5.   Market incentives and advertising disclosure regulations

Daniel Ershov

1   INTRODUCTION

Influencers – popular social media users who monetise the attention they receive through commercial arrangements with brands – are often portrayed as "sincere" or "authentic" advertisers.[1] However, it is important to remember that they are economic agents whose behaviour responds to incentives. At a high level, an influencer chooses how much content to post, what type of content to post (sponsored content produced in partnership with brands, or organic content), and whether or not to disclose any sponsored content they posted. When choosing a particular action, influencers weigh its benefits and costs relative to alternatives. For example, when deciding whether to post an additional piece of sponsored content or organic non-paid content, creators weigh the benefits of additional income from payments from brands, against the potential cost of reduced attention and trust from followers. Ultimately, they are not unlike traditional content creators, such as TV stations or magazines, which mix advertising content and "organic" non-paid content. Legally, the boundaries of how influencers are defined, or the regulations that govern them, are murky,[2] but the average influencer still responds to incentives.

---

[1]   Sophie Elmhirst, "'It's genuine, you know?': why the online influencer industry is going 'authentic'" (*The Guardian*, 5 April 2019) <https:// www .theguardian .com/ media/ 2019/ apr/ 05/ its -genuine -you -know -why -the -online -influencer-industry-is-going-authentic> accessed 23 June 2024.

[2]   See, for example, Catalina Goanta and Isabelle Wildhaber, "In the business of influence: Contractual practices and social media content monetisation" (2019) Schweizerische Zeitschrift für Wirtschafts-und Finanzmarktrecht, SZW, 4; Catalina Goanta and Sofia Ranchordás, "The regulation of social media influencers: An introduction" in Catalina Goanta and Sofia Ranchordás (eds), *The Regulation of Social Media Influencers* (Edward Elgar 2020).

Changes in regulations associated with sponsored content inevitably affect the incentives of influencers and their decision-making. A particular example is the stronger proposed regulations on the disclosure of advertising content on social media, such as the recent 2023 French consumer protection laws. These laws restrict the set of products that influencers can advertise (e.g., no tobacco products or medical devices), as well as specifying precise disclosure terms that should be posted with all sponsored content, at the risk of facing fines, or being forced to post a "black banner" stating that they failed to comply with the rules on their account.[3] If influencers successfully disclose existing undisclosed advertising, following this law, it may affect the relationship between influencers and their followers. For example, it could negatively affect trust due to a greater understanding of the quantity of sponsored content in their feed. At the same time, knowing they will have to disclose all or some of their sponsored content, influencers may change their posting strategy *ex ante*. They may post more or less content of any kind, or change the quality of the content that they post.

The goals of most regulations targeting influencers are clear – protecting consumers and maximising their well-being. French social media advertising regulations clearly state that their goal is to protect consumers and social media users from fraud or malpractice.[4] Similarly, the US Federal Trade Commission's (FTC) jurisdiction over advertising falls under Section 5 of the FTC Act, which allows the FTC to levy fines that compensate consumers for any harms experienced from misleading, deceptive, or improperly disclosed endorsements.[5] However, it is crucial to carefully think about how any changes in regulations will affect influencer incentives ahead of time. This chapter surveys the recent literature on influencers as economic agents, and how their incentive schemes and actions influence the implementation of advertising

---

[3] Angelique Chrisafis "French social media influencers feel the heat over new law on paid content" (*The Guardian*, 11 November 2023) <https:// www .theguardian .com/ world/ 2023/ nov/ 11/ france -social -media -influencers -feel -heat -law-paid-content> accessed 23 June 2024.

[4] Proposition de loi n°790 <https:// www .assemblee -nationale .fr/ dyn/ 16/ textes/l16b0790_proposition-loi> accessed 23 June 2024.

[5] Federal Trade Commission "FTC's Endorsement Guides: What People Are Asking" (2023), <https:// www .ftc .gov/ business -guidance/ resources/ ftcs -endorsement-guides-what-people-are-asking> accessed 23 June 2024.

disclosure regulations.[6] In particular, the focus is on three recently published or forthcoming economics papers.[7]

Each of these papers formally studies a different aspect of the relationship between influencers and followers, and a different mechanism through which changes in disclosure requirements could affect this relationship. As is standard in economics and marketing, their arguments are formalised through game-theoretical mathematical modelling. Each paper develops a situation, a "game", where "players" (i.e., influencers or followers) make decisions that affect one another. Each game defines the actions that players could take, and the effects or outcomes of these actions on the players – the "payoffs" players could receive. Both the actions and payoffs are defined explicitly through mathematical expressions. Each game is then solved to discover the "equilibrium", the combination of actions from which no player has an incentive to deviate. For example, this may be the condition under which an influencer chooses to post sponsored undisclosed content (over organic content), and the follower chooses to pay attention to that content (over not paying attention to that content). Although these models are necessarily abstract, their goal is to capture the fundamental features of reality and generate actionable insights into the decision processes of the players involved.

The remainder of this chapter will survey each of the three papers in turn, using simple, non-technical language to outline their theoretical arguments

---

6   The literature on influencers in economics and management is rapidly growing. Additional recent related theoretical papers see Mohamed Mostagir and James Siderius, 'Strategic reviews' (2023) 69 Management Science 904; Amy Pei and Dina Mayzlin, 'Influencing social media influencers through affiliation' (2022) Marketing Science 593. For recent published empirical work see Christian Hughes, Vanitha Swaminathan and Gillian Brooks, 'Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns' (2019) Journal of Marketing 78; Zike Cao and Rodrigo Belo 'Effects of explicit sponsorship disclosure on user engagement in social media influencer marketing', MIS Quarterly, *forthcoming*. This literature also relates to an older literature on firm-driven word-of-mouth communications, see Dina Mayzlin, 'Promotional chat on the Internet' (2006) Marketing Science 155; and David Godes and Dina Mayzlin 'Firm-created word-of-mouth communication: Evidence from a field test' (2009) Marketing Science 721.

7   Matthew Mitchell 'Free ad(vice): internet influencers and disclosure regulation' (2021) RAND Journal of Economics 3; Itay P. Fainmesser and Andrea Galeotti 'The market for online influence' (2021) American Economic Journal: Microeconomics 332; Daniel Ershov and Matthew Mitchell 'The effects of advertising disclosure regulations on social media: Evidence from Instagram' RAND Journal of Economics, forthcoming.

or empirical methodology. Then, this chapter will discuss and compare the three approaches and evaluate what empirical evidence we have to support or dismiss any of the arguments. Finally, the chapter will conclude with some speculative thoughts about how to improve advertising regulations.

2   REPUTATION BUILDING

Mitchell's paper[8] considers the dynamics in the repeated interactions between a representative (i.e., average) influencer and a representative follower. Their setting considers an influencer making the same decision over a number of "periods" (i.e., posting opportunities) – whether to post sponsored or organic content. The two types of posts offer the influencer different benefits (i.e., revenues), and the follower different "utility" (i.e., satisfaction from reading these). A sponsored post produces higher revenues for the influencer but has zero utility for the follower, as compared with an organic post, which generates no revenues for the influencer but some positive utility for the follower. Both types of posts come at some cost to the influencer, representing the time and effort required to make creative content online. Clearly, this introduces the central trade-off for the influencer – how much content should they post that is better for them but not good for the followers, and when should they post more of that content? This trade-off is present for many influencers, whose followers often actively dislike the ads they post.[9]

Once the influencer chooses to post a particular type of post, the follower chooses whether to read it. In a world without regulations, the follower does not know whether a post is sponsored or organic ahead of reading it. However, they have information about all previous posts the influencer has posted and whether these have been sponsored or organic. The key idea in this chapter is that ahead of choosing whether to pay attention to any new post, the follower considers this past information and forms a prediction about whether the next post from the influencer is going to be sponsored or organic. The influencer knows that this is the follower's thought process. This exacerbates the previ-

---

8   Mitchell (n 7).
9   For instance, Kylie Jenner's followers often mock or criticise her ads for various drinks or supplements. For examples, see Heather Gardner, 'We're confused — why is near billionaire Kylie Jenner promoting detox tea on Instagram?' (*Yahoo News*, 31 July 2018) <https:// www .yahoo .com/ entertainment/ confused -near -billionaire -kylie -jenner -promoting -detox -tea -instagram -185140304 .html> accessed 23 June 2024; Katie Francis, 'Kardashian fans mock Kylie Jenner for new drink ad and agree she's 'not selling it' in new photos of star in tight dress' (*The Sun,* 14 May 2023) <https://www.the-sun.com/entertainment/8114983/kardashian -kylie-jenner-mock-drink-ad-tight-dress/> accessed 23 June 2024.

ously discussed tension in the model – although the influencer would ideally post only sponsored content, they cannot do so. The follower will start to anticipate that they will only see sponsored content from the influencer, and will stop paying attention as a result.

What comes out of the "equilibrium" of the model is the need for the influencer to establish a good reputation with the follower. The influencer's optimal strategy is to initially provide the follower with a large volume of high-quality organic content, building their reputation, and establishing follower predictions that content will continue to be organic. Then, once the reputation is established, the influencer can cash in on that reputation (and on the follower's expected attention) by posting a large amount of sponsored content. Then, as the influencer's reputation dwindles and the follower is on the verge of not paying attention, the influencer resumes providing organic content. Interestingly, in the model, the follower knows and understands this strategy and chooses to pay attention to the influencer's posts even during the sponsorship-heavy periods. The reason for this is that they know that the influencer will have to revert to posting more organic content in the future, and the influencer posts just enough content to keep the follower's attention. This setup captures some of the fundamental dynamics of the market – the vast majority of Kim Kardashian's followers do not stop following her once she posts some sponsored content, because they know that she will resume posting content they enjoy soon. They anticipate being advertised some products that Kim Kardashian is promoting in return. Moreover, influencers early on in their careers "invest" more in higher-quality content with less advertising. They do so both to convince followers that they are worth following for the long run, and also to convince advertisers that they have enough follower attention to monetise.

Disclosure regulations enter the model by reducing the benefits that the influencer receives from posting a sponsored post. This is meant to capture the lower attention that an average disclosed sponsored post by influencers receives from their followers.[10] Interestingly, in this setting a reduction in the benefits of a sponsored post will not necessarily reduce the total amount of sponsored content influencers create. The reason for this is that the influencer has to cover the costs they incurred in gaining follower attention through organic posts. Therefore, as per-post benefits fall, the influencer wants to create more sponsored content to earn revenue and recover the investment they previously put into organic content. This is the case for established influencers, but there is also a distinct effect on new influencers. The reputation influencers build with their followers through their organic content is predicated on the

---

[10]   See additional evidence in Ershov and Mitchell (n 7), discussed below.

future benefits they will receive when they start "cashing in" through sponsored content. As a result, a reduction in the benefits coming from sponsored content reduces the expected future benefits for new influencers and disincentivises them from producing additional organic content earlier on. Therefore, advertising disclosure regulations in this model can both reduce the amount of organic content in the market and increase the amount of sponsored content.

Mitchell proposes several alternative policies that regulators could follow to maximise follower well-being – i.e., to increase organic content and minimise sponsored content that followers do not like. The most important suggestion is to have stricter disclosure guidelines and enforcement for new influencers as compared with established influencers. Put another way, more sponsored content by less established influencers will be disclosed, and less content by more established influencers will be disclosed. The idea is based on the mechanism described above. Laxer standards for established influencers maximise the benefits new influencers anticipate receiving from sponsored content in the future, which improves their incentives to post organic content earlier on. The stricter disclosure guidelines for less established influencers also dissuade them from producing more sponsored content. Interestingly, this is the opposite of sponsored content regulations in most countries. Most of the time, regulations and enforcement are very strict for the "top" (most popular) influencers. