The market for online influence

By Itay P. Fainmesser and Andrea Galeotti*

Recent developments in social media have morphed the age-old practice of paying influential individuals for product endorsements into a multibillion-dollar industry, extending well beyond celebrity sponsorships. We develop a parsimonious model in which influencers trade off the increased revenue they obtain from paid endorsements with the negative impact that these have on their followers' engagement and, therefore, on the price influencers receive from marketers.
The model provides testable predictions that match suggestive evidence on pricing of paid endorsements, reveals a novel type of inefficiency that emerges in this market, and clarifies the role of search technology and advice transparency in shaping market activity. In particular, we show that recent policies that make paid endorsements more transparent can backfire, whereas an increase in the effectiveness of the search technology that matches followers to influencers has both direct and strategic positive welfare effects.

From the beginning of time, people sought advice and recommendations from others. Individuals and firms have long tapped into the resulting networks of influence to promote products, services, and agendas by eliciting endorsements from influential individuals. The rise of social media led to the proliferation of *social media influencers*—individuals who produce online content, often focused on one product category, and have followers who engage with and trust their recommendations. Influencers create blogs and are active on Instagram, YouTube, and other online platforms. They attract the attention of marketers, who are willing to pay them to endorse brands and products.[1] This practice, referred to as *influencer marketing*, is growing and becoming a sizable component of the marketing budget of many brands.[2] The fast growth of the market has also attracted the

* Fainmesser: The Johns Hopkins Carey Business School and The Department of Economics, The Johns Hopkins University, Baltimore, MD 21202 (email: Itay_Fainmesser@jhu.edu); Galeotti: London Business School, London, NW1 4SA, United Kingdom (email: agaleotti@london.edu). For their insightful comments and their advice, we thank Nageeb Ali, Kfir Eliaz, Glenn Ellison, Jack Fanning, Ben Golub, Manuel Hermosilla, Laura Hosken, Matt Jackson, Wei Li, Elie Ofek, Jacopo Perego, Devesh Raval, Al Roth, Douglas Smith, Eduard Talamàs, Ruxian Wang, Leeat Yariv, participants of the Retreat on Networks, Information, and Social Economics, and the Sixth European Meeting on Networks in Barcelona, and seminar participants at Brown University, The Johns Hopkins University, London Business School, Paris School of Economics, Pennsylvania State University, and the Federal Trade Commission. We thank Simon Skipka and Jianhui Li for excellent research assistance. Andrea Galeotti is grateful to the European Research Council for support through ERC-consolidator grant (award no. 283454) and to the European University Institute for support through the Research Council internal grant.

[1] This is different from display advertising, often provided by large platforms, in which authenticity does not play a role. For more on display advertising, ad-targeting, ad-skipping, and ad-blocking, see the references in Section V.

[2] For example, recent articles in *The New York Times* ("Inside the Mating Rituals of Brands and

attention of regulatory authorities and consumer protection organizations, and recent transparency-oriented interventions in the US and Europe ask influencers to clearly indicate marketer-sponsored content.

Marketing practitioners broadly divide influencers into several categories based on their followership.[3] Celebrities often have millions of followers; influencers who have audiences of more than a couple hundred thousands but who do not quite reach a celebrity status are often referred to as *macro-influencers*. The emerging *micro-influencers* usually have access to 3,000-100,000 followers, and the smaller *nano-influencers* have as few as several hundred followers. It is not surprising that different influencers are paid vastly different amounts for their endorsements. What is more interesting is that the per-post-per-follower price decreases in the number of followers that an influencer attracts. For example, an influencer with 100k followers often receives per post more than half the amount received by an influencer with 200k.[4] This is the opposite of the effect in traditional media (e.g., the price per viewer for a TV ad is the highest during the Super Bowl; see Chwe (2013)). This unique phenomena, in which endorsements from influencers with smaller audiences are sought after by advertisers, has been called the "rise of the micro-influencers" by marketing practitioners, who also noted that micro-influencers are able to remain credible in the eyes of their followers and generate high engagement with their products' recommendations. This is in contrast with macro-influencers and celebrities who have many more followers, yet often generate much lower engagement per follower.[5]

We develop a model of the market for online influence and answers the following questions:

- How should influencers optimally design their strategies? What are their resulting equilibrium profits?
- What are the attributes of the market for online influence that lead to the so-called "rise of the micro-influencers?" Under what conditions is this phenomena sustainable?
- What are the roles of search quality and advice transparency in shaping market activity and market performance? What policies effectively enhance followers' experiences, and increase profits and total welfare?

Our model of the market for online influence consists of followers, influencers, and marketers. There are many influencers providing product recommendations.

Online Stars") and *Forbes* ("7 Predictions on the Future of Influencer Marketing"), and a recent 2018 report from eMarketer, depict a strong shift of marketing resources away from traditional advertising practices towards influencer marketing. We discuss insights from these articles in more details in the Online Appendix.

[3]For a lists of social influencers ranked by followership, see also https://ranking.influencer.world/.

[4]See the Online Appendix for more details.

[5]A more detailed discussion and suggested evidence on this phenomenon is reviewed in a section of the Online Appendix.

Influencers differ in the standalone benefit they provide to followers. This heterogeneity captures influencers' skills in writing, taking pictures, understanding market trends, and "status" such as being a celebrity offline. Each influencer produces a fixed amount of recommendations and chooses how many should be marketer-sponsored and how many will be organic. Organic posts recommend high quality products, but sponsored posts may also recommend low quality products

Once a follower is matched to a particular influencer, her utility is the sum of the influencer's standalone benefit and the value that she obtains from his recommendations. We start with a model in which followers cannot tell if a recommendation is organic or sponsored; in other words, influencers' advice is opaque.

The heterogeneity in standalone benefits across influencers, together with their choice of sponsored vs. organic recommendations, leads to a ranking of influencers based on the service quality (utility) they provide to followers. We assume that followers are more likely to be matched to highly ranked influencers, but there are search frictions, which capture the many aspects of the technology that governs the search and matching of followers to influencers.[6] For a marketer, the value of having an influencer endorsing its product depends on how many followers the influencer has and on the surplus that the marketer can obtain from each follower.

In equilibrium, each influencer chooses the amount of sponsored recommendations he puts up for sale to maximize his profit, taking as given the choices of other influencers.[7]

Our first set of results characterizes the unique equilibrium of this model. We show that influencers with higher standalone benefit supply more sponsored posts, have more followers, and receive a higher per-post price from marketers, but a lower per-reader-per-post price. Put differently, the influencers with the highest standalone benefits become macro-influencers or celebrities, despite putting the largest amount of sponsored content. Influencers with intermediate standalone benefit become micro-influencers: relative to macro-influencers they supply less sponsored content, get paid less per-post but get paid more per-follower-per-post.

The mechanism behind this result is the imperfect competition amongst influencers for followers. A marginal increase in the influencer's standalone benefit is therefore passed through to followers' utility only in part, whereas the influencer extracts part of the remaining surplus by increasing his amount of sponsored content. As sponsored content dilutes the recommendation quality, the per-post-

[6]In having an induced quality ranking for influencers, we abstract from potential horizontal differentiation. We do so for two reasons: First, our analysis puts emphasis on existing search frictions. We note that given the current quality of search engines, there are rarely search frictions in finding a specific category and instead, users find the search for the "good" influencer more challenging. Second, our analysis is intended to highlight certain features of the market for influence that are different from traditional media, one of which is the significant diversity across influencers on a vertical dimension. Strategic horizontal product positioning has been studied extensively in earlier work on traditional media markets—see Section V for details.

[7]The analysis extends to the case in which influencers are also intrinsically motivated; e.g., they also obtain some utility from the fact that they have followers.

per-follower price declines in the influencer's standalone benefit.

The equilibrium allocation of sponsored content is inefficient. Celebrities and macro-influencers over-provide sponsored content, whereas micro-influencers (and nano-influencers) under-provide sponsored content. Influencers do not fully internalize the negative externalities that posting sponsored content imposes on their followers, and this creates excessive sponsored content. This classical inefficiency can be corrected with taxes and subsidies. A much less obvious source of inefficiency is that influencers do not internalize that their choice of recommendations affect the matching frictions: it alters the relative utilities that different influencers give to their followers and therefore the matching quality between influencers and followers. In the efficient allocation, nano- and micro-influencers would create sponsored content, whereas macro-influencers would not. This content allocation alleviates matching frictions by increasing the likelihood that followers are matched to macro-influencers (those providing higher standalone benefit). This inefficiency cannot be corrected with standard taxes and subsidies. It can, however, be alleviated by investing in improving the effectiveness of the search technology as we explain next.

Our second set of results clarifies the market implications of a change in the search technology's efficiency. We interpret such change as the consequence of any investment that improves the underlying institution in screening influencers' advice and transmitting this information to followers. The key observation is that an increase in search-technology efficiency creates competitive pressures amongst influencers because small differences in the utility they provide to followers now create large differences in their followership size. Influencers react to competition by reducing their supply of sponsored content. Consequently, the distribution of followers becomes more skewed towards macro-influencers, who start earning even higher profits at the expense of the lower ranked micro-influencers. Both forces lead to an increase of followers' aggregate surplus and total welfare. Put differently, improving the search technology will lead to a welfare improvement through the demise of the micro-influencers.

Finally, we derive results on the effect of transparency regulations on the market for online influence. Recently, competition and media authorities in different countries–such as the US Federal Trade Commission (FTC) in 2016, AGM (Italy's state competition authority) in 2017, and Landesmedienanstalten (Germany's state media authority) in 2017– have instructed influencers to clearly mark sponsored content.[8] The policy alleviates asymmetric information between influencers and followers and it has a positive direct effect: Followers can focus on

[8]This is in line with regulation forbidding covert ads in traditional media. The FTC's rationale for its intervention, taken from its *Endorsements Guides* is: "Say you're planning a vacation. You do some research and find a glowing review on someone's blog that a particular resort is the most luxurious place he has ever stayed. If you knew the hotel had paid the blogger hundreds of dollars to say great things about it or that the blogger had stayed there for several days for free, it could affect how much weight you'd give the blogger's endorsement. The blogger should, therefore, let his readers know about that relationship."

organic recommendations and, if they have better outside options, they can ignore sponsored recommendations. However, transparency also affects the incentives of influencers to create organic content and, as we show now, this strategic effect may overturn the direct one.

We show that when transparency is introduced, the price that an influencer receives to endorse a brand becomes less elastic with respect to his choice of sponsored content. This is so for two reasons. First, when recommendations are transparent, an increase in the fraction of sponsored posts affects less the utility of followers, as they can now exercise their outside options instead of following recommendations they know to be sponsored. In turn, the influencer's audience—and therefore the per-post price—become less sensitive to the composition of sponsored versus organic content. Second, an increase in the fraction of sponsored posts no longer affects followers' assessments of the authenticity of sponsored recommendations, which are now known to be sponsored. As a result, under transparency, followers' willingness to pay for recommended products and marketers' willingness to pay per-follower-per-sponsored-post are independent of the composition of sponsored versus organic content.

Since the equilibrium per-post price is less elastic with respect to the level of sponsored content under transparency, the marginal cost for an influencer to supply an additional sponsored post is lower, thereby increasing the supply of sponsored posts by all influencers. This decreases the aggregate followers' utility and total welfare, and thus confounds the positive direct effect of the policy. In our model this strategic effect is of first order; we discuss in Section III some practical economic effects that could attenuate this effect. The contribution of this exercise is to point out a basic unintended consequence of transparency in this market.

Our reading of the existing debate is that possible counterintuitive strategic effects of transparency are not yet taken into consideration by practitioners and policy makers.[9] There is now some evidence that this effect may be important. Ershov and Mitchell (2020) test the prediction of our theory and of Mitchell (2020) (see discussion below). They collected data from top instagram influencers in Germany and Spain from 2014 to 2019, a period in which Germany experienced changes in disclosure regulation for social media sponsorship. Using a difference-in-difference approach, Ershov and Mitchell (2020) document that new regulation has increased the proportion of sponsored posts published by influencers from 19% to 26%, whereas the total number of posts has not changed.

Perhaps most related to our work is the literature, dating back to Brin and

[9] The idea that transparency in economic interactions may trigger unwanted strategic effects is also studied in principal-agent models. In the canonical model of Hölmstrom (1979), transparency always improves aggregate welfare. However, Prat (2005), and Dasgupta, Prat and Verardo (2011) show that in a career-concern model, in which an agent type is her ability to correctly identify the optimal action, transparency over agents' actions may lead to conformity and suboptimal action choices. Those authors' motivation, analysis, and underlying forces are very different from ours. For recent work on the effect of transparency on individual and collective decision making see Ali and Bénabou (2020).

Page (1998), on the conflict between advertising and advice on the Internet. Brin and Page (1998) focus on search engines and highlight the difficulty of having unbiased advertising-funded search engines. Mitchell (2020) formalizes this idea in the context of a dynamic relationship between an influencer and a follower. The influencer chooses a mix of organic and sponsored posts to maximize advertising income and intrinsic utility from being influential, whereas the follower forms expectations with respect to the authenticity of the influencer's recommendations and seeks to maximize surplus from product purchases. We take a reduced-form approach to model the long-term relationship between a follower and an influencer and instead focus on analyzing the equilibrium in a market with many influencers, followers, and marketers. Pei and Mayzlin (2017) consider the regulated environment in which content is transparent and show that in order to preserve the value of recommendations, marketers generally prefer not to establish exclusivity relationship with influencers, even at the cost of the influencer posting less favorable reviews for their products. In contrast, we abstract from considerations of exclusivity and analyze both opaque and transparent recommendation content.

The work on the conflict between advertising and advice on the Internet is predated by an extensive literature that points out the limitations of free advice, often paid for by commissions and kickbacks, as in the case of physicians' recommendations of drugs and treatments and of advice given by financial intermediaries (see Inderst and Ottaviani (2012) and references therein). Our paper focuses on the competition between intermediaries (influencers), whereas the aforementioned literature focuses on (a) competition between marketers to be recommended by an advisor (e.g., Inderst and Ottaviani (2012)) or (b) the direct relationship between a powerful advisor and a marketer (e.g., Fulghieri, Strobl and Xia (2014)).[10] These modeling choices reflect differences between the underlying markets, including (i) the number and accessibility of advisors/influencers, (ii) the way consumers choose to take advice from a financial advisor or physician as opposed to an Instagramer, and (iii) the resulting market power of advisors/influencers.

Our paper also relates to the literatures on advertising provision in traditional media, curation algorithms and news aggregation in online platforms, two-sided platforms, and informative advertising. We review the contribution of our paper with respect to existing work on these topics in Section IV.

Section I develops the model, Section II characterizes the equilibrium, discusses inefficiencies, and derives comparative statics, and Section III assesses policy interventions. Section V concludes. Proofs are relegated to the Appendix.

I. Model: The market for online influence

The market for online influence consists of followers, online influencers, and marketers. Online influencers can be viewed as small platforms. They create product

[10]In earlier work, Lizzeri (1999) studies the optimal information revelation rule by information intermediaries and how the ability of such intermediaries to capture surplus from marketers varies with competition.

recommendations, thereby connecting marketers and followers. Marketers pay influencers to endorse their products, with the hope of boosting their demand. Followers read product recommendations, which influence their purchase decisions. Followers search for influencers and the matching is often mediated by search engines or alike and entails frictions. In what follows, we provide a parsimonious model that links all these elements and allows for a study of the market for online influence.

A. *Influencers*

There is a continuum of influencers of mass 1. Influencers are heterogeneous in the standalone benefit that they give to followers. Influencer-specific standalone benefits $\theta = (\theta_i)_{i \in [0,1]}$ are uniformly distributed on $[0, 1]$. We interpret the standalone benefit as the value that followers obtain from following the influencer regardless of the influencer's product recommendation and endorsements. Hence, a high standalone benefit describes an influencer with a high status/celebrity, or an influencer that is able to create value beyond the specific information conveyed about products.

A primary function of influencers is to create content with product recommendations. We consider an economy with a continuum of mass $\overline{m}$ of products. A fraction τ of products is of high-quality and their consumption value is 1. The remaining fraction of products is of low-quality and their consumption value is 0. The total amount of each influencer's content is normalized to 1 unit of product recommendations, which corresponds to recommendations about a mass $\overline{r} << \overline{m}$ of products (that is, $\overline{r}/\overline{m} \approx 0$). The strategy of influencer i is to choose the proportion of sponsored recommendations, $s_i \in [0, 1]$ and the proportion of organic recommendations $1 - s_i$. In particular:

- The $1 - s_i$ units of organic recommendations are created as follows: the influencer searches for products at random and costlessly screens each product he finds. The influencer disregards low quality products. Hence, this process stops when the influencer has found a fraction $1 - s_i$ of high quality products.
- The influencer auctions s_i units of sponsored content to marketers in a uniform-price auction. The influencer does not observe the quality of the products of the marketers bidding in the auction and does not screen those products.

An important assumption in this formulation is that influencers have better accuracy when screening products for organic posts than for sponsored posts.[11]

[11] The formulation assumes that the cost for the influencers to perfectly screen products for organic content is zero and that influencers are not able to screen products for sponsored content with any accuracy level. These extreme assumptions are made for convenience, but our analysis can be extended trivially to allow for screening costs, interior screening accuracy levels, and heterogeneity in influencers screening abilities. See for example the extension on endogenous content curation in the Online Appendix.

We suggest two complementary reasons for this difference. First, there are products that the influencer has an intrinsic benefit from trying and by trying them receives signals about their quality. The influencer can fill the organic content with these products. Second, a contract for a sponsored post is often tied, explicitly or implicitly, to a target date for the post because it is coordinated with the marketer's overall campaign. Screening products quickly is costly and imperfect because some issues come up only over time and with extended use under different conditions. On the other hand, an influencer can independently begin to use many products and over time release organic content that the influencer feels comfortable with its screening level.

We will show that, in equilibrium, an influencer always sells his entire quota s_i. In the meantime, denote by p_i the price outcome of the auction. If a marketer purchases an endorsement, the influencer receives p_i, creates the content and posts it. Influencer i's profit is:

$$\Pi_i(s_i, p_i) = p_i s_i. \tag{1}$$

This description and payoff function aim at capturing the different nature of organic content – "genuine" influencer's product recommendations that generate no revenue for the influencer – and sponsored recommendations – products' endorsements in exchange for monetary transfers. That is, the fraction of organic content by influencer i can be thought of as the level of i's authenticity. Extending influencers' motives beyond purely monetary ones does not alter our results. A strategy profile of the influencers is $\mathbf{s}$; it specifies s_i for each influencer i.[12]

B. Followers

There is a unit mass of identical followers. Consider a follower who is matched to influencer i. The follower receives utility from the influencer's standalone benefit θ_i. Furthermore, the follower benefits from following influencer i's recommendations. We postulate that the follower observes s_i but does not observe the quality of each recommendation. The follower, however, formulates an expectation, that we denote by τ_i^{SC}, of the proportion of influencer i's sponsored endorsements that are of high quality. Given s_i and expectation τ_i^{SC}, the expected value to the follower of each i's recommendation is $1 - s_i + s_i\tau_i^{SC}$. We assume that in the interaction with the marketer, the follower extracts $1 - \beta$ of this surplus, and the marketer extracts the rest. Overall, the follower expected utility from this match

[12] A way to incorporate the case in which influencers have a direct benefit from followership is to postulate that $\Pi_i(\mathbf{s}) = p_i s_i + \omega n_i(\mathbf{s})$, where $n_i(\mathbf{s})$ is the number of followers of influencer i. As we shall see, even when $\omega = 0$, influencers care about followership because it increases the price they receive per sponsored post. The additional terms obtained when $\omega > 0$ does not affect the qualitative results and conclusions we present. We therefore develop the analysis based on the influencers' objective captured by Expression 1.

is:[13]

$$q_i(s_i;\tau_i^{SC}) = \underbrace{\theta_i}_{\text{standalone benefit}} + \underbrace{[s_i\tau_i^{SC} + (1-s_i)](1-\beta)}_{\text{Recommendations' value}}. \tag{2}$$

We will require that, in equilibrium, the expectation τ_i^{SC} are correct. With some abuse of notation, the follower's utility from the match with influencer i under correct expectation will be denoted by $q_i(s_i)$.

A simplifying assumption in expression 2 is that an influencer's standalone benefit enters linearly in followers' utility and does not affect the value to the consumers of his recommendations. This eliminate situations in which a consumer may like a product just because it has been sponsored by a specific influencer. We extend the analysis to this case in the Online Appendix.

C. Search technology

We assume that followers are more likely to be matched with influencers who provide greater overall quality (utility). The degree of this correlation depends on the efficiency of the search technology. Our model aims to capture, in a static way, a reputation mechanism that aggregates followers' experiences and uses them to create high-quality matching. Formally, the probability density that a follower is matched to influencer i, given $\mathbf{s}$ and θ, is:

$$n_i(\mathbf{s}) = \frac{q_i(s_i)^\alpha}{q(\mathbf{s})},$$

where $q(\mathbf{s}) = \int_j [q_j(s_j)]^\alpha d_j$. The search technology is a continuous variant of the classical urn-ball matching function, in which influencers take the role of balls. In contrast with the standard urn-ball matching function, balls have a difference prominence and the prominence of influencer i depends on the quality of the service (the utility) he provides to followers.

The extent to which higher follower's utility translates into higher prominence, and therefore a higher follower base, depends on $\alpha \geq 0$, which captures the "efficiency" of the search technology. When α is close to zero, the ranking of influencers with respect to $\{q_i\}$ is irrelevant for the matching. Hence, the distribution of followers across influencers is uniform. This describes a situation in which influencers who provide better services to followers are not distinguishable from influencers who provide mediocre services. At the other extreme, when α tends to infinity, each follower is matched, with very high probability, with the highest-ranked influencer; that is, the influencer with highest q_i. There are differ-

[13]This specification is in line with traditional models of platforms and the media market, such as Anderson and Coate (2005).

ent rationales for these search frictions and we refer the reader to our discussion in Section I.F and to the related literature, discussed in Section IV.

D. Marketers

There is a mass $\overline{m}$ of marketers (or firms), each representing a distinct product. Marketers buy product endorsements from influencers. All marketers that approach influencer i observe his standalone benefit θ_i and his offered amount of sponsored content s_i. The marketers place sealed-bids to purchase an endorsement from the influencer. Once all bids are placed, the marketers that placed the top s_i bids are selected with ties broken uniformly at random. Influencer i posts sponsored endorsements on their behalf for a price p_i, which equals the highest losing bid.[14]

The profit of marketer m that purchases an endorsement from influencer i is then:

$$V_{m,i}(\mathbf{s}, p_i) = \underbrace{n_i(\mathbf{s})[s_i\tau_i^{SC} + 1 - s_i]\beta}_{\text{Expected revenue from } i\text{'s endorsement}} - \underbrace{p_i}_{\text{Price for } i\text{'s endorsement}} .$$

That is, the marketer's expected revenue from the sponsored post depends on the number of followers of influencer i's, $n_i(\mathbf{s})$, the willingness to pay of each follower who will meet the marketer via the influencer, which is simply $s_i\tau_i^{SC} + (1 - s_i)$, and the bargaining power of the marketer *viz.* the followers, which we have assumed to be $\beta \in (0, 1)$.

E. Timing and equilibrium concept

The decisions taken by market participants and the timing of these decisions is summarized as follows:

- First, influencers simultaneously choose the shares of organic and sponsored content.
- Second, each influencer is approached by a mass $> \overline{r}$ of marketers, selected uniformly at random from the set of all marketers, that observe his standalone benefit and quantity of SC offered, and bid for SC in the auction.
- Third, sponsored and organic content is created, and the search technology matches followers to influencers. Followers want to buy a unit of products. A follower matched to influencer i purchases products following i's recommendations. The surplus division between marketers and followers follows the simple reduced form we describe above.

[14] This auction format is standard in the auction literature and is commonly called the uniform-price auction, see also Krishna (2009).

We are interested in the Subgame Perfect Equilibrium in weakly undominated strategies of the market. An important assumption that allows to keep the analysis tractable is that there are many more products than any one influencer can ever endorse, i.e., $\overline{r} << \overline{m}$. This assumption delivers two main simplifications:

First, the products endorsed organically by influencer i make an infinitesimally small fraction of the products in the economy. This implies that for all influencers the fraction of high-quality products recommended in their sponsored content equals the fraction of high-quality product in the economy as a whole. Hence, the equilibrium condition that followers have correct expectations about the fraction of high quality products of influencer i's sponsored recommendations means that $\tau_i^{SC} = \tau$.

Second, standard results in auction theory imply that it is a weakly dominant strategy for marketers to bid their true value (see Krishna 2002). The assumption that $r << m$ implies that the ex-ante probability that any given product, be it high or low quality, is picked by influencer i for an organic endorsement is zero. Therefore, all marketers who are unsuccessful to commission a post with influencer i have an identical continuation payoff, which we normalize to zero. Using that $\tau_i^{SC} = \tau$, we obtain that the bid of marketer m for a sponsored post by influencer i is $b_{m,i} := n_i(\mathbf{s})[s_i\tau + 1 - s_i]\beta$.

Therefore, without loss of generality, an equilibrium in our setup can be defined as follows:

Definition 1. An equilibrium is a strategy profile $\mathbf{s}^*$ such that for every influencer i, the fraction of sponsored content s_i^* maximizes his profit (expression 1) given a) the supply of sponsored content $\mathbf{s}_{-i}^*$ and b) that every marketer m that is matched with influencer i bids its true value $b_{m,i} = n_i(\mathbf{s})[s_i\tau + 1 - s_i]\beta$.[15]

Throughout the analysis, we maintain the assumption that $\tau < 1/2$; this assumption assures interiority of equilibrium and it is made only for expositional reasons.

F. *Discussion*

We comment on the main assumptions of the model and elaborate on the role they play in the analysis that follows.

Search technology. Our model postulates search frictions leading to suboptimal matching between followers and influencers. There are different rationales for these search frictions

• *Mediated search.* Almost all online searches are mediated by search or curation algorithms and the technology can be imperfect.[16] Our underlying assumption is

[15] Notably, we omit the requirement that the strategy s_i^* is weakly undominated. This is without loss of generality because, as we show below, the equilibrium is unique.

[16] The importance of search frictions in online markets gave rise to an active area of research in the intersection of marketing, management, computer science, and economics. We briefly review this

that the matching technology tries to match followers to influencers who provide the highest quality service to followers. The parameter α captures underlying imperfections in the technology and may be controlled by search engines platforms or by platforms that host influencers.

• *Asymmetric information.* Search frictions could also be manifestations of an asymmetric information problem that does not allow influencers who provide a high quality service to signal this to followers. In this case, a low α represents institutions that are not able to screen influencers with respect to the quality of the service they provide to followers, e.g., a customer review system that can easily be manipulated.

• *Search costs.* Finally, high search costs for followers – or any behavioral bias that prevents followers from screening influencers – could also be modeled using low values of α.

Sequentiality of moves and observability. We assume that influencers' choices $\mathbf{s}$ are observed by marketers and followers before they make their decisions. This assumption captures two important aspects.

• *Price per sponsored post contingent on follower base.* Since marketers observe s_i, the price that a marketer pays to influencer i depends on the *realized* number of influencer i's followers, $n_i(\mathbf{s})$. An equivalent model is one in which marketers do not observe s_i, but marketers and influencers agree on a price per sponsored post that is contingent on the influencer's realized number of followers, or alternatively on the number of clicks.

• *Reputation.* The assumption that followers observe s_i implies that followers quickly adjust their expected value of each influencer's recommendation when s_i changes. This is a way to model an underlying reputation mechanism: If an influencer deviates by creating more sponsored posts, as compared to what followers have conjectured, consumers will systematically experience a lower utility than expected when visiting the influencer. We are assuming that this information propagates more quickly to other consumers than the influencer can change the supply of sponsored posts.[17]

A consequence of this assumption is that influencers will balance sponsored content with organic content in order to avoid losing reputation and trust among followers, thereby attracting a sizeable follower base. This is in line with anecdotal evidence and discussions amongst practitioners. Incidentally, the assumption that followers are aware of the amount of sponsored content posted by some influencers has been used successfully in European courts to argue that unmarked sponsored content does not constitute unfair advertising.[18] Remark 1 in Section II outlines

literature in Section IV. We also refer the reader to Tadelis (2016) for a survey of feedback systems in online platforms and a discussion of the possible bias in feedback systems and reputation systems.

[17]Analyzing a model of a recommender reputation game in the presence of unobservable side payments, Mostagir and Siderius (2020) show that followers' beliefs converge to the correct ones with probability 1.

[18]For example, at the end of April 2019, the Munich Regional Court dismissed a civil lawsuit filed by a Berlin organization against the influencer Cathy Hummels. The court argued that informed Internet users would know that Hummels pursues commercial interests with her Instagram profile. For more, see

the consequences of relaxing this assumption for equilibrium outcomes.

Influencers' actions and motivations. In our model influencers only decide on their mix of sponsored and organic content. We understand this is a limited description of what influencers choose. For example, influencers can choose to invest more or less costly effort in curating organic recommendations. Similarly, it might be possible for influencers to invest more or less in insuring high-quality sponsored content, for example, by actively engaging in a costly search for sponsors.[19] In an Online Appendix, we discuss how content curation can be incorporated into our model, and argue that such additions do not undermine the tradeoff considered in this paper. Moreover, our model abstracts from potential nuances in reviews or even ratings. Instead, we consider influencers who can give thumbs-up to a small mass of products. For work on the informational value of reviews and ratings, see Chevalier and Mayzlin (2006), Godes and Mayzlin (2009), Fainmesser, Lauga and Ofek (2019) and references therein.

Opaque content. We assume that influencers do not disclose whether a specific product's recommendation is sponsored or not. In this sense, the content produced by influencers is opaque for followers. As we discussed in the Introduction, competition authorities in different countries have recently introduced new legislation or emphasized the application of existing legislation to influencer marketing, in an attempt to increase market transparency. We study the effect of influencers' transparency in Section III.

Sponsored-content pricing. There is little data available on the price determination mechanism in the online market for endorsements. For expositional purposes we assume that influencers auction their endorsements to marketers. In previous iterations of this work we derived identical results assuming instead perfect competition between marketers or posted prices. One advantage of the auction mechanism is that it makes it more straightforward to endogenize the expected quality of a product recommended in a sponsored post, or τ_i^{SC}. In Section IV, we illustrate how this makes the connection between our work and the literature on informative advertising more transparent.

II. Equilibrium analysis

To derive the equilibrium we impose the zero profit condition for marketers and then derive the optimal mix of sponsored and organic content. Solving for the zero profit condition we obtain:

$$V_{m,i}(\mathbf{s}) = n_i(\mathbf{s})[s_i\tau + 1 - s_i]\beta - p_i = 0 \quad \Leftrightarrow \quad p_i(\mathbf{s}) = n_i(\mathbf{s})[1 - s_i(1-\tau)]\beta. \tag{3}$$

https://www.tagesschau.de/wirtschaft/influencer-gesetz-101.html

[19]For a glimpse into the mutual search for a match by influencers and marketers, see a recent article in The New York Times titled "Inside the Mating Rituals of Brands and Online Stars."

Hence, the profits to influencer i are:

$$\Pi_i(\mathbf{s}) \quad = \quad p_i(\mathbf{s})s_i = n_i(\mathbf{s})[1 - s_i(1-\tau)]\beta s_i.$$

The optimal mix of sponsored and organic content solves:

$$\max_{s_i \in [0,1]} n_i(\mathbf{s})[1 - s_i(1-\tau)]\beta s_i.$$

The influencer's trade-off when designing his strategy is summarized by:

$$(4) \qquad \frac{\partial \Pi_i(\mathbf{s})}{\partial s_i} = \underbrace{n_i(\mathbf{s})[1 - s_i(1-\tau)]\beta}_{\uparrow \text{ Sponsored content's revenue}} + \overbrace{\underbrace{\frac{\partial n_i(\mathbf{s})}{\partial s_i}[1 - s_i(1-\tau)]s_i\beta}_{\downarrow \text{ Number of followers}} - \underbrace{n_i(\mathbf{s})s_i\beta(1-\tau)}_{\downarrow \text{ Recommendations' surplus}}}^{\downarrow \text{ Price per sponsored post}}$$

Influencer i's marginal profits can be decomposed into three terms, as described in Expression 4. When influencer i substitutes organic content with sponsored content, his revenue increases. This reflects the opportunity costs of organic content or, alternatively, the marginal benefit to the influencer of increasing s_i (first term of Expression 4). The marginal cost of increasing s_i is the decrease in the price per sponsored post. The remaining two terms describes this effect. First, an increase in s_i decreases the utility that followers receive from influencer i. His number of followers therefore declines, so the price marketers are willing to pay influencer i goes down (second term of Expression 4). Second, an increase in s_i decreases how much followers value influencer i's recommendation, so marketers' revenue from advertising via influencer i goes down. In turn, this pushes down the price p_i (last term of Expression 4).

Remark 1 (Relaxing the assumption of observability of s_i.). Suppose that the influencers make decisions simultaneously to marketers and followers. That is, the strategy profile of the influencers $\mathbf{s}$ is not observed by marketers and influencers before they make their decisions. Then, the first two terms in Expression 4 would be equal to zero, and, in the only equilibrium, influencers create only sponsored content; i.e., $s_i = 1$ for all i ||

Note that each influencer has negligible effect on the aggregate distribution of the utility that influencers provide to followers. That is, the choice of influencer i does not alter $q(\mathbf{s}) = \int_j [q_j(s_j)]^\alpha dj$. This implies that influencers with the same ability face the same maximization problem. Since their payoffs are strictly concave in the amount of sponsored content, in equilibrium, influencers with the same ability will adopt the same strategy.[20] In a symmetric profile $\mathbf{s}$, we denote by $s(\theta)$, $q(\theta)$, and $n(\theta)$ the corresponding $(s_i, q_i(s_i), n_i(\mathbf{s}))$ for each influencer i with $\theta_i = \theta$. Developing Expression 4, we obtain that the equilibrium conditions for a symmetric and interior equilibrium read:

[20]This result extends to a discrete version of our model, because influencers' actions are strategic complements.

$$\beta n(\theta)\left[\underbrace{[1-s(\theta)(1-\tau)]}_{\uparrow\text{ Sponsored content's revenue}} - \overbrace{\left(\underbrace{s(\theta)\alpha(1-\tau)\frac{(q(\theta)-\theta)}{q(\theta)}}_{\downarrow\text{ Number of followers}} + \underbrace{s(\theta)(1-\tau)}_{\downarrow\text{ Recommendations' surplus}}\right)}^{\downarrow\text{ price per-posted content}}\right] = 0$$

or

$$q^*(\theta)[1-2s^*(\theta)(1-\tau)] = \alpha(1-\tau)(1-\beta)[1-s^*(\theta)(1-\tau)]s^*(\theta). \tag{5}$$

The RHS is increasing in $s(\theta)$ and equals 0 at $s(\theta)=0$. The LHS is a decreasing function of $s(\theta)$ and is positive for all $s(\theta) \leq \frac{1}{2(1-\tau)}$. Hence, there is a unique equilibrium, and, the assumption that $\tau < 1/2$ guarantees that equilibrium decisions are interior. We obtain the following characterization:

Proposition 1. *In the unique equilibrium every influencer with standalone benefit θ posts sponsored content $s^*(\theta)$ that satisfies condition 5 and receives a price per sponsored post $p^*(\theta) = n^*(\theta)[1-s^*(\theta)(1-\tau)]\beta$. Furthermore:*

1. *Influencers with higher standalone benefit select higher levels of sponsored content, generate higher utility for followers, and have more followers, i.e., $s^*(\theta)$, $q^*(\theta)$, and $n^*(\theta)$ are all increasing in θ.*
2. *Influencers with higher standalone benefit command higher prices per sponsored post, but lower prices per follower per sponsored post; i.e., $p^*(\theta)$ is increasing in θ, but $p^*(\theta)/n^*(\theta)$ is decreasing in θ.*

We first provide the economic intuitions for part 1 and part 2 of Proposition 1. We then interpret the results around the idea of micro-influencers.

Because the matching technology favors influencers who provide a better service to followers, the elasticity of the number of followers to sponsor content is lower for influencers with a higher θ. It follows that influencers with high θ create more sponsored content.[21] This, in turn, lowers the quality of the service that those high-θ influencers offer to followers. Despite posting more sponsored content, high-θ influencers still provide a better service to followers and so equilibrium follower size is positively correlated with an influencer's standalone benefit. However, due to the equilibrium response of sponsored content decisions, these correlations are weaker than the correlations that would obtain, had the level of sponsored content been kept constant across influencers.

The correlation between the standalone benefit and the fraction of sponsored content has a striking implication for the equilibrium price of sponsored content.

[21] This is consistent with the empirical study of Ershov and Mitchell (2020) who shows that the fraction of sponsored content is higher for influencers with more experience.

Since high-θ influencers have more followers, their price per sponsored post is higher. However, the price per follower per sponsored post that an influencer receives depends only on the quality of his recommendations, i.e., $s^*(\theta)\tau+1-s^*(\theta)$. Since influencers with a high status supply more sponsored content, followers trust them less and so the price per follower per sponsored post declines with the influencer's standalone benefit.[22] That is, while high-θ influencers influence greater audiences, they influence each of their followers to a lesser extent than lower-θ influencers.

Proposition 1 provides a micro-foundation for the so called "rise of the micro-influencers". Micro-influencers are influencers who reach many followers and yet are able to maintain their authenticity and, therefore, to engage their followers. In contrast with the very small influencers (nano-influencers), micro-influencers have a sizeable follower base. In contrast with macro-influencers, they have a smaller follower base, but they provide higher quality recommendations. Micro-influencers manage to compete with macro-influencers to attract followers because they post less sponsored content and because of the underlying search frictions. In addition, the expected value of a micro-influencer's recommendation is higher than the one offered by macro-influencers, as market participants understand that the latter is more tempted to post sponsored content. This pushes up the price for posting via micro-influencers.

To have a better sense of these effects we construct an example with $\tau = 1/3$, $\beta = 1/2$ and $\alpha = 20$. Figure 1 summarizes equilibrium outcomes: we plot, as a function of influencers' standalone benefit, the cumulative distribution of followers to influencers, the profits to influencers, the price per sponsored post and the price per follower per sponsored post.

Around 83% of influencers attract a total of just 10% of followers. These nano-influencers receive the highest price per follower per sponsored post (on average 0.46), reflecting that they are trusted the most by followers. However, since they have very few followers, marketers are willing to pay only a small price per sponsored post and so their profits are very low. We then have micro-influencers: around 15% of influencers attract around 60% of followers. Micro-influencers receive a higher price per sponsored post than nano-influencers, and their price per follower per sponsored post is only slightly lower than nano-influencers (on average 0.44). With a sizeable follower base and a large price per follower per sponsored post, the average profit to micro-influencers is 800% higher than the average profits of nano-influencers. We finally have the macro-influencers and celebrieties, who constitutes the top 2% of influencers, attract 30% of all followers and obtain an average profit which is slightly more than double of the average profit of micro-influencers.

The observation that the price per follower per sponsored post decreases in the

[22]Another way to understand this is by looking at the pass-through to followers' utility for a marginal increase in influencer's standalone benefit. Influencers compete for followers but this competition is not perfect. A marginal increase in θ is therefore passed through to followers only in part; the other part is extracted by influencers who increase the amount of sponsored content.

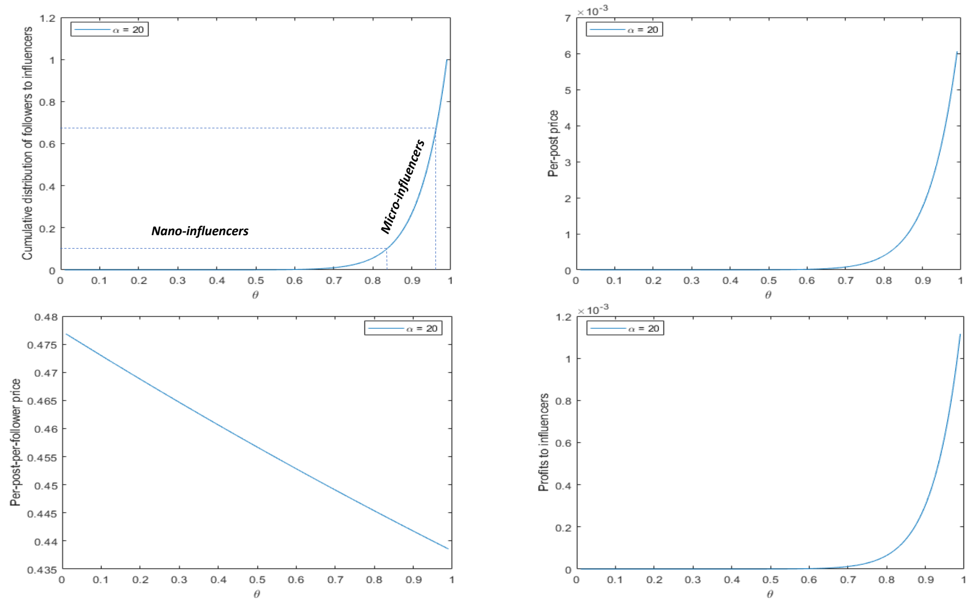

Figure 1. Nano and Micro influencers: $\tau = 1/3$, $\beta = 1/2$ and $\alpha = 20$.

followership of an influencer is the opposite of the patterns observed in traditional media (and predicted by models of traditional media). Two observations clarify this distinction. The first observation is that traditional media plays an important role in coordinating consumers' behavior and therefore the value for a consumer to follow a specific media platform depends on its network size. Such network effects naturally pushes toward a positive relationship between the price per follower per sponsored post and followership. We have not included this effect in our model. We are not aware of empirical studies estimating network effects on the consumer side generated by influencers. Those effects are probably present for some product categories, but we conjecture that the size of most influencers is too small to generate substantial network effects.

Second, the absence of network effects is not sufficient to generate the prediction that the price per follower per sponsored post decreases in the followership of an influencer. In our model, it is the opacity of the recommendations that generates this prediction. To see why, note first that an influencer who posts a higher fraction of sponsored content has a lower average quality of a recommendation. This is true regardless of opacity. However, if followers can distinguish sponsored content from organic recommendations, they ignore the average quality of a recommendation when they assess the quality of a sponsored post. On the other hand, when followers cannot distinguish sponsored content from organic recommendations, they assess all recommendations (including sponsored ones) according to the average recommendation quality. As a result, when content is opaque, influencers with high standalone benefits, who attract more followers and post a significant fraction of sponsored content, end up with less effective sponsored posts than influencers who have less followers. Section III shows formally that when content opacity is removed (as it is in traditional media), the price per follower per sponsored post is independent of the followership of an influencer. In this case, even small network effects will imply a positive relationship.

A. *Inefficiencies*

This "rise of the micro-influencers" is a symptom of a novel source of inefficiency in the market for online influence. To be specific, our model highlights three sources of inefficiency. The first two are standard: first, conditional on being matched, an influencer and a follower have misaligned preferences, with the follower preferring zero sponsored content. Second, competition between influencers is imperfect for any finite α. As a result, influencers do not compete away their incentives to post sponsored content. If these were the only sources of inefficiency, an efficient market allocation would tolerate no sponsored content, and could be implemented with a simple tax.

However, there is a third, more novel, source of inefficiency, which we call *technological inefficiency*. Because of search frictions, followers are not always matched to the influencers who provide them with the highest utility. This inefficiency is exacerbated by influencers' equilibrium choices. Since macro-influencers post more sponsored content, influencers become less heterogeneous with respect to the utilities they offer to their followers, which, in turn makes the assignment of followers to influencers even less efficient: it detracts followers from macro-influencers and increases the number of followers of micro- and nano-influencers.

Formally, for a symmetric allocation of sponsored content $s(\theta)$ and corresponding readership $n(\theta)$, the expression for total surplus is:

$$TS = \int n(\theta)[\theta + 1 - s(\theta)(1-\tau)]d\theta.$$

Proposition 2. *Consider a planner choosing the allocation of sponsored content $s(\theta)$ to maximize total surplus. There exists a threshold $\widetilde{\theta}$ such that for every $\theta < \widetilde{\theta}$, $s^{FB}(\theta) = 1$ and for every $\theta > \widetilde{\theta}$, $s^{FB}(\theta) = 0$. Furthermore, $\widetilde{\theta}$ is increasing in α, tends to 0 as α tends to 0, and tends to 1 as α tends to infinity.*

The best way to alleviate the technological inefficiencies is to tolerate positive levels of sponsored content and use it in a way that amplifies the heterogeneity across influencers with respect to their utility to followers. Specifically, the planner allocates sponsored content to low-ability influencers and only organic content to high-ability influencers. In this way, high-ability influencers provide a much greater utility to followers and low-ability influencers a much lower utility to followers. By creating these strong asymmetries, followers are directed to high-ability influencers and away from low-ability influencers. This effect is illustrated in Figure 2.

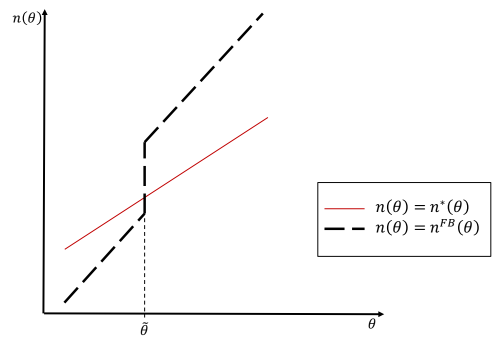

Figure 2. Efficient followership distribution *viz.* equilibrium followership distribution

This qualitative discrepancy between the distributions of sponsored content in equilibrium versus in the welfare maximizing solution is not one that can be fixed by a simple tax or subsidy. In the next section, we show that an effective way to alleviate the inefficiencies in the market for online content is to improve the search technology.

B. The role of the search technology

We now study the implications for the market for influence of a change in the efficiency of the search technology; i.e., a change in α (see Sections I.C and I.F for interpretation of this parameter). This exercise provides testable predictions on how potentially observable outcomes, such as the distribution of followers across influencers or the mix of sponsored and organic content, are expected to differ across markets with a different level of alpha. It also informs managers of platforms hosting influencers about the effects of investing in technologies that improve the matching between their core users and influencers.

To appreciate the role of the parameter α note that two extreme market structures emerge when α converges to 0 or to infinity. When α tends to 0, the distribution of followers across influencers does not depend on content choices. It is as if each influencer has a loyal base of followers. In this case, each influencer acts as a monopolist and the only cost of supplying more sponsored content is that followers will value the influencer's recommendation less. In this case, the fraction of sponsored content of each influencer converges to $1/[2(1-\tau)]$. At the other extreme, when α converges to infinity, an influencer can get all followers by offering the highest utility. In this case, influencers compete á la Bertrand for followers, and in equilibrium they produce only organic content (i.e., $s^*(\theta)$ converges to 0 for all θ and the equilibrium outcome becomes efficient).

More generally, an increase in efficiency of the search technology creates more competition amongst influencers. As a result influencers reduce sponsored content, thus offering greater utility to followers. This is summarized in the next proposition.

Proposition 3. *An increase in the efficiency of the search technology decreases the sponsored content for each influencer, thus increasing the utility that each influencer provides to his followers; i.e., $s^*(\theta)$ decreases and $q^*(\theta)$ increases for all θ.*

An increase in search-technology efficiency also affects the distribution of followers across influencers. Since influencers provide different utility to followers, this indirect change affects the different measures of market performance. The expressions for followers' welfare and influencers' welfare are, respectively:

$$W_F = \int n^*(\theta)q^*(\theta)d\theta \text{ and } W_I = \frac{\beta}{1-\beta}\int n^*(\theta)[q^*(\theta)-\theta]s^*(\theta)d\theta.$$

Proposition 4. *An increase in the efficiency of the search technology leads to:*

1. *A first-order stochastic dominance (FOSD) shift in the distribution of followers across influencers with standalone benefit $\theta \in [0,1]$, i.e., if $\alpha > \alpha'$, then $\int_0^x n^*(\theta,\alpha)d\theta \leq \int_0^x n^*(\theta,\alpha')d\theta$ for all $x \in [0,1]$;*

2. *An increase in the profit of influencers with a large standalone benefit and a decrease in the profit of influencers with a low standalone benefit, i.e., there exists $\bar{\bar{\theta}} \in (0,1]$ such that for every $\theta \leq \bar{\bar{\theta}}$, the profit for an influencer with ability θ decreases and for every $\theta > \bar{\bar{\theta}}$, the profit for an influencer with ability θ increases; and*

3. *An increase in aggregate followers' equilibrium surplus and total equilibrium surplus.*

Figure 3 illustrates how influencers' profits and the distribution of followers across influencers change with α. An increase in α increases the follower base of the macro-influencers and decreases the followers' size of nano-influencers. The direct effect follows from the fact that influencers with higher θ have higher $q^*(\theta)$ and therefore an increase in α will increase their number of followers at the expense of low-θ influencers. However, there is also an indirect effect. Each influencer decreases the amount of sponsored content, which changes the slope of $q^*(\theta)$. When we start from a high α, the highest-θ influencers compete fiercely for followers. An increase in α, then, leads high-θ influencers to reduce their level of sponsored content the most. As a result, the function $q^*(\theta)$ becomes steeper, which implies that the search technology will direct followers more often to influencers with high standalone benefit. In this case, the indirect effect complements the direct effect. However, when α is low to begin with, an increase in α leads low-θ influencers to decrease sponsored content the most. In this case, the profile $q^*(\theta)$ becomes flatter, which makes the search technology more noisy. This effect crowds out, at least in part, the direct effect of an increase in α. Part 1 of Proposition 4 points out that this strategic effect never overturns the direct effect.

The first-order stochastic shift of $n^*(\cdot)$, due to an increase in α, leads to an increase in the profit of macro-influencers, and this comes to the disadvantage of the nano- and potentially also micro-influencers (Part 2). Finally, because an increase in α leads to an FOSD shift in the distribution of followers and because $q^*(\theta)$ is increasing in θ, followers' aggregate surplus increases. In addition, when α increases, influencers lower $s^*(\theta)$, so $q^*(\theta)$ increases for all θ (Proposition 3), which reinforces the former effect so that, overall, followers' welfare increase. The same intuition and logic applies to total welfare.

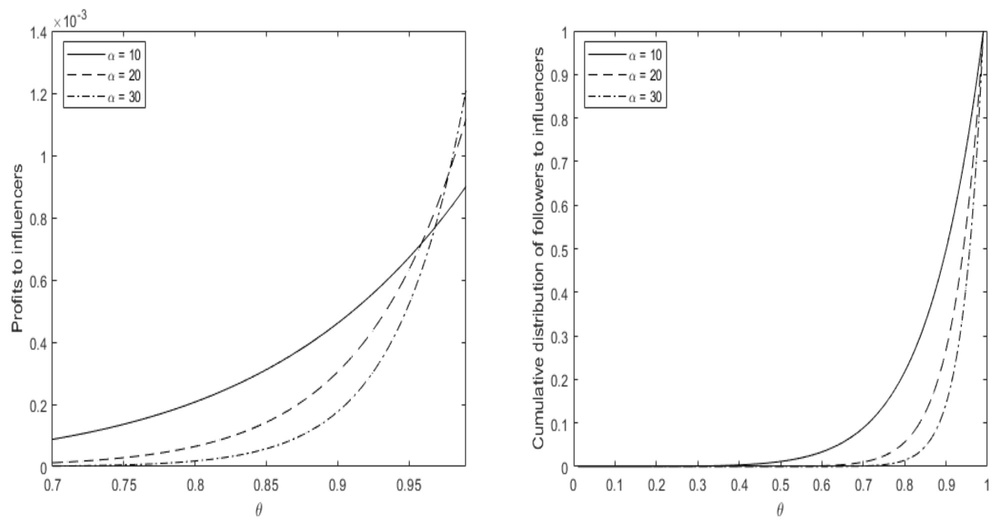

FIGURE 3. THE EFFECT OF A CHANGE IN THE SEARCH-TECHNOLOGY EFFICIENCY ON THE DISTRIBUTION OF FOLLOWERS ACROSS INFLUENCERS AND INFLUENCERS' PROFITS; $\tau = 1/4$, $\beta = 1/2$.

III. Content transparency

In the Introduction we document several attempts of competition authorities to increase transparency by instructing influencers to clearly indicate sponsored content. In this section we extend the benchmark model to study the effect of content transparency on the strategy adopted by influencers and on market performance.

A rationale for enforcing transparency in the content published by influencers is to prevent the bundling of sponsored and organic content, thus allowing followers to ignore sponsored recommendations. With transparent content, followers recognize that the value of a sponsored recommendation is $(1-\beta)\tau$; a follower will ignore such recommendation when she has a more valuable outside option. To introduce this effect, we consider followers who are heterogeneous with respect to an outside option c_i. Only followers with an outside option $c_i < \tau(1-\beta)$ will follow sponsored recommendations; followers will always follow organic recommendations (as their value is $1 \times (1-\beta)$). We assume that $\{c_i\}_{i\in[0,1]}$ are drawn independently from some distribution F in the support $\left[0, \frac{1}{2}(1-\beta)\right]$. We denote by $\hat{x}$ a variable of interest in the model with transparency and by x the same variable prior to the intervention.

Remark 2 (Outside option without transparency). The restriction that $c \le \frac{1}{2}(1-\beta)$ implies that, in our benchmark model (without transparency), followers never exercise their outside option. Hence, we can understand the implication of transparency by comparing the equilibrium outcomes under transparency and in our benchmark model. ||

Let $\gamma \in [0,1]$ be the probability that $c < \tau(1-\beta)$. Let $C = \frac{1}{1-\beta}\int_{\tau(1-\beta)}^{\frac{1}{2}(1-\beta)} c\ dF$ and note that $C \in [\tau, 1/2]$. The expected utility of a follower matched with influencer i (prior to the realization of the outside option), and the marketer's profit from buying a sponsored post from influencer i are, respectively:

$$\begin{aligned}\hat{q}_i(s_i) &= \theta_i + (1-s_i)(1-\beta) + s_i[\tau\gamma + (1-\gamma)C](1-\beta),\\ \hat{V}_{m,i}(\mathbf{s}, p_i) &= \gamma\hat{n}_i(\mathbf{s})\tau\beta - p_i.\end{aligned}$$

The direct effects of the content transparency policy are the following: Followers with high enough outside options will substitute products from sponsored recommendations with more profitable alternatives, and therefore, for the same level of sponsored content s_i, the follower's utility from influencer i under transparency is larger than the utility she obtains prior to the intervention, i.e., $\hat{q}_i(s_i) > q_i(s_i)$, where $q_i(s_i)$ is given by Expression 2. Furthermore, under transparency, only a fraction γ of followers will follow the sponsored recommendation and each now believes that the product has an expected value of τ. Hence, for the same level of sponsored content and for the same per-post price, the marketer's profit is lowered by the introduction of transparency, i.e., $\hat{V}_{m,i}(\mathbf{s}, p_i) < V_{m,i}(\mathbf{s}, p_i)$, where $V_{m,i}(\mathbf{s}, p_i)$ is given by Expression 3.

In equilibrium, the price $\hat{p}_i$ is such that $\hat{V}_{m,i}(\hat{\mathbf{s}}) = 0$. The equilibrium price under transparency and, for comparison, the equilibrium price prior to the intervention are:

$$\underbrace{\hat{p}_i(\hat{\mathbf{s}}) = \hat{n}_i(\hat{\mathbf{s}})\gamma\tau\beta}_{\text{price with transparency}} \quad \text{and} \quad \underbrace{p_i(\mathbf{s}) = n_i(\mathbf{s})[s_i\tau + 1 - s_i]\beta}_{\text{price prior to intervention}}.$$

Note that under transparency the price per sponsored post only depends on the sponsored content s_i via the number of followers $\hat{n}_i(\hat{\mathbf{s}})$.[23] Furthermore, $\hat{n}_i(\hat{\mathbf{s}})$ reacts less to an increase in sponsored posts because it does not decreases as much the utility of followers with high outside options. These two observations subsume the key implication of transparency: *the equilibrium price per sponsored post of an influencer becomes less sensitive to his supply of sponsored content.* Recall that the profit to influencer i with standalone benefit θ_i is:

$$\hat{\Pi}_i(\hat{\mathbf{s}}) \quad = \quad \hat{p}_i(\hat{\mathbf{s}})\hat{s}_i.$$

As transparency decreases the elasticity of $\hat{p}_i(\hat{\mathbf{s}})$ with respect to $\hat{s}_i$, the intervention will increase the amount of sponsored content of each influencer. The next proposition summarizes these observations:

Proposition 5. *Under transparency, there is a unique equilibrium, which is symmetric. The introduction of transparency increases the equilibrium level of sponsored content for each influencer; that is, $\hat{s}(\theta) > s^*(\theta)$ for all θ.*

The strategic effect described in Proposition 5 confounds the positive direct effect of content transparency summarized above. First, a follower with a low outside option matched to a specific influencer will now see more sponsored posts, and second, the matching between followers and influencers change. Our next proposition shows that the strategic response of influencers to transparency is so strong that, for given follower-influencer matches, it erodes all the direct benefits associated with content transparency. It also shows that both for very efficient and very inefficient search technologies, content transparency policies decrease influencers' and total welfare.

Proposition 6. *Relative to the case without transparency, the introduction of transparency implies that:*

1. *A follower's average expected utility from a match to an influencer with ability θ decreases. That is, $\hat{q}(\theta) < q^*(\theta)$ for all θ.*
2. *There exists $0 < \underline{\alpha} < \overline{\alpha}$ such that if $\alpha < \underline{\alpha}$ or $\alpha > \overline{\alpha}$, then both followers' welfare and total surplus decrease.*

When α is sufficiently low, the assignment of followers to influencers is, both with and without transparency, (roughly) uniform. Hence, changes in the relative service quality across influencers do not affect welfare much, and part 1 of Proposition 6 implies that content transparency policies hurt both followers' and aggregate surplus. When α is high, we can show that the introduction of the transparency policy decreases the heterogeneity across influencers with respect to the utility they provide to influencers, i.e., the profile $\hat{q}(\theta)$ is flatter than

[23]Note also that with transparency the price per follower per sponsored post is independent of the followership of the influencer.

$q^*(\theta)$. This means that under content transparency the assignment of followers to influencers becomes less efficient. This effect reinforces the effect in part 1 of Proposition 6.[24]

The main objective of this exercise is to point out that transparency introduces some adverse strategic effects and can backfire in those markets. The recent empirical work from Ershov and Mitchell (2020) provide some support for importance of this strategic effect. Qualitatively, our theoretical result is robust to changes in the specification of the model. For example, the strategic effect does not rely on the uniform distribution assumption of influencers' standalone benefits, or the fact that influencers only care about advertising revenue. At the same time, introducing direct influencers' benefit from followership, or having consumers with outside offers larger than $\frac{1}{2}(1-\beta)$, will decrease the quantitative effect of this strategic effect. Similarly, adding other realistic elements of these markets, such as introducing horizontal differentiation between influencers, may increase or decrease the quantitive effect, but will not eliminate it. We therefore argue that whether this adverse strategic effect of transparency dominates the potential positive effects is an important empirical question.

IV. Related literature

Traditional media. There is an established literature on content and advertising provision in traditional media. Anderson and Coate (2005) study competition in broadcasting, Peitz and Valletti (2008) compare pay-TV and free-to-air media platforms, and Wilbur (2008) provides an empirical model of television advertising and estimates viewers' aversion to advertising; see Anderson and Gabszewicz (2006) for a survey. More recent work, such as Ambrus, Calvano and Reisinger (2016) and Athey, Calvano and Gans (2018), studies media competition in advertising markets with multi-homing users. In this literature, content (programming) has entertainment value and does not include organic recommendations. Advertising is, therefore, never considered authentic and is modelled using a nuisance cost function.[25] In contrast, an important channel in our model is that followers are interested in influencers' recommendations and sponsored content may or may not be "hidden" amongst organic recommendations.

[24]We solved numerically the effect of transparency on followers' welfare, total surplus and influencers' aggregate profits for intermediate levels of α. We note that $\overline{\alpha}$ is bounded from above by a value that is uniquely determined by γ, C, τ and β. Our numerical analysis consisted of an extensive search in the admissible space of (γ, C, τ, β). For each point selected in this set, we checked the effect of transparency for α ranging from 0 to the upper bound of α. The code for this analysis is available upon request from the authors. The numerical analysis points out that the negative effect of transparency on total surplus and followers' surplus also applies to intermediate levels of α. It also provides evidence that aggregate profits to influencers decrease substantially when content is forced to be transparent.

[25]Even when advertising is not considered a nuisance, its benefits are not studied in the context of recommendation authenticity: Gabszewicz, Laussel and Sonnac (2001) present a model of newspaper competition in which newspaper readers do not find advertisements a nuisance because ads can be ignored in a written medium. Rysman (2004) studies a model of the market for Yellow Pages directories in which readers like advertisements.

A second significant difference is that traditional media markets are concentrated and therefore the literature has focused on oligopolistic competition; that is, there are only a few platforms matching the two sides of the market. Online influencers, in contrast, have low entry costs leading to an environment in which (a) influencers are abundant and (b) search frictions are important in shaping the competition amongst them.

Online frictions. The importance of search frictions on online markets has stimulated research at the intersection of management science, computer science, and economics. Search frictions in social media lead many platforms to develop curation algorithms to help populate consumers' feeds. Curation algorithms are, in essence, selection and ranking algorithms that help users (followers) search for the most relevant content (influencers). Early papers study how to better design such algorithms (see also Shardanand and Maes (1995) and Linden, Smith and York (2003)), whereas the more recent literature studies their effect on the content produced (see Latzer et al. (2016) for a survey). For example, Berman and Katona (2016) consider the impact of three curation algorithms on the quality of content created by producers. Su, Sharma and Goel (2016) analyze Twitter's "Who to follow" system that gives users suggestions for which other users to follow. Our comparative statics, with respect to search-technology efficiency, aim at capturing this technological innovation and studying its interaction with market forces.

In related literature on news aggregators, Dellarocas, Katona and Rand (2013) andRoos, Mela and Shachar (2020) find that one effect of news aggregators is increased competition amongst content creators' websites. We find a parallel of that effect when we analyse the effects of improvements in the search and matching technologies. Athey and Mobius (2012), Chiou and Tucker (2017), and Calzada and Gil (2020) find empirical support for the hypothesis that news aggregators serve as a complement to news websites and that they are especially beneficial to niche content providers.[26] In the market we study, all content providers (influencers) are niche and our analysis shows that an improvement in the search technology, akin to a better aggregator, may benefit the high-quality content providers, but hurt the low-quality ones.

Two-sided platforms. A two-sided market is one in which the participants on each side care about the number of participants on the other side, so that there are bilateral network externalities. Hence, each influencer in our model is effectively a two-sided platform whose recommendations facilitate the matching between marketers and followers. There is by now an extensive literature on two-sided platforms (see also Rochet and Tirole (2003), Armstrong (2006), and references therein). However, this literature often focuses on the case of positive

[26]In contrast, de Cornière and Sarvary (2018) find that when the aggregator bundles the news with aggregator unique content, as is the case with Facebook, the aggregator generally harms news outlets and can lead to increase or decrease in news quality and overall news consumption. Kranton and McAdams (2019) study social networks, which percolate and help verify information quality. They show that a denser network can lead to an increase or decrease in the quality of content created, depending on the payoff function of content producers.

externalities between the two sides hosted on the platform, whereas in our case marketers impose negative externalities on followers. As a result, the emphasis on equilibrium selection that is most commonly studied in this literature is not an issue in our economic environment in which there is a unique equilibrium.

Informative advertising. A well known result in the literature on informative advertising is that the average quality of a product featured in a costly advertisement is higher than the average quality of a randomly selected product, even when the content of the advertisement itself is uninformative (see Milgrom and Roberts (1986)). This result can rise in our model if we introduce ex ante heterogeneity of marketers with respect to their profit function. In particular, if marketers of high quality products benefit, on average, more from an endorsement. Such higher expected returns to endorsements are motivated in the literature using repeated purchases, which is also the deriving force in the result of Milgrom and Roberts (1986). Interestingly, as long as some low quality products are featured in sponsored content, we get that sponsored content still features products that are, on average, of lower quality of the ones featured in organic content, and our results go through.

V. Conclusion

We develop a model of market interactions between influencers, followers, and marketers. Our model provides testable predictions on the joint distribution of price per sponsored post and numbers of followers and detects a novel source of inefficiency in this market. We then study how an improvement in the technology that matches followers to influencers affects these market outcomes. Finally, we use the model to reflect on how recent competition and media authorities' interventions in these markets may affect market interactions and outcomes.

An aspect of the market, from which we abstracted, involves the interactions of influencers, followers, and marketers with the platforms hosting them. Influencers are two-sided platforms bringing together followers and marketers. However, influencers and followers are also hosted by a third party. For concreteness, consider the platform Instagram. It does not charge influencers and followers and does not get a cut of the fee that influencers receive from marketers. Rather, its business model is to obtain revenue from display advertisements that marketers place directly on the platform.[27] Hence, the relationship between Instagram and its clients is complex. On the one hand, Instagram competes with the influencers it hosts for attracting advertising revenue from marketers. On the other hand, the

[27]There is a large body of work on display advertising. Evans (2008) provides an early study of the market structure of the online advertising industry, focusing on display advertising and the large platforms that provide it. Bergemann and Bonatti (2011) provide a more comprehensive analysis of the market for offline and online ads, taking into account online markets' greater ability to target audiences, and Goldfarb and Tucker (2011*a*,*b*) find that targeting online display advertisements is highly effective. More recent work on ads targeting includes Deng and Mela (2018) and references within. Followup studies consider the effect of ad skipping and ad-blocking (see also Kumar (2018) and Tuchman, Nair and Gardete (2018).

attractiveness of Instagram for marketers is related to the presence of influencers and followers, whereas the attractiveness of influencers for marketers depend on the quality of the Instagram platform. We believe that our basic model could be extended to these under-studied interactions.

We also abstract from the mechanisms underlying the search technology, α. This is important because our model shows that improving the search technology may be an effective way to enhance surplus and reduce the inefficiencies in the market for online influence. One way of improving the search technology is through the development of better search and matching algorithms used by platforms. Another is by increasing the amount of data used by such algorithms. Recent events teach us that this is not without cost, and that the way consumer privacy is regulated may have significant welfare implications. In Fainmesser, Galeotti and Momot (2020), we propose a framework that allows us to study the tradeoff between the social costs and benefits from increasing data collection by platforms.

REFERENCES

Ali, S Nageeb, and Roland Bénabou. 2020. "Image versus information: Changing societal norms and optimal privacy." *American Economic Journal: Microeconomics*, 12(3): 1–49.

Ambrus, Attila, Emilio Calvano, and Markus Reisinger. 2016. "Either or both competition: A" two-sided" theory of advertising with overlapping viewerships." *American Economic Journal: Microeconomics*, 8(3): 189–222.

Anderson, Simon P, and Jean J Gabszewicz. 2006. "The media and advertising: a tale of two-sided markets." *Handbook of the Economics of Art and Culture*, 1: 567–614.

Anderson, Simon P, and Stephen Coate. 2005. "Market provision of broadcasting: A welfare analysis." *The Review of Economic Studies*, 72(4): 947–972.

Armstrong, Mark. 2006. "Competition in two-sided markets." *The RAND Journal of Economics*, 37(3): 668–691.

Athey, Susan, and Markus Mobius. 2012. "The impact of news aggregators on internet news consumption: The case of localization." Mimeo.

Athey, Susan, Emilio Calvano, and Joshua S Gans. 2018. "The impact of consumer multi-homing on advertising markets and media competition." *Management Science*, 64(4): 1574–1590.

Bergemann, Dirk, and Alessandro Bonatti. 2011. "Targeting in advertising markets: implications for offline versus online media." *The RAND Journal of Economics*, 42(3): 417–443.

Berman, Ron, and Zsolt Katona. 2016. "The Impact of Curation Algorithms on Social Network Content Quality and Structure." *NET Institute Working Paper*, 16(08).

Brin, Sergey, and Lawrence Page. 1998. "The anatomy of a large-scale hypertextual web search engine." *Computer Networks and ISDN Systems*, 30(1/7): 107–117.

Calzada, Joan, and Ricard Gil. 2020. "What Do News Aggregators Do? Evidence from Google News in Spain and Germany." *Marketing Science*, 39(1): 134–167.

Chevalier, Judith A, and Dina Mayzlin. 2006. "The effect of word of mouth on sales: Online book reviews." *Journal of marketing research*, 43(3): 345–354.

Chiou, Lesley, and Catherine Tucker. 2017. "Content aggregation by platforms: The case of the news media." *Journal of Economics & Management Strategy*, 26(4): 782–805.

Chwe, Michael Suk-Young. 2013. *Rational ritual: Culture, coordination, and common knowledge.* Princeton University Press.

Dasgupta, Amil, Andrea Prat, and Michela Verardo. 2011. "Institutional trade persistence and long-term equity returns." *The Journal of Finance*, 66(2): 635–653.

de Cornière, Alexandre, and Miklos Sarvary. 2018. "Social Media and News: Attention Capture via Content Bundling." Working Paper.

Dellarocas, Chrysanthos, Zsolt Katona, and William Rand. 2013. "Media, aggregators, and the link economy: Strategic hyperlink formation in content networks." *Management Science*, 59(10): 2360–2379.

Deng, Yiting, and Carl F Mela. 2018. "TV viewing and advertising targeting." *Journal of Marketing Research*, 55(1): 99–118.

Ershov, Daniel, and Matthew Mitchell. 2020. "The Effects of Influencer Advertising Disclosure Regulations: Evidence From Instagram." *Proceedings of the 21st ACM Conference on Economics and Computation*, 73–74.

Evans, David S. 2008. "The economics of the online advertising industry." *Review of Network Economics*, 7(3).

Fainmesser, Itay P., Andrea Galeotti, and Ruslan Momot. 2020. "Digital Privacy." Working Paper.

Fainmesser, Itay P., Dominique Lauga, and Elie Ofek. 2019. "Ratings, Reviews, and the Marketing of New Products." Working Paper.

Fulghieri, Paolo, Günter Strobl, and Han Xia. 2014. "The economics of solicited and unsolicited credit ratings." *The Review of Financial Studies*, 27(2): 484–518.

Gabszewicz, Jean J, Dider Laussel, and Nathalie Sonnac. 2001. "Press advertising and the ascent of the Pensée Unique." *European Economic Review*, 45(4-6): 641–651.

Godes, David, and Dina Mayzlin. 2009. "Firm-created word-of-mouth communication: Evidence from a field test." *Marketing Science*, 28(4): 721–739.

Goldfarb, Avi, and Catherine E Tucker. 2011*a*. "Online display advertising: Targeting and obtrusiveness." *Marketing Science*, 30(3): 389–404.

Goldfarb, Avi, and Catherine E Tucker. 2011*b*. "Privacy regulation and online advertising." *Management Science*, 57(1): 57–71.

Hölmstrom, Bengt. 1979. "Moral hazard and observability." *The Bell Journal of Economics*, 74–91.

Inderst, Roman, and Marco Ottaviani. 2012. "Competition through commissions and kickbacks." *American Economic Review*, 102(2): 780–809.

Kranton, Rachel, and David McAdams. 2019. "Social Networks and the Market for News." Working paper.

Krishna, Vijay. 2009. *Auction theory.* Academic press.

Kumar, Rajeev. 2018. "Individual Incentive Based Reward Policy for an Ad Exchange to Mitigate the Threat of Ad Blocking." *International Journal of Information Systems and Social Change (IJISSC)*, 9(3): 23–35.

Latzer, Michael, Katharina Hollnbuchner, Natascha Just, and Florian Saurwein. 2016. "The economics of algorithmic selection on the Internet." In *Handbook on the Economics of the Internet.* Edward Elgar Publishing.

Linden, Greg, Brent Smith, and Jeremy York. 2003. "Amazon. com recommendations: Item-to-item collaborative filtering." *IEEE Internet computing*, 7(1): 76–80.

Lizzeri, Alessandro. 1999. "Information revelation and certification intermediaries." *The RAND Journal of Economics*, 214–231.

Milgrom, Paul, and John Roberts. 1986. "Price and advertising signals of product quality." *Journal of Political Economy*, 94(4): 796–821.

Mitchell, Matthew. 2020. "Free Ad(vice): Internet Influencers and Disclosure Regulation." *The RAND Journal of Economics*, Forthcoming.

Mostagir, Mohamed, and James Siderius. 2020. “Strategic Reviews.” Working Paper.

Pei, Amy, and Dina Mayzlin. 2017. “Paid vs Independent Product Recommendation By Bloggers.” Working paper.

Peitz, Martin, and Tommaso M Valletti. 2008. “Content and advertising in the media: Pay-tv versus free-to-air.” *International Journal of Industrial Organization*, 26(4): 949–965.

Prat, Andrea. 2005. “The wrong kind of transparency.” *American Economic Review*, 95(3): 862–877.

Rochet, Jean-Charles, and Jean Tirole. 2003. “Platform competition in two-sided markets.” *Journal of the European Economic Association*, 1(4): 990–1029.

Roos, Jason MT, Carl F Mela, and Ron Shachar. 2020. “The effect of links and excerpts on internet news consumption.” *Journal of Marketing Research*, 57(3): 395–421.

Rysman, Marc. 2004. “Competition between networks: A study of the market for yellow pages.” *The Review of Economic Studies*, 71(2): 483–512.

Shardanand, Upendra, and Pattie Maes. 1995. “Social information filtering: algorithms for automating word of mouth.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, 210–217.

Su, Jessica, Aneesh Sharma, and Sharad Goel. 2016. “The effect of recommendations on network structure.” *Proceedings of the 25th international conference on World Wide Web*, 1157–1167.

Tadelis, Steven. 2016. “Reputation and feedback systems in online platform markets.” *Annual Review of Economics*, 8: 321–340.

Tuchman, Anna E, Harikesh S Nair, and Pedro M Gardete. 2018. “Television ad-skipping, consumption complementarities and the consumer demand for advertising.” *Quantitative Marketing and Economics*, 16(2): 111–174.

Wilbur, Kenneth C. 2008. “A two-sided, empirical model of television advertising and viewing markets.” *Marketing science*, 27(3): 356–378.

Appendix

Proof of Proposition 1.

Krishna (2009) establishes that in a sealed-bid unit-demand uniform-price auction, it is a weakly dominant strategy for all bidders to bid their true value. Combined with the assumption that each influencer is approached by a mass$\geq \overline{r}$,

it follows in our setup that marketers make zero profit and that for all i, $\tau_i^{SC} = \tau$. We next prove that the induced game between influencers has a unique equilibrium as characterized by proposition 1.
Existence and uniqueness. Recall that

$$\frac{\partial \Pi_i(\mathbf{s})}{\partial s_i} = \frac{\beta q_i(s_i)^{\alpha-1}}{q(\mathbf{s})}\left[q_i(s_i)(1-s_i(1-\tau)) - s_i\left[\alpha(1-\tau)(1-\beta)(1-s_i(1-\tau)) + (1-\tau)q_i(s_i)\right]\right]$$

and therefore $\frac{\partial \Pi_i(\mathbf{s})}{\partial s_i} \geq 0$ if, and only if,

$$q_i(s_i)(1-2s_i(1-\tau)) \geq s_i\alpha(1-\tau)(1-\beta)(1-s_i(1-\tau)).$$

Define $\hat{s} = \frac{1}{2(1-\tau)}$ and note that $\hat{s} \leq 1$ because $\tau \leq 1/2$. Note also that the LHS of the above inequality equals $\theta + (1-\beta)$ at $s_i = 0$ and 0 at $s_i = \hat{s}$, and it is decreasing in s_i. The RHS is 0 at $s_i = 0$, it equals $\alpha(1-\beta)/4$ at $s_i = \hat{s}$ and it is increasing in $s_i \in [0, \hat{s}]$. So, there is a unique solution s_i and $s_i \in (0, \hat{s})$.[28] So, influencer i with ability θ_i chooses s_i so that

$$q_i(s_i)(1-2s_i(1-\tau)) = s_i\alpha(1-\tau)(1-\beta)(1-s_i(1-\tau)).$$

Since there is a unique solution to this equation, influencers with the same θ will choose the same strategy. Hence, the equilibrium is symmetric. Notice also that the equilibrium price to influencers with ability θ is derived by the following zero profit condition $n(\mathbf{s})[s(\theta)\tau + b(\theta)]\beta - p(\theta) = 0$.

Part 1. First, to see that $s^*(\theta)$ is increasing in θ, consider equilibrium condition 5 and note that the LHS shifts up as θ increases. The RHS is independent of θ. Second, to see that $q^*(\theta)$ is increasing in θ, rewrite equilibrium condition 5 as follows

$$q^*(\theta) = \alpha(1-\tau)(1-\beta)\frac{[1-s^*(\theta)(1-\tau)]s^*(\theta)}{1-2s^*(\theta)(1-\tau)}. \tag{A1}$$

The LHS is $q^*(\theta)$, and, holding $s^*(\theta)$ fixed, it shifts up when θ increases, whereas the RHS is independent of θ. Third, since $q^*(\theta)$ is increasing in θ, it follows immediately that $n^*(\theta)$ is increasing in θ.

Part 2. Note that $p^*(\theta)/n^*(\theta) = \beta[1-s^*(\theta)(1-\tau)]$ and since $s^*(\theta)$ is increasing in θ it follows that $p^*(\theta)/n^*(\theta)$ is decreasing in θ.

We conclude by showing that $p^*(\theta) = \beta n^*(\theta)[1-s^*(\theta)(1-\tau)]$ is increasing in θ. To see this note that the claim is true whenever $q^*(\theta)[1-s^*(\theta)(1-\tau)]$ is increasing in θ, and this follows by inspection of equilibrium condition $A1$: multiply the LHS

[28]The case in which $\tau > 1/2$ will lead to a similar characterization. The only difference is that influencers with sufficiently high θ will only select sponsored content. We restrict the analysis to $\tau \leq 1/2$ so that we do not take into account the possibility of a corner solution for some influencers, and this makes the analysis easy to present.

and RHS of A1 by $[1 - s^*(\theta)(1-\tau)]$ and then note that the LHS is increasing in θ and it is decreasing in $s^*(\theta)$ and that the RHS is independent of θ and it is increasing in $s^*(\theta)$. This concludes the proof of proposition 1.

Proof of Proposition 2. Recall that total surplus reads:

$$\begin{aligned} TS &= W^r + W^b + W^f = \int n(\theta)[\theta + s(\theta)\tau + b(\theta)]d\theta \\ &= \frac{\int q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^\alpha d\theta}. \end{aligned}$$

Next, note that TS increases in $s(\theta')$ if

$$\frac{\frac{\partial\left(q(\theta')^\alpha[\theta'+1-s(\theta')(1-\tau)]\right)}{\partial s(\theta')}}{\int q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta} > \frac{\frac{\partial q(\theta')^\alpha}{\partial s(\theta')}}{\int_0^1 q(\theta)^\alpha d\theta}$$

or

$$\frac{\frac{1}{1-\beta}\left(\frac{1+\alpha}{\alpha}q(\theta') - \beta\theta'\right)}{\int q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta} < \frac{1}{\int_0^1 q(\theta)^\alpha d\theta}$$

and decreases otherwise. Note further that if the inequality holds for θ' for some $s(\theta')$ then it holds for any larger $s(\theta')$, and if the reverse inequality holds for θ' for some $s(\theta')$ then it holds for any smaller $s(\theta')$. Thus, for any θ' the planner will choose $s(\theta') \in \{0,1\}$, which maps to $q(\theta') \in \{\theta' + \tau(1-\beta), \theta' + (1-\beta)\}$.

To be more specific, the planner will choose $s(\theta') = 0$ (equivalently, $q(\theta') = \theta' + (1-\beta)$) if

$$\frac{q(\theta')^\alpha[\theta' + 1 - s(\theta')\,(1-\tau)]|_{s(\theta')=0} - q(\theta')^\alpha[\theta' + 1 - s(\theta')\,(1-\tau)]|_{s(\theta')=1}}{\int_0^1 q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta} > \frac{q(\theta')^\alpha|_{s(\theta')=0} - q(\theta')^\alpha|_{s(\theta')=1}}{\int_0^1 q(\theta)^\alpha d\theta}$$

and choose $s(\theta') = 1$ (equivalently, $q(\theta') = \theta' + \tau(1-\beta)$) otherwise. The condition simplifies to

$$\frac{(\theta' + (1-\beta))^\alpha\,[\theta'+1] - (\theta' + \tau(1-\beta))^\alpha\,[\theta' + 1 - (1-\tau)]}{\int_0^1 q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta} > \frac{(\theta' + (1-\beta))^\alpha - (\theta' + \tau(1-\beta))^\alpha}{\int_0^1 q(\theta)^\alpha d\theta}$$

or

$$\frac{(\theta' + (1-\beta))^\alpha\,[\theta'+1] - (\theta' + \tau(1-\beta))^\alpha\,[\theta' + 1 - (1-\tau)]}{(\theta' + (1-\beta))^\alpha - (\theta' + \tau(1-\beta))^\alpha} > \frac{\int_0^1 q(\theta)^\alpha[\theta + 1 - s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^\alpha d\theta} \tag{A2}$$

where the RHS is independent of θ' and the LHS increases in θ'. To see that the LHS increases in θ' we write it as follows

$$1 + \theta' + (1-\tau)\,\frac{(\theta' + \tau(1-\beta))^\alpha}{(\theta' + (1-\beta))^\alpha - (\theta' + \tau(1-\beta))^\alpha}$$

which is increasing in θ' if the following is decreasing in θ'

$$\frac{(\theta' + (1-\beta))^{\alpha} - (\theta' + \tau(1-\beta))^{\alpha}}{(\theta' + \tau(1-\beta))^{\alpha}}$$

or if

$$\frac{(\theta' + (1-\beta))}{(\theta' + \tau(1-\beta))}$$

is decreasing in θ', which is always the case because $\tau < 1$. This completes the proof that there is a threshold $\widetilde{\theta}$ as required.

We next show how the threshold $\widetilde{\theta}$ changes with α. To show that the threshold $\widetilde{\theta}$ is increasing in α, it is sufficient to show that inequality A2 is easier to satisfy when α decreases. That will work if for example, the LHS decreases in α and the RHS increases in α. That the RHS increases in α is immediate because $q(\theta)^{\alpha}$ and $[\theta + 1 - s(\theta)\,(1-\tau)]$ both increase in θ. We next wan to show that

$$\frac{(\theta' + (1-\beta))^{\alpha}\,[\theta' + 1] - (\theta' + \tau(1-\beta))^{\alpha}\,[\theta' + 1 - (1-\tau)]}{(\theta' + (1-\beta))^{\alpha} - (\theta' + \tau(1-\beta))^{\alpha}}$$

decreases in α.

$$\frac{\partial\left(\frac{(\theta+(1-\beta))^{\alpha}(\theta+1)-(\theta+\tau(1-\beta))^{\alpha}(\theta+1-(1-\tau))}{(\theta+(1-\beta))^{\alpha}-(\theta+\tau(1-\beta))^{\alpha}}\right)}{\partial\alpha} =$$
$$= (1-\tau)\,(\theta+\tau-\beta\tau)^{\alpha}\,(\theta-\beta+1)^{\alpha}\,\frac{\ln(\theta+\tau-\beta\tau) - \ln(\theta-\beta+1)}{\left((\theta-\beta+1)^{\alpha} - (\theta+\tau-\beta\tau)^{\alpha}\right)^{2}}$$

which is negative if and only if $\ln(\theta+\tau-\beta\tau) - \ln(\theta-\beta+1) < 0$, or $(1-\beta)\,\tau < 1-\beta$, which always holds.

Next, we show that if $\alpha \to 0$ then inequality A2 always holds. The inequality can be rewritten as follows

$$\frac{(\theta' + (1-\beta))^{\alpha}\,[\theta' + 1] - (\theta' + \tau(1-\beta))^{\alpha}\,[\theta' + 1 - (1-\tau)]}{\int_0^1 q(\theta)^{\alpha}[\theta + 1 - s(\theta)\,(1-\tau)]d\theta} > \frac{(\theta' + (1-\beta))^{\alpha} - (\theta' + \tau(1-\beta))^{\alpha}}{\int_0^1 q(\theta)^{\alpha} d\theta}$$

and substituting $\alpha = 0$ we get

$$\frac{(1-\tau)}{\int_0^1 [\theta + 1 - s(\theta)\,(1-\tau)]d\theta} > 0,$$

which always holds.

Next, we show that if $\alpha \to \infty$ then inequality A2 holds only when $\theta' = 1$. We begin by showing that inequality A2 holds for $\theta' = 1$ regardless of α. When $\theta' = 1$ the inequality becomes:

$$\frac{2\,(2-\beta))^{\alpha} - [1+\tau]\,(1 + \tau(1-\beta))^{\alpha}}{(2-\beta)^{\alpha} - (1 + \tau(1-\beta))^{\alpha}} > \frac{\int_0^1 q(\theta)^{\alpha}[\theta + 1 - s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^{\alpha} d\theta}$$

note that

$$\frac{\int_0^1 q(\theta)^\alpha[\theta+1-s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^\alpha d\theta} < \frac{\int_0^1 2q(\theta)^\alpha d\theta}{\int_0^1 q(\theta)^\alpha d\theta} = 2$$

and

$$\frac{2\,(2-\beta)^\alpha - [1+\tau]\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - (1+\tau(1-\beta))^\alpha} > \frac{2\,(2-\beta)^\alpha - 2\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - (1+\tau(1-\beta))^\alpha} = 2.$$

Finally, we show that

$$\lim_{\alpha\to\infty} \frac{\int_0^1 q(\theta)^\alpha[\theta+1-s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^\alpha d\theta} = 2$$

and

$$\lim_{\alpha\to\infty} \frac{2\,(2-\beta)^\alpha - [1+\tau]\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - (1+\tau(1-\beta))^\alpha} = 2.$$

Hence, if $\alpha\to\infty$ the inequality A2 holds only when $\theta' = 1$.
To see the first limit

$$\begin{aligned}
\lim_{\alpha\to\infty} \frac{\int_0^1 q(\theta)^\alpha[\theta+1-s(\theta)\,(1-\tau)]d\theta}{\int_0^1 q(\theta)^\alpha d\theta} &= \lim_{\alpha\to\infty} \frac{\ln(q(1))\,q(1)^\alpha[1+1] - \ln(q(0))\,q(0)^\alpha[1-(1-\tau)]}{\ln(q(1))\,q(1)^\alpha - \ln(q(0))\,q(0)^\alpha} \\
&= \lim_{\alpha\to\infty} \frac{2\ln(q(1))\,q(1)^\alpha - \tau\ln(q(0))\,q(0)^\alpha}{\ln(q(1))\,q(1)^\alpha - \ln(q(0))\,q(0)^\alpha} \\
&= 2
\end{aligned}$$

where the last equlity holds because $s\,(0) = 1$ and $q(0) < 1$. We now prove the second limit

$$\begin{aligned}
\lim_{\alpha\to\infty} \frac{2\,(2-\beta)^\alpha - (1+\tau)\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - (1+\tau(1-\beta))^\alpha} &= \lim_{\alpha\to\infty}\left(1+\frac{(2-\beta)^\alpha - \tau\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - (1+\tau(1-\beta))^\alpha}\right) \\
&= \lim_{\alpha\to\infty}\left(1+\frac{1}{\frac{(2-\beta)^\alpha-(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha-\tau(1+\tau(1-\beta))^\alpha}}\right) \\
&= \lim_{\alpha\to\infty}\left(1+\frac{1}{1+\frac{-(1-\tau)(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha-\tau(1+\tau(1-\beta))^\alpha}}\right) \\
&= 1+\lim_{\alpha\to\infty}\left(\frac{1}{1+\frac{-(1-\tau)(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha-\tau(1+\tau(1-\beta))^\alpha}}\right) \\
&= 2
\end{aligned}$$

where the last inequality holds because

$$\lim_{\alpha\to\infty} \frac{-\,(1-\tau)\,(1+\tau(1-\beta))^\alpha}{(2-\beta)^\alpha - \tau\,(1+\tau(1-\beta))^\alpha} = 0$$

Proof of proposition 3. Consider equilibrium condition 5. The RHS increases in α (specifically, an increase in α rotates leftward the RHS), whereas the LHS does not change with α. Recalling that the LHS increases in $s^*(\theta)$ and the RHS decreases in $s^*(\theta)$, we get that $s(\theta)$ declines in α for all θ. Furthermore, as α goes to 0, the RHS goes to zero for all θ and so $s(\theta)$ goes to $\frac{1}{2(1-\tau)}$ for all θ. As α goes to ∞, the RHS goes to ∞ unless $s(\theta)$ goes to 0 for all θ. Next, by definition $q(\theta) = \theta + [1 - s(\theta)(1-\tau)](1-\beta)$; hence, an increase in α decreases $s(\theta)$ and therefore $q(\theta)$ increases for all θ.

Proof of proposition 4.

Part 1. The result of part 1 follows by showing that there exists a $\overline{\theta} \in (0,1)$ such that an increase in α leads to an increase in $n(\theta)$ for all $\theta > \overline{\theta}$ and a decrease in $n(\theta)$ for all $\theta < \overline{\theta}$. We now prove this claim. Since $q^*(\theta)$ increases in θ, it follows that $n^*(\theta)$ is increasing in θ. Since total readership is fixed, it is sufficient to show that in equilibrium

$$\frac{\frac{dq^*(\theta)^\alpha}{d\alpha}}{q^*(\theta)^\alpha}$$

increases in θ, or equivalently, to prove that

$$\frac{\frac{dq^*(\theta)}{d\alpha}}{q^*(\theta)}$$

increases in θ (to see why this is equivalent, note that for any well behaved function f, $\frac{\frac{df(x)^\alpha}{dx}}{f(x)^\alpha} = \frac{\alpha f(x)^{\alpha-1}\frac{df(x)}{dx}}{f(x)^\alpha} = \alpha\frac{\frac{df(x)}{dx}}{f(x)}$).

Next, recall that in equilibrium

$$q^*(\theta) = \alpha(1-\tau)(1-\beta)\frac{[1-s^*(\theta)(1-\tau)]s^*(\theta)}{1-2s^*(\theta)(1-\tau)}.$$

With some abuse of notation we denote $s = s^*(\theta)$ for the remainder of the proof when the dependence on θ is clear from the context. We get that

$$\begin{aligned}
\frac{dq^*(\theta)}{d\alpha} &= (1-\tau)(1-\beta)\left(\frac{(1-s(1-\tau))\,s}{1-2s(1-\tau)} + \alpha\frac{\partial\left(\frac{(1-s)(1-\tau))s}{1-2s(1-\tau)}\right)}{\partial s}\frac{\partial s}{\partial\alpha}\right)\\
&= (1-\tau)(1-\beta)\left(\frac{(1-s(1-\tau))\,s}{1-2s(1-\tau)} + \alpha\frac{(2s^2\tau^2 - 4s^2\tau + 2^2 + 2s\tau - 2s + 1)}{(2s\tau - 2s + 1)^2}\frac{\partial s}{\partial\alpha}\right)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\frac{dq(\theta)}{d\alpha}}{q(\theta)} &= \frac{(1-\tau)(1-\beta)\left(\frac{(1-s(1-\tau))s}{1-2s(1-\tau)} + \alpha\frac{(2s^2\tau^2-4s^2\tau+2s^2+2s\tau-2s+1)}{(2s\tau-2s+1)^2}\frac{\partial s}{\partial\alpha}\right)}{\alpha(1-\tau)(1-\beta)\frac{[1-s(\theta)(1-\tau)]s(\theta)}{1-2s(\theta)(1-\tau)}}\\
&= \frac{\left((1-s(1-\tau))\,s + \alpha\frac{(2s^2\tau^2-4s^2\tau+2s^2+2s\tau-2s+1)}{(1-2s(1-\tau))}\frac{\partial s}{\partial\alpha}\right)}{\alpha\,(1-s(1-\tau))\,s}\\
&= \frac{1}{\alpha} + \frac{1}{(1-s(1-\tau))}\frac{\frac{\partial s}{\partial\alpha}}{s} + \frac{2s\,(1-\tau)^2}{(1-2s(1-\tau))\,(1-s(1-\tau))}\frac{\partial s}{\partial\alpha}
\end{aligned}$$

We note that $\frac{2s(1-\tau)^2}{(1-2s(1-\tau))(1-s(1-\tau))}$ and $\frac{1}{(1-s(1-\tau))}$ increase in s and therefore in θ. Therefore, to prove that $\frac{\frac{dq(\theta)}{d\alpha}}{q(\theta)}$ increases in θ it is sufficient to show that $\frac{\frac{\partial s}{\partial \alpha}}{s}$ increases in θ.

To move forward we rewrite the equilibrium condition for s as follows:

$$(\theta + (s\tau + 1 - s)(1-\beta))(1 - 2s(1-\tau)) = \alpha(1-\tau)(1-\beta)(1 - s(1-\tau))s$$

and apply the implicit function theorem to get

$$\frac{\frac{\partial s}{\partial \alpha}}{s} = -(1-\beta)\frac{1 - s(1-\tau)}{2\theta - 4s + \alpha - 3\beta - 2s\alpha + 4s\beta + 4s\tau - \alpha\beta + 2s\alpha\beta + 2s\alpha\tau - 4s\beta\tau - 2s\alpha\beta\tau + 3}$$

which is increasing in θ if

$$\frac{1 - s(1-\tau)}{2\theta - 4s + \alpha - 3\beta - 2s\alpha + 4s\beta + 4s\tau - \alpha\beta + 2s\alpha\beta + 2s\alpha\tau - 4s\beta\tau - 2s\alpha\beta\tau + 3}$$

decreases in θ. We know that $1 - s(1-\tau)$ decreases in θ, and therefore it is sufficient to show that

$$2\theta - 4s + \alpha - 3\beta - 2s\alpha + 4s\beta + 4s\tau - \alpha\beta + 2s\alpha\beta + 2s\alpha\tau - 4s\beta\tau - 2s\alpha\beta\tau + 3$$

increases in θ. This is true because $-4s + \alpha - 3\beta - 2s\alpha + 4s\beta + 4s\tau - \alpha\beta + 2s\alpha\beta + 2s\alpha\tau - 4s\beta\tau - 2s\alpha\beta\tau + 3$ is increasing in s which increases in θ.

Part 2. We first note that

$$\Pi^*(\theta') = \frac{q^*(\theta')^\alpha[1 - s^*(\theta')(1-\tau)]\beta s^*(\theta')}{\int_0^1 q^*(\theta)^\alpha d\theta}$$

Next, we show that

$$\Psi = \frac{\frac{d\left(q^*(\theta)^\alpha[1-s^*(\theta)(1-\tau)]\beta s^*(\theta)\right)}{d\alpha}}{q^*(\theta)^\alpha[1 - s^*(\theta)(1-\tau)]\beta s^*(\theta)}$$

is increasing in θ, which will imply that there is a threshold $\bar{\bar{\theta}} \in [0,1]$ such that for every $\theta \leq \bar{\bar{\theta}}$, the utility of a blogger with ability θ is decreasing in α, and for every $\theta > \bar{\bar{\theta}}$, the utility of a blogger with ability θ is increasing in α. To that end, with some abuse of notation we denote $s = s^*(\theta)$ and $q = q^*(\theta)$ and note that

$$\begin{aligned}\frac{d\left(q^\alpha[1 - s(1-\tau)]\beta s\right)}{d\alpha} &= \beta\left(\alpha q^{\alpha-1}\frac{dq}{d\alpha}[1 - s(1-\tau)]s - q^\alpha\frac{\partial s}{\partial \alpha}(1-\tau)s + q^\alpha[1 - s(1-\tau)]\frac{\partial s}{\partial \alpha}\right)\\ &= \beta q^{\alpha-1}\left(\alpha\frac{dq}{d\alpha}[1 - s(1-\tau)]s - q\frac{\partial s}{\partial \alpha}(1-\tau)s + q[1 - s(1-\tau)]\frac{\partial s}{\partial \alpha}\right)\end{aligned}$$

and therefore

$$\begin{aligned}\Psi &= \frac{\alpha\frac{dq}{d\alpha}[1-s(1-\tau)]s+q[1-s(1-\tau)]\frac{\partial s}{\partial\alpha}-q\frac{\partial s}{\partial\alpha}(1-\tau)s}{q[1-s(1-\tau)]s}\\ &= \frac{\alpha\frac{dq}{d\alpha}[1-s(1-\tau)]s+q\frac{\partial s}{\partial\alpha}(1-2s(1-\tau))}{q[1-s(1-\tau)]s}\\ &= \alpha\frac{\frac{dq}{d\alpha}}{q}+\frac{\frac{\partial s}{\partial\alpha}(1-2s(1-\tau))}{[1-s(1-\tau)]s}.\end{aligned}$$

Next recall that in proposition 3 and its proof we showed that $\frac{\frac{dq(\theta)}{d\alpha}}{q(\theta)}$ and $\frac{\frac{\partial s}{\partial\alpha}}{s}$ increase in θ. Therefore, to show that Ψ increase in θ it is sufficient to show that $\frac{(1-2s(1-\tau))}{[1-s(1-\tau)]}$ increases in θ, which always holds because $\frac{(1-2s(1-\tau))}{[1-s(1-\tau)]}$ decreases in s.

To prove that the threshold is strictly positive we note that from proposition 3 we know that there exists $\overline{\theta}\in(0,1)$ such that $n^*(\theta)$ is increasing in α if $\theta>\overline{\theta}$ and decreasing otherwise. It is then sufficient to note that $[s(\theta)\tau+b(\theta)]\beta s(\theta)$ is increasing in s and thus decreasing in α.

Part 3. Part 1 together with the observation that $q(\theta)$ is increasing in θ (see part 1 of Proposition 1), implies that aggregate welfare of followers increases in α, keeping the function $q(\theta)$ constant. Furthermore, in view of proposition 2 we know that when α increases $q(\theta)$ increases for all θ. Hence, aggregate welfare of followers increases even further. These two observations are easily adapted to total surplus.

Proof of Proposition 5. We start with the first part of the proposition. The equilibrium price to influencer i is

$$\hat{p}_i(\hat{\mathbf{s}})=\gamma\hat{n}_i(\hat{\mathbf{s}})\tau\beta. \tag{A3}$$

The profits to influencer i, by choosing $\hat{s}_i$, are

$$\hat{\Pi}_i(\hat{\mathbf{s}}) = \hat{p}_i(\hat{\mathbf{s}})\hat{s}_i=\gamma\hat{n}_i(\hat{\mathbf{s}})\tau\beta\hat{s}_i,$$

Influencer i selects $\hat{s}_i$ in order to

$$\max_{\hat{s}_i}\gamma\tau\beta\hat{n}_i(\theta,\hat{\mathbf{s}})\hat{s}_i$$

We have that

$$\frac{\partial\hat{\Pi}_i(\hat{\mathbf{s}},\hat{p}_i)}{\partial\hat{s}_i}=\gamma\tau\beta\hat{n}_i(\hat{\mathbf{s}})+\gamma\tau\beta\frac{\partial\hat{n}_i(\hat{\mathbf{s}})}{\partial\hat{s}_i}\hat{s}_i. \tag{A4}$$

In an interior equilibrium, influencer i with ability $\theta_i=\theta$ will select $\hat{s}_i=\hat{s}(\theta)$ so that $\frac{\partial\hat{\Pi}_i(\hat{\mathbf{s}})}{\partial\hat{s}_i}|_{\hat{s}_i=\hat{s}(\theta)}=0$. Developing expression 4 for a symmetric an interior equilibrium we have that

$$\frac{\partial\hat{\Pi}_i(\hat{\mathbf{s}})}{\partial\hat{s}_i}|_{\hat{s}_i=\hat{s}(\theta)}=0$$

if and only if

$$\theta_i + (1 - \hat{s}_i (1 - \gamma\tau - (1-\gamma) C)) (1-\beta) - \alpha \hat{s}_i (1 - \gamma\tau - (1-\gamma) C) (1-\beta) = 0. \tag{A5}$$

Hence

$$\hat{s}(\theta) = \min \left\{ \frac{\theta + 1 - \beta}{(1-\beta)(\alpha+1)(1 - \gamma\tau - (1-\gamma) C)}, 1 \right\}.$$

We now turn to the second part of the Proposition: We prove that $s^*(\theta) < \hat{s}(\theta)$ for all θ. Absent transparency $s^*(\theta) \in (0,1)$ for all θ. Hence, if for a specific θ, transparency leads to $\hat{s}(\theta) = 1$, then the claim holds for influencers with ability θ. Suppose, next, that after the policy $\hat{s}(\theta) \in (0,1)$ for some θ. We know that

$$\hat{s}(\theta) = \frac{\theta + 1 - \beta}{(1-\beta)(\alpha+1)(1 - \gamma\tau - (1-\gamma)C)}.$$

Furthermore, from the FOC above,

$$\hat{q}(\theta) - \alpha\hat{s}(\theta)(1-\beta)(1 - \gamma\tau - (1-\gamma)C) = 0,$$

and since $C \geq \tau$,

$$\hat{q}(\theta) - \alpha\hat{s}(\theta)(1-\beta)(1 - \gamma\tau - (1-\gamma)C) > \hat{q}(\theta) - \alpha\hat{s}(\theta)(1-\beta)(1-\tau).$$

Hence

$$\hat{q}(\theta) - \alpha\hat{s}(\theta)(1-\beta)(1-\tau) < 0$$

Now, take the FOC prior intervention for the same θ influencer, we have that

$$[q(\theta) - \alpha s(\theta)(1-\beta)(1-\tau)][1 - s(\theta)(1-\tau)] - s(\theta)q(\theta)(1-\tau) = 0$$

but if we evaluate this at the post intervention $\hat{s}$ and so $\hat{q}(\theta)$ we see that the first term is negative and therefore the all expression is negative. Concavity of the objective function implies that the $s(\theta)$ prior intervention must be lower than the one post intervention.

Proof of proposition 6.

Part 1. Define $\Delta(\theta) = q^*(\theta) - \hat{q}(\theta)$. In what follows we show that $\Delta(\theta) > 0$ for all θ, which is part 1 of the proposition. In addition we show that there exists $\breve{\theta} \in [0,1]$ such that $\Delta(\theta)$ increases in θ for all $\theta < \breve{\theta}$ and decreases in θ for all $\theta > \breve{\theta}$, where $\breve{\theta} \in [0,1]$ is the equilibrium threshold under transparency, such that influencers with $\theta > \breve{\theta}$ post only sponsored content and the other influencers post at least some organic content. Formally, $\breve{\theta}$ is determined as the influencer $\theta = \breve{\theta}$ so that $\hat{s}(\breve{\theta}) = 1$ and $\hat{s}(\theta) < 1$ for all $\theta < \breve{\theta}$; if such θ does not exist, then set $\breve{\theta} = 1$.

We first show that both of the proposition's claims are true for all $\theta \leq \breve{\theta}$, then we show that the claims are true for all $\theta \geq \breve{\theta}$.

Step 1. Consider a $\theta \leq \breve{\theta}$.

Step 1.a. We first derive an explicit expression $\hat{q}(\theta)$. Recall that

$$\hat{q}(\theta) = \theta + 1 - \beta - \hat{s}(\theta)(1-\beta)[1 - \tau\gamma - (1-\gamma)C]$$

and interiority implies that

$$\hat{s}(\theta) = \frac{\theta + 1 - \beta}{(1-\beta)(\alpha+1)(1-\gamma\tau-(1-\gamma)C)}$$

and so

$$\hat{q}(\theta) = \theta + 1 - \beta - \frac{\theta + 1 - \beta}{(\alpha+1)}$$

or

$$\hat{q}(\theta) = [\theta + 1 - \beta]\frac{\alpha}{\alpha+1}$$

Step 1.b. Recall that

$$q(\theta) = \theta + (1-\beta)[s(\theta)\tau + 1 - s(\theta)].$$

or

$$q(\theta) = \theta + (1-\beta) - (1-\beta)s(\theta)(1-\tau).$$

Define $\Delta(\theta) = q^*(\theta) - \hat{q}(\theta)$ and note that

$$\Delta(\theta) = \frac{\theta + 1 - \beta}{1+\alpha} - (1-\beta)(1-\tau)s^*(\theta).$$

Step 1.c. We show that $\Delta(\theta)$ is increasing in θ (and so this prove the second part of the proposition for all $\theta \leq \breve{\theta}$). To see this note that

$$\frac{d\Delta(\theta)}{d\theta} = \frac{1}{1+\alpha} - (1-\beta)(1-\tau)\frac{ds^*(\theta)}{d\theta} > 0$$

if and only if

$$\frac{ds^*(\theta)}{d\theta} < \frac{1}{(1+\alpha)(1-\beta)(1-\tau)}$$

To show that the above inequality holds, we return to the unregulated market FOC

$$q^*(\theta) - \alpha(1-\tau)(1-\beta)\frac{s^*(\theta)[1-s^*(\theta)(1-\tau)]}{[1-2s^*(\theta)(1-\tau)]} = 0$$

with some rearranging we get

$$\frac{ds^*(\theta)}{d\theta} = \frac{1}{(1-\beta)(1-\tau)(1+\alpha) + \frac{\alpha(1-\tau)(1-\beta)}{[1-2s^*(\theta)(1-\tau)]^2}[2s^*(\theta)(1-\tau)[1-s^*(\theta)(1-\tau)]]}$$

To complete the proof of this step, note that the second term of the denominator of $\frac{ds^*(\theta)}{d\theta}$ is positive because prior to intervention $s^*(\theta) < 1/(2(1-\tau))$. Hence, $\frac{ds^*(\theta)}{d\theta} < \frac{1}{(1+\alpha)(1-\beta)(1-\tau)}$ as required.

Step 1.d. We now conclude and show that $\Delta(\theta) > 0$ for all $\theta \leq \breve{\theta}$. Since $\Delta(\theta)$ is increasing in θ for all $\theta \leq \breve{\theta}$, we just need to show that $\Delta(0) > 0$. To see this note that using the FOC for $s^*(\theta)$ and specializing it for $s^*(0)$ we obtain that

$$s^*(0) = \frac{1}{(2+\alpha)(1-\tau)}$$

and so

$$\Delta(0) = \frac{1-\beta}{1+\alpha} - (1-\beta)(1-\tau)s^*(0) = \frac{1-\beta}{1+\alpha} - \frac{(1-\beta)}{(2+\alpha)} > 0$$

This concludes the proof that $q(\theta) > \hat{q}(\theta)$ for all $\theta \leq \breve{\theta}$.

Step 2. Consider a $\theta \geq \breve{\theta}$.

Step 2.a. We first derive an explicit expression $\hat{q}(\theta)$. Recall that $\hat{s}(\theta) = 1$ and so

$$\hat{q}(\theta) = \theta + (1-\beta)[\tau\gamma + (1-\gamma)C]$$

Step 2.b. Recall that

$$q(\theta) = \theta + (1-\beta)[s(\theta)\tau + 1 - s(\theta)].$$

Define $\Delta(\theta) = q^*(\theta) - \hat{q}(\theta)$ and note that

$$\Delta(\theta) = (1-\beta)[s^*(\theta)\tau + 1 - s^*(\theta) - \tau\gamma - (1-\gamma)C]$$

Step 2.c. Since $s^*(\theta)$ is increasing in θ, it follows that $\Delta(\theta)$ is decreasing in θ, for all $\theta \geq \breve{\theta}$.

Step 2.d. We now conclude and show that $\Delta(\theta) > 0$ for all $\theta \geq \breve{\theta}$. Since $\Delta(\theta)$ is decreasing in θ for all $\theta \leq \breve{\theta}$, we just need to show that $\Delta(1) > 0$. To see this note that $s^*(1) \leq \frac{1}{2(1-\tau)}$ and so

$$\Delta(1) = (1-\beta)[1 - s^*(1)(1-\tau) - \tau\gamma - (1-\gamma)C] \geq (1-\beta)[\frac{1}{2} - \tau\gamma - (1-\gamma)C] \geq 0$$

where the last inequality follows because, since $\tau < 1/2$ and $C \leq 1/2$, then $\tau\gamma + (1-\gamma)C \leq 1/2$. This concludes the proof of part 1 of proposition 6.

Part 2. Note that followers' welfare prior and post policy read:

$$W_F = \int n^*(\theta)q^*(\theta)d\theta \text{ and } \hat{W}_F = \int \hat{n}(\theta)\hat{q}(\theta)d\theta$$

We first claim that when $\alpha = 0$ the introduction of transparency decreases followers' welfare and total surplus. To see that, it is sufficient to note that when $\alpha = 0$, for all θ, $n^*(\theta) = \hat{n}(\theta)$ and $q^*(\theta) < \hat{q}(\theta) = 1$. By continuity, the result holds for all $\alpha < \overline{\alpha}$ for some $\overline{\alpha} > 0$.

Next note that the assumption $(1+\alpha)(1-\gamma\tau-(1-\gamma)C) \geq \frac{2-\beta}{1-\beta}$ implies that $\breve{\theta}=1$ and so part 1 implies that $\Delta(\theta)=q^*(\theta)-\hat{q}(\theta)$ is increasing in θ for all θ. Hence, $\hat{q}(\theta)$ is flatter than $q(\theta)$ and so the distribution of readership $n^*(\cdot)$ FOSD the distribution of readership post policy $\hat{n}(\cdot)$. Hence, since $q^*(\theta)$ is increasing in θ,we obtain that

$$W_F = \int n^*(\theta)q^*(\theta)d\theta > \int \hat{n}(\theta)q^*(\theta)d\theta.$$

We now use part 1 (i.e., $\Delta(\theta)=q^*(\theta)-\hat{q}(\theta)>0$ for all θ), to conclude that

$$W_F = \int n^*(\theta)q^*(\theta)d\theta > \int \hat{n}(\theta)q^*(\theta)d\theta > \int \hat{n}(\theta)\hat{q}(\theta)d\theta = \hat{W}_F.$$

We now turn to total surplus. Recall that

$$TS = \int n^*(\theta)[\theta+1-s^*(\theta)+s^*(\theta)\tau]$$

and

$$\hat{TS} = \int \hat{n}(\theta)[\theta+1-\hat{s}(\theta)+\hat{s}(\theta)(\tau\gamma+(1-\gamma)C]$$

It is immediate to check that $q^*(\theta) > \hat{q}(\theta)$ for all θ implies that

$$\theta+1-s^*(\theta)+s^*(\theta)\tau > \theta+1-\hat{s}(\theta)+\hat{s}(\theta)(\tau\gamma+(1-\gamma)C$$

for all θ. And so, replicating the same steps for the readers' welfare, we obtain:

$$\begin{aligned} TS &= \int n^*(\theta)[\theta+1-s^*(\theta)+s^*(\theta)\tau] > \int \hat{n}(\theta)[\theta+1-s^*(\theta)+s^*(\theta)\tau] \\ &> \int \hat{n}(\theta)[\theta+1-\hat{s}(\theta)+\hat{s}(\theta)(\tau\gamma+(1-\gamma)C] = \hat{TS}. \end{aligned}$$