (chosen by editors)—two distinct aspects of news quality.

Each owner j decides the resources to make investigations about the state. This includes decisions such as funds made available to journalists to do research, and the size and quality of the editorial staff. Once the basic infrastructure of the news office is determined, the owner stays away from editorial decisions concerning the selection or presentation of news stories.

We assume that perfect accuracy is not feasible: $\gamma_j \in [0, \overline{\gamma}]$ for some $\overline{\gamma} < 1$. The cost of investigation $C_j(\gamma_j)$ is increasing and convex in γ_j. The owners choose the accuracy of their news signals simultaneously. These decisions, once made, are known to all news consumers and editors. This assumption captures the idea that the accuracy of news outlets depends on their long term investments, which are generally well observed to players in the media market.[24] The strategy of owners is summarized by the vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_J)$.

News outlet j receives revenue from advertisements, which is assumed for simplicity to be proportional to the attention received and is denoted by $S_j z_j$, where z_j is the amount of attention from news consumers given to that outlet and $S_j > 0$ is the revenue per unit of attention. The objective of media owner j is to choose media quality γ_j to maximize profits,

$$\Pi_j = S_j z_j - C_j(\gamma_j).$$

The timing of this enriched game is summarized in Figure 4. Owners of media firms simultaneously choose the accuracy $\boldsymbol{\gamma}$ of their news outlets, which becomes commonly known to editors and news consumers. Then editors and news consumers play the sender-receiver game described in Section 3.

4.2. Monopolistic Competition and the Discontinuous Demand Function

When the owner of a news firm tries to attract attention by investing in accuracy to raise h_j (defined in equation (12)), total influence H_G will also rise because H_G is an adjusted weighted average of h_j. However, the effect of a change in h_j on H_G is of

[23] It is common that editorial and business decisions are made separately in news firms nowadays. The standard practice is that both the chief editor and managerial head, such as CEO, are appointed independently by trustees, so that the editorial decisions are not compromised by commercial interests. *The Economist* magazine is a case in point. See `http://www.economistgroup.com/results_and_governance/trustees.html`. Alternatively, Sobbrio (2014) studies a contrasting scenario of the media market with "citizen journalists," in which owners can pick editors, and shows how such a market functions. Our works differ, but are complementary to each other.

[24] The assumption is made to abstract away the reputation concern, which has been thoroughly studied in the media literature (e.g., Gentzkow and Shapiro 2006).

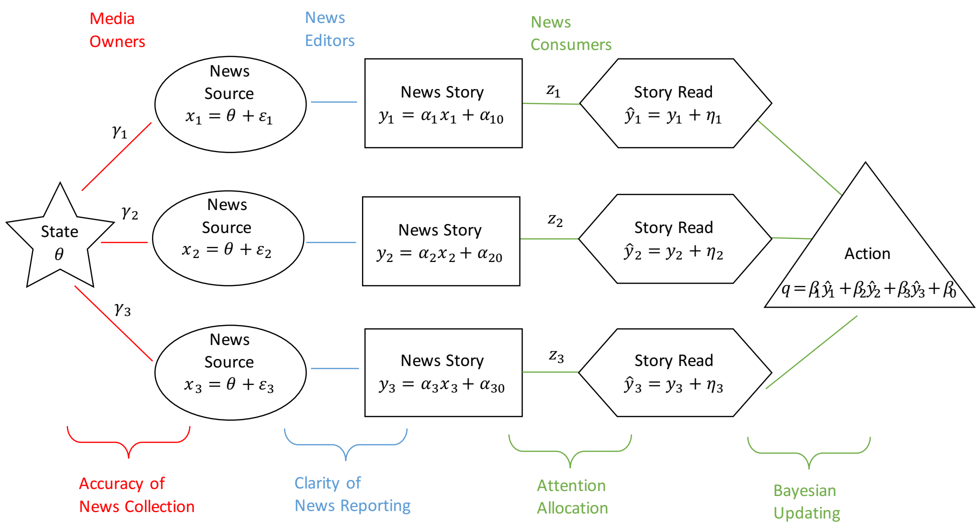


Figure 4. *Timing of the model with an example of three media firms.*

order $1/|G|$. When the number of news firms in the active media group G is large, this effect is small. The proliferation of news firms in the modern news market means that a change in each individual firm only has a tiny impact on the industry. It is therefore realistic to assume that each media owner takes the overall media environment as given and ignores the impact of his own action on the news industry as a whole.

Such an assumption deserves elaboration. First, it is commonly adopted in monopolistic competition models. In product markets, a producer of differentiated goods takes the aggregate price as given and chooses the price of his own variety, while ignoring the impact of his own price on the price level. Similarly, in our model, a large number of news firms are producing differentiated products (news stories that provide conditionally independent information about the state), and the total influence H_G serves as an aggregate variable (akin to the aggregate price level in product markets) that summarizes the overall media environment. Second, the market structure of monopolistic competition allows us to characterize the model analytically. In a setting where firms do take into account their negligible influence on the aggregate of a large market, this model is not analytically tractable but the results remain qualitatively the same. In online Appendix C, we study one such case where owners take the choices of the accuracy of other owners' as given and demonstrate that our results are robust and not specific to the market structure of monopolistic competition.

In this section, we establish that a monopolistic competitive equilibrium exists, in which each firm optimizes by taking total influence as given and the aggregation of their optimal choices is consistent with the conjectured aggregate. We start by characterizing the "demand function" for each news firm, that is, how total influence at the industry level and investment at the firm level affect the amount of attention that each firm receives, which arises from the sender-receiver game played by editors and news consumers.[25] Recall that, given total influence H_G, the key equation (16) determines attention z_j given to each news outlet j. It is clear from Figure 1 that equation (16) either admits two solutions or no solution. When there are two solutions, we focus on the larger one because it is a locally "stable" root and gives intuitive comparative statics results.

Since the left-hand-side of (16) is increasing in γ_j, there exists a critical $\underline{\gamma}_j$ (which depends on H_G and other parameters) such that a solution to the equation does not exist if γ_j is below this value. When accuracy γ_j is low, Proposition 1 shows that the influence of news outlet j is low too, because editor j offers less clear report (lower α_j), and readers also put less weight of its report (lower β_j). Further, the marginal benefit of attention is proportional to influence (see equation (15)). Therefore, the marginal benefit of paying attention to a news outlet with low γ_j can be below the marginal cost of attention. As a result, news consumers ignore news outlets with low accuracy, i.e., attention received by firm j is zero when its accuracy is lower than $\underline{\gamma}_j$. Such a discontinuity in demand function would not arise if the reporting strategy of editors is exogenous.

Formally, we define the demand function, $z_j = D_j(\gamma_j, H_G)$, to be news consumers' attention to outlet j, given its accuracy γ_j and total media influence H_G:

$$D_j(\gamma_j, H_G) = \begin{cases} \max\left\{ z_j : \frac{\gamma_j(h_j(z_j)-H_G)}{1-\gamma_j}\frac{1}{z_j} = \frac{\sqrt{p}}{\chi} \right\} & \text{if } \gamma_j \geq \underline{\gamma}_j; \\ 0 & \text{otherwise.} \end{cases}$$

Figure 5 shows the demand functions corresponding to two different values of H_G. This demand function inherits all the intuitive comparative statics from the equilibrium of the sender-receiver game and the attention allocation decision of news consumers.

Lemma 4. *The demand function $D_j(\gamma_j, H_G)$ is discontinuous at $\gamma_j = \underline{\gamma}_j$. When attention to news outlet j is positive, it increases in the accuracy of its news sources ($\partial D_j/\partial\gamma_j > 0$); it decreases in the marginal cost of its editor ($\partial D_j/\partial\phi_j < 0$); it decreases in the marginal cost*

[25] We use the term "demand function" to highlight the analogy with product markets, in which consumers' utility maximization problem gives rise to demand for different goods as functions of all prices. In our model, consumers' attention allocation problem gives rise to different amounts of attention to media outlets, which is used to generate revenue for the media firm.

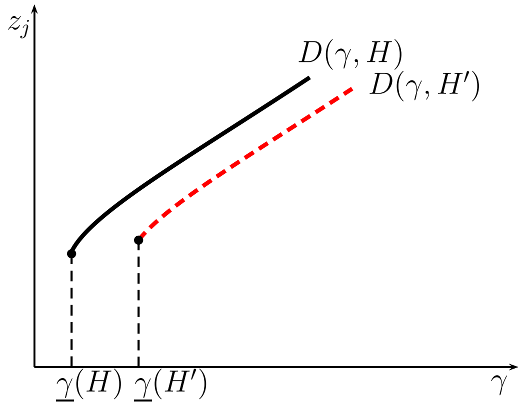


Figure 5. *The demand curve $D_j(\gamma_j, H)$ is discontinuous at $\gamma_j = \underline{\gamma}_j(H)$. An increase in total media informativeness (from H to H') shifts down the demand curve: consumers pay less attention to each news outlet as total media informativeness increases.*

of attention ($\partial D_j/\partial p < 0$); and it decreases when the total influence of the media industry is higher ($\partial D_j/\partial H_G < 0$).

The key result is that attention to news outlet j falls when the total influence H_G of the news industry is higher. It emerges for two reasons. First is attention diversion: a more informative news industry means that news outlet j faces competition from better substitutes when news consumers allocate their attention. Second is free riding: a more informative news industry creates more incentive for editor j to rely on other editors' reports to inform the public, which causes him to write stories with lower α_j that attracts less attention. In this model, the strategic substitutability among media firms is endogenous, and this result is broadly consistent with empirical findings about media substitutability.[26]

Lemma 4 also establishes that news firms can get more attention (and hence more advertising revenue) by investing in higher-quality news sources. This result is consistent with observations on the news industry (Peitz and Reisinger 2015, p. 449).

