Coordination Motives and Competition for Attention in Information Markets[*]

Simone Galperti and Isabel Trevino[†]

UCSD

March 30, 2020

Abstract

People seek to learn about world events, but often also what others know about those events—for instance, to coordinate their actions. News sources rely on monetizing people's attention to thrive in the market. We find that competition for attention leads to a homogeneous supply of information: News sources are equally accurate in reporting events and equally clear in conveying their reports. This occurs even though people would demand different accuracies and clarities. The type of supplied sources depends on a subtle interaction between the cost structure of producing information and people's coordination motives. We also find that by becoming the "currency" whereby people pay for information, attention causes novel market inefficiencies, whose form and size depend on people's coordination motives. We investigate supply-side policies tackling such inefficiencies.

Keywords: information supply, coordination, payoff interdependence, attention, accuracy, clarity, market inefficiency.

JEL classification: C72, D62, D83, L10.

---

[*]We thank for comments and suggestions S. Nageeb Ali, Dirk Bergemann, Rahul Deb, John Duffy, Joseph Engelberg, Renato Gomes, Marina Halac, Garrett Johnson, Stephen Morris, Alessandro Pavan, Joel Sobel, Philipp Strack, Xavier Vives, Rakesh Vohra, and Joel Watson and seminar participants at ESSET (Gerzensee), Yale, Princeton, Caltech, Queen's University, Brown, Boston University, Bocconi, and Collegio Carlo Alberto. John Rehbeck provided excellent research assistance. All remaining errors are ours.

[†]Department of Economics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA, 92093. E-mail addresses: sgalperti@ucsd.edu; itrevino@ucsd.edu. First draft: July 2016

1 Introduction

People seek information to learn about world events, but often also what others know about those events—for instance, to respond to each other's actions. Political protesters look for mutually known news that can catalyze rallies. Stock traders look for information others may not know, so as to anticipate price swings. Friends may simply look for topics of common interest to have meaningful conversations. Shared knowledge matters in many economic and social settings and is influenced by mass media (Chwe (2013)). Indeed, "significant market events generally occur only if there is similar thinking among large groups of people, and the news media are essential vehicles for the spread of ideas." (Shiller (2015), p. 101) Recognizing that news sources help people learn not only facts, but also others' beliefs about those facts, is then crucial for understanding information markets in environments where people may care about coordinating their actions with others.

Prominent research has studied how the desires to learn facts and others' beliefs, stemming from coordination motives, shape the demand for information. However, it has treated the sources of information as exogenous. Thus, little is known about how those desires affect the supply of information. To fill this gap, in this paper we endogenize the information supply in a full-fledged competitive market. We emphasize an aspect that sets information apart from other commodities: It consumes attention while being consumed. Since attention can be "monetized"—usually by selling it to advertisers—it is a key revenue source for information providers, who fiercely compete for it.[1]

We offer two sets of results. First, competition for attention leads to a homogeneous supply of information: News sources are equally accurate in reporting events and equally clear in conveying their reports. This happens even when consumers would value accessing heterogeneous sources. The type of supplied information depends on a subtle interaction between the cost structure of producing information and the consumers' fact- and belief-learning desires. Second, despite perfect competition the supply of information can be inefficient. We ascribe this inefficiency to two causes specific to information markets: The role of information sources as devices to learn others' beliefs and the role of attention as the "currency" whereby consumers pay for information.

To model the consumers' fact- and belief-learning desires, we follow the paradigm of coordination games with incomplete information developed by Morris and Shin (2002),

[1] The importance of competition for attention when it comes to information sources has been stressed by many scholars across disciplines, including Simon (1971), Benkler (2006), Lanham (2006), Sunstein (2009), Anderson and De Palma (2012), Davenport and Beck (2013), Webster (2014), and Wu (2017).

Angeletos and Pavan (2007), and Hellwig and Veldkamp (2009). For example, consider a society of identical agents—the information consumers—who have to choose which policy to support. Each assigns a weight $1-\gamma$ to supporting the best policy, which depends on some unknown state of the world (the 'facts'), and a weight γ to supporting what others support. Hereafter, we refer to γ as the *coordination motive*: $\gamma > 0$ reflects strategic complementarities and can be interpreted as conformism, while $\gamma < 0$ reflects strategic substitutabilities and can be interpreted as anti-conformism. Each agent wants to learn the state as well as the beliefs of others about it in order to predict their choices according to their coordination motive.

The consumers have access to multiple sources of information, broadly interpreted as traditional and online news outlets, social media, blogs, etc.[2] Realistically, they may interpret each source differently, which may lead to miscoordination. To model this, we follow Dewan and Myatt (2008, 2012) and Myatt and Wallace (2012). Each source produces a signal about the state that has some source-specific noise, which captures the signal's *accuracy*: A smaller noise means that the signal contains better information about the state. A consumer observes a source's signal with some additional individual-specific noise. If this noise is smaller, it means that she understands better the source's content. This jointly depends on the source's *clarity*—namely, how easily it conveys its content—and how much costly attention the consumer pays to it. Concretely, a newspaper may publish an in-depth (accurate) or superficial (inaccurate) report on some proposed policy and may write it in plain English (clear) or technical jargon (unclear). Given this, how much the consumers understand the report depends on how much attention they devote to it.

Our model endogenizes the quantity of information sources and their choice of accuracy and clarity. Each source is a distinct profit-maximizing supplier. Revenues come only from attention, where its marginal dollar value may depend on the amount of attention a source receives and on its accuracy and clarity. We shut down the usual price revenues to better understand the competition for attention, which seems understudied yet important in the Internet era.[3] In the spirit of perfect competition, entry is free and

[2]**?**, Gentzkow and Shapiro (2011), Webster (2014), and Kennedy and Prat (2019) find that many people get their news from multiple sources.

[3]A zero price for accessing information may be optimal for some news outlets, which are similar to two-sided platforms. This seems the case for online outlets facing a strong demand for advertising (**?**). According to Webster (2014), "advertising now supports much of the world's media, both online and off." Newman et al. (2017) find that an overwhelming majority of people does not pay for online news. For an analysis of information pricing see, e.g., Admati and Pfleiderer (1986, 1990), Sarvary (2011), Babaioff et al. (2012), Bergemann and Bonatti (2015), Hörner and Skrzypacz (2016), and Bergemann et al. (2018).

sources act non-strategically.[4] Upon entering, each source commits to the accuracy and clarity of its signal. Higher accuracy and clarity cost more to produce. We call each accuracy-clarity pair a type. After the sources enter, the consumers observe their types, allocate attention, update beliefs based on their signals, and choose their actions. Given how consumers allocate attention, in equilibrium all sources have to make zero profits.

Our first contribution is to characterize the competitive equilibrium and its dependence on the coordination motive γ. Generically, in equilibrium all sources supply the same signal type. This homogeneity arises even though, at least for small enough γ, multiple signal types would receive attention if supplied.[5] Nonetheless, the properties of how the consumers allocate attention render one type always more profitable than all the others. As more—possibly different—sources enter the market, each receives less attention, forcing all but the most profitable type out of business. This type depends on γ and a specific measure of production cost: roughly, on the cost of producing a signal type per unit of its accuracy (i.e., content) normalized by its clarity. Depending on how this measure changes with clarity, the equilibrium type becomes clearer as agents' desire to coordinate with others increases (γ rises), or it always has the highest feasible clarity independently of γ. The equilibrium accuracy depends on γ only through its effect on clarity—for instance, if being clearer requires to be less accurate. The equilibrium type does not depend on other aspects of the consumers, such as their attention cost or prior knowledge about the state. These only determine how much attention they are willing to pay overall and so to the equilibrium quantity of sources.

Our results qualify and sometimes overturn the literature's predictions on where consumers get their information. Fixing its sources, Dewan and Myatt (2008, 2012) and Myatt and Wallace (2012) showed that, as γ falls, the consumers shift attention from the clearest sources to the opaque ones. The former always receive positive attention, but the latter may receive more attention for low enough γ. This is because clearer content tends to be interpreted more similarly and so it becomes more public among its consumers. Intuitively, public information is less worthy of costly attention for consumers who want to match the state but not the actions of others. As a result, opaque sources—which are more private—receive more attention as γ falls. In short, we find that the endogeneity of supply can magnify the attention reallocation mechanism identified in the literature.

[4]It is certainly worth studying industry structures other than perfect competition. This is, however, an important starting point for comparing other markets with information markets and for isolating inefficiencies specific to the latter. Also, news markets—especially in the broad sense used here—tend to be fairly competitive (see Footnote 12).

[5]Consistent with this result, Boczkowski (2010) argues that "the rise of homogenization in the news has led to a state of affairs that neither journalists nor consumers like but feel powerless to alter." (p. 6)

At the same time, we highlight that this mechanisms is only half of the story, the other half being the cost structure of producing information. This can overturn the literature prediction that the consumers' information set always involves the clearest signals (and possibly others) and that it depends on γ. In fact, in some environments the equilibrium supply may involve only the lowest feasible clarity; for others, it can be independent of γ.

Our analysis also offers some broader insights. First, competition for attention need not promote the sources offering the *highest* return to attention, here measured by clarity. This challenges the common wisdom (Davenport and Beck (2013)) that in the "attention economy" higher clarity is always better. In fact, it can *hurt* attention revenues and help unclear sources. We highlight how this depends on the consumers' coordination motive. A second insight relates to the ongoing transition from hard (accurate but unclear) to soft (clear but inaccurate) news. Our results suggest that stronger social conformism (i.e., higher γ) and innovations in news production might be drivers of this phenomenon—not scarcer attention per se. Our emphasis on clarity contrasts with other studies of information markets (e.g., Sarvary (2011)), which focus on accuracy because they ignore the consumers' belief-learning desires. A third insight is to draw a distinction between quantity and diversity of information sources, where competition for attention pushes towards abundance but homogeneity. This echoes Webster's (2014) point that "the media marketplace [is] less diverse than its sheer numerical abundance might suggest" (p. 16) and "competition doesn't do much to improve the diversity of news products." (p. 58) This should not be confused with the variety of news *topics*, which can be explained by people's interests. Our analysis can be viewed as focusing on one specific topic.

Our second contribution is to unveil market inefficiencies due to the very nature of information being the traded good—as opposed to lack of competition, for instance. It is well known that payoff interdependences between consumers can distort the demand for information (Morris and Shin (2002), Angeletos and Pavan (2007), Pavan (2014)). We show how these distortions shape the supply of information. More importantly, we find that even if the information demand is undistorted, the competitive equilibrium can be Pareto inefficient. Efficiency holds if the equilibrium clarity is at the highest feasible level; if not, shifting supply to clearer types can improve welfare. We investigate how supply-side policies can do so by incentivizing entrants to choose specific signal types. These complement the demand-side policies studied by Angeletos and Pavan (2009).

Our analysis identifies two distinctive culprits of market inefficiencies. One is the consumers' coordination motive, which adds to information sources the role of coordination devices and, depending on its sign, can favor inefficiently low clarity. The second is the unique fact that, in contrast to other commodities, attention can become the "currency" whereby consumers pay for information. Unlike the frictionless price mechanism, costly attention does not adjust freely to allow the trading sides to internalize costs and benefits of producing information. This seems a general point that matters for all contexts where people get news "for free" in exchange for their attention.

Related Literature. A rich literature studies information in Gaussian-quadratic games where agents care about coordination. Morris and Shin (2002) stress the role of information as a coordination device and show that more accurate public information can *decrease* welfare, as it may lead agents to react less to informative private signals. Angeletos and Pavan (2007) study how the use and social value of information depend on the coordination motive.[6] In these papers the agents do not choose which information to acquire.

Subsequent research has endogenized information acquisition. Hellwig and Veldkamp (2009) show that it inherits the strategic motives from the underlying coordination game. Colombo et al. (2014) study the effect of coordination motives on efficiency in the acquisition and use of information and the social value of free public information when private information is costly. Myatt and Wallace (2012) allow information sources to differ in accuracy and clarity and characterize the resulting rich patterns of attention allocation. Pavan (2014) studies how this attention allocation can be inefficient. All these papers assume exogenous and heterogeneous sources. We call this assumption into question by finding that competition for attention pushes towards a homogeneous information supply. This literature has also shown that decentralized information acquisition need not serve social interests; we show that its supply need not either.

Few papers endogenize information sources in coordination games; none analyze competitive information markets and their efficiency. The closest papers to ours are Dewan and Myatt (2008, 2012).[7] In these models the number of sources and their accuracy are fixed, each can choose its clarity, and there is no cost of producing signals. Thus, even though their sources compete for attention, their model is not suited to study how such competition ultimately shapes the information supply in a full-fledged market, taking into account the costs and benefits of information. By contrast, we show when the competitive-equilibrium clarity is increasing in γ and how this depends on the cost structure of producing information. Finally, by considering production costs, we are able to

[6] For earlier seminal studies of inefficiencies related to information in markets, see Vives (1988) and related papers in Vives (2010).

[7] Other papers include Cornand and Heinemann (2008), Myatt and Wallace (2014), and Chahrour (2014). For a review of the broader literature on information design see Bergemann and Morris (2019).

define and analyze the efficiency of information markets.

A vast literature has studied the phenomenon of media bias and its political-economy consequences.[8] Media bias refers to the deliberate and systematic distortion of information, so it is different and complementary to the focus of the present study. The idea of competition for attention and the distinction between accuracy and clarity are mostly absent from that literature, which asks whether competition—broadly defined—promotes media independence, timeliness, and unbiasedness. An exception is Chen and Suen (2017), who study a model of news markets where each outlet's owner chooses accuracy to attract attention and its editor chooses a reporting bias that trades off pursuing her agenda and helping readers make informed decisions. Their readers do not care about coordination, thereby removing the key belief-learning motive studied here.

2 A Model of Information Markets

2.1 Information Demand

To model the demand for information we use the framework of Dewan and Myatt (2008, 2012) and Myatt and Wallace (2012) (hereafter, DMW). There is a unit mass of ex-ante identical agents, the *information consumers*, indexed by $\ell \in [0, 1]$, who play a simultaneous-move game. The timeline of each consumer's decisions is as follows:

1. Consumer ℓ chooses a vector of attention allocations $z_\ell \in \mathbb{R}^n_+$, where n is the number of available information sources and $z_{i\ell}$ the amount of attention ℓ pays to source i.
2. Consumer ℓ observes a vector of signals $x_\ell \in \mathbb{R}^n$, one from each source. Each $x_{i\ell}$ provides information about the underlying state of the world θ. The informativeness of these signals depends on z_ℓ as explained below.
3. Consumer ℓ chooses a signal-contingent action $a_\ell \in \mathbb{R}$. The function $A_\ell : \mathbb{R}^n \to \mathbb{R}$ describes this action for each realization of x_ℓ.

Consumer ℓ's payoff depends on how close a_ℓ is to θ and to the average action in the population, denoted by $\overline{a} = \int_0^1 a_\ell d\ell$, and on the attention cost, $C(z_\ell)$:

$$u_\ell = -(1-\gamma)(a_\ell - \theta)^2 - \gamma(a_\ell - \overline{a})^2 - C(z_\ell). \tag{1}$$

The parameter γ captures each consumer's desire to align ($\gamma > 0$) or misalign ($\gamma < 0$) her behavior with others relative to matching the state. We refer to these as the *coordination*

[8] See Gentzkow and Shapiro (2008) and Prat and Stromberg (2013) for literature reviews.

and *fundamental* motives, which drive how much the consumers care about learning the facts (i.e., θ) relative to others' beliefs. As usual, $\gamma < 1$ ensures that best replies are not explosive (Angeletos and Pavan (2007)). The cost of acquiring information, $C(z_\ell)$, depends only on the total amount of attention allocated, not on how it is divided between sources:

$$C(z_\ell) = c\left(\sum_i z_{i\ell}\right),$$

where $c : \mathbb{R} \to \mathbb{R}$ is strictly increasing, continuously differentiable, and convex. That is, the marginal cost of attention is the same between sources and its return may differ between them, but this will be captured by specific properties of each source introduced later.

The consumers have a common prior about the state, given by $\theta \sim N(0, \kappa_0^2)$. DMW assume a diffuse prior with zero precision. While this is without loss of generality for their analysis, we need a proper prior for our welfare analysis to be well defined.

We now describe the *information sources* formally. Each signal $x_{i\ell}$ is of the form

$$x_{i\ell} = \theta + \eta_i + \varepsilon_{i\ell}, \quad \text{where } \eta_i \sim N(0, \kappa_i^2) \text{ and } \varepsilon_{i\ell} \sim N\left(0, \frac{\xi_i^2}{z_{i\ell}}\right),$$

where the noise terms are independently distributed across consumers and sources. The interpretation is that each source i comes with some "sender" noise η_i that captures the quality or *accuracy* of an underlying signal $\overline{x}_i = \theta + \eta_i$. This accuracy is measured by the precision $1/\kappa_i^2$. If consumer ℓ pays attention to source i, he does so imperfectly due to the "receiver" noise $\varepsilon_{i\ell}$, i.e., by observing $x_{i\ell} = \overline{x}_i + \varepsilon_{i\ell}$. This noise reflects the precision of the communication between source i and consumer ℓ. This jointly depends on the *clarity*, $1/\xi_i^2$, with which source i conveys its information and on the attention that consumer ℓ pays to it, $z_{i\ell}$. Intuitively, we can think of $z_{i\ell}$ as how much time consumer ℓ listens to source i: The longer she listens, the better she understands the message $\overline{x}_i$. Choosing $z_{i\ell} = 0$ means ignoring source i, as $x_{i\ell}$ becomes pure noise.[9] Along these lines, we can think of clarity $1/\xi_i^2$ as the return, in terms of content extraction, of each unit of attention.

Fixing the available sources, the demand for information that each faces is given by how much attention it receives from the consumers. This is pinned down by the Perfect Bayesian equilibrium of the game among the consumers, which DMW characterize and will be reviewed in Section 3. We refer to it as the *consumer equilibrium*.

[9] To avoid ambiguity, when κ_i falls we will say that signal i "becomes more accurate." Similarly, when ξ_i falls we will say that signal i "becomes more clear."

2.2 Information Supply

We depart from DMW by endogenizing the supply of information in a full-fledged competitive market. We first describe the general setup and then discuss its main assumptions.

Information production works as follows. Each source is a distinct supplier that can produce one unit of the good "information," namely, a signal with some accuracy and clarity. We can interpret accuracy as a reporting style (summary vs. in depth report) and clarity as a communication technology (broadcast vs. print).[10] Each source can choose among multiple levels of accuracy and clarity and each accuracy-clarity pair defines a signal *type*. For tractability, there are finitely many types, indexed by $t = 1, \ldots, T$, with $\kappa_t > 0$ and $\xi_t > 0$ for all t. Producing a signal of type t costs $h(\kappa_t, \xi_t) > 0$, but communicating it has zero marginal cost—a typical cost structure for information goods in the digital age (Hamilton (2004)). More accurate and clear signals cost more to produce: h is strictly *decreasing* in κ_t and ξ_t. For simplicity, $h(\kappa_t, \xi_t) \neq h(\kappa_{t'}, \xi_{t'})$ if $t \neq t'$. Each source can choose to not produce information at no cost. We refer to a source producing a signal of type t as a t-source.

Each source gets its revenues from the attention it captures and we assume that it cannot charge a price for information. Given this, source i's profit from producing a signal of type t is

$$r\left(\int_0^1 z_{i\ell} d\ell \,\middle|\, \kappa_t, \xi_t\right) - h(\kappa_t, \xi_t).$$

We only assume that $r(\cdot|\kappa_t, \xi_t)$ is strictly increasing and $r(0|\kappa_t, \xi_t) = 0$ for all t.

To define the notion of perfect competition and equilibrium in the overall market, we assume that there is an arbitrarily large number of potential sources and that entry and exit are free. When making its decisions, each source takes as given the existing sources and the consumers' behavior as described by the consumer equilibrium. A source enters and chooses to produce a signal of type t if and only if this is profitable. Let the number of t-sources be q_t. We say that $\mathbf{q}^e = (q_1^e, \ldots, q_T^e)$ is a *competitive equilibrium* if, given $\mathbf{q}^e$, no active source wants to exit and no new source wants to enter, for every type of signals. Characterizing such equilibria is the goal of this paper.

Discussion of the Model

Our setup aims to capture two key aspects of information supply: Information sources

[10] We are agnostic on which technology is clearer. This is ultimately an empirical question and may depend on the type of news (for example, financial vs. general interest), which of the human senses the technologies rely on, and their fleeting or permanent nature. The importance of news technologies is emphasized by Prat and Stromberg (2013), Webster (2014), and Newman et al. (2017).

compete with each other and attracting attention is central to this competition.[11] To do this in the cleanest way, we assume perfect competition and that the sources earn revenues only from attention, thereby shutting down price revenues. Regarding the latter, while price competition has been extensively studied, competition for attention has not. Yet, many outlets supply news for free and rely on attention to make money from advertisement. Suppliers may also compete in prices, of course. This can add a trade-off between price and attention revenues (as in Crampes et al. (2009)), which may obfuscate our paper's main points. Also, the discreteness of the "price cost" of acquiring an information source can lead to multiplicity of consumer equilibria as in Hellwig and Veldkamp (2009). A study of these interesting issues is beyond the scope of this paper.

Although it is certainly worth considering other industry structures, perfect competition is an important starting point. First, it is a useful benchmark for comparing information markets with other markets. One of our goals is to study whether information markets are efficient and how this depends on the nature of information as a good and on the consumers' coordination motive. Perfect competition lets us isolate the effects of these features from other expected inefficiencies that might arise, for example, due to lack of competition. Second, even though in reality news markets are not perfectly competitive, they tend to be fairly competitive—at least in developed countries.[12] Also, we interpret information sources broadly, including traditional news outlets as well as online outlets, social media, blogs, etc. In this respect, Webster (2014) argues that "perhaps the most astonishing thing about digital media is their numerical abundance."

In terms of the sources' objective, note that, when choosing their signal, they have no information about the state and no preference over which consumer attends to which source or how she uses information. One interpretation is that they supply information to attract "eye balls" and sell them to advertisers. Monotonicity of the revenue function r is intuitive: The more time consumers spend attending a source, the more ads it can show them, and the more it can charge to advertisers.[13] Except for this minimal property, r can be quite general. For instance, a non-linear r may capture that the price advertisers pay varies with how much attention a source gets. This price can result as an equilibrium in the ads market, which we leave implicit, and can depend on the attention other sources

[11] These are key aspects of the "attention economy" according to Davenport and Beck (2013) and Webster (2014).

[12] Kennedy and Prat (2019) find concentration indices of news markets that are fairly small relative to usual standards (see also Noam (2009)).

[13] Webster (2014), Ch. 4, describes measures that media use to assess audiences which are consistent with this story. Davenport and Beck (2013) describe the "stickiness" of online news outlets—namely, their ability to grab and keep attention—as the time spent on a site and the number of visits and viewed pages per person.

get. In this model, however, any change in supply (i.e., $\mathbf{q}$) affects the attention received by all types of sources (see Lemma 1). Given this, the assumption here is that the attention received by t-sources is a sufficient statistic to determine its value for advertisers. This value—and so r—may also depend on a source's accuracy and clarity: More ads may clutter a web page, lowering its clarity, or compete for time with content in a documentary show, lowering its accuracy.

3 Attention Allocation

We briefly review how the consumers allocate their attention, as shown by DMW. As standard in this literature, DMW focus on linear symmetric equilibria—so we drop the subscript ℓ from z_ℓ and A_ℓ. DMW show that there is a unique consumer equilibrium that satisfies symmetry and linearity. Lemma 1 characterizes it, slightly extending DMW's arguments to account for the quantity of each type of sources.

Lemma 1 (Attention Allocation)**.** *Fix* $\mathbf{q} = (q_1, \ldots, q_T)$*. If* $q_t > 0$*, the equilibrium allocation of attention to each* t*-source satisfies*

$$z_t(\mathbf{q}) = \frac{\xi_t \max\{K(\mathbf{q}) - \xi_t, 0\}}{(1-\gamma)\kappa_t^2}, \tag{2}$$

where $K(\mathbf{q})$ *has the following properties: (i) it is a continuous function of* $\mathbf{q}$*; (ii) it decreases strictly as the coordination motive or the prior precision rises; (iii) for every* t*, it decreases as the accuracy or the quantity of* t*-sources rises (strictly if* $z_t(\mathbf{q}) > 0$*); (iv) for every* t*, it is non-monotonic in the clarity of* t*-sources; (v) it converges to zero as the prior becomes infinitely precise (i.e.,* $\lim_{\kappa_0^2 \to 0} K(\mathbf{q}) = 0$*).*

DMW provide intuition for the properties of $K(\mathbf{q})$, except of course its dependence on the quantity of sources and the prior. Intuitively, the more sources enter the market, the harder it is for each to attract attention. Also, if consumers can know more about the state from their prior, they will pay less attention to each source: An informative prior weakens their incentives to acquire new information to learn about the state and it already helps them infer others' beliefs.[14]

Therefore, the prior has to be sufficiently imprecise for some source to receive attention (property (v)). We make the stronger assumption that each source type has a chance of

[14] Consumers' paying attention to all sources of either type is clearly driven by the symmetry of the model. Although this is in part at odds with reality, some evidence shows significant overlaps in audiences across news sources (Gentzkow and Shapiro (2011), Webster (2014), Newman et al. (2017)).

receiving enough attention to make positive profits. Define

$$H(\kappa_t, \xi_t) = r^{-1}(h(\kappa_t, \xi_t)|\kappa_t, \xi_t). \tag{3}$$

We can interpret $H(\kappa_t, \xi_t)$ as the production cost of t-sources denominated in units of attention, taking into account the ability of t-sources to monetize attention (i.e., $r(\cdot|\kappa_t, \xi_t)$).

Assumption 1 (Strong Demand)**.** *The consumers' prior satisfies*

$$\kappa_0^2 > \sqrt{c'(0)}\left[\frac{\xi_t}{1-\gamma} + \frac{H(\kappa_t, \xi_t)}{\xi_t/\kappa_t^2}\right] \text{ for all } \ t = 1, \ldots, T.$$

This condition is akin to assuming that the intercept of a demand curve is large enough to induce positive trade in equilibrium. To see this, suppose only t-sources can be produced (i.e., $q_{t'} = 0$ for $t' \neq t$). One can show (see equation (11) in Appendix A) that as $q_t \to 0$, $K(\mathbf{q})$ converges to the upper bound

$$\overline{K} = \frac{(1-\gamma)\kappa_0^2}{\sqrt{c'(0)}}.$$

Assumption 1 ensures that for every t

$$r\left(\frac{\xi_t(\overline{K} - \xi_t)}{(1-\gamma)\kappa_t^2}\middle| \kappa_t, \xi_t\right) > h(\kappa_t, \xi_t).$$

That is, the consumers are willing to pay enough attention to t-sources for the first entrants to earn a profit.

For future reference, we review another property of the attention allocation. As the coordination motive γ increases, the consumers shift attention towards the clearest sources from the unclear ones (Myatt and Wallace (2012), Proposition 5). This reflects the former's comparative advantage as coordination devices.[15] As the consumers pay more attention to source i, they interpret its content more similarly: $x_{i\ell}$ and $x_{i\ell'}$ become more correlated ($corr(x_{i\ell}, x_{i\ell'}|\theta) = \frac{\kappa_i^2}{\kappa_i^2 + \xi_i^2/z_i}$). This content then becomes more public, helping coordination. Clearer sources are easier to interpret and so get more attention as γ rises. That unclear sources have an advantage when the consumers don't want to coordinate with others may seem counterintuitive, but it has a simple logic that relies on attention being costly. Each consumer now wants to respond in opposite directions to public information suggesting, for instance, that θ is high *and* thus everybody else's action is high. As a result, such information barely moves her action and so has little value. By contrast, an unclear source conveys less information about others' actions

[15] That clear sources are better coordination devices seems to be a general property that holds beyond the Gaussian framework considered here (see Chwe (2013)).

(i.e., it is more private). Thus, the consumer responds more to its content about θ, which renders acquiring it worth the cost. Note that even though accuracy affects the correlation between $x_{i\ell}$ and $x_{i\ell'}$, it is not the main driver of how the consumers shift attention as γ rises.

4 Equilibrium Supply of Information

4.1 Entry Decisions

Consider now the decision of information sources to enter the market and which signal type to produce. Recalling (3), for every $\mathbf{q}$ a new t-source enters if and only if

$$z_t(q_t + 1, \mathbf{q}_{-t}) \geq H(\kappa_t, \xi_t),$$

where $\mathbf{q}_{-t}$ is the vector excluding q_t. This immediately implies the following.

Lemma 2. $\mathbf{q}^e$ *is a competitive equilibrium if and only if, for all* $t = 1, \ldots, T$*, we have* $z_t(q_t^e, \mathbf{q}_{-t}^e) \geq H(\kappa_t, \xi_t)$ *when* $q_t^e > 0$ *and* $z_t(q_t^e + 1, \mathbf{q}_{-t}^e) < H(\kappa_t, \xi_t)$.

These conditions characterize all equilibria, but are inconvenient to use with discrete quantities. As usual, one more entry can cause profits to jump strictly below zero. To avoid such issues, textbook analysis of competitive markets typically assumes that each supplier is small so that equilibria can be characterized by zero-profit conditions. We can adopt a similar approach by letting q_t be a real number for all t. The interpretation is again that each source is small relative to the whole market, which is consistent with the idea of perfect competition. Given this, the equilibrium conditions become

$$q_t^e > 0 \Rightarrow z_t(q_t^e, \mathbf{q}_{-t}^e) = H(\kappa_t, \xi_t), \quad t = 1, \ldots, T, \tag{4}$$

$$q_t > q_t^e \Rightarrow z_t(q_t, \mathbf{q}_{-t}^e) < H(\kappa_t, \xi_t), \quad t = 1, \ldots, T. \tag{5}$$

This transition to continuous quantities should be viewed as just a convenient way to avoid integer problems. However, some readers may detect a conceptual issue with respect to DMW's setup: Since each source provides an independent signal, allowing their number to grow arbitrarily could eventually lead the consumers to fully learn the state. We avoid this issue and maintain the spirit of the model by appropriately rescaling accuracies in the transition to the continuum. Given this, $1/\kappa_t^2$ can be viewed as the *rate* at which t-sources entering the market contribute to the amount of content they provide to the consumers. We present the details in Appendix B.

4.2 Type and Quantity of Information and the Role of Coordination Motives

We can now characterize the equilibrium supply of information. Our goal is to shed light on how it depends on the sources' characteristics and the consumers' preferences—in particular, their coordination motive. These preferences are reflected in specific properties of the attention functions, which, importantly, are derived (not assumed) from primitives about how agents use information in a variety of social and economic contexts.

Using expression (2), we have that $z_t(q_t, \mathbf{q}_{-t}) \geq H(\kappa_t, \xi_t)$ if and only if

$$K(\mathbf{q}) \geq m(\gamma, t) \equiv \xi_t + (1-\gamma)\frac{H(\kappa_t, \xi_t)}{\xi_t/\kappa_t^2}.$$

Define

$$m(\gamma) = \min_{t=1,\ldots,T} m(\gamma, t) \quad \text{and} \quad T(\gamma) = \arg\min_{t=1,\ldots,T} m(\gamma, t). \tag{6}$$

Proposition 1. *For every γ, there exists a competitive equilibrium $\mathbf{q}^e$. The equilibrium satisfies the following properties: (i) $q_t^e = 0$ for $t \notin T(\gamma)$; (ii) $\mathbf{q}^e \neq \mathbf{0}$ and solves*

$$c'\left(\sum_{t\in T(\gamma)} q_t^e H(\kappa_t, \xi_t)\right)\left[\frac{m(\gamma)}{(1-\gamma)\kappa_0^2} + \sum_{t\in T(\gamma)} q_t^e \frac{H(\kappa_t, \xi_t)}{\xi_t}\right]^2 = 1; \tag{7}$$

and (iv) if $T(\gamma) = \{t^e\}$, there exists a unique $q_{t^e}^e$ that solves (7).

The intuition is simple. Profitable source types attract more entrants, which lowers $K(\mathbf{q})$ and so the attention revenues of every source in the market. Eventually, this battle for attention forces all but the the most competitive source types (i.e., those in $T(\gamma)$) out of business. We explain shortly what determines this relative competitiveness.

Interestingly, only one source type is supplied in equilibrium. Indeed, $T(\gamma)$ contains only one element, generically, with respect to γ. This supply homogeneity occurs even though every type can, a priori, attract enough attention to be profitable (Assumption 1). Moreover, at least for small γ, multiple types would receive attention if provided (Lemma 1). Yet, some types receive more attention and so are more profitable and draw more entrants. Put differently, one source type dominates the market not because the consumers do not want to attend to other sources, but because competition renders these other types unprofitable.

Turning to the equilibrium dependence on the consumer's preferences, we first show that for the coordination motive to matter, clarity must vary across source types. Note that if $\xi_t = \xi_{t'}$ but $\frac{H(\kappa_t,\xi_t)}{1/\kappa_t^2} < \frac{H(\kappa_{t'},\xi_{t'})}{1/\kappa_{t'}^2}$, by condition (6) no source will ever choose type

t' in equilibrium. The quantity $\frac{H(\kappa,\xi)}{1/\kappa^2}$ is simply the *average* cost—in attention units—of producing accuracy $1/\kappa^2$, holding clarity fixed. Thus, it measures how efficient a t-source is at producing content. As a result, if sources can use only accuracy to attract attention, the most efficient type at producing content wins the whole market. This type is independent of γ—even though accuracy affects the correlation of signals between consumers and so their ability to coordinate.

Corollary 1 (Competition via Accuracy)**.** *Suppose all source types have the same clarity (i.e., $\xi_t = \xi$ for all t). Then, only the types that minimize the average cost of accuracy are supplied: For all γ,*

$$T(\gamma) = \left\{ t : \frac{H(\kappa_t, \xi)}{1/\kappa_t^2} = \min_{t'=1,\dots,T} \frac{H(\kappa_{t'}, \xi)}{1/\kappa_{t'}^2} \right\}.$$

A higher γ reduces the equilibrium total accuracy (i.e., $\sum_{t\in T(\gamma)} q_t^e/\kappa_t^2$).[16]

The total accuracy measures how much content is supplied overall. It falls as γ rises because this reduces the consumers' willingness to pay attention and, therefore, how many sources the market can sustain in equilibrium. Note that if the average cost $\frac{H(\kappa,\xi)}{1/\kappa^2}$ is well behaved, the equilibrium type is again unique, generically. For instance, this is the case if $\frac{H(\kappa,\xi)}{1/\kappa^2}$ is strictly monotonic or U-shaped in κ.

When clarity differs across types of sources, it opens a channel for the coordination motive to affect the equilibrium type. To examine this, we need some preliminary observations. By the logic of Corollary 1, we can assume that every $t \in T$ with clarity $1/\xi_t^2$ is associated with a unique accuracy level $1/[\kappa(\xi_t)]^2$ (i.e., the feasible level that minimizes $\frac{H(\kappa,\xi_t)}{1/\kappa^2}$). We call $1/[\kappa(\xi)]^2$ the *efficient* accuracy for clarity $1/\xi^2$. Using this, define

$$\hat{H}(\xi_t) = \frac{H(\kappa(\xi_t), \xi_t)}{\xi_t/[\kappa(\xi_t)]^2}.$$

The denominator of $\hat{H}$ is the ratio of the efficient accuracy and the (square root of the) clarity of t-sources. This measures how much content they provide, adjusted by how accessible it is. Thus, $\hat{H}(\xi_t)$ is the efficient average cost—in attention units—of the *clarity-adjusted* content of t-sources. For simplicity, suppose $\hat{H}(\xi_t) \neq \hat{H}(\xi_{t'})$ if $\xi_t \neq \xi_{t'}$.

[16] This last part follows from observing that, by Proposition 1, $\mathbf{q}^e$ must satisfy

$$c'\left(\frac{H(\kappa_t,\xi)}{1/\kappa_t^2}\sum_{t\in T(\gamma)}\frac{q_t^e}{\kappa_t^2}\right)\left[\frac{\xi}{(1-\gamma)\kappa_0^2} + \frac{H(\kappa_t,\xi)}{\xi/\kappa_t^2}\left(\frac{1}{\kappa_0^2} + \sum_{t\in T(\gamma)}\frac{q_t^e}{\kappa_t^2}\right)\right]^2 = 1. \tag{8}$$

Since a higher γ increases the left-hand side of this equation, it must lead to a lower $\sum_{t\in T(\gamma)} \frac{q_t^e}{\kappa_t^2}$.

Proposition 2. *In equilibrium, all sources supply the clearest type of signals for every γ if and only if this type minimizes the efficient average cost of clarity-adjusted content $\hat{H}(\xi_t)$. Otherwise, as γ rises, the equilibrium sources supply types with higher $\hat{H}(\xi_t)$.*[17]

This result uncovers a tight and general relation between the information-production costs—as measured by $\hat{H}$—and the coordination motive in determining the equilibrium type of sources.

Whether a higher γ leads to a higher equilibrium accuracy or clarity depends on finer properties of the information-production costs. We start from the equilibrium clarity.

Corollary 2 (Coordination Motive and Equilibrium Clarity)**.**
Case 1: Suppose clearer signal types cost more to produce: $\hat{H}(\xi_t) < \hat{H}(\xi_{t'})$ if and only if $\xi_t > \xi_{t'}$. In equilibrium, the sources supply clearer signals as γ rises. They all supply the least (most) clear signal for γ sufficiently small (large).[18]
Case 2: Suppose clearer signal types cost less to produce: $\hat{H}(\xi_t) < \hat{H}(\xi_{t'})$ if and only if $\xi_t < \xi_{t'}$. In equilibrium, all sources supply the clearest signal for every γ.

To gain intuition, recall that to coordinate their actions with others, the consumers want to observe the same information as others (Hellwig and Veldkamp (2009)). Clearer sources help in this endeavor, as they make it easier to extract a given content. Thus, clearer sources earn larger attention revenues at the expense of less clear sources as γ rises. But this is only half of the story. It turns out that if clearer sources are cheaper to produce (according to $\hat{H}$), they are also more profitable for every γ, so they always win the whole market. By contrast, if higher clarity costs more to produce (according to $\hat{H}$), the consumers have to value it sufficiently (i.e., γ has to be sufficiently large) for clearer sources to be the most profitable. For smaller γ, instead, less clear sources will control the whole market since the consumers do not seek sources that help them know what others know.

It may be puzzling that low-clarity sources can win the competition for attention. After all, they offer a low return to attention and can never fully steal attention from high-clarity sources (see DMW). But the effect of γ in conjunction with the cost $\hat{H}$ helps explain this puzzle and highlights that the role of information sources as coordination devices is important to understand their supply. That the *least* clear type of sources can dominate the whole market is also in sharp contrast with DMW's prediction that agents always consume some high-clarity information.

[17] The proof takes care of the non-generic cases with multiple possible equilibrium types.

[18] Note that $\xi_t > \xi_{t'}$ implies $\hat{H}(\xi_t) < \hat{H}(\xi_{t'})$ if $\xi_t > \xi_{t'}$ implies $\frac{H(\kappa(\xi_t),\xi_t)}{1/[\kappa(\xi_t)]^2} < \frac{H(\kappa(\xi_{t'}),\xi_{t'})}{1/[\kappa(\xi_{t'})]^2}$, for instance. Again, the proof takes care of the non-generic cases with multiple possible equilibrium types.

The monotonicity properties of the efficient, adjusted, average cost $\hat{H}$ are ultimately an empirical question. Since $\hat{H}$ is expressed in attention units, our theory also emphasizes taking into account how the ability to monetize attention depends on accuracy and clarity. To build intuition, suppose there exists an intrinsic production trade-off between accuracy and clarity, namely, the efficient accuracy falls as clarity rises (i.e., $\kappa(\xi)$ is decreasing in ξ).[19] Concretely, each news source may be reporting on the desirability of some public policy. On the one hand, they may want to convey whether the policy is desirable in simple and accessible words (high clarity), which requires to leave out details about the policy itself (low accuracy). On the other hand, they can publish an in-depth report explaining all arguments in favor and against the policy (high accuracy), which however requires rich data, intricate graphs, technical jargon, and subtle logics (low clarity). In this case, $\hat{H}$ is decreasing if and only if the efficient accuracy falls sufficiently fast as clarity rises. Thus, if the accuracy-clarity trade-off is not too severe, $\hat{H}$ should be increasing.

A production trade-off between accuracy and clarity opens a channel for the coordination motive to also affect the equilibrium accuracy of the information sources. By contrast, if for instance the production cost were separable between accuracy and clarity, the equilibrium accuracy would always be independent of γ.

Corollary 3 (Coordination Motive and Equilibrium Accuracy)**.** *Suppose the efficient accuracy is decreasing in clarity (i.e., $\kappa(\xi)$ is decreasing in ξ).*
Case 1: If clearer signal types cost more to produce according to $\hat{H}$, then in equilibrium the sources supply less accurate signals as γ rises.
Case 2: If clearer signal types cost less to produce according to $\hat{H}$, then in equilibrium all sources supply a low-accuracy signal, independently of γ.

Abstracting from details, high-clarity-low-accuracy and low-clarity-high-accuracy sources may resemble what the news industry calls “soft” and “hard” news. The last decades have seen much discussion that soft news has been replacing hard news. One cause may be that the consumers’ attention has become scarcer, which may penalize hard news. Our model suggests another story—here the attention cost does not affect the equilibrium type of sources (as shown below). The demise of hard news may be due to rising social conformism (i.e., γ) or technological changes lowering the efficient average cost of soft news.

Finally, the equilibrium type of sources does not depend on the consumers’ prior and

[19] A sufficient condition for this is that the marginal cost of increasing accuracy (i.e., lowering κ) rises sufficiently fast as we increase clarity (i.e., lower ξ): If H is differentiable, this sufficient condition is $H_{\kappa\xi}(\kappa,\xi) > -\frac{2H_{\xi}(\kappa,\xi)}{\kappa}$ for all (κ,ξ).

attention cost. These aspects affect their overall willingness to pay attention to acquire information, not how to divide attention between sources. Indeed, expression (7) implies the following.

Corollary 4. *Fix any equilibrium selection that involves a unique type t^e for all γ. Then, $q^e_{t^e}$ falls as the prior becomes more precise or the marginal attention cost uniformly rises.*[20]

5 Inefficiencies in Information Markets

5.1 Undistorted Demand for Information

Our goal is to understand whether *information* markets exhibit some specific inefficiencies. They can certainly exhibit inefficiencies common with other markets, caused for instance by a lack of competition. Ruling out standard inefficiencies by considering perfect competition as a benchmark allows us to focus on our goal.

Payoff interdependencies between consumers of information can already distort its demand. For fixed sources (i.e., $\mathbf{q}$), Angeletos and Pavan (2007) and Pavan (2014) characterize these distortions by comparing how the consumers actually acquire and use information and the acquisition and use that would maximize their ex-ante utility. This criterion takes as constraint that information cannot be transferred between consumers. Their characterization uses a general quadratic payoff function $\hat{u}(a_\ell, \overline{a}, \theta, \sigma)$, where $\sigma^2 \equiv \int_0^1 [a_\ell - \overline{a}]^2 d\ell$ is the action dispersion in the population. Appendix A.1 lists the assumptions on $\hat{u}$ common in the literature, which subsume u in (1) and other applications (e.g., Morris and Shin (2002)). While $\hat{u}$ and u are strategically equivalent, they can cause different distortions. Section 5.2 reviews the possible distortions and shows how they affect the overall market outcome.

We start from a more novel question: If the demand exhibits no distortions, will the resulting competitive equilibrium $\mathbf{q}^{e*}$ be efficient? That is, starting from $\mathbf{q}^{e*}$, does there exist $\mathbf{q}$ that increases either the consumers' expected payoff or the profit of some source, without reducing the payoff of any other party? The demand is undistorted if and only if the partial derivatives of $\hat{u}$ satisfy $\hat{u}_{a\overline{a}} = -\hat{u}_{\overline{a}\overline{a}}$, $\hat{u}_{\overline{a}\theta} = 0$, $\hat{u}_{\sigma\sigma} = 0$, and $\hat{u}_{\overline{a}}(0, 0, 0, 0) = 0$ (see Section 5.2 and Angeletos and Pavan (2007) and Pavan (2014)). For instance, this is the case for u in (1).

[20] In the case of competition via accuracy (Corollary 1), the equilibrium total accuracy (i.e., $\sum_{t \in T(\gamma)} q^e_t / \kappa^2_t$) falls as $1/\kappa^2_0$ rises or c' uniformly rises (see expression (8)).

Proposition 3. *Suppose the acquisition and use of information is undistorted. If in the competitive equilibrium all sources choose the clearest signal type, then the equilibrium is Pareto efficient.*

The argument proceeds as follows. By Lemma 4 in the appendix, under undistorted demand we can express the consumers' expected payoff (up to a constant and positive scalar multiplication) as

$$V(\mathbf{q}) \;=\; -K(\mathbf{q})\sqrt{c'\left(Z(\mathbf{q})\right)} - c\left(Z(\mathbf{q})\right),$$

where $Z(\mathbf{q}) \equiv \sum_{t=1}^{T} q_t z_t(\mathbf{q})$ and $K(\mathbf{q})$ and $z_t(\mathbf{q})$ are calculated as in Lemma 1 setting γ equal to $\gamma^* \equiv -\frac{\hat{u}_{a\overline{a}}}{\hat{u}_{aa}}$. Intuitively, $V(\mathbf{q})$ takes this simple form because, even if $\mathbf{q}$ involves different quantities of different types of sources, the optimal allocation of attention equates the benefits and costs of attending to them. Using this and recalling that at $\mathbf{q}^{e*}$ all sources make zero profits, we consider the problem[21]

$$\max_{\mathbf{q}\in\mathbb{R}^T_+} V(\mathbf{q}) \quad \text{s. t.} \quad q_t[r(z_t(\mathbf{q})|\kappa_t,\xi_t) - h(\kappa_t,\xi_t)] \geq 0, \quad t = 1,\ldots,T.$$

Note that V is decreasing in $K(\mathbf{q})$ and in the total attention paid by the consumers, $Z(\mathbf{q})$. Competitive pressures already minimize $K(\mathbf{q})$, that is, $K(\mathbf{q}^{e*}) \leq K(\mathbf{q})$ for all $\mathbf{q} \neq \mathbf{0}$ that satisfy the profit constraints. If $\mathbf{q}^{e*}$ also minimizes $Z(\mathbf{q})$ subject to all constraints, $\mathbf{q}^{e*}$ is efficient. The key step of the proof is the following. Suppose the clearest type is the most profitable, but $\mathbf{q}$ involves positive supply of other types, which also earn non-negative profits. It is then possible to shift supply towards the clearest type and, in so doing, reduce $Z(\mathbf{q})$ without violating the profit constraints. Thus, in this case the market forces and the welfare goals are aligned.

Perhaps surprisingly, this alignment can fail despite perfect competition.

Proposition 4. *Suppose the acquisition and use of information is undistorted. Also, suppose the equilibrium type of sources t^{e*} is not the clearest ($\xi_t < \xi_{t^{e*}}$ for some $t \in T$). Then, everything else equal, the equilibrium is not Pareto efficient provided that ξ_t and $c'(0)$ are sufficiently small.*

Intuitively, under these conditions there is a way to shift supply from t^{e*}-sources to t-sources that reduces the consumers' total attention while ensuring non-negative profits. This is non-trivial because the shift does not involve simply a marginal change in $\mathbf{q}^{e*}$, which would cause t^{e*}-sources to earn negative profits.

[21] This problem has a solution because $w_t(\mathbf{q})$ and $z_t(\mathbf{q})$ are continuous functions for every t. Moreover, since $z_t(\mathbf{q})$ is strictly decreasing in q_t when $z_t(\mathbf{q}) > 0$, we can restrict attention to $q_t \in [0, \overline{q}_t]$ for some $\overline{q}_t < +\infty$ for every t. Existence of a solution follows by standard compactness and continuity arguments.

Thus, despite perfect competition and undistorted demand, information markets can be inefficient. We ascribe this to two distinctive causes. The first is that information sources can serve as coordination devices, which matters for the consumers' belief-learning motive. Indeed, the equilibrium can be inefficient only if this motive is sufficiently weak so that they prefer unclear sources. The second cause is that costly attention becomes the "currency" whereby consumers pay for information. This is a key difference from standard markets where consumers pay with "money" through the frictionless price mechanism. The friction represented by the attention cost may be viewed as an unavoidable externality imposed on information consumers by its suppliers, who do not always internalize it. These causes are fundamentally different from the drivers of inefficient entry usually emphasized in the industrial-organization literature.[22]

This role of attention as the currency of information markets can be further understood with the following observation. In our model, entrants of type t essentially generate a horizontal supply curve of level $h(\kappa_t, \xi_t)$ (akin to a constant marginal cost of production for society). Therefore, if they faced a standard demand curve, the competitive equilibria would be efficient. This is because the price mechanism would convey to the trading sides the social marginal cost and benefit of the traded good. The ability of information sources to monetize attention, however, can be seen as giving rise to a non-standard demand schedule, which we show causes inefficiency.

5.2 Distorted Demand for Information

We now consider settings where the demand for information is distorted. Angeletos and Pavan (2007) and Pavan (2014) classify the distortions as follows and discuss applications where they arise:

• *complete-information externalities (CIE)*: Even under complete information, each consumer may not internalize how her action affects everybody else. Their equilibrium action rule is $\alpha(\theta) = \overline{\nu} + \nu\theta$, where $\nu = -\frac{\hat{u}_a(0,0,0,0)}{\hat{u}_{aa}+\hat{u}_{a\overline{a}}}$ and $\overline{\nu} = -\frac{\hat{u}_{a\theta}}{\hat{u}_{aa}+\hat{u}_{a\overline{a}}}$. The first-best action rule satisfies $\alpha^*(\theta) = \overline{\nu}^* + \nu^*\theta$, where $\overline{\nu}^* = -\frac{\hat{u}_a(0,0,0,0)+\hat{u}_{\overline{a}}(0,0,0,0)}{\hat{u}_{aa}+2\hat{u}_{a\overline{a}}+\hat{u}_{\overline{a}\overline{a}}}$ and $\nu^* = -\frac{\hat{u}_{a\theta}+\hat{u}_{\overline{a}\theta}}{\hat{u}_{aa}+2\hat{u}_{a\overline{a}}+\hat{u}_{\overline{a}\overline{a}}}$.

• *socially optimal degree of coordination (SOC)*: Under incomplete information, each consumer may not internalize how aligning her action with others' affects dispersion (i.e., $Var[a - \overline{a}]$) and non-fundamental volatility (i.e., $Var[\overline{a} - \alpha^*]$). To fix this, the

[22] In contrast to this paper, classic IO theories ascribe inefficient entry to the fact that firms are not price-takers, incur fixed set-up costs (scale economies), and serve consumers with heterogeneous tastes (Dixit and Stiglitz (1977); Mankiw and Whinston (1986); Anderson et al. (1995)). In Anderson and Coate (2005), the result that broadcasters supply programs inefficiently depends—among other things—on their having market power and acting strategically.

first-best action strategy A^* requires to use the coordination weight

$$\gamma^* = 1 - \frac{\hat{u}_{aa} + 2\hat{u}_{a\overline{a}} + \hat{u}_{\overline{a}\overline{a}}}{\hat{u}_{aa} + \hat{u}_{\sigma\sigma}},$$

while the consumers use $\gamma = -\frac{\hat{u}_{a\overline{a}}}{\hat{u}_{aa}}$. A standard assumption is that $\gamma^* < 1$.[23]

• *social aversion to dispersion (SAD)*: By assumption, the dispersion σ^2 has a non-strategic effect on the consumers' payoff, so they do not take it into account when allocating attention. A standard assumption is that the social preference for dispersion, measured by $\hat{u}_{aa} + \hat{u}_{\sigma\sigma}$, is negative.

Importantly, these distortions arise from the consumer's preferences. Hence, their existence and effects apply for every set of available information sources.

The next result characterizes how these distortions affect the supply of information. Let $\mathbf{q}^e$ be the equilibrium under the distorted acquisition and use of information (i.e., those driven by α and γ). Let $\mathbf{q}^{e*}$ be the equilibrium under the first-best acquisition and use of information (i.e., those driven by α^*, γ^*, and $\hat{u}_{\sigma\sigma}$).

Proposition 5. *Suppose both $\mathbf{q}^e$ and $\mathbf{q}^{e*}$ involve only one source type in positive supply.*
(1) CIE distortions do not affect the equilibrium type of sources (i.e., $t^e = t^{e*}$ generically when $\gamma = \gamma^*$ and $\hat{u}_{\sigma\sigma} = 0$). Moreover, $q^{e*}_{t^{e*}} > q^e_{t^e}$ if and only if $|\nu^*| > |\nu|$.
(2) SAD distortions do not affect the equilibrium type of sources (i.e., $t^e = t^{e*}$ generically when $\alpha = \alpha^*$ and $\gamma = \gamma^*$). Moreover, $q^{e*}_{t^{e*}} > q^e_{t^e}$ if and only if $\hat{u}_{\sigma\sigma} < 0$.
(3) SOC distortions can affect the equilibrium type of sources (given $\alpha = \alpha^*$ and $\hat{u}_{\sigma\sigma} = 0$). In this case, if $\gamma < \gamma^*$ (resp. $\gamma > \gamma^*$), then $\hat{H}(\xi_{t^e}) < \hat{H}(\xi_{t^{e*}})$ (resp. $\hat{H}(\xi_{t^e}) > \hat{H}(\xi_{t^{e*}})$). If instead $t^e = t^{e*}$, then $q^{e*}_{t^{e*}} < q^e_{t^e}$ if and only if $\gamma < \gamma^*$.

Part (1) and (2) follow from the characterization of the attention allocation in Appendix A.1 ($\mathbf{q}^{e*}$ is obtained by replacing $(\alpha, \gamma, \hat{u}_{aa})$ with $(\alpha^*, \gamma^*, \hat{u}_{aa} + \hat{u}_{\sigma\sigma})$). Intuitively, when $|\nu^*| > |\nu|$ or $\hat{u}_{\sigma\sigma} < 0$ the first-best use of information involves a stronger incentive to acquire information, either to respond more to the state or to reduce dispersion. This stronger demand would render entry profitable for more sources. Part (3) follows from Proposition 2 and can be illustrated by Case 1 in Corollary 2. The case of $t^e = t^{e*}$ follows from applying equation (7) when only q_{t^e} and $q_{t^{e*}}$ are positive and by noting that $m(\gamma) = m(\gamma^*)$ if $t^e = t^{e*}$. Intuitively, when all sources are of the same type, a stronger coordination motive induces the consumers to let their action depend more on their common prior and less on their signals. This renders information less valuable, depressing its demand and so its supply.

[23] Another standard assumption is that $\hat{u}_{aa} + 2\hat{u}_{a\overline{a}} + \hat{u}_{\overline{a}\overline{a}} < 0$ and $\hat{u}_{aa} + \hat{u}_{\sigma\sigma} < 0$, which ensure that the first-best attention allocation is unique and bounded (see Pavan (2014)).

Supply-side Policies Tackling Demand Distortions

A distorted demand can add inefficiencies to the competitive equilibrium—even when only the clearest type of sources is supplied. This raises the question of whether policies aimed at the supply of information (i.e., $\mathbf{q}$) can improve welfare. For instance, subsidies or quotas may be used to redirect entrants to produce specific signal types.[24] To provide an answer, we characterize how marginal changes in $\mathbf{q}$ affect the consumers' expected payoff under a distorted demand, denoted by V^d. We focus on the case of linear attention cost for tractability.

We consider each kind of distortion separately.

Proposition 6. *Suppose the demand exhibits only CIE distortions (i.e., $\gamma = \gamma^*$ and $\hat{u}_{\sigma\sigma} = 0$) and t-sources can receive attention at $\mathbf{q}$ (i.e., $K(\mathbf{q}) > \xi_t$). If $\nu < \nu^*$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$. If $\nu > \nu^*$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$ if and only if t-sources are sufficiently clear.*

Intuitively, if the planner would like the consumers to respond more to the state (i.e., $\nu^* > \nu$), adding more sources of any type is always useful. This is because the consumers will spread their attention more thinly across the larger number of sources, that is, they will pay less attention to all sources. This in turn decreases the publicity of each signal and renders the signals from each source less correlated between consumers. As a result, it becomes harder for the consumers to coordinate their actions, which incentivizes them to use their information to respond more to the state. Put differently, the consumers benefit from an overload of their attention channel. If instead the planner would like the consumers to respond less to the state (i.e., $\nu^* < \nu$), then only adding sufficiently clear sources improves welfare. Such sources facilitate the extraction of information for each unit of attention and thus help the consumers to coordinate their action to the others' actions, therefore responding less to the state.

Proposition 7. *Suppose the demand exhibits only SOC distortions (i.e., $\alpha = \alpha^*$ and $\hat{u}_{\sigma\sigma} = 0$) and t-sources can receive attention at $\mathbf{q}$ (i.e., $K(\mathbf{q}) > \xi_t$). If $\gamma < \gamma^*$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$ if and only if t-sources are sufficiently clear. If $\gamma > \gamma^*$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$ if and only if t-sources are sufficiently unclear.*

Intuitively, if the planner cares more about coordination than the consumers do ($\gamma^* > \gamma$), adding more sources that are sufficiently clear and possibly eliminate sources that are sufficiently unclear helps because clearer sources allow the consumers to better coordinate with each other. The intuition is reversed for the case of $\gamma > \gamma^*$.

[24] Angeletos and Pavan (2009) study policies that can remove distortions in the demand for information.

Proposition 8. *Suppose the demand exhibits only SAD distortions (i.e., $\alpha = \alpha^*$ and $\gamma = \gamma^*$) and t-sources can receive attention at $\mathbf{q}$ (i.e., $K(\mathbf{q}) > \xi_t$). There exists $\overline{u}_{\sigma\sigma} > 0$ with the following property. If $\hat{u}_{\sigma\sigma} < \overline{u}_{\sigma\sigma}$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$ if and only if t-sources are sufficiently clear. If $u_{\sigma\sigma} > \overline{u}_{\sigma\sigma}$, then $\frac{\partial V^d(\mathbf{q})}{\partial q_t} > 0$ if and only if t-sources are sufficiently unclear.*

Intuitively, if dispersion harms consumer welfare (or does not benefit it enough), then adding more sources that are sufficiently clear and possibly eliminating sources that are sufficiently unclear helps because clearer sources result in more correlated information among consumers, leading to more correlated actions and so less dispersion. The opposite holds when more dispersion benefits welfare.

Propositions 6–8 only consider local interventions involving a single type of sources. Nonetheless, they suggest how supply-side policies may improve the welfare of information consumers. One thing to keep in mind is that, starting from $\mathbf{q}^e$, promoting the choice of t-signals may require interventions in support of other types of sources already in the market. This is because the new t-sources will steal attention from the existing sources, thereby possibly pushing some out of business. If instead welfare-improving policies call for reducing the quantity of the equilibrium type of sources, doing so also increases their profits.

It is worth noting an important difference between markets with distorted and undistorted demand for information. In the latter case, the consumers' marginal value for one extra entrant is positive independently of its type and the existing supply (see the proof of Proposition 3). By contrast, that marginal value can be negative under distorted demand. Of course, we know that in strategic settings more information does not always lead to better outcomes. Our results go beyond this point by identifying which type of information improves outcomes.

6 Concluding Remarks

Information sources serve two roles: They allow agents to learn about relevant events as well as about others' knowledge of those events, which is essential for responding to strategic externalities in social and economic contexts. We show how these two roles affect information markets when suppliers compete for consumers' attention. Competition pushes towards a homogenization of information—not in terms of topics, but in terms of how accurately and clearly each topic is delivered. This is driven by a subtle interaction between the consumers' desire to coordinate and the cost of producing information. Also, information markets can be inefficient when attention becomes the "currency" whereby consumers pay for information. This cause of inefficiency is distinct from distortions in the demand for information and suppliers' market power.

We close with two remarks. First, to focus on the consequences of competition for attention, we intentionally ruled out price competition. Though in many settings it seems optimal for information providers to set zero prices, this is not always the case. Our results can be seen as a first pass in understanding how the dual role of information sources affects market outcomes. It would be interesting to examine the trade-off between attention and price revenues, its solution under competition, and whether the price mechanism can restore efficiency in information markets. We conjecture that as long as attention provides a share of the revenues—which is often the case in reality—the inefficiencies we found should persist to some degree.

Second, our model can be extended in several ways to capture other aspects of information markets. For instance, in reality consumers are heterogeneous—in their preferred actions, conformism, topics of interest, and cost of acquiring information. Some of these extensions do not change the thrust of our analysis. Dewan and Myatt (2008, 2012) point out that the state θ can simply represent the average of the agents' ideal actions. Angeletos and Pavan (2009) allow the state to have private and common components. Different topics and social groups sometimes give rise to distinct markets, where each can be studied using our model. Other extensions are more intricate. Different coordination weights or attention costs in the same market have deep consequences on the demand for information, which remain unexplored. Either way, consumer heterogeneity is likely to result in a heterogeneous information supply. This should not be surprising. The point remains that, within market segments, competition for attention creates a force towards homogeneity and inefficiencies.

A Appendix

A.1 Proof of Lemma 1

We present the proof of Lemma 1 following Myatt and Wallace (2012) and Pavan (2014). We do this to make the paper self contained and because we use this characterization in the rest of the analysis. Also, we adopt the more general payoff function

$$\hat{u}(a_\ell, \overline{a}, \theta, \sigma) - \hat{c}\left(\sum_i z_{i\ell}\right),$$

where $\hat{u}$ satisfies the following properties (Angeletos and Pavan (2007)): (i) $\hat{u}$ is a second-order polynomial, (ii) $\sigma^2 \equiv \int_0^1 [a_\ell - \overline{a}]^2 d\ell$ is the action *dispersion* and has only a non-strategic externality effect, so in terms of partial derivatives $\hat{u}_{a\sigma} = \hat{u}_{\overline{a}\sigma} = \hat{u}_{\theta\sigma} = 0$ and $\hat{u}_\sigma(a, \overline{a}, 0, \theta) = 0$ for all $(a, \overline{a}, \theta)$, (iii) $\hat{u}_{aa} < 0$, (iv) $-\hat{u}_{a\overline{a}}/\hat{u}_{aa} < 1$, and (v) $\hat{u}_{a\theta} \neq 0$. By (i), the agents have linear best responses. By (ii), $\hat{u}$ is additively separable in σ^2, whose coefficient will be denoted by $\hat{u}_{\sigma\sigma}/2$. By (iv), the slope of best-response functions is less than 1, delivering uniqueness of equilibrium actions (Angeletos and Pavan (2007)). These assumptions allow for rich payoff externalities. Note that the function u in (1) is a special case of this class of functions, because

$$u(a_\ell, \overline{a}, \sigma, \theta) = -a_\ell^2 + 2\gamma a_\ell \overline{a} - \gamma \overline{a}^2 + 2(1-\gamma) a_\ell \theta - (1-\gamma)\theta^2.$$

As Lemma 3 below shows, $\hat{u}$ and u are strategically equivalent. They are, however, not equivalent for welfare analysis (Angeletos and Pavan (2007); Pavan (2014)), which matters for our analysis in Section 5.

Fix $\mathbf{q} = (q_1, \ldots, q_T)$. Enumerate the sources so that i runs from $i = 0$ (i.e., the prior) to $i = n = \sum_{t=1}^T q_t$. By linearity of A_ℓ, we can write $A_\ell(x_\ell) = \sum_{i=0}^n w_{i\ell} x_{i\ell}$ for some vector of weights $w_\ell \in \mathcal{R}^n$. We can view the prior as a signal with accuracy $1/\kappa_0^2$, infinite precision (i.e., $\xi_0 = 0$), and with $x_{0\ell} = 0$—namely, the prior mean. Using this, Myatt and Wallace (2012) and Pavan (2014) show the following (see their Proposition 1).

Lemma 3. *Define* $\gamma = -\hat{u}_{a\overline{a}}/\hat{u}_{aa}$, $\nu = -\frac{\hat{u}_{a\theta}}{\hat{u}_{aa} + \hat{u}_{a\overline{a}}}$, *and* $c(z_\ell) = \frac{2}{\nu^2 |\hat{u}_{aa}|} \hat{c}(z_\ell)$. *In the unique linear symmetric equilibrium, the influence* w_i *of the* i*-th source and the attention* z_i *paid to it satisfy*

$$w_i = \frac{\psi_i}{\sum_{j=0}^n \psi_j} \qquad \text{and} \qquad z_i = \frac{\xi_i w_i}{\sqrt{c'(\sum_{j=0}^n z_j)}} \qquad \text{and} \qquad \psi_i = \frac{1}{(1-\gamma)\kappa_i^2 + \xi_i^2/z_i}, \tag{9}$$

and where $\psi_i = 0$ *for any source which is ignored (so that* $z_i = w_i = 0$*).*

Since the prior "signal" has infinite precision, no consumer needs to allocate positive attention to it ($z_0 = 0$), yet it receives positive weight ($w_0 > 0$) as in standard updating from normal signals.

Using equation (9) and substituting, we obtain that for $i \neq 0$ such that $z_i > 0$

$$z_i = \frac{\xi_i (K - \xi_i)}{(1-\gamma)\kappa_i^2}, \qquad \text{where } K \equiv \frac{1}{\sqrt{c'(\sum_{j=0}^n z_j)} \left[\sum_{j=0}^n \psi_j\right]}. \tag{10}$$

Substituting in the expression for ψ_i in equation (9), we get

$$\begin{aligned}\sum_{j=0}^{n}\psi_j &= \frac{1}{(1-\gamma)\kappa_0^2}+\sum_{j=1}^{n}\frac{1}{(1-\gamma)\kappa_j^2+\xi_j^2/z_j}\\ &= \frac{1}{(1-\gamma)\kappa_0^2}+\sum_{j:z_j>0}\frac{1}{(1-\gamma)\kappa_j^2+\xi_j^2\frac{(1-\gamma)\kappa_j^2}{\xi_j(K-\xi_j)}}\\ &= \frac{1}{(1-\gamma)\kappa_0^2}+\frac{1}{(1-\gamma)K}\sum_{j=1}^{n}\frac{\max\{K-\xi_j,0\}}{\kappa_j^2}.\end{aligned}$$

Using this in the expression for K in (10) and rearranging implies that K has to satisfy

$$c'\left(\sum_{j=1}^{n}\frac{\xi_j\max\{K-\xi_j,0\}}{(1-\gamma)\kappa_j^2}\right)\left[\frac{K}{(1-\gamma)\kappa_0^2}+\sum_{j=1}^{n}\frac{\max\{K-\xi_j,0\}}{(1-\gamma)\kappa_j^2}\right]^2=1.$$

By convexity of c, the left-hand side of this equation is strictly increasing in K, which implies that there exists a unique solution K^*. Since q_t of the n sources are identical for every $t=1,\ldots,T$, we can rewrite the last condition as

$$c'\left(\sum_{t=1}^{T}q_t\frac{\xi_t\max\{K-\xi_t,0\}}{(1-\gamma)\kappa_t^2}\right)\left[\frac{K}{(1-\gamma)\kappa_0^2}+\sum_{t=1}^{T}q_t\frac{\max\{K-\xi_t,0\}}{(1-\gamma)\kappa_t^2}\right]^2=1. \qquad (11)$$

It is immediate that every t-source receives the same amount of attention given by

$$z_t=\frac{\xi_t\max\{K^*-\xi_t,0\}}{(1-\gamma)\kappa_t^2}.$$

Consider now the properties of K^*. First, the left-hand side of (11) is strictly increasing in γ and in $1/\kappa_0^2$, so K^* is strictly decreasing in γ and in $1/\kappa_0^2$. Second, the left-hand side of (11) is increasing in q_t and in $1/\kappa_t^2$ (strictly if $z_t>0$), so K^* is decreasing in q_t and in $1/\kappa_t^2$ (strictly if $z_t>0$). Clearly, K^* is non-monotonic in ξ_t for any t. Finally, since K^* is bounded below by zero, it must converge to some $\underline{K}^*$ as $\kappa_0^2\to 0$. It is immediate to see that $\underline{K}^*=0$.

A.2 Proof of Proposition 1

Suppose $\mathbf{q}^e$ is an equilibrium, $t\notin T(\gamma)$, and $q_t^e>0$. Then, by (4), it must be that $K(\mathbf{q}^e)=m(\gamma,t)>m(\gamma,t')$ for some $t'\in T(\gamma)$. Since $K(\mathbf{q})$ is a continuous function, there exists $q_{t'}>q_{t'}^e$ such that $K(q_{t'},\mathbf{q}_{-t'}^e)>m(\gamma,t')$, which violates (5). This contradiction proves property *(i)*.

We now prove that an equilibrium exists. For every $t\in T(\gamma)$, $m(\gamma,t)>\xi_t$ and so

$z_t > 0$ provided that $K(\mathbf{q}) \geq m(\gamma)$. Consider any vector $\mathbf{q}$ that satisfies $q_{t'} = 0$ if $t' \notin T(\gamma)$ and

$$c'\left(\sum_{t\in T(\gamma)} q_t \frac{\xi_t(m(\gamma)-\xi_t)}{(1-\gamma)\kappa_t^2}\right)\left[\frac{m(\gamma)}{(1-\gamma)\kappa_0^2} + \sum_{t\in T(\gamma)} q_t \frac{(m(\gamma)-\xi_t)}{(1-\gamma)\kappa_t^2}\right]^2 = 1,$$

which implies $K(\mathbf{q}) = m(\gamma)$ by construction. For all $t \in T(\gamma)$, condition (4) holds and condition (5) holds because K is strictly decreasing when $z_t > 0$. For $t' \notin T(\gamma)$, condition (5) holds because $K(\mathbf{q})$ is decreasing and $m(\gamma, t') > K(\mathbf{q})$. Thus, we have an equilibrium. Using the expression of $m(\gamma, t) = m(\gamma)$ for every $t \in T(\gamma)$, we can simplify the last expression to

$$c'\left(\sum_{t\in T(\gamma)} q_t H(\kappa_t, \xi_t))\right)\left[\frac{m(\gamma)}{(1-\gamma)\kappa_0^2} + \sum_{t\in T(\gamma)} q_t \frac{H(\kappa_t, \xi_t))}{\xi_t}\right]^2 = 1.$$

By Assumption 1, any vector $(q_t)_{t\in T(\gamma)}$ solving this equation must contain some positive entry. This proves property *(ii)*.

To see property *(iii)*, suppose that $T(\gamma) = \{t^e\}$ and $q_t = 0$ for all $t \neq t^e$. Note that t^e-sources receive strictly positive attention for all $K(q_{t^e}, \mathbf{0}) \geq m(\gamma)$, where $K(q_{t^e}, \mathbf{0})$ is the unique solution to the condition

$$c'\left(q_{t^e}\frac{\xi_{t^e}(K-\xi_{t^e})}{(1-\gamma)\kappa_{t^e}^2}\right)\left[\frac{K}{(1-\gamma)\kappa_0^2} + q_{t^e}\frac{K-\xi_{t^e}}{(1-\gamma)\kappa_{t^e}^2}\right]^2 = 1.$$

Since $K(q_{t^e}, \mathbf{0})$ is strictly decreasing in q_{t^e} and $\lim_{q_{t^e}\to+\infty} K(q_{t^e}, \mathbf{0}) = \xi_{t^e}$, there exists a unique $q^e_{t^e}$ for which $K(q^e_{t^e}, \mathbf{0}) = m(\gamma)$. Using $m(\gamma) = \xi_{t^e} + (1-\gamma)\frac{\kappa^2_{t^e}}{\xi_{t^e}} H(\kappa_{t^e}, \xi_{t^e})$, $q^e_{t^e}$ has to satisfy

$$c'\left(q^e_{t^e} H(\kappa_{t^e}, \xi_{t^e})\right)\left[\frac{\xi_{t^e}}{(1-\gamma)\kappa_0^2} + \left(\frac{\kappa_{t^e}^2}{\kappa_0^2} + q^e_{t^e}\right)\frac{H(\kappa_{t^e}, \xi_{t^e})}{\xi_{t^e}}\right]^2 = 1.$$

A.3 Proof of Proposition 2 and Corollary 2

Label types so that $t < t'$ implies $\hat{H}(\xi_t) < \hat{H}(\xi_{t'})$ and let $\overline{t}^e(\gamma)$ and $\underline{t}^e(\gamma)$ be the highest and lowest possible equilibrium types for every γ. Recall that generically $\overline{t}^e(\gamma) = \underline{t}^e(\gamma)$. Note that

$$m(\gamma, t) = \xi_t + (1-\gamma)\hat{H}(\xi_t).$$

Now,

$$T(\gamma) = \arg\min_{t\in\{1,\ldots,T\}} m(\gamma, t) = \arg\max_{t\in\{1,\ldots,T\}} \{(\gamma-1)\hat{H}(\xi_t) - \xi_t\}.$$

It follows by standard monotone comparative statics that the set $T(\gamma)$ is increasing in the strong set order. Therefore, as γ increases, $\bar{t}(\gamma) = \max T(\gamma)$ and $\underline{t}(\gamma) = \min T(\gamma)$ also increase. Recall that, generically, $T(\gamma) = \{t\}$. Therefore, by linearity of the objective function defining $T(\gamma)$, there exists a connected interval $(\underline{\gamma}_t, \overline{\gamma}_t) \subseteq (-\infty, 1)$ such that $T(\gamma) = \{t\}$ for all $\gamma \in (\underline{\gamma}_t, \overline{\gamma}_t)$. Note that some type t may never belong to $T(\gamma)$, which we can let $\underline{\gamma}_t = \overline{\gamma}_t$ to that $(\underline{\gamma}_t, \overline{\gamma}_t) = \varnothing$. By the monotonicity of $\min T(\gamma)$ and $\max T(\gamma)$, we must have that $\overline{\gamma}_{t-1} = \underline{\gamma}_t$ and $\overline{\gamma}_t = \underline{\gamma}_{t+1}$ for $t = 2, \ldots, T-1$.

Now suppose that type $\hat{t}$ satisfies $\hat{H}(\xi_{\hat{t}}) < \hat{H}(\xi_t)$ and $\xi_{\hat{t}} < \xi_t$ for all $t \neq \hat{t}$. It follows that, for every $t \neq \hat{t}$, we have $m(\gamma, \hat{t}) < m(\gamma, t)$ for every γ. Therefore, $T(\gamma) = \{\hat{t}\}$ for all γ. Conversely, suppose that if $\hat{t}$ satisfies $\xi_{\hat{t}} < \xi_t$ for all $t \neq \hat{t}$, then $\hat{H}(\xi_{\hat{t}}) > \hat{H}(\xi_{t'})$ for some t'. Then, $m(1, \hat{t}) < m(1, t)$ for all $t \neq \hat{t}$, which implies that $T(\gamma) = \{\hat{t}\}$ for γ sufficiently close to 1. However, there exists γ sufficiently small such that $m(\gamma, \hat{t}) > m(\gamma, t')$, which implies that $\hat{t} \notin T(\gamma)$. This shows that the equilibrium type depends on γ.

Now, for every γ let $\overline{\xi}^e(\gamma)$ and $\underline{\xi}^e(\gamma)$ correspond to the lowest and highest possible clarity of the type of sources offered in equilibrium. Generically, $\overline{\xi}^e(\gamma) = \underline{\xi}^e(\gamma)$. Suppose $\xi_t > \xi_{t'}$ implies $\hat{H}(\xi_t) < \hat{H}(\xi_{t'})$. By our convention, it must then be that $t < t'$. The monotonicity properties of $\overline{t}^e(\gamma)$ and $\underline{t}^e(\gamma)$ are inherited by $\overline{\xi}^e(\gamma)$ and $\underline{\xi}^e(\gamma)$. Clearly, $m(1, t) > m(1, t')$ if and only if type t is less clear than type t'. Thus, $\underline{\gamma}_T < 1$ and only the clearest type is supplied for γ sufficiently large. Conversely, $\overline{\gamma}_1 > -\infty$ because $\lim_{\gamma \to -\infty} m(\gamma, 1) > \lim_{\gamma \to -\infty} m(\gamma, t)$ for all $t > 1$.

A.4 Calculation of V^d, V, and their Derivatives

The parameters used here are defined in Section 5.2. Note that when $\nu = \nu^*$, $\gamma = \gamma^*$, and $\hat{u}_{\sigma\sigma} = 0$, the expression of V^d below equals the expression of V in the main text.

Lemma 4. *Fix $\mathbf{q} = (q_1, \ldots, q_T)$. Let $Z(\mathbf{q}) = \sum_{t=1}^T q_t z_t(\mathbf{q})$. Each consumer's expected surplus equals (up to a constant and scalar multiplication) to*

$$\begin{aligned} V^d(\mathbf{q}) \;=\; & -K(\mathbf{q})\sqrt{c'\left(Z(\mathbf{q})\right)} - c\left(Z(\mathbf{q})\right) \\ & - \frac{1}{(1-\gamma)^2} c'\left(Z(\mathbf{q})\right)(\gamma - \gamma^*) \left[\frac{[K(\mathbf{q})]^2}{\kappa_0^2} + \sum_{t=1}^T q_t \frac{[\max\{K(\mathbf{q}) - \xi_t, 0\}]^2}{\kappa_t^2} \right] \\ & + \frac{1}{(1-\gamma)^2} c'\left(Z(\mathbf{q})\right) \left(\frac{\nu - \nu^*}{\nu} \right) \frac{2(1-\gamma^*)[K(\mathbf{q})]^2}{\kappa_0^2} \\ & + \frac{\hat{u}_{\sigma\sigma}}{|\hat{u}_{aa} + \hat{u}_{\sigma\sigma}|} c\left(Z(\mathbf{q})\right). \end{aligned}$$

Proof. Consider first the model with finite, integer, numbers of sources for $t = 1, \ldots, T$.

Following Angeletos and Pavan (2007) and Pavan (2014), the consumers' expected welfare under the equilibrium strategy $A(\cdot;\mathbf{q})$ and given attention allocation $z(\mathbf{q})$ is

$$V^d(\mathbf{q}) \equiv \mathbb{E}[\hat{u}(a,\overline{a},\sigma,\theta)|A(\cdot),z(\mathbf{q}),\mathbf{q}] = \mathbb{E}[W(\alpha,0,\theta)] + \mathcal{L}(\mathbf{q}) - \hat{c}(Z(\mathbf{q})),$$

where $W(\overline{a},0,\theta) \equiv \hat{u}(\overline{a},\overline{a},0,\theta)$ and

$$\begin{aligned}\mathcal{L}(\mathbf{q}) \;=\; & \frac{\hat{u}_{aa}+2\hat{u}_{a\overline{a}}+\hat{u}_{\overline{a}\overline{a}}}{2}Var[\overline{a}-\alpha|A(\cdot;\mathbf{q}),z(\mathbf{q}),\mathbf{q}] \\ & +\frac{\hat{u}_{aa}+\hat{u}_{\sigma\sigma}}{2}Var[a-\overline{a}|A(\cdot;\mathbf{q}),z(\mathbf{q}),\mathbf{q}] \\ & -Cov[\overline{a}-\alpha,W_{\overline{a}}(\alpha,0,\theta)|A(\cdot;\mathbf{q}),z(\mathbf{q}),\mathbf{q}].\end{aligned}$$

Following Angeletos and Pavan (2007) and ignoring the constant $\mathbb{E}[W(\alpha,0,\theta)]$, we have

$$\begin{aligned}V^d(\mathbf{q}) \;=\; & \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}\sum_{t=1}^{T} q_t\frac{[w_t(\mathbf{q})]^2\xi_t^2}{z_t(\mathbf{q})} \\ & +\frac{(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}\left\{\kappa_0^2[w_0(\mathbf{q})]^2+\sum_{t=1}^{T} q_t[w_t(\mathbf{q})]^2\kappa_t^2\right\} \\ & -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2\left(\frac{\nu-\nu^*}{\nu}\right)\kappa_0^2[w_0(\mathbf{q})]^2-\hat{c}\left(Z(\mathbf{q})\right).\end{aligned}$$

Using $w_0(\mathbf{q}) = 1-\sum_{t=1}^{T} q_t w_t(\mathbf{q})$,

$$w_t(\mathbf{q}) = \frac{z_t(\mathbf{q})}{\xi_t}\sqrt{c'\left(Z(\mathbf{q})\right)},$$

and

$$\sqrt{c'\left(Z(\mathbf{q})\right)} = \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}+\sum_{t'=1}^{T} q_{t'}\frac{z_{t'}(\mathbf{q})}{\xi_t'}\right]^{-1},$$

we get

$$\begin{aligned}V^d(\mathbf{q}) \;=\; & \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}c'\left(Z(\mathbf{q})\right)Z(\mathbf{q})-\hat{c}\left(Z(\mathbf{q})\right) \\ & +\frac{(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}\left\{\kappa_0^2\left[1-\frac{\sum_{t=1}^{T} q_t\frac{z_t(\mathbf{q})}{\xi_t}}{\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}+\sum_{t'=1}^{T} q_{t'}\frac{z_{t'}(\mathbf{q})}{\xi_t'}}\right]^2\right\} \\ & +\frac{(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}\left\{c'\left(Z(\mathbf{q})\right)\sum_{t=1}^{T} q_t\frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2\right\} \\ & -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2\left(\frac{\nu-\nu^*}{\nu}\right)\kappa_0^2\left[1-\frac{\sum_{t=1}^{T} q_t\frac{z_t(\mathbf{q})}{\xi_t}}{\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}+\sum_{t'=1}^{T} q_{t'}\frac{z_{t'}(\mathbf{q})}{\xi_t'}}\right]^2\end{aligned}$$

$$
\begin{aligned}
&= \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \sum_{t=1}^{T} Z(\mathbf{q}) - \hat{c}(Z(\mathbf{q})) \\
&\quad +(1-\gamma)\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \left\{ \kappa_0^2 \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 + \sum_{t=1}^{T} q_t \frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2 \right\} \\
&\quad +(\gamma-\gamma^*)\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \left\{ \kappa_0^2 \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 + \sum_{t=1}^{T} q_t \frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2 \right\} \\
&\quad -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2 \left(\frac{\nu-\nu^*}{\nu}\right) \kappa_0^2 c' \left(\sum_{t=1}^{T} q_t z_t(\mathbf{q})\right) \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 \\
&= \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \left\{ \frac{[K(\mathbf{q})]^2}{(1-\gamma)\kappa_0^2} + \sum_{t=1}^{T} q_t \frac{z_t(\mathbf{q})}{\xi_t} \left[(1-\gamma)\kappa_t^2 \frac{z_t(\mathbf{q})}{\xi_t} + \xi_t\right] \right\} \\
&\quad +(\gamma-\gamma^*)\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \left\{ \kappa_0^2 \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 + \sum_{t=1}^{T} q_t \frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2 \right\} \\
&\quad -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2 \left(\frac{\nu-\nu^*}{\nu}\right) \kappa_0^2 c'(Z(\mathbf{q})) \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 - \hat{c}(Z(\mathbf{q})).
\end{aligned}
$$

Now, note that

$$
\begin{aligned}
\sum_{t=1}^{T} q_t \frac{z_t(\mathbf{q})}{\xi_t} \left[(1-\gamma)\kappa_t^2 \frac{z_t(\mathbf{q})}{\xi_t} + \xi_t\right] &= \sum_{t=1}^{T} q_t \frac{\max\{K(\mathbf{q})-\xi_t, 0\}}{(1-\gamma)\kappa_t^2} \left[\max\{K(\mathbf{q})-\xi_t, 0\} + \xi_t\right] \\
&= K(\mathbf{q}) \sum_{t=1}^{T} q_t \frac{\max\{K(\mathbf{q})-\xi_t, 0\}}{(1-\gamma)\kappa_t^2} \\
&= K(\mathbf{q}) \left[\frac{1}{\sqrt{c'(Z(\mathbf{q}))}} - \frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right].
\end{aligned}
$$

where the last equality uses (11). Therefore,

$$
\begin{aligned}
V^d(\mathbf{q}) &= \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} K(\mathbf{q}) \sqrt{c'(Z(\mathbf{q}))} - \hat{c}(Z(\mathbf{q})) \\
&\quad +(\gamma-\gamma^*)\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c'(Z(\mathbf{q})) \left\{ \kappa_0^2 \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 + \sum_{t=1}^{T} q_t \frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2 \right\} \\
&\quad -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2 \left(\frac{\nu-\nu^*}{\nu}\right) \kappa_0^2 c'(Z(\mathbf{q})) \left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2.
\end{aligned}
$$

Now recall that $\hat{c}(\cdot) = \frac{|\hat{u}_{aa}|\nu^2}{2} c(\cdot)$. Thus, by summing and subtracting $\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} c(Z(\mathbf{q}))$, we get

$$
V^d(\mathbf{q}) = \frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2} \left\{ K(\mathbf{q}) \sqrt{c'(Z(\mathbf{q}))} + c(Z(\mathbf{q})) \right\}
$$

$$
\begin{aligned}
& +(\gamma-\gamma^*)\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}c'\left(Z(\mathbf{q})\right)\left\{\kappa_0^2\left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2+\sum_{t=1}^{T}q_t\frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2\right\} \\
& -(1-\gamma^*)(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2\left(\frac{\nu-\nu^*}{\nu}\right)\kappa_0^2c'\left(Z(\mathbf{q})\right)\left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 \\
& -\hat{u}_{\sigma\sigma}\frac{\nu^2}{2}c\left(Z(\mathbf{q})\right) \\
= & -\frac{(\hat{u}_{aa}+\hat{u}_{\sigma\sigma})\nu^2}{2}\left\{-K(\mathbf{q})\sqrt{c'\left(Z(\mathbf{q})\right)}-c\left(Z(\mathbf{q})\right)\right. \\
& -(\gamma-\gamma^*)c'\left(Z(\mathbf{q})\right)\left\{\kappa_0^2\left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2+\sum_{t=1}^{T}q_t\frac{z_t^2(\mathbf{q})}{\xi_t^2}\kappa_t^2\right\} \\
& +2(1-\gamma^*)\left(\frac{\nu-\nu^*}{\nu}\right)\kappa_0^2c'\left(Z(\mathbf{q})\right)\left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}\right]^2 \\
& \left.+\frac{\hat{u}_{\sigma\sigma}}{\hat{u}_{aa}+\hat{u}_{\sigma\sigma}}c\left(Z(\mathbf{q})\right)\right\}.
\end{aligned}
$$

Thus, we proved the result for every vector $\mathbf{q}$ of non-negative integers. As argued in Appendix B, it is possible to allow q_t to be any non-negative real number for all t without changing the meaning of the analysis. □

We calculate the derivative of $V(\mathbf{q})$ with respect to q_t assuming that the acquisition and use of information is efficient and t-sources receive positive attention (i.e., $K(\mathbf{q}) > \xi_t$).

Lemma 5. *Fix* $\mathbf{q}$. *Suppose that* $\gamma=\gamma^*$, $\nu=\nu^*$, *and* $\hat{u}_{\sigma\sigma}=0$. *If* $K(\mathbf{q})>\xi_t$, *then*

$$
\frac{\partial V(\mathbf{q})}{\partial q_t}=c'\left(Z(\mathbf{q})\right)\frac{[K(\mathbf{q})-\xi_t]^2}{(1-\gamma)\kappa_t^2}.
$$

Proof. Fix $\mathbf{q}$ and the corresponding threshold $K(\mathbf{q})$ in Lemma 1. Let $T^+(\mathbf{q})=\{t: K(\mathbf{q})>\xi_t\}$, that is, this is the set of types that can receive positive attention at $\mathbf{q}$. Suppose first that t satisfies $q_t>0$ and so $z_t(\mathbf{q})>0$. We start by showing that $z_t(\mathbf{q})$ is continuously differentiable in a neighborhood $(q_t-\delta, q_t+\delta)$. Recall that

$$
z_t(q_t,\mathbf{q}_{-t})=\frac{\xi_t[K(q_t,\mathbf{q}_{-t})-\xi_t]}{(1-\gamma)\kappa_t^2},
$$

so we need to show that $K(q_t,\mathbf{q}_{-t})$ is continuously differentiable in a neighborhood $(q_t-\delta, q_t+\delta)$. From the proof of Lemma 1, recall that $K(q_t,\mathbf{q}_{-t})$ is the unique solution to

$$
1 \;=\; c'\left(\sum_{t'\in T^+(\mathbf{q}):t'\neq t,q_t>0}q_{t'}\frac{\xi_{t'}(K-\xi_{t'})}{(1-\gamma)\kappa_{t'}^2}+q_t\frac{\xi_t(K-\xi_t)}{(1-\gamma)\kappa_t^2}\right)
$$

$$\times \left[\frac{K}{(1-\gamma)\kappa_0^2} + \sum_{t' \in T^+(\mathbf{q}): t' \neq t, q_t > 0} q_{t'} \frac{K - \xi_{t'}}{(1-\gamma)\kappa_{t'}^2} + q_t \frac{K - \xi_t}{(1-\gamma)\kappa_t^2} \right]^2 .$$

By the Implicit Function Theorem, there exists a neighborhood $(q_t - \delta, q_t + \delta)$ such that the function $K(\cdot, \mathbf{q}_{-t})$ is continuously differentiable. For future reference, in the case of linear attention cost, we have

$$\frac{\partial K(q_t, \mathbf{q}_{-t})}{\partial q_t} = -\frac{K(q_t, \mathbf{q}_{-t}) - \xi_t}{(1-\gamma)\kappa_t^2} \left[\frac{1}{(1-\gamma)\kappa_0^2} + \sum_{t'=1}^{T} \frac{q_{t'} \mathbb{I}\{K(\mathbf{q}) > \xi_{t'}\}}{(1-\gamma)\kappa_{t'}^2} \right]^{-1} . \tag{12}$$

Now suppose that $t \in T^+(\mathbf{q})$, but $q_t = 0$. Recall that $K(q_t, \mathbf{q}_{-t})$ is continuous, so there exists a neighborhood $(0, \delta)$ of q_t such that $T^+(q_t, \mathbf{q}_{-t}) = T^+(\mathbf{q})$ for all $q_t \in (0, \delta)$. Given this, a similar argument to the one before shows that there exists a neighborhood $(0, \delta')$ of $q_t = 0$ such that $K(\cdot, \mathbf{q}_{-t})$ is continuously differentiable on it.

Now consider $V(\mathbf{q})$. From Lemma 1, recall that

$$w_t(q_t, \mathbf{q}_{-t}) = \frac{z_t(q_t, \mathbf{q}_{-t})}{\xi_t} \sqrt{c'(Z(\mathbf{q}))}. \tag{13}$$

Therefore, $w_t(q_t, \mathbf{q}_{-t})$ inherits the differentiability properties of $z_t(q_t, \mathbf{q}_{-t})$ derived before. It follows that $V(q_t, \mathbf{q}_{-t})$ is continuously differentiable with respect to $q_t > 0$ in a neighborhood $(q_t - \delta, q_t + \delta)$ and with respect to $q_{t'}$ in a neighborhood $(0, \delta')$ of $q_{t'} = 0$ for every $t \in T^+(\mathbf{q})$.

We can now invoke the Envelope theorem to calculate the derivative of $V(\mathbf{q})$ on these neighborhoods. Suppose first that t satisfies $q_t > 0$. Over the neighborhood $(q_t - \delta, q_t + \delta)$ we have

$$\begin{aligned} \frac{\partial V(q_t, \mathbf{q}_{-t})}{\partial q_t} &= 2 w_t(q_t, \mathbf{q}_{-t}) \,(1-\gamma)\, \kappa_0^2 \left[1 - \sum_{t'=1}^{T} q_{t'} w_{t'}(q_t, \mathbf{q}_{-t}) \right] \\ &\quad - [w_t(q_t, \mathbf{q}_{-t})]^2 \left[(1-\gamma)\kappa_t^2 + \frac{\xi_t^2}{z_t(q_t, \mathbf{q}_{-t})} \right] - c'(Z(\mathbf{q})) z_t(q_t, \mathbf{q}_{-t}). \end{aligned}$$

Now recall (13) to substitute and express $\frac{\partial V(q_t, \mathbf{q}_{-t})}{\partial q_t}$ as equal to

$$\begin{aligned} & 2 \frac{z_t(q_t, \mathbf{q}_{-t})}{\xi_t} c'(Z(\mathbf{q})) \,(1-\gamma)\, \kappa_0^2 \left[\frac{1}{\sqrt{c'(Z(\mathbf{q}))}} - \sum_{t'=1}^{T} q_{t'} \frac{z_{t'}(q_t, \mathbf{q}_{-t})}{\xi_{t'}} \right] \\ & - c'(Z(\mathbf{q})) \frac{z_t(q_t, \mathbf{q}_{-t})}{\xi_t} \left[(1-\gamma)\kappa_t^2 \frac{z_t(q_t, \mathbf{q}_{-t})}{\xi_t} + \xi_t \right] - c'(Z(\mathbf{q})) z_t(q_t, \mathbf{q}_t). \end{aligned}$$

Using

$$\frac{1}{\sqrt{c'(Z(\mathbf{q}))}} = \frac{K(q_t, \mathbf{q}_{-t})}{(1-\gamma)\kappa_0^2} + \sum_{t'=1}^{T} \frac{q_{t'}}{\xi_{t'}} z_{t'}(q_t, \mathbf{q}_{-t}),$$

we get

$$\begin{aligned}\frac{\partial V(q_t, \mathbf{q}_{-t})}{\partial q_t} &= -c'(Z(\mathbf{q}))\frac{z_t(q_t, \mathbf{q}_t)}{\xi_t}\left\{(1-\gamma)\frac{z_t(q_t, \mathbf{q}_{-t})}{\xi_t/\kappa_t^2} - 2(K(q_t, \mathbf{q}_{-t}) - \xi_t)\right\} \\ &= c'(Z(\mathbf{q}))\frac{[K(q_t, \mathbf{q}_{-t}) - \xi_t]^2}{(1-\gamma)\kappa_t^2}.\end{aligned}$$

By a similar argument, we can obtain the same expression for every other $t' \in T^+(\mathbf{q})$ in a neighborhood $(0, \delta')$ of $q_{t'} = 0$. $\square$

A.5 Proof of Proposition 3

The profit constraints are equivalent to

$$q_t\left[K(\mathbf{q}) - \xi_t - (1-\gamma)\frac{H(\kappa_t, \xi_t)}{\xi_t/\kappa_t^2}\right] \geq 0.$$

For every $\mathbf{q}$ such that $K(\mathbf{q}) = \hat{K} \geq K(\mathbf{q}^{e*})$, let

$$T(\hat{K}) = \left\{t : \hat{K} \geq \xi_t + (1-\gamma)\frac{H(\kappa_t, \xi_t)}{\xi_t/\kappa_t^2}\right\}.$$

We first show that, for every $\mathbf{q}$ such that $K(\mathbf{q}) = \hat{K} \geq K(\mathbf{q}^{e*})$, $\mathbf{q}$ minimizes

$$\sum_{t=1}^{T} q_t \frac{(\hat{K} - \xi_t)\xi_t}{(1-\gamma)\kappa_t^2}$$

if and only if $\xi_t > \min_{t' \in T^+(\hat{K})} \xi_{t'}$ implies that $q_t = 0$. To see this, recall that if $K(\mathbf{q}) = \hat{K}$, then $\mathbf{q}$ must satisfy

$$c'\left(\sum_{t \in T(\hat{K})} \zeta_t(q_t)\right)\left[\frac{\hat{K}}{(1-\gamma)\kappa_0^2} + \sum_{t \in T(\hat{K})} \frac{\zeta_t(q_t)}{\xi_t}\right]^2 = 1. \tag{14}$$

where $\zeta_t(q_t) = q_t \frac{(\hat{K}-\xi_t)\xi_t}{(1-\gamma)\kappa_t^2}$ for every $t \in T^+(\hat{K})$. By contradiction, suppose that $\mathbf{q}$ does not satisfy the claimed property. Then, there exists $t', t'' \in T^+(\hat{K})$ such that $\xi_{t'} < \xi_{t''}$ and $\zeta_{t''}(q_{t''}) > 0$. Consider the adjustment to $\zeta_{t'}(\tilde{q}_{t'}) = \zeta_{t'}(q_{t'}) + \varepsilon$, $\zeta_{t''}(\tilde{q}_{t''}) = \zeta_{t''}(q_{t''}) - \varepsilon > 0$ for some $\varepsilon > 0$, and $\tilde{q}_t = q_t$ for all other t, so that the total allocated attention does not change. Since $\frac{1}{\xi_{t'}} > \frac{1}{\xi_{t''}}$ we have that

$$\sum_{t\in T(\hat{K})} \frac{\zeta_t(\tilde{q}_t)}{\xi_t} = \sum_{t\in T(\hat{K})} \frac{\zeta_t(q_t)}{\xi_t} + \varepsilon\left(\frac{1}{\xi_{t'}} - \frac{1}{\xi_{t''}}\right) > \sum_{t\in T(\hat{K})} \frac{\zeta_t(q_t)}{\xi_t}.$$

Thus, there exists $\varepsilon > 0$ sufficiently small that, holding $\tilde{q}_t$ for $t \neq t''$ and everything else fixed, the $q_{t''}$ that solves expression (14) must be strictly lower, which means that the total allocated attention must be strictly lower.

Now suppose that in equilibrium only the clearest type $\hat{t}$ is supplied, which implies that $K(\mathbf{q}^{e*}) = \xi_{\hat{t}} + (1-\gamma)\frac{\kappa_{\hat{t}}^2}{\xi_{\hat{t}}}H(\kappa_{\hat{t}}, \xi_{\hat{t}})$ and so $\hat{t}$-sources always make non-negative profits. We just established that for every $K(\mathbf{q}) \geq K(\mathbf{q}^{e*})$ the clearest type is always profitable and

$$\sum_{t\in T^+(K(\mathbf{q}))} q_t \frac{(K(\mathbf{q}) - \xi_t)\xi_t}{(1-\gamma)\kappa_t^2} \geq q_{\hat{t}} \frac{(K(\mathbf{q}) - \xi_{\hat{t}})\xi_{\hat{t}}}{(1-\gamma)\kappa_{\hat{t}}^2} \geq q_{\hat{t}} H(\kappa_{\hat{t}}, \xi_{\hat{t}}).$$

From Lemma 5, we have that

$$\frac{\partial V(q_{\hat{t}}, \mathbf{0})}{\partial q_{\hat{t}}} \propto c'\left(q_{\hat{t}} z_{\hat{t}}(q_{\hat{t}}, \mathbf{0})\right) \frac{[K(q_{\hat{t}}, \mathbf{0}) - \xi_{\hat{t}}]^2}{(1-\gamma)\kappa_{\hat{t}}^2},$$

so $V(q_{\hat{t}}, \mathbf{0})$ is strictly increasing on $[0, q_{\hat{t}}^{e*}]$. Therefore, $\mathbf{q}^{e*} = (q_{\hat{t}}^{e*}, \mathbf{0})$ is the unique maximizer of $V(\mathbf{q})$ among all $\mathbf{q}$ that involve only the clearest type $\hat{t}$. Since any $\mathbf{q}$ such that $K(\mathbf{q}) \geq K(q_{\hat{t}}^{e*}, \mathbf{0})$ leads to a weakly lower $V(\mathbf{q})$ than does a $\mathbf{q}'$ that involves only $\hat{t}$ and satisfies $K(\mathbf{q}') = K(\mathbf{q})$, it follows that for any such $\mathbf{q}$ we have $V(\mathbf{q}) < V(q_{\hat{t}}^{e*}, \mathbf{0})$. Therefore, $(q_{\hat{t}}^{e*}, \mathbf{0})$ is Pareto efficient.

A.6 Proof of Proposition 4

Suppose only t^{e*}-sources are supplied in equilibrium and that t^{e*} is not the clearest feasible type. For t such that $\xi_t < \xi_{t^{e*}}$, consider the quantity $\overline{q}_t$ that allows t-sources to break even if no other type of sources is supplied. That is, $z_t(\overline{q}_t, \mathbf{0}) = H(\kappa_t, \xi_t)$ or equivalently

$$K(\overline{q}_t, \mathbf{0}) = \xi_t + (1-\gamma)\frac{\kappa_t^2}{\xi_t}H(\kappa_t, \xi_t),$$

which requires that

$$c'(\overline{q}_t H(\kappa_t, \xi_t)) \left[\frac{K(\overline{q}_t, \mathbf{0})}{(1-\gamma)\kappa_0^2} + \frac{\overline{q}_t H(\kappa_t, \xi_t)}{\xi_t}\right]^2 = 1. \tag{15}$$

Using the expression of $V(\mathbf{q})$, we have that $V(\overline{q}_t, \mathbf{0}) > V(q_{t^{e*}}^{e*}, \mathbf{0})$ if and only if

$$c(\overline{q}_t H(\kappa_t, \xi_t)) - c(q_{t^{e*}}^{e*} H(\kappa_{t^{e*}}, \xi_{t^{e*}}))$$

is strictly smaller than

$$K(q^{e*}_{t^{e*}},\mathbf{0})\sqrt{c'(q^{e*}_{t^{e*}}H(\kappa_{t^{e*}},\xi_{t^{e*}}))}-K(\overline{q}_t,\mathbf{0})\sqrt{c'(\overline{q}_tH(\kappa_t,\xi_t))}.$$

Note that $K(\overline{q}_t,\mathbf{0})>K(q^{e*}_{t^{e*}},\mathbf{0})$—otherwise, t would be the clearest type that can be supplied in equilibrium. Thus, $V(\overline{q}_t,\mathbf{0})>V(q^{e*}_{t^{e*}},\mathbf{0})$ requires that $\overline{q}_tH(\kappa_t,\xi_t)<q^{e*}_{t^{e*}}H(\kappa_{t^{e*}},\xi_{t^{e*}})$, which can occur only if ξ_t is sufficiently small.

Now suppose that we lower ξ_t while adjusting $H(\kappa_t,\xi_t)$ so that $K(\overline{q}_t,\mathbf{0})$ does not change. By (15), such changes must cause $\overline{q}_tH(\kappa_t,\xi_t)$ to fall. In particular, $\lim_{\xi_t\downarrow 0}\overline{q}_tH(\kappa_t,\xi_t)=0$. Given this, if the marginal attention cost is not constant, we have that $V(\overline{q}_t,\mathbf{0})>V(q^{e*}_{t^{e*}},\mathbf{0})$ provided that ξ_t and $c'(0)$ are sufficiently small. If the marginal attention cost is constant at c', $V(\overline{q}_t,\mathbf{0})>V(q^{e*}_{t^{e*}},\mathbf{0})$ if and only if

$$q^{e*}_{t^{e*}}H(\kappa_{t^{e*}},\xi_{t^{e*}})\left[\frac{1}{\xi_{t^{e*}}}-\frac{\sqrt{c'}}{(1-\gamma)\kappa_0^2}\right]<\overline{q}_tH(\kappa_t,\xi_t)\left[\frac{1}{\xi_t}-\frac{\sqrt{c'}}{(1-\gamma)\kappa_0^2}\right]. \tag{16}$$

Expression (15) implies that

$$\overline{q}_tH(\kappa_t,\xi_t)=\xi_t\left[\frac{1}{\sqrt{c'}}-\frac{K(\overline{q}_t,\mathbf{0})}{(1-\gamma)\kappa_0^2}\right],$$

and similarly

$$q^{e*}_{t^{e*}}H(\kappa_{t^{e*}},\xi_{t^{e*}})=\xi_{t^{e*}}\left[\frac{1}{\sqrt{c'}}-\frac{K(q^{e*}_{t^{e*}},\mathbf{0})}{(1-\gamma)\kappa_0^2}\right].$$

Recall that $\kappa_0^2>\sqrt{c}\left[\frac{\xi_t}{1-\gamma}+\frac{\kappa_t^2}{\xi_t}H(\kappa_t,\xi_t)\right]$ for every t by Assumption 1, so both right-hand sides are strictly positive. Thus, substituting in (16), we need

$$\frac{(1-\gamma)\kappa_0^2-\sqrt{c'}K(q^{e*}_{t^{e*}},\mathbf{0})}{(1-\gamma)\kappa_0^2-\sqrt{c'}K(\overline{q}_t,\mathbf{0})}<\frac{(1-\gamma)\kappa_0^2-\xi_t\sqrt{c'}}{(1-\gamma)\kappa_0^2-\xi_{t^{e*}}\sqrt{c'}}.$$

Since $K(q^{e*}_{t^{e*}},\mathbf{0})<K(\overline{q}_t,\mathbf{0})$, the left-hand side is strictly greater than 1. Therefore, we must have that ξ_t is sufficiently smaller than $\xi_{t^{e*}}$.

In either case, under both $(\overline{q}_t,\mathbf{0})$ and $(q^{e*}_{t^{e*}},\mathbf{0})$ all sources make zero profits and $V(\overline{q}_t,\mathbf{0})>V(q^{e*}_{t^{e*}},\mathbf{0})$. Thus, $(\overline{q}_t,\mathbf{0})$ Pareto dominates $(q^{e*}_{t^{e*}},\mathbf{0})$.

A.7 Proof of Propositions 6–8

Let $\hat{c}'>0$ be the constant marginal cost of attention. We want calculate $\frac{\partial V^d(\mathbf{q})}{\partial q_t}$ under the condition that $K(\mathbf{q})>\xi_t$. We have

$$\frac{\partial V^d(\mathbf{q})}{\partial q_t} \;=\; c'\frac{[K(\mathbf{q})-\xi_t]^2}{(1-\gamma)\kappa_t^2}-(\gamma-\gamma^*)c'\frac{[K(\mathbf{q})-\xi_t]^2}{(1-\gamma)^2\kappa_t^2}-\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\frac{\xi_t(K(\mathbf{q})-\xi_t)}{(1-\gamma)^2\kappa_t^2}$$

$$
\begin{aligned}
&-2\frac{\gamma-\gamma^*}{1-\gamma}c'\left[\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}+\sum_{t'=1}^{T}q_{t'}\frac{\max\{K(\mathbf{q})-\xi_{t'},0\}}{(1-\gamma)\kappa_{t'}^2}\right]\frac{\partial K(\mathbf{q})}{\partial q_t}\\
&+\left(\frac{\nu-\nu^*}{\nu}\right)\frac{4(1-\gamma^*)c'}{(1-\gamma)^2\kappa_0^2}K(\mathbf{q})\frac{\partial K(\mathbf{q})}{\partial q_t}\\
&+\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\left\{\sum_{t'=1}^{T}q_{t'}\frac{\xi_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{(1-\gamma)\kappa_{t'}^2}\right\}\frac{\partial K(\mathbf{q})}{\partial q_t}.
\end{aligned}
$$

Now use the fact that

$$
\frac{1}{\sqrt{c'}}=\frac{K(\mathbf{q})}{(1-\gamma)\kappa_0^2}+\sum_{t'=1}^{T}q_{t'}\frac{\max\{K(\mathbf{q})-\xi_{t'},0\}}{(1-\gamma)\kappa_{t'}^2}
$$

to simplify the expression to

$$
\begin{aligned}
\frac{\partial V^d(\mathbf{q})}{\partial q_t} \quad = \quad & \frac{K(\mathbf{q})-\xi_t}{(1-\gamma)\kappa_t^2}\left\{[K(\mathbf{q})-\xi_t]\left[1-\frac{\gamma-\gamma^*}{1-\gamma}\right]c'+\frac{\hat{u}_{\sigma\sigma}c'\xi_t}{(1-\gamma)|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\right\}\\
&-2\frac{\gamma-\gamma^*}{1-\gamma}\sqrt{c'}\frac{\partial K(\mathbf{q})}{\partial q_t}\\
&+\left(\frac{\nu-\nu^*}{\nu}\right)\frac{2(1-\gamma^*)c'}{(1-\gamma)^2\kappa_0^2}2K(\mathbf{q})\frac{\partial K(\mathbf{q})}{\partial q_t}\\
&+\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\left\{\sum_{t'=1}^{T}q_{t'}\frac{\xi_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{(1-\gamma)\kappa_{t'}^2}\right\}\frac{\partial K(\mathbf{q})}{\partial q_t}.
\end{aligned}
$$

Recall that $\partial K(\mathbf{q})/\partial q_t < 0$ whenever $K(\mathbf{q}) > \xi_t$. Using expression (12), we have that

$$
\begin{aligned}
\frac{\partial V^d(\mathbf{q})}{\partial q_t} \quad \propto \quad & [K(\mathbf{q})-\xi_t]\left[1-\frac{\gamma-\gamma^*}{1-\gamma}\right]c'+\frac{\hat{u}_{\sigma\sigma}c'\xi_t}{(1-\gamma)|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\\
&+(\gamma-\gamma^*)\frac{2\sqrt{c'}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}\\
&-\left(\frac{\nu-\nu^*}{\nu}\right)\frac{\frac{4(1-\gamma^*)c'K(\mathbf{q})}{(1-\gamma)\kappa_0^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}\\
&-\left(\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\right)\frac{\sum_{t'=1}^{T}q_{t'}\frac{\xi_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}.
\end{aligned}
$$

Note that only the first and last line really depend on the source type; the other three terms are common to all types.

Using the last expression, the conclusions of Proposition 6 and 7 are immediate. For

Proposition 8, we have

$$
\begin{aligned}
\frac{\partial V^d(\mathbf{q})}{\partial q_t} \quad \propto \quad & [K(\mathbf{q})-\xi_t]c' \\
& + \frac{\hat{u}_{\sigma\sigma}c'}{(1-\gamma)|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\left[1-\frac{(1-\gamma)\frac{q_t}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}\right]\xi_t \\
& - \left(\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\right)\frac{\sum_{t'\neq t} q_{t'}\frac{\xi_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}} \\
= \quad & K(\mathbf{q})c'-\xi_t c'\left\{1-\frac{\hat{u}_{\sigma\sigma}}{(1-\gamma)|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\left[\frac{\frac{1}{\kappa_0^2}+\sum_{t'\neq t}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}+\gamma\frac{q_t}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}\right]\right\} \\
& - \left(\frac{\hat{u}_{\sigma\sigma}c'}{|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|}\right)\frac{\sum_{t'\neq t} q_{t'}\frac{\xi_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}.
\end{aligned}
$$

Note that the term in brackets multiplying $\xi_t c'$ is positive if and only if

$$
\hat{u}_{\sigma\sigma}\left[\frac{\frac{1}{\kappa_0^2}+\sum_{t'\neq t}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}+\gamma\frac{q_t}{\kappa_{t'}^2}}{\frac{1}{\kappa_0^2}+\sum_{t'=1}^{T}\frac{q_{t'}\mathbb{I}\{K(\mathbf{q})>\xi_{t'}\}}{\kappa_{t'}^2}}\right] < (1-\gamma)|\hat{u}_{aa}+\hat{u}_{\sigma\sigma}|
$$

This always holds when $\hat{u}_{\sigma\sigma}<0$. For $\hat{u}_{\sigma\sigma}>0$ it requires that $|\hat{u}_{\sigma\sigma}|$ be sufficiently small.

B Appendix: Continuum of Sources

From the characterization of the consumer equilibrium (Lemma 3 and 1), it is easy to see that what ultimately matters for each consumer is how much attention she allocates to the group of t-suppliers *as a whole*—this amount is then divided evenly among all its members. Indeed, let $Z_t = q_t z_t$ and $W_t = q_t w_t$ and recall that $q_0 = 1$, where $t = 0$ denotes the prior. We can then rewrite the expression (9) as

$$
W_t = \frac{\Psi_t}{\sum_{t'=0}^{T}\Psi_{t'}} \qquad \text{and} \qquad Z_t = \frac{\xi_t W_t}{\sqrt{c'(\sum_{t'=0}^{T} Z_{t'})}} \qquad \text{and} \qquad \Psi_t = \frac{1}{(1-\gamma)\kappa_t^2/q_t + \xi_t^2/Z_t}.
$$

Similarly, we can write the equilibrium action function A_ℓ of every consumer ℓ as

$$
A_\ell(X_{1\ell},\ldots,X_{T\ell}) = \sum_{t=1}^{T} W_t X_{t\ell},
$$

where $X_{t\ell} = \frac{1}{q_t} \sum_{\{i:\kappa_i=\kappa_t,\xi_i=\xi_t\}} x_{i\ell}$ for all t.[25] Note that the random variable $X_{t\ell}$ is normally distributed with mean zero and variance $\kappa_0^2 + \frac{\kappa_t^2}{q_t} + \frac{\xi_t^2}{Z_{t\ell}}$. Thus, we can view each consumer as basing her actions on the sufficient statistic $X_{t\ell}$ and so choosing how much attention to allocate to the source of $X_{t\ell}$. This variable summarizes all the information conveyed by the group of t-sources, whose joint signal has accuracy q_t/κ_t^2 and clarity $1/\xi_t^2$.

The condition characterizing K^* in expression (11) is consistent with this interpretation, as it can be written as

$$c'\left(\sum_{t=1}^{T} \frac{\xi_t \max\{K-\xi_t,0\}}{(1-\gamma)\kappa_t^2/q_t}\right)\left[\frac{K}{(1-\gamma)\kappa_0^2} + \sum_{t=1}^{T} \frac{\max\{K-\xi_t,0\}}{(1-\gamma)\kappa_t^2/q_t}\right]^2 = 1. \tag{17}$$

Also, note that

$$Z_t = \frac{\xi_t \max\{K^*-\xi_t,0\}}{(1-\gamma)\kappa_t^2/q_t}.$$

We now want to argue that we can let q_t be any positive real number for every t and continue to use the above equations to describe the consumers' behavior, while preserving the economic meaning of the model. To this end, first note that if q_t is a positive rational number (i.e., $q_t \in \mathbb{Q}_+$), then we can write q_t/κ_t^2 as $\frac{1}{\zeta^t \kappa_t^2}\chi^t$ where $\chi^t, \zeta^t \in \mathbb{Z}_+$. Moreover, if we take $\zeta \in \mathbb{Z}_+$ sufficiently large, we can approximate all q_t in $\mathbb{Q}_+$ with $\frac{\chi^t}{\zeta}$ for some $\chi^t \in \mathbb{Z}_+$ for all t. In this case, $\frac{1}{\zeta\kappa_t^2}\chi^t$ can be interpreted as a situation where there are χ^t t-sources (which is a positive integer) in the market and the informational content of each potential source is $\frac{1}{\zeta\kappa_t^2}$. Note that $\frac{1}{\zeta\kappa_t^2}$ decreases to zero as ζ increases, which can be interpreted as saying that the informational content of each source becomes arbitrarily small when there is a large number of potential sources. This is consistent with a notion of perfect competition in information markets defined as the property that each source is "small" in terms of the amount of information it can provide. Since $\mathbb{Q}_+$ is dense in $\mathbb{R}_+$, by letting ζ become arbitrarily large and χ^t adjust correspondingly across t's, in this way we can approximate every $\mathbf{q} \in \mathbb{R}_+^T$ and hence every level of q_t/κ_t^2 across t's. In the limit, the interpretation of $1/\kappa_t^2$ is that it measures the *rate* at which t-sources entering the market contribute to the total amount of information that they provide to the consumers.

References

Admati, A. R. and P. Pfleiderer (1986). A monopolistic market for information. *Journal of Economic Theory 39*(2), 400–438.

[25]Recall that the prior has mean zero and so it disappears in the expression of A_ℓ.

Admati, A. R. and P. Pfleiderer (1990). Direct and indirect sale of information. *Econometrica 58*(4), 901–928.

Anderson, S. P. and S. Coate (2005). Market provision of broadcasting: A welfare analysis. *Review of Economic studies 72*(4), 947–972.

Anderson, S. P. and A. De Palma (2012). Competition for attention in the information (overload) age. *The RAND Journal of Economics 43*(1), 1–25.

Anderson, S. P., A. De Palma, and Y. Nesterov (1995). Oligopolistic competition and the optimal provision of products. *Econometrica*, 1281–1301.

Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica 75*(4), 1103–1142.

Angeletos, G.-M. and A. Pavan (2009). Policy with dispersed information. *Journal of the European Economic Association 7*(1), 11–60.

Babaioff, M., R. Kleinberg, and R. Paes Leme (2012). Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pp. 92–109. ACM.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Bergemann, D. and A. Bonatti (2015). Selling cookies. *American Economic Journal: Microeconomics 7*(3), 259–94.

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

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

Boczkowski, P. J. (2010). *News at work: Imitation in an age of information abundance.* University of Chicago Press.

Chahrour, R. (2014). Public communication and information acquisition. *American Economic Journal: Macroeconomics 6*(3), 73–101.

Chen, H. and W. Suen (2017). Competition for attention in the news media market. *mimeo, University of Hong Kong.*

Chwe, M. S.-Y. (2013). *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.

Colombo, L., G. Femminis, and A. Pavan (2014). Information acquisition and welfare. *Review of Economic Studies 81*(4).

Cornand, C. and F. Heinemann (2008). Optimal degree of public information dissemination. *Economic Journal 118*(528), 718–742.

Crampes, C., C. Haritchabalet, and B. Jullien (2009). Advertising, competition and entry in media industries. *Journal of Industrial Economics 57*(1), 7–31.

Davenport, T. H. and J. C. Beck (2013). *The attention economy: Understanding the new currency of business*. Harvard Business Press.

Dewan, T. and D. P. Myatt (2008). The qualities of leadership: Direction, communication, and obfuscation. *American Political Science Review 102*(03), 351–368.

Dewan, T. and D. P. Myatt (2012). On the rhetorical strategies of leaders: Speaking clearly, standing back, and stepping down. *Journal of Theoretical Politics 24*(4), 431–460.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review 67*(3), 297–308.

Gentzkow, M. and J. M. Shapiro (2008). Competition and truth in the market for news. *Journal of Economic Perspectives 22*(2), 133–154.

Gentzkow, M. and J. M. Shapiro (2011). Ideological segregation online and offline. *Quarterly Journal of Economics 126*(4), 1799–1839.

Hamilton, J. (2004). *All the news that's fit to sell: How the market transforms information into news*. Princeton University Press.

Hellwig, C. and L. Veldkamp (2009). Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies 76*(1), 223–251.

Hörner, J. and A. Skrzypacz (2016). Selling information. *Journal of Political Economy 124*(6), 1515–1562.

Kennedy, P. J. and A. Prat (2019). Where do people get their news? *Economic Policy 34*(97), 5–47.

Lanham, R. A. (2006). *The economics of attention: Style and substance in the age of information*. University of Chicago Press.

Mankiw, N. G. and M. D. Whinston (1986). Free entry and social inefficiency. *The RAND Journal of Economics*, 48–58.

Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review 92*(5), 1521–1534.

Myatt, D. P. and C. Wallace (2012). Endogenous information acquisition in coordination games. *Review of Economic Studies 79*(1), 340–374.

Myatt, D. P. and C. Wallace (2014). Central bank communication design in a Lucas-Phelps economy. *Journal of Monetary Economics 63*, 64–79.

Newman, N., R. Fletcher, A. Kalogeropoulos, D. A. L. Levy, and R. K. Nielsen (2017). Digital news report 2017. *Reuters Institute*.

Noam, E. M. (2009). *Media Ownership and Concentration in America*. Oxford University Press.

Pavan, A. (2014). Attention, coordination and bounded recall. *Mimeo, Northwestern University*.

Prat, A. and D. Stromberg (2013). The political economy of mass media. In *Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress*, Volume 50, pp. 135. Cambridge University Press.

Sarvary, M. (2011). *Gurus and Oracles: The Marketing of Information*. MIT Press.

Shiller, R. J. (2015). *Irrational exuberance*. Princeton university press.

Simon, H. (1971). Computers, communications and the public interest. *Computers, communications, and the public interest. Johns Hopkins Press, Baltimore*, 40–41.

Sunstein, C. R. (2009). *Republic. com 2.0*. Princeton University Press.

Vives, X. (1988). Aggregation of information in large cournot markets. *Econometrica*, 851–876.

Vives, X. (2010). *Information and learning in markets: the impact of market microstructure*. Princeton University Press.

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press.

Wu, T. (2017). *The attention merchants: The epic scramble to get inside our heads*. Vintage.