4.3. Equilibrium Analysis

Given the properties of the demand function $D_j(\gamma_j, H_G)$, the equilibrium levels of investment in this market are defined as follows.

Definition 1. *A monopolistic competitive equilibrium is described by an active media group $G \subseteq \{1, \ldots, J\}$ and a profile $\{\gamma^*, H_G^*\}$ such that:*

[26] For example, Gentzkow (2007) finds that online and print versions of news sources are significant substitutes instead of complements, once consumer heterogeneity is properly controlled for. Wallsten (2015) also finds that increased attention spent on the internet, such as obtaining news, is associated with less attention to television.

1. *(a) For each* $j \in G$, $\gamma_j^* = g_j(H_G^*) > 0$, *where*

$$g_j(H) = \arg\max_{\gamma_j} S_j D_j(\gamma_j, H) - C_j(\gamma_j);$$

   and (b) for each $j \notin G$, $\gamma_j^* = 0$.

2. *Given the accuracy profile* $\{g_j(H_G^*)\}_{j \in G}$, *there exists an equilibrium in the sender-receiver game with endogenous attention allocation in which the active media group is* G *and total influence of the media is* H_G^*, *i.e.,* $H_G^* = \kappa_G(H_G^*)$, *where*

$$\kappa_G(H) = \frac{\sum_{j \in G} \frac{g_j(H)}{1-g_j(H)} \left(1 - \frac{\chi}{\sigma_\theta} \sqrt{\frac{\phi_j}{D_j(g_j(H),H) g_j(H)}}\right)}{1 + \sum_{j \in G} \frac{g_j(H)}{1-g_j(H)}}.$$

Given an equilibrium profile $\{\gamma^*, H_G^*\}$, the corresponding equilibrium attention allocation is $z_j^* = D_j(\gamma_j^*, H_G^*)$ for $j \in G$. Equilibrium reporting strategy α_j^* and equilibrium action weight β_j^* are specified by $\hat{\alpha}_j$ and $\hat{\beta}_j$ of the solution to the sender-receiver game described in Proposition 1, evaluated at $\gamma_j = \gamma_j^*$ and $z_j = z_j^*$. If $j \notin G$, then $z_j^* = \alpha_j^* = \beta_j^* = 0$.

To study the existence and properties of equilibrium, we first provide a result for the optimal investment function $g_j(H)$ and for the aggregator function $\kappa_G(H)$.

Lemma 5. *For each media firm* j, *there exists* $\overline{p}_j$ *such that, when the marginal cost of attention* p *is lower than* $\overline{p}_j$, *optimal investment in accuracy* $g_j(H)$ *decreases in* H, p, *and* ϕ_j. *Further, if* $p \leq \min\{\overline{p}_j : j \in G\}$, *then the aggregator function* $\kappa_G(H)$ *decreases in* H, p, *and* ϕ_j.

Lemma 4 already establishes that $\partial D_j / \partial H < 0$. To show that optimal investment decreases in H, we need to establish that H lowers the marginal benefit from investment, i.e., $\partial^2 D_j / \partial H \partial \gamma_j < 0$. Lemma 5 shows that this is indeed true when the marginal cost of attention is sufficiently low.

An increase in H lowers the aggregated total influence $\kappa_G(H)$—both through a direct channel where lower accuracy, $g_j(H)$, reduces influence, and through an indirect channel where lower attention, $D_j(g_j(H), H)$, reduces influence. In other words, given a higher aggregate influence, owners of all the firms in the group G reduce their investment in accuracy and therefore, attract less attention, resulting in a lower total influence of news industry. The property that the aggregator function $\kappa_G(H)$ decreases in H gives rise to the result that the monopolistic competition equilibrium is unique for given G, provided that an equilibrium can be supported with active media group G.

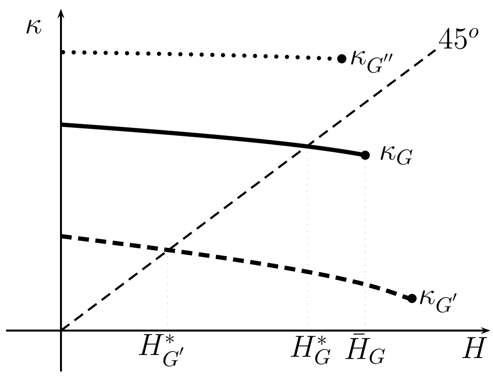


Figure 6. *The function $\kappa_G(H)$ is well-defined on the interval $[0, \overline{H}_G]$, and is downward sloping when the marginal cost of attention is low. The fixed point of $\kappa_G(\cdot)$ is an equilibrium. There is also an equilibrium with active media group G' for any $G' \subset G$, but there may not be an equilibrium with active media group G'' if $G'' \supset G$.*

Proposition 5. *Fix any set of media firms $G \subseteq \{1, \ldots, J\}$, and assume that the cost functions of the active firms are sufficiently convex. There exists $\overline{p}_G$ such that for any $p \leq \overline{p}_G$, there is a monopolistic competitive equilibrium in which the active media group is G and the corresponding profile $\{\gamma^*, H^*_G\}$ is uniquely determined. Further, for any $G' \subset G$, an equilibrium with active media group G' also exists, with $H^*_{G'} < H^*_G$; but there may not be an equilibrium with active media group G'', if $G'' \supset G$.*

In Figure 6, the aggregator function $\kappa_G(\cdot)$ is illustrated with the solid downward-sloping curve. It is continuous when the aggregate H is smaller than some $\overline{H}_G < 1$ and undefined beyond $\overline{H}_G$. To ensure its continuity on $[0, \overline{H}_G]$, we require that the cost functions of the active firms are sufficiently convex (specifically, $C''_j(\gamma_j)/C'_j(\gamma_j) \geq \underline{d}(\gamma_j)$ for some function $\underline{d}(\cdot)$ and for all $j \in G$), so that the owners' profit maximization problems are quasi-concave. When $H > \overline{H}_G$, some firm $j \in G$ becomes inactive and drops out. The firm j chooses not to invest because the accuracy threshold for information production $\underline{\gamma}_j$ increases in H. However, when attention is cheap and abundant, $\overline{H}_G$ is sufficiently close to 1 and therefore there exists a unique equilibrium. Generally, when the cost of attention is sufficiently low, there is an equilibrium with any subset of J firms being active.

To understand the second part of Proposition 5, recall two important features of this framework. First, the discontinuity of the demand curve $D_j(\gamma_j, H)$, arising from the endogenous reporting strategy of editors, dictates that there is a lower bound for news accuracy, only above which news consumers pay a strictly positive amount of attention to outlet j. Second, information acquisition choices of news consumers and producers complement each other. These two mechanisms combined may prevent an equilibrium with a larger active media group (say, $G'' \supset G$) from being supported.

When firms in G'' are making positive investments in accuracy, news consumers can only pay a little attention and assign a small action weight to each news outlet. But this tilts editors' trade-off against informing the public. Lower α_j causes consumers to pay even less attention to news outlets, whose owners, in turn, have less incentive to invest. The firms' accuracy and consumers' attention choices reinforce each other, which produce a downward spiral. For a large enough aggregate influence, a set of firms in group G'' cease operating and no equilibrium can be supported with all firms making investment. The dotted line in Figure 6 illustrates such a situation. Such a result that the media market cannot support an arbitrarily large set of news outlets, is a generalization of the point elaborated in Section 3.4.[27]

Following the same logic, equilibrium exists for any smaller group $G' \subset G$. News consumers can simply ignore news firms in the set $G \setminus G'$, and those firms also quit producing news when they receive no attention at all, which is a consequence of the coordination failure in the sender-receiver game of Section 3.2. The equilibrium with the smaller active media group is less informative ($H^*_{G'} < H^*_G$) than the one with a larger group, as shown in Figure 6.[28]

4.4. New Entry and Media Competition

As we have shown in the previous section and in Section 3.4, the news media market may reach a saturation point where any lager set of media firms cannot be supported in equilibrium. In this section, we analyze the effect of entry on the existing firms and market aggregate, when there is still room for new entries into the market.

We begin with the question: Is increased competition (specifically, a larger number of media firms) beneficial to news consumers? The answer to this question is by no means settled in the theoretical literature. First, in the media bias literature, Gentzkow and Shapiro (2006) show that competition can strengthen the incentive of building a reputation and therefore discipline media bias effectively. In contrast, Mullainathan and Shleifer (2005) show that stronger competition may exacerbate bias because consumers consumers prefer hearing news that is likely to confirm their priors. Second, closer to our work, in the news provision literature, Perego and Yuksel (2018) characterize a novel mechanism with news topics selection and show that the competition

[27] Our result in Proposition 4 and 5 that there is an upper limit for the number of firms that can be supported in equilibrium, resembles that in Sutton (1991), in which the number of firms reaches a limit even when market size is arbitrarily large. However, the underlying mechanisms are different. In Sutton (1991), such a result arises because the sunk cost (advertising expenditure) prior to entry is endogenous. In contrast, an upper bound for the number of firms exists in our model because of the endogenous quality of news and the complementarity of news production and consumption. Further, we assume that there is no sunk cost for news production.

[28] It is also possible that one equilibrium active media group is neither a subset nor a superset of another equilibrium active media group.

can be indeed welfare decreasing.

In this paper, we focus on an alternative dimension, i.e., investment in news gathering and news quality (as opposed to bias and topics). One common concern among media practitioners is actually how competition affects the incentive for firms to invest in news production. Commentators typically worry that there exists a vicious cycle of intense media competition: new entrants would compress the demand and therefore the production budget for each news producer, "which compromises the quality ..., further reducing the audience and alienat[ing] the advertisers" (Keen 2007, p. 33). Becker et al. (2009, p. 376) also wrote, "As competition among news providers becomes extreme, the organization's financial commitment to quality news is expected to decline, as will the market performance of the organization. The quality and diversity of news content should fall, as will journalists' wages, the size and quality of the editorial staff, and the numbers of bureaus and subscriptions to wire services and other external sources of content." In the following, we start with some stylized facts regarding the consequences of competition in the current media environment and then demonstrate how they can be reconciled in this framework.

The rise of news outlets in the digital sector has left many existing and established organizations struggling, which is probably one of the most striking phenomena in the recent development of the news market. For example, the newspaper industry in the United States lost 70 percent of advertising revenues since 2000 (Chandra and Kaiser 2015). Due to the financial strains that beset news organizations, the total number of reporters, editors and other journalists fell from a peak of 56,400 in 2000 to 32,900 in 2014, a decline of more than 40 percent.[29] The new entrants also had a hard time attracting resources for their operation, and some media commentators (e.g., Grubisich 2005; and Carpenter 2009) have expressed strong reservations about the financial viability of these operations as well as their commitment to news quality and journalism standards.

Such a sharp decrease in input to news production contributed to the general trend in the public's perception about the quality of individual news outlets: a steady decline in news accuracy. In the annual Media Consumption Survey conducted by Pew Research Center, respondents are asked whether they believe that the news organizations "get the facts straight" and are "willing to admit their mistakes." There is a clear trend that the fraction of respondents who offer a positive answer has been dwindling in the last three decades; see Figure 7.

Despite the substantial decline in quality of individual news outlets over time, Americans are spending more time on news (Kohut et al. 2010). The 2010 wave of

[29] Data obtained from the American Society of News Editors, Newsroom Employment Census 2015.

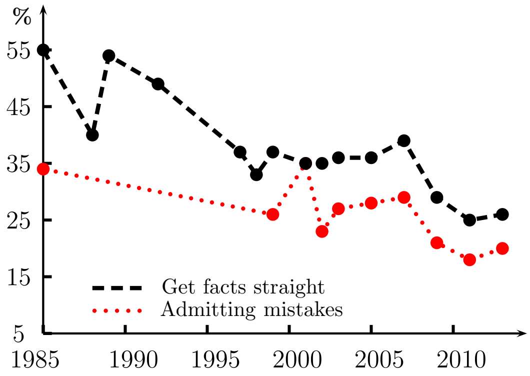


Figure 7. *The fraction of respondents who believe that news organizations are providing accurate news reports in general steadily declines. Source: Media Consumption Survey, 2013.*

Pew Media Consumption Survey shows that, for an average American, the total time of getting news in a given day has risen from 57 minutes in 2000 to 67 minutes in 2006, and to 70 minutes in 2010. That upward trend is largely driven by news consumption online, which offsets the mild decline in time spent with news offline. This measure does not take into account time spent getting news on cell phones or other digital devices; otherwise, the increase may be even sharper. As a result of longer time with news, Americans are better informed. The 2014 wave of Pew Media Survey reveals that 62 percent of respondents claim that they are better informed about local news compared with five years ago, 75 percent claim so about national news and 74 percent about international news.

It seems to be interesting that Americans spend more time on news and become better informed, while they believe that the quality of news has fallen. Increased competition may not necessarily lead to either of these changes. Moreover, the proliferation of news outlets by itself need not cause news consumers to spend more time with them. Nevertheless, our model of competition for attention can reconcile both trends we describe. The following proposition summarizes the key results.

Proposition 6. *Compare the most informative equilibrium with no entry to that when a new firm e is introduced. In the latter equilibrium: (a) total media influence H^* is higher; (b) total attention spent on news, $z_e^* + \sum_j z_j^*$, is higher; and (c) incumbent firm j that remains active invests less in accuracy γ_j^*, produces news reports with lower α_j^*, and receives less attention z_j^* and action weight β_j^* from news consumers.*

For ease of exposition, consider that all J firms are active in equilibrium with total influence H^* prior to the new entry, and all $J+1$ firms are active after the firm e enters the market (Proposition 6 does not rely on this assumption). Upon new entry, we have $\kappa^e(H^*) > \kappa(H^*) = H^*$, where $\kappa^e(\cdot)$ and $\kappa(\cdot)$ are the aggregator functions with and without entry, respectively. It follows that the fixed point of $\kappa^e(\cdot)$ is higher than H^*.

In our model, total attention paid to news media is proportional to total influence H^*. Therefore, when there is new entry into the industry, news consumers also pay more attention to all the news media combined. This result is consistent with the observed trend that Americans are spending more time on news.

In Lemma 5, we have already shown that, if the marginal cost of attention is low enough, a higher total informativeness H^* in the media industry discourages investment in news accuracy—thanks to the direct effect of a diversion of news consumers' attention to other news media, and to the indirect strategic effect due to more severe free-riding among news editors. The decrease in γ_j^* in our model corresponds to a reduction in the precision of the facts obtained in the news gathering process. Our model prediction is consistent with the trend shown in Figure 7.

4.5. Discussion

Extensive vs. intensive margins. The previous section shows that the proliferation of media outlet induces a trade-off between quality at the individual firm level and total quantity of news at the industry level. But the terms of this trade-off are unambiguous: overall media informativeness H^* necessarily increases. In other words, a more competitive media industry also produces a more informed citizenry. We stress that this conclusion follows from a setup in which news consumers endogenously allocate attention to multiple firms, and would not obtain if we do not study the attention allocation problem. To illustrate this point further, we outline below a comparable case in which firms compete at the extensive margin.

Consider a discrete choice model in which news consumer i has a fixed amount of attention (normalized to 1) and chooses to consume news from only one media outlet. The value from choosing news outlet j is $v_{ij} = v_j + \omega_{ij}$, where $v_j = -\sigma_\theta^2/(1+\tau_j) - p/\alpha_j^2$, and ω_{ij} is an idiosyncratic preference for news outlet j that follows the extreme value distribution. The total amount of attention (from all news consumers) given to news outlet j is $e^{v_j}/\sum_k e^{v_k}$. We can show that the marginal benefit from investing in news accuracy in this logit demand model decreases when there are more firms in the industry. Thus, each news outlet becomes less informative. Because each news consumer only gets news from one firm, his action will deviate further from the state. As a result, $\text{Cov}[Q,\theta]$ also falls. This result is opposite to that of our model, in which $H^* = \text{Cov}[Q,\theta]$ rises when there is new entry into the industry. In other words, media competition improves the aggregate informativeness of the industry only when firms compete at the intensive margin (i.e., competing for attention); and news consumers as a whole may be worse off when firms compete only at the extensive margin (i.e., competing for audience).

Correlated information production. In our benchmark model, we assume that the facts obtained by media firms are conditionally independent. However, journalists from competing news outlets may share common news sources—they may interview similar sets of witnesses or consult overlapping groups of experts. Thus, the news gathering process is likely to produce source materials that are correlated across media firms even conditional on the true state. We can embed this concern into another comparable extension of the benchmark model and examine how correlation in news production may affect the decisions of media owners, editors and news consumers, as well as the total influence of the news industry.

Assume that the news source for firm j is a signal $x_j = \theta + \epsilon_j + \zeta_j$, where $\epsilon_j \sim N(0, \sigma^2_{\epsilon j})$ is independent across firms as before. The noise $\zeta_j \sim N(0, \sigma^2_\zeta)$, however, is correlated across firms. Specifically, let $\mathrm{Cov}[\zeta_j, \zeta_k] = \rho\sigma^2_\zeta$, where $\rho \in [0,1]$ indicates the degree of correlation and let $R = 1 + \rho\sigma^2_\zeta/\sigma^2_\theta$. Define $\gamma_j \equiv \sigma^2_\theta/(\sigma^2_\theta + \sigma^2_{\epsilon j} + \sigma^2_\zeta)$ as the accuracy of x_j. We assume that the owner of firm j's investment only affects $\sigma^2_{\epsilon j}$, so that the accuracy of another firm is not affected by firm j's decision.

In such a setup, when the degree of correlation is high (i.e., when R is large), editor j expects other editors to write stories that are similar to his own news source x_j. His incentive to write a story that closely reflects x_j to inform the public is diminished. Therefore, the reporting strategy α_j chosen by editor j is decreasing in R, that is, the free-riding problem is exacerbated by a higher correlation in information production. Furthermore, when the news stories are conditionally correlated, these stories become jointly less informative about the state. Not only is the posterior variance of the news consumers larger, the marginal benefit from paying attention to reduce this variance is also lower. In other words, a higher value of R tends to lower news consumers' attention and the action weights they put on the news stories. That in turn also pushes down the marginal revenue for firm owners, undermining their incentive to invest. As a result, the total influence of the media also declines when the correlation is higher. These mechanisms combined contribute to the following result.

Proposition 7. *Suppose all firms are identical (i.e., ϕ_j, S_j, and $C_j(\cdot)$ do not vary with j). In a symmetric equilibrium, a higher correlation in news production reduces the total informativeness of the media industry. The influence of and the attention given to each media firm also fall.*

Proposition 7 rationalizes why the production of original content and independent journalistic investigation are encouraged in the news market and why they are beneficial to news consumers.

While the result that correlated signals may reduce the overall informativeness of the media may seem intuitive, such a feature is not particularly prominent in existing models of media economics or sender-receiver games. Economic models of the media industry often emphasize the “cross-checking” effect: a strategic sender is less likely to hide or distort his facts if he thinks the receiver can get similar facts from other sources (Gentzkow and Shapiro 2006). In such models, the cross-checking effect tends to be stronger when signals are more correlated.[30] In fact, when signals are perfectly correlated, multiple-sender models can often support perfect information revelation (Krishna and Morgan 2001). In contrast, our approach highlights the endogenous quality of signals and shows that the high correlation among signals undermines the incentives of senders to invest.

5. Conclusion

Probably, in the internet era, few industries have been transformed and restructured by new entrants more profoundly than the news industry. New media such as *The Huffington Post* and *Politico* have achieved record-breaking growth in readership, attention, as well as revenue in the last few years. On the one hand, they grabbed the market shares of existing mainstream news media and left them struggling for survival; on the other hand, they expanded the base of news consumers and offered fresh perspectives to them so that people are spending more time on news than ever before. In this paper, we propose a framework to analyze such a media environment and study various issues related to media competition.

We intentionally focus on news provision (gathering and presentation) and the impact of multi-homing news consumption and deliberately restrain from addressing some important issues in media economics, such as media bias and the interactions between media and politics. Our model, however, lays the groundwork for investigating those issues in the current landscape of media market.

Our work also makes theoretical contributions to the literature of endogenous information acquisition and sender-receiver games. We extend the former by allowing both accuracy and clarity of underlying signals to be chosen in response to information acquisition of receivers. We enrich the latter by developing a workable approach to characterize sender-receiver games in which a large number of heterogeneous senders who possess non-identical private information attempt to influence the same set of decision makers. This model can be useful in other information markets that feature monopolistic competition.

[30] In a different context, Chan et al. (2007) and Damiano et al. (2008) show that schools (or credit-rating agencies) are less likely to distort their grades (or credit ratings) when their private information is more correlated.

References

Acemoglu, D. and M. K. Takahashi (2008). "Aggregate Comparative Statics." *Games and Economic Behavior 81*: 27–49.

Ambrus, A. and S. Takahashi (2008). "Multi-Sender Cheap Talk with Restricted State Spaces." *Theoretical Economics 3*(1): 1–27.

Anderson, C. (2009). *Free: The Future of a Radical Price.* New York: Random House.

Anderson, S. P. and J. McLaren (2012). "Media Mergers and Media Bias with Rational Consumers." *Journal of the European Economic Association 10*(4): 831–859.

Battaglini, M. (2002). "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica 70*(4): 1379–1401.

Becker, L. B., C. A. Hollifield, A. Jacobsson, E.-M. Jacobsson, and T. Vlad (2009). "Is More Always Better? Examining the Adverse Effects of Competition on Media Performance." *Journalism Studies 10*(3): 368–385.

Bernhardt, D., S. Krasa, and M. Polborn (2008). "Political Polarization and the Electoral Effects of Media Bias." *Journal of Public Economics 92*(5-6): 1092–1104.

Bhattacharya, S. and A. Mukherjee (2013). "Strategic Information Revelation when Experts Compete to Influence." *RAND Journal of Economics 44*(3): 522–544.

Carpenter, S. (2009). "News Quality Differences in Online Newspaper and Citizen Journalism Sites." Chap. 6 in J. Rosenberry and B. St. John, III, eds., *Public Journalism 2.0*. London: Routledge.

Chan, W., H. Li, and W. Suen (2007). "A Signaling Theory of Grade Inflation." *International Economic Review 48*(3): 1065–1090.

Chan, J. and D. F. Stone (2013). "Media Proliferation and Partisan Selective Exposure." *Public Choice 156*(3–4): 467–490.

Chan, J. and W. Suen (2008). "A Spatial Theory of News Consumption and Electoral Competition." *Review of Economic Studies 75*(3): 699–728.

Chan, J. and W. Suen (2009). "Media as Watchdogs: The Role of News Media in Electoral Competition." *European Economic Review 53*(7): 799–814.

Chandra, A. and U. Kaiser (2015). "Newspapers and Magazines." Chap. 9 in S. J. Anderson, J. Waldfogel, and D. Strömberg, eds., *Handbook of Media Economics*, Vol. 1A. Amsterdam: North-Holland.

Chen, H., Y. Luo, and G. Pei (2015). "Attention Misallocation, Social Welfare and Policy Implications." *Journal of Economic Dynamics and Control* 59: 37–57.

Damiano, E., H. Li, and W. Suen (2008). "Credible Ratings." *Theoretical Economics 3*(3): 325–365.

Davenport, T. H. and J. C. Beck (2013). *The Attention Economy: Understanding the New Currency of Business*. Boston: Harvard Business Review Press.

Dewan, T. and D. P. Myatt (2008). "The Qualities of Leadership: Direction, Communication, and Obfuscation." *American Political Science Review 102*(3): 351–368.

Dewatripont, M. and J. Tirole (2005). "Modes of Communication." *Journal of Political Economy 113*(6): 1217–1238.

Dixit, A. K. and J. E. Stiglitz (1977). "Monopolistic Competition and Optimal Product Diversity." *American Economic Review 67*(3): 297–308.

Dubey P., Haimanko, O. and A. Zapechelnyuk (2006). "Strategic Complements and Substitutes, and Potential Games." *Games and Economic Behavior 54*(1): 77–94.

Galperti, S. and I. Trevino (2018). "Shared Knowledge and Competition for Attention in Information Markets." Working paper, University of California San Diego.

Gentzkow, M. (2007). "Valuing New Goods in a Model with Complementarity: Online Newspapers." *American Economic Review 97*(3): 713–744.

Gentzkow, M. and E. Kamenica (2015). "Disclosure of Endogenous Information." Working paper, University of Chicago.

Gentzkow, M. and J. M. Shapiro (2006). "Media Bias and Reputation." *Journal of Political Economy 114*(2): 280–316.

Gentzkow, M. and J. M. Shapiro (2011). "Ideological Segregation Online and Offline." *Quarterly Journal of Economics 126*(4): 1799–1839.

Gentzkow, M., J. M. Shapiro, and D. F. Stone (2016). "Media Bias in the Marketplace: Theory." Chap. 14 in S. J. Anderson, J. Waldfogel, and D. Strömberg, eds., *Handbook of Media Economics*, Vol. 1B, Amsterdam: North-Holland.

Grubisich, T. (2005). "Grassroots Journalism: Actual Content vs. Shining Ideal." *USC Annenberg Online Journalism Review*, October 6.

Hellwig, C., S. Kohls, and L. Veldkamp (2012). "Information Choice Technologies." *American Economic Review 102*(3): 35–40.

Hellwig, C. and L. Veldkamp (2009). "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies 76*(1): 223–251.

Kartik, N., F. X. Lee, and W. Suen (2017). "Investment in Concealable Information by Biased Experts." *RAND Journal of Economics 48*(1): 24–43.

Keen, A. (2007). *The Cult of the Amateur.* New York: Currency.

Kohut, A., C. Doherty, M. Dimock, and S. Keeter (2010). "Americans Spending More Time Following the News." Pew Research Center report.

Krishna, V. and J. Morgan (2001). "A Model of Expertise." *Quarterly Journal of Economics 116*(2): 747–775.

Lee, A. M. (1947). *The Daily Newspaper in America.* New York: Macmillan.

Matsa, D. (2010). "Competition and Product Quality in the Supermarket Industry." *Quarterly Journal of Economics 126*(3): 1539–1591.

Mondria, J. and C. Quintana-Domeque (2013). "Financial Contagion and Attention Allocation." *Economic Journal 123*(568): 429–454.

Mullainathan, S. and A. Shleifer (2005). "The Market for News." *American Economic Review 95*(4): 1031–1053.

Myatt, D. P. and C. Wallace (2012). "Endogenous Information Acquisition in Coordination Games." *Review of Economic Studies 79*(1): 340–374.

Napoli, P. M. (2013). *Audience Economics: Media Institutions and the Audience Marketplace*. New York: Columbia University Press.

Nimark, K. and S. Pitschner (2018). "News Media and Delegated Information Choice." Working Paper, Cornell University.

Novshek, W. (1985). "On the Existence of Cournot Equilibrium." *Review of Economic Studies 52*(1): 86–98.

Peitz, M. and M. Reisinger (2015). "The Economics of Internet Media." Chap. 10 in S. J. Anderson, J. Waldfogel, and D. Strömberg, eds., *Handbook of Media Economics*, Vol. 1A, Amsterdam: North-Holland.

Perego, J. and S. Yuksel (2018). "Media Competition and Social Disagreement." Working paper, University of California Santa Barbara.

Prat, A. and D. Strömberg (2013). "The Political Economy of Mass Media." Chap. 3 in D. Acemoglu, M. Arellano and E. Dekel, eds., *Advances in Economics and Econometrics: Tenth World Congress*, Vol. 2. Cambridge: Cambridge University Press.

Schmalensee, R. (1979). "Market Structure, Durability, and Quality: A Selective Survey." *Economic Inquiry 17*(2): 177–196.

Sims, C. A. (2003). "Implications of Rational Inattention." *Journal of Monetary Economics 50*(3): 665–690.

Sobbrio, F. (2014). "Citizen-Editors' Endogenous Information Acquisition and News Accuracy." *Journal of Public Economics 113*: 43–53.

Spence, A. M. (1975). "Monopoly, Quality, and Regulation." *Bell Journal of Economics 6*(2): 417–429.

Sun, M. and F. Zhu (2013). "Ad Revenue and Content Commercialization: Evidence from Blogs." *Management Science 59*(10): 2314–2331.

Sutton, J. (1991). *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge: MIT Press.

Webster, J. G. (2014). *The Marketplace of Attention: How Audiences Take Shape in a Digital Age*. Cambridge: MIT Press.

Wallsten, S. (2015). "What Are We Not Doing When We're Online?" Chap. 2 in A. Goldfarb, S.M. Greenstein, and C.E. Tucker, eds., *Economic Analysis of the Digital Economy*. Chicago: University of Chicago Press.

White, A. and K. Jain (2018). The Attention Economy of Online Advertisement. *Available at SSRN 1694849*.

Appendix

Proof of Lemma 1. For $j = 1, \ldots, J$, define $t_j \equiv \gamma_j \beta_j^2 / (\beta_j^2 + \phi_j)$. Multiply equation (8) by β_j and subtract $t_j \alpha_j \beta_j$ from both sides of the equation to get

$$(1 - t_j)\alpha_j \beta_j = t_j (1 - \sum_k \alpha_k \beta_k).$$

Divide both sides by $1 - t_j$ and sum over all j, we obtain:

$$\sum_k \alpha_k \beta_k = \frac{\sum_k t_k / (1 - t_k)}{1 + \sum_k t_k / (1 - t_k)}.$$

Thus,

$$\alpha_j \beta_j = \frac{t_j / (1 - t_j)}{1 + \sum_k t_k / (1 - t_k)},$$

which is equivalent to equation (9). The comparative statics results are obtained by taking derivatives of (9) with respect to the relevant variables. ∎

Proof of Lemma 2. By the Gaussian updating formula, the posterior expectation of θ is

$$\mathrm{E}[\theta \mid \hat{y}_{1i}, \ldots, \hat{y}_{Ji}] = \frac{1}{1 + \sum_k \tau_k} \mu + \sum_j \frac{\tau_j}{1 + \sum_k \tau_k} \frac{\hat{y}_{ji} - \alpha_{j0}}{\alpha_j}.$$

Comparing coefficients with the linear action strategy $q_i = \beta_0 + \sum_j \beta_j \hat{y}_{ji}$ gives the formula (11) for β_j. The comparative statics results are obtained by taking derivatives of (11) with respect to the relevant variables. ∎

Proof of Proposition 1. For any j, the equilibrium values of α_j and β_j must satisfy equations (9) and (11). Comparing these two equations gives

$$\tau_j = \frac{\gamma_j \beta_j^2}{(1 - \gamma_j)\beta_j^2 + \phi_j} = \frac{t_j}{1 - t_j},$$

where we adopt the definition $t_j \equiv \gamma_j \beta_j^2 / (\beta_j^2 + \phi_j)$. Use (10) for the relative precision τ_j, this equation reduces to:

$$\frac{\sigma_\theta^2 z_j}{\chi^2} \alpha_j^2 = \frac{\gamma_j t_j}{\gamma_j - t_j}.$$

Multiply both sides by β_j^2 and use equation (9), we obtain:

$$\frac{\sigma_\theta^2 z_j}{\chi^2} \left(\frac{t_j / (1 - t_j)}{1 + \sum_k t_k / (1 - t_k)} \right)^2 = \beta_j^2 \frac{\gamma_j t_j}{\gamma_j - t_j} = \frac{\phi_j \gamma_j t_j^2}{(\gamma_j - t_j)^2},$$

where the second equality follows because the definition of t_j implies $\beta_j^2 = \phi_j t_j/(\gamma_j - t_j)$. Define

$$H \equiv \frac{\sum_k t_k/(1-t_k)}{1+\sum_k t_k/(1-t_k)}.$$

Then the earlier equation reduces to

$$\frac{t_j^2}{(1-t_j)^2}(1-H)^2 = \frac{\chi^2}{\sigma_\theta^2}\frac{\phi_j}{\gamma_j z_j}\frac{\gamma_j^2 t_j^2}{(\gamma_j - t_j)^2} = (1-h_j)^2\frac{\gamma_j^2 t_j^2}{(\gamma_j - t_j)^2},$$

where the second equality is implied by the definition of h_j in equation (12). Obviously, $t_j = 0$ is a solution to the above equation, which would entail $\hat{\alpha}_j = \hat{\beta}_j = 0$. Since $\gamma_j < 1$, the equation admits a non-zero solution if and only if $h_j > H$. For a non-trivial equilibrium, suppose there is a non-empty subset G of media outlets such that $t_j > 0$ if $j \in G$ and $t_j = 0$ otherwise. The non-zero solution to the equation is:

$$t_j = \frac{\gamma_j\left(h_j - H\right)}{(1-H) - \gamma_j\left(1-h_j\right)}.$$

Use such value of t_j for $j \in G$ and use $t_j = 0$ for $j \notin G$ to substitute into the definition of H, we can solve for H to obtain $H = H_G$, as is given by equation (13).

For $j \in G$, we can recover the equilibrium value of $\hat{\beta}_j$ from the non-zero solution t_j using the definition of t_j. This yields:

$$\hat{\beta}_j^2 = \frac{(h_j - H)\phi_j}{(1-\gamma_j)(1-h_j)}.$$

Substitute this value of β_j into equation (9) to get:

$$\hat{\alpha}_j^2 = \frac{\gamma_j^2(h_j - H)(1-h_j)}{(1-\gamma_j)\phi_j}.$$

Multiplying these two equations gives equation (14). ■

Proof of Lemma 3. From equation (13), we can write

$$H_G = \frac{1+\sum_{j\in G'}\frac{\gamma_j}{1-\gamma_j}}{1+\sum_{j\in G}\frac{\gamma_j}{1-\gamma_j}}H_{G'} + \frac{\sum_{j\in G\setminus G'}\frac{\gamma_j}{1-\gamma_j}h_j}{1+\sum_{j\in G}\frac{\gamma_j}{1-\gamma_j}}.$$

Therefore,

$$H_G - H_{G'} = \frac{\sum_{j\in G\setminus G'} \frac{\gamma_j}{1-\gamma_j}(h_j - H_{G'})}{1 + \sum_{j\in G} \frac{\gamma_j}{1-\gamma_j}} > \frac{\sum_{j\in G\setminus G'} \frac{\gamma_j}{1-\gamma_j}(H_G - H_{G'})}{1 + \sum_{j\in G} \frac{\gamma_j}{1-\gamma_j}},$$

where the inequality follows because $h_j > H_G$ for all $j \in G$. If $H_G - H_{G'}$ is non-positive, the above inequality is a contradiction. Thus, we must have $H_{G'} < H_G$. ■

Proof of Proposition 2. We start by defining an aggregating function for G:

$$\kappa_G(H) \equiv \frac{\sum_{j\in G} \frac{\gamma_j}{1-\gamma_j}(1 - \frac{\chi}{\sigma_\theta}\sqrt{\frac{\phi_j}{z_j(H)\gamma_j}})}{1 + \sum_{j\in G} \frac{\gamma_j}{1-\gamma_j}},$$

where $z_j(H)$ is the larger solution to equation (16), given any $H < 1$. Given any $H < 1$, there exists a $\tilde{p}_j$ for each outlet j such that the maximum of the left-hand side of equation (16) is $\sqrt{\tilde{p}_j}/\chi$, because it is single-peaked. Let $\tilde{p}_G \equiv \min\{\hat{p}_j : j \in G\}$. Further, the value of the left-hand side of equation (16) approaches zero from above as z_j becomes sufficiently large. Therefore, for any $p \leq \tilde{p}_G$, $z_j(H)$ is well-defined and positive for all j. This also means that $\kappa_G(H)$ is well-defined.

Suppose that there are n firms in the set G. Let $\tilde{\gamma} \equiv \max_j\{\gamma_j\}$, and let $\tilde{H} \in (0,1)$ satisfy

$$n\frac{\tilde{\gamma}}{1-\tilde{\gamma}}(1-\tilde{H}) - \tilde{H} = 0.$$

We have

$$\kappa_G(\tilde{H}) - \tilde{H} < \frac{n\frac{\tilde{\gamma}}{1-\tilde{\gamma}}}{1 + n\frac{\tilde{\gamma}}{1-\tilde{\gamma}}} - \tilde{H} = 0.$$

The first inequality follows because $\gamma_j \leq \tilde{\gamma}$ and $h_j < 1$; and the second equality follows from the definition of $\tilde{H}$. Moreover, for $p \leq \tilde{p}_G$, $\kappa_G(0)$ is well-defined and positive. It is obvious that $\kappa_G(H)$ decreases in H, since $z_j(H)$ decreases in H. Therefore, the value of the fixed point H_G^* that satisfies $H_G^* = \kappa(H_G^*)$ is uniquely determined, from which we also obtain $z_j^* = z_j(H_G^*)$ by equation (16). ■

Proof of Proposition 3. From the proof of Proposition 1, the equilibrium value of α for an active firm must satisfy:

$$\alpha^2 = \frac{\gamma^2(h - H_G)(1-h)}{(1-\gamma)\phi}.$$

The key equation (16) also requires:

$$\frac{\gamma^2}{\phi(1-\gamma)}(h - H_G)(1-h)^2 = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

These two equations imply that

$$1 - h = \frac{\sqrt{p}\chi}{\alpha^2\sigma_\theta^2}.$$

Recall that the left-hand-side of the key equation above is increasing then decreasing in h, and that the equilibrium h is the larger root to the key equation. Therefore, in any equilibrium,

$$h \geq \arg\max_h \,(h - H_G)(1-h)^2 = \frac{1+2H_G}{3}.$$

If equilibrium reporting strategy is $\overline{\alpha}$ and equilibrium total influence is H_G^*, we must have

$$1 - \frac{1+2H_G^*}{3} \geq \frac{\sqrt{p}\chi}{\overline{\alpha}^2\sigma_\theta^2},$$

which establishes the upper bound stated in the proposition. ■

Proof of Proposition 4. Let n be the number of firms in an active media group G. From the definition of H_G, we have

$$h - H_G = \frac{h}{1 + n\frac{\gamma}{1-\gamma}}.$$

Substitute this into the key equation to obtain:

$$\frac{\gamma^2}{\phi(1-\gamma)}\frac{h(1-h)^2}{1+n\frac{\gamma}{1-\gamma}} = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

The maximum value of $h(1-h)^2$ is $4/27$. So an upper bound on the number of active firms that can be supported in any equilibrium is the largest integer $\overline{n}$ such that

$$\frac{4}{27}\frac{\gamma^2}{\phi(1-\gamma)}\frac{1}{1+\overline{n}\frac{\gamma}{1-\gamma}} \geq \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

It is obvious that such $\overline{n}$ increases in γ and decreases in ϕ. ■

Proof of Lemma 4. We rewrite the key equation (16) using the definition of h_j from equation (12):

$$\frac{\gamma_j^2}{\phi_j(1-\gamma_j)}\left(h_j - H_G\right)\left(1-h_j\right)^2 = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

The left-hand-side of this equation attains a maximum at $h_j = (1 + 2H_G)/3$. We define $\underline{\gamma}_j$ to be the value of γ_j for which equation (16) holds at such h_j, i.e.,

$$\frac{4}{27}\frac{\underline{\gamma}_j^2}{\phi_j(1-\underline{\gamma}_j)}(1-H_G)^3 = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

Therefore, at $\gamma_j = \underline{\gamma}_j$, $D(\cdot, H_G)$ jumps up from 0 to $z^* > 0$, where z^* satisfies $h_j(z^*) = (1 + 2H_G)/3$.

When a solution to the key equation (16) exists, the left-hand-side is locally decreasing in z_j at the larger root. Because the left-hand-side increases in γ_j, decreases in ϕ_j, and decreases in H_G, the comparative statics results follow from the implicit function theorem. Similarly, because the right-hand-side of equation (16) increases in p, we have $\partial D_j/\partial p < 0$. ■

Proof of Lemma 5. Claim 1 in online Appendix A establishes that $\partial^2 D_j/\partial H \partial \gamma_j < 0$ when p is sufficiently small. Therefore the function $\Pi_j = S_j D_j(\gamma_j, H) - C_j(\gamma_j)$ is submodular in γ_j and H. This immediately implies that $g_j(H)$ decreases in H. Similarly, $\partial^2 D_j/\partial p \partial \gamma_j < 0$ and $\partial^2 D_j/\partial \phi_j \partial \gamma_j < 0$ when p is sufficiently small, which imply that $g_j(H)$ decreases in p and ϕ_j.

From equations (13) and (12), we see that H_G increases in γ_j and z_j. The aggregator function $\kappa_G(H)$ is simply H_G evaluated at $\gamma_j = g_j(H)$ and $z_j = D_j(g_j(H), H)$. Since $g_j(H)$ decreases in H and $D_j(g_j(H), H)$ decreases in H when p satisfies the condition stated in the lemma, we have $\kappa_G(H)$ decreases in H. Similar reasoning shows that $\kappa_G(H)$ decreases in p and ϕ_j. ■

Proof of Proposition 5. Suppose there are n firms in the set G. Define $\hat{H} \in (0, 1)$ such that

$$n\frac{\overline{\gamma}}{1-\overline{\gamma}}(1-\hat{H}) - \hat{H} = 0.$$

Provided that $\kappa_G(\hat{H})$ is well-defined, we have

$$\kappa_G(\hat{H}) - \hat{H} < \frac{n\frac{\overline{\gamma}}{1-\overline{\gamma}}}{1 + n\frac{\overline{\gamma}}{1-\overline{\gamma}}} - \hat{H} = 0.$$

The first inequality follows because $\gamma_j \leq \overline{\gamma}$ and $h_j < 1$; and the second equality follows from the definition of $\hat{H}$.

Let

$$\Pi_j^*(H, p) = \max_{\gamma_j \in [0, \overline{\gamma}]} S_j D_j(\gamma_j, H; p) - C_j(\gamma_j),$$

where we include the dependence on p explicitly into the demand function $D_j(\cdot)$. Then $\kappa_G(\hat{H})$ is well-defined if and only if $\Pi_i^*(\hat{H}, p) > 0$ for all $j \in G$. Since $\Pi_j^*(\hat{H}, p)$ is continuous and weakly decreases in p, it goes to infinity as p goes to 0, and is equal to 0 when p is sufficiently high. Therefore, there is a $\hat{p}_j$ such that $\Pi_j^*(\hat{H}, p) > 0$ for $p < \hat{p}_j$. Let $\hat{p}_G = \min\{\hat{p}_j : j \in G\}$. Then, for any $p < \hat{p}_G$, $\kappa_G(\hat{H})$ is well-defined with $\kappa_G(\hat{H}) < \hat{H}$.

Next, because $D_j(\gamma_j, H; p)$ is decreasing in H by Lemma 4, $\Pi_j^*(H, p) > 0$ implies $\Pi_j^*(0, p) > 0$. Thus, for $p < \hat{p}_G$, $\kappa_G(0)$ is well-defined, and it satisfies $\kappa_G(0) > 0$.

Finally, when $\partial \Pi_j / \partial \gamma_j = 0$, we have

$$\begin{aligned}
\frac{\partial^2 \Pi_j}{\partial \gamma_j^2} &= S_j \frac{\partial D_j}{\partial \gamma_j} \frac{\partial^2 D_j / \partial \gamma_j^2}{\partial D_j / \partial \gamma_j} - C_j'(\gamma_j) \frac{C_j''(\gamma_j)}{C_j'(\gamma_j)} \\
&\leq S_j \frac{\partial D_j}{\partial \gamma_j} \left(\underline{d}(\gamma_j) - \frac{C_j''(\gamma_j)}{C_j'(\gamma_j)} \right) < 0,
\end{aligned}$$

where the first inequality follows from Claim 2 in online Appendix A, and the second inequality follows because $C_j(\cdot)$ is sufficiently convex. This establishes that $\Pi_j(\gamma_j, H)$ is quasi-concave in γ_j. Since $\Pi_j(\gamma_j, H)$ is also continuous in H, its maximizer $g_j(H)$ is continuous on $[0, \hat{H}]$. Thus, $\kappa_G(H)$ is also continuous on $[0, \hat{H}]$. It follows that a fixed point exists such that $\kappa_G(H_G^*) = H_G^*$. Furthermore, if we let $\overline{p}_G = \min\{\hat{p}_G, \overline{p}_j : j \in G\}$, then for any $p < \overline{p}_G$, Lemma 5 implies that $\kappa_G(H)$ is decreasing in H. The value of the fixed point H_G^* is uniquely determined.

For any H, if $\kappa_G(H)$ is well-defined, then $\kappa_{G'}(H)$ is well-defined. Moreover, by Lemma 3, we have $H_G^* = \kappa_G(H_G^*) > \kappa_{G'}(H_G^*)$. Because $\kappa_{G'}(0) > 0$, there exists $H_{G'}^* < H_G^*$ such that $\kappa_{G'}(H_{G'}^*) = H_{G'}^*$. ■

Proof of Proposition 6. Take any equilibrium with no entry and suppose the active media group in that equilibrium is G. After firm e enters, G is still an equilibrium (firm e is simply inactive in this equilibrium). But when there is an equilibrium with active media group $G \cup \{e\}$, Proposition 5 says that $H_{G \cup \{e\}}^* > H_G^*$. Therefore, total influence in the most informative equilibrium must be higher.

Recall that equation (15) implies that, in equilibrium, the attention z_j^* given to each outlet is proportional to its influence $\alpha_j^* \beta_j^*$. Aggregating this equation over all firms in the active media group shows that total attention is proportional to total influence. Thus, a higher H^* means that total attention, i.e., $z_e^* + \sum_j z_j^*$, is also higher.

For the incumbent firms, $\gamma_j^* = g_j(H^*)$. Since $g_j(\cdot)$ is decreasing, γ_j^* falls. The attention given to firm j is $z_j^* = D_j(\gamma_j^*, H^*)$, where $D_j(\cdot)$ is increasing in the first argument and decreasing in the second. Thus, z_j^* falls.

From the proof of Proposition 1, the equilibrium reporting strategy is given by

$$\alpha_j^2 = \frac{\gamma_j^2 (h_j - H)(1 - h_j)}{(1 - \gamma_j)\phi_j}.$$

Furthermore, we can rewrite the key equation (16) using the definition of h_j to get

$$\frac{\gamma_j^2}{\phi_j(1-\gamma_j)} \left(h_j - H\right) \left(1 - h_j\right)^2 = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

Combining these two equations, we obtain:

$$\alpha_j^2 = \frac{1}{1-h_j} \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

We have shown that both z_j^* and γ_j^* fall in equilibrium, which implies that the equilibrium value of h_j decreases. Therefore, equilibrium α_j^* also decreases.

Finally, we can write the key equation (16) as:

$$\frac{h_j - H}{1 - \gamma_j} = \frac{\sqrt{p}}{\chi} \frac{z_j}{\gamma_j}.$$

New entry lowers h_j^* and γ_j^* and raises H^*, so the right-hand-side of the above equation decreases. This implies that z_j^*/γ_j^* decreases. From the proof of Proposition 1, equilibrium action weight is given by

$$\beta_j^2 = \frac{(h_j - H)\phi_j}{(1-\gamma_j)(1-h_j)} = \frac{\sqrt{p}}{\chi} \frac{z_j}{\gamma_j} \frac{\phi_j}{1-h_j},$$

where the second inequality follows from the key equation (16). Since both z_j^*/γ_j^* and h_j^* fall in equilibrium, β_j^* also falls. ■

Proof of Proposition 7. Following the proof of Lemma 1, we can use the best-response of α_j against $\boldsymbol{\alpha}_{-j}$ to solve for the fixed point $\boldsymbol{\alpha}$:

$$\alpha_j = \frac{1}{R\beta_j} \frac{t_j/(1-t_j)}{1 + \sum_k t_k/(1-t_k)},$$

where $t_j \equiv R\gamma_j\beta_j^2/(\beta_j^2 + \phi_j)$. Moreover, news consumers assign action weights using Bayes' rule:

$$\beta_j = \frac{1}{R\alpha_j} \frac{\tau_j}{1 + \sum_k \tau_k}.$$

We can solve for t_j using similar steps as described in the proof of Proposition 1 to obtain:

$$t_j = \frac{\left(1 - \frac{1-h_j}{1-RH}\right) R\gamma_j}{1 - \frac{1-h_j}{1-RH} R\gamma_j},$$

where

$$H \equiv \frac{1}{R}\left(\frac{\sum_j \frac{R\gamma_j}{1-R\gamma_j} h_j}{1 + \sum_j \frac{R\gamma_j}{1-R\gamma_j}}\right).$$

The influence of media outlet j is:

$$\alpha_j \beta_j = \frac{\gamma_j}{1 - R\gamma_j}\left(h_j - RH\right).$$

News consumers allocate attention by maximizing

$$V = -\sigma_\theta^2 \left(1 - \frac{1}{R}\frac{\sum_j \tau_j}{1 + \sum_j \tau_j}\right) - \sum_j \frac{p}{\alpha_j^2} z_j.$$

Combining the first-order conditions with Bayes' rule and the formula for influence $\alpha_j \beta_j$, we derive the counterpart to the key equation (16) in the benchmark model:

$$\frac{\gamma_j (h_j - RH)}{1 - R\gamma_j}\frac{1}{z_j} = \frac{\sqrt{p}}{\chi}.$$

Define $D_j(\gamma_j, H)$ as the larger solution to z_j in this key equation (and let $D_j(\gamma_j, H) = 0$ if it has no solution). The derivative of the left-hand-side of the above with respect to R has the same sign as $\gamma_j h_j - H$. In a symmetric equilibrium,

$$\gamma h - H = \frac{-(J-1)\gamma h}{1 + J\frac{R\gamma}{1-R\gamma}} < 0,$$

(where we have dropped the subscript for media firms). Moreover, the derivative of the left-hand-side of the key equation with respect to z_j is negative at the larger root. It follows from the implicit function theorem that $\partial D_j / \partial R < 0$.

Claim 3 in online Appendix A shows that, if the marginal cost of attention is low enough, then $\partial^2 D_j / \partial R \partial \gamma_j < 0$. This in turn implies that $g_j(H)$ decreases in R.

In a symmetric equilibrium, let

$$\kappa(H) = \frac{1}{R}\left(\frac{J\frac{Rg(H)}{1-Rg(H)}\left(1 - \frac{\chi}{\sigma_\theta}\sqrt{\frac{\phi}{D(g(H),H)g(H)}}\right)}{1 + J\frac{Rg(H)}{1-Rg(H)}}\right).$$

One can verify that

$$\frac{\partial \kappa}{\partial R} = -(J-1)J\left(\frac{\frac{g(H)}{1-Rg(H)}}{1+J\frac{Rg(H)}{1-Rg(H)}}\right)^2\left(1--\frac{\chi}{\sigma_\theta}\sqrt{\frac{\phi}{D(g(H),H)g(H)}}\right) < 0.$$

Furthermore, we have $\partial\kappa/\partial g > 0$ and $\partial\kappa/\partial D > 0$. Therefore,

$$\frac{\mathrm{d}\kappa}{\mathrm{d}R} = \frac{\partial \kappa}{\partial R} + \frac{\partial \kappa}{\partial g}\frac{\partial g}{\partial R} + \frac{\partial \kappa}{\partial D}\left(\frac{\partial D}{\partial g}\frac{\partial g}{\partial R} + \frac{\partial D}{\partial R}\right) < 0.$$

We conclude that the fixed point of $\kappa(\cdot)$ falls when R increases. Since $H^* = J\alpha^*\beta^*$, the influence of each firm also falls. Since attention is proportional to influence in equilibrium, attention to each firm z^* falls. ∎

Online Appendix to
Competition for Attention in the News Media Market

HENG CHEN and WING SUEN

April 23, 2019

(Not intended for publication)

A. Technical Materials

Claim 1. There exists $\overline{p}_j$ such that if $p \leq \overline{p}_j$, then $\partial D_j/\partial\gamma_j$ decreases in p, ϕ_j and H.

Proof. Implicit differentiation of the key equation (16) shows that

$$\begin{aligned} f_z^2 D_{\gamma H} &= (f_{zz}D_H + f_{zH}) f_\gamma - \left[f_{\gamma z}(-f_H) + f_z f_{\gamma H}\right] \\ &= \gamma \frac{3}{4}\frac{2}{z}(-1)(1-H)\frac{m\left[\left(\frac{1}{2} - \frac{1}{1-\gamma}\right)m + \frac{1}{1-\gamma}\right]}{3m-2} \\ &\quad + \frac{1}{z}\gamma(1-H)\left[\frac{5-\gamma}{4(1-\gamma)}m - \frac{1}{1-\gamma}\right], \end{aligned}$$

where we let $m \equiv (1-h)/(1-H)$. Further manipulation shows that $D_{\gamma H}$ has the same sign as:

$$\frac{\left(\frac{9}{1-\gamma}\right)m^2 - \left(\frac{14-\gamma}{1-\gamma}\right)m + \frac{4}{1-\gamma}}{6m-4}.$$

The above expression is negative for any γ_j when m is equal to 0. By continuity, it is negative for any γ_j when m is small. Observe that the key equation (16) can also be written as:

$$\frac{\gamma_j^2}{\phi_j(1-\gamma_j)}(1-H_G)^3(1-m)m^2 = \frac{\sqrt{p}\chi}{\sigma_\theta^2}.$$

The smaller solution in m (which corresponds to the larger root in z_j), is decreasing in p, and goes to 0 as p goes to 0. Thus, for p sufficiently small, m is small, which implies that $\partial^2 D_j/\partial H\partial\gamma_j < 0$.

Similarly, $\partial^2 D_j/\partial p\partial\gamma_j$ has the same sign as:

$$-3\left(\gamma_j+1\right)m^2 + \left(\gamma_j+10\right)m - 4;$$

and $\partial^2 D_j/\partial\phi_j\partial\gamma_j$ has the same sign as:

$$-3(\gamma_j+1)m^2 + 4(\gamma_j+1)m - 2\gamma_j.$$

When p is small, m is small, and both of the above expressions are negative. ■

Claim 2. For any j and any H, there exists $\underline{d}(\gamma)$ such that

$$\frac{\partial^2 D_j(\gamma,H)/\partial\gamma_j^2}{\partial D_j(\gamma,H)/\partial\gamma_j} \leq \underline{d}(\gamma) < \infty.$$

Proof. Let

$$f(z,\gamma) \equiv \gamma(h(z)-H) - z(1-\gamma)\frac{\sqrt{p}}{\chi}.$$

The demand function is given by the larger root in z to the equation $f(z,\gamma) = 0$. Implicit differentiation gives:

$$\begin{aligned} f_z D_\gamma + f_\gamma &= 0, \\ f_{zz} D_\gamma^2 + 2 f_{z\gamma} D_\gamma + f_z D_{\gamma\gamma} + f_{\gamma\gamma} &= 0, \end{aligned}$$

where we write D_γ and $D_{\gamma\gamma}$ for the first and second derivatives of D_j with respect to γ_j. Thus,

$$\frac{D_{\gamma\gamma}}{D_\gamma} = \frac{f_{zz} f_\gamma}{f_z^2} - \frac{2 f_{z\gamma}}{f_z} + \frac{f_{\gamma\gamma}}{f_\gamma}.$$

Let $m \equiv (1-h)/(1-H)$. The restriction that the demand function is the larger root to $f(z,\gamma) = 0$ is equivalent to $h \geq (1+2H)/3$ (see the proof of Lemma 4), which is equivalent to requiring $m \leq 2/3$. Writing the derivatives of f in terms of m, we obtain:

$$\begin{aligned} \frac{D_{\gamma\gamma}}{D_\gamma} &= \frac{\left(\frac{-3\gamma}{4z^2} m(1-H)\right)\left(\left(\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)\right)(1-H)\right)}{\left(\frac{-\gamma}{2z}(2-3m)(1-H)\right)^2} \\ &\quad - \frac{2\left(\frac{1}{z}\left(\frac{1}{4}m + \frac{\gamma}{1-\gamma}(1-m)\right)(1-H)\right)}{\frac{-\gamma}{2z}(2-3m)(1-H)} + \frac{\frac{1}{4\gamma}m(1-H)}{\left(\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)\right)(1-H)} \\ &= \frac{\frac{-3}{\gamma}m\left(\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)\right)}{(2-3m)^2} + \frac{\frac{4}{\gamma}\left(\frac{1}{4}m + \frac{\gamma}{1-\gamma}(1-m)\right)}{2-3m} + \frac{\frac{1}{4\gamma}m}{\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)} \\ &< \frac{(8-15m)\left(\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)\right)}{\gamma(2-3m)^2} + \frac{\frac{1}{4\gamma}m}{\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)}. \end{aligned}$$

Both the first term and the second term in the final expression are bounded above for

$m \in [0, 2/3]$. So if we let

$$\underline{d}(\gamma) = \max_{m\in[0,2/3]} \frac{\frac{-3}{\gamma}m\left(\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)\right)}{(2-3m)^2} + \frac{\frac{4}{\gamma}\left(\frac{1}{4}m + \frac{\gamma}{1-\gamma}(1-m)\right)}{2-3m} + \frac{\frac{1}{4\gamma}m}{\frac{1}{2}m + \frac{1}{1-\gamma}(1-m)},$$

then the claim is established ■

Claim 3. In the model with correlated information production, if the price of attention is sufficiently low, then $\partial D_j/\partial\gamma_j$ decreases in R.

Proof. From the proof of Proposition 7, equilibrium attention $D_j(\cdot)$ is the solution in z_j to the following key equation:

$$\frac{\gamma_j(h_j(z_j) - RH)}{1 - R\gamma_j}\frac{1}{z_j} = \frac{\sqrt{p}}{\chi}.$$

Implicit differentiation of this equation shows that $\partial^2 D_j/\partial R\partial\gamma_j$ has the same sign as:

$$\left[\frac{-\frac{5}{4}(1-h_j) + (h_j - RH)}{\frac{1}{2}(1-h_j) - (h_j - RH)} + \frac{\gamma_j(h_j - RH)}{H - \gamma_j h_j}\right]\left[\frac{1}{2} + \frac{h_j - RH}{(1-R\gamma_j)(1-h_j)}\right] + \frac{3}{4}.$$

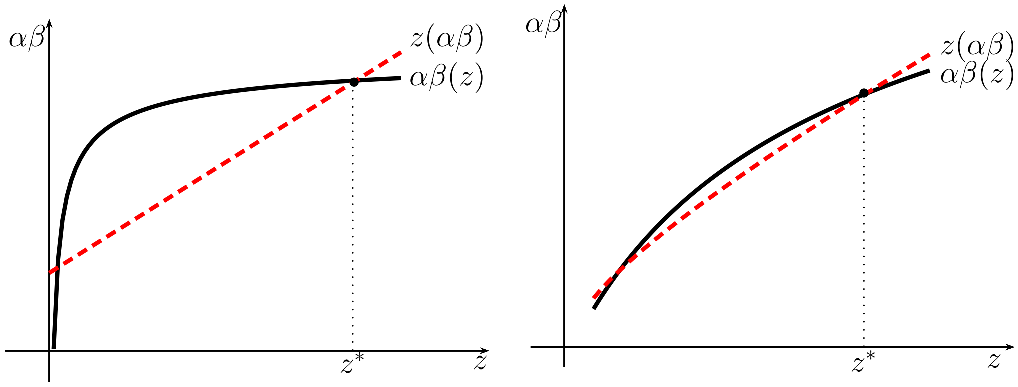
When p is small enough, h_j goes to 1. The first bracketed term goes to $-1 + 1/(J-1)$, which is negative. The second bracketed term approaches $+\infty$. Therefore, $\partial^2 D_j/\partial R\partial\gamma_j$ is negative. ■

B. Discussion on the Marginal Cost of Attention

In the benchmark model, we assume that the marginal cost of giving attention to media outlet j consists of two components: p, the opportunity cost of paying attention to media; and $1/\alpha_j^2$, obscurity or the difficulties to process news stories from media outlet j. In the main text, we claim that allowing the marginal cost to vary in α_j only simplifies our analysis and does not affect our results qualitatively. We elaborate on this point in this section.

Consider an alternative model with constant marginal cost, p. In this case, the characterization of the sender-receiver game is not affected and Proposition 1 holds. But in this specification, the first order condition for the consumer's attention allocation becomes:

$$\frac{\tau_j}{1 + \sum_k \tau_k} = z_j\frac{\sqrt{p}}{\chi}\alpha_j.$$



(a) Marginal cost of attention, p/α_j^2

(b) Marginal cost of attention, p

Figure 8. *The specifications of the marginal cost of attention. When the marginal cost varies in α_j, attention given to media outlet j is in proposition to its influence, illustrated by the red dashed line in Figure 8(a). When the marginal cost is constant, attention increases in the influence of media outlet j, illustrated by the red dashed curve in Figure 8(b). Both cases lead to similar qualitative results.*

Combining the equation above with equation (14), we derive the counterpart of the key equation (16):

$$\frac{\gamma_j(h_j(z_j) - H_G)}{1 - \gamma_j} = z_j \frac{\sqrt{p}}{\chi} \alpha_j.$$

Similar to the benchmark case, the larger solution to this equation implicitly determines the demand function of attention for media outlet j. The left-hand-side is the media outlet j's influence which is a function of z_j. That is exactly the same as the benchmark case. The right-hand-side gives how attention to media outlet j is related to its influence. The proof of Proposition 1 shows that α_j increases in z_j. Therefore, attention given to media outlet j increases in its influence. The difference from the benchmark case is that the relationship is not linear anymore. We illustrate such a difference in Figure 8. The two cases are similar qualitatively, but it is more intuitive and technically convenient to allow the marginal cost of attention to vary in α_j and to work with the linear case.

C. Discussion on the Market Structure

In the benchmark model, we assume that an individual firm j ignores its own impact on the aggregate influence H while making a decision on its accuracy γ_j. Allowing the firm owners to take into account their own impact on the market aggregate only complicates our analysis and reduces tractability of the model, but does not affect our results qualitatively. In other words, the setting of monopolistic competition characterized in the main text is a close approximation of the Nash equilibrium played by firm owners, where owners take the choices of others' as given. We elaborate on this

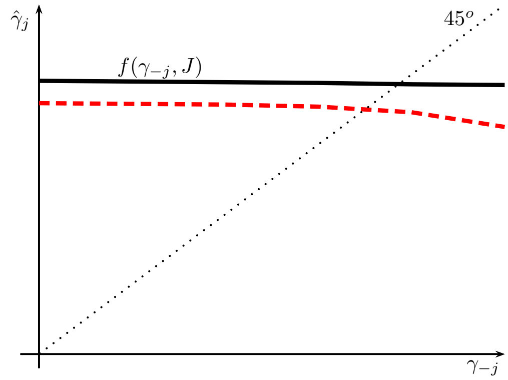


Figure 9. *The (symmetric) Nash equilibrium of a market with homogeneous firms. The horizontal axis is the accuracy γ_{-j} of other firms, and the vertical axis shows the accuracy $\hat{\gamma}_j$ chosen by firm j. The downward sloping curves depict the best response functions for different values of J. Equilibrium accuracy γ^* is given by the intersection of the best-response function with the 45-degree line. When the number of firms increases, f shifts down, which results in a lower γ^*. The solid line illustrates the f function for $J = 3$, while the dashed line for $J = 4$.*

point in this section.

Consider an alternative model with J homogenous firms, in which each owner chooses its own accuracy non-cooperatively, while taking the decisions of other owners' as given. In such a model, the sender-receiver game between editors and news consumers remains unchanged. We rewrite the key equation (16) as a system of equations to emphasize the interdependence of firm j's accuracy γ_j and other firms' accuracy γ_{-j}.

$$\begin{aligned} \left(\frac{\gamma_j}{1-\gamma_j}\right)\left(h_j - \frac{\frac{\gamma_j}{1-\gamma_j}h_j+(J-1)\frac{\gamma_{-j}}{1-\gamma_{-j}}h_{-j}}{1+\frac{\gamma_j}{1-\gamma_j}+(J-1)\frac{\gamma_{-j}}{1-\gamma_{-j}}}\right)\left((1-h_j)\frac{\sigma_\theta}{\chi}\right)^2\frac{\gamma_j}{\phi} &= \frac{\sqrt{p}}{\chi}, \\ \left(\frac{\gamma_{-j}}{1-\gamma_{-j}}\right)\left(h_{-j} - \frac{\frac{\gamma_j}{1-\gamma_j}h_j+(J-1)\frac{\gamma_{-j}}{1-\gamma_{-j}}h_{-j}}{1+\frac{\gamma_j}{1-\gamma_j}+(J-1)\frac{\gamma_{-j}}{1-\gamma_{-j}}}\right)\left((1-h_{-j})\frac{\sigma_\theta}{\chi}\right)^2\frac{\gamma_{-j}}{\phi} &= \frac{\sqrt{p}}{\chi}. \end{aligned} \tag{17}$$

The solution to the key equation system (17) determines h_j and h_{-j}, from which we can recover z_j and z_{-j} through equation (12). We let the attention z_j thus obtained be denoted by $D_j(\gamma_j, \gamma_{-j}, J)$.

The demand function $D_j(\gamma_j, \gamma_{-j}, J)$ increases in γ_j but decreases in γ_{-j} and J. For a cost function described in the main text, the accuracy of owner j can be solved as the best response to γ_{-j}. The best-response function is given by:

$$f(\gamma_{-j}, J) = \arg\max_{\gamma_j} S_j D_j(\gamma_j, \gamma_{-j}, J) - C_j(\gamma_j).$$

We illustrate the best-response function f function in Figure 9. The (symmetric) equilibrium accuracy γ^* satisfies $\gamma^* = f(\gamma^*, n)$.

The function f is downward sloping, because, similar to the Claim 1, $D_{\gamma_j \gamma_{-j}}$ is negative when p is sufficiently small. The function f shifts down when J increases, because $D_{\gamma_j J}$ is negative. These two features combined give rise to the result that entry of a new firm leads to a lower equilibrium accuracy. We experiment with the number of firms J and find that such results hold even for a very small number such as $J = 3$. We also compute the aggregate influence H^* in equilibrium, which can not be studied analytically, and verify that it increases in J. This is the same result as is established in Proposition 6 in a monopolistic competition setting. Our findings are robust to a variety of choices of cost functions and parameter values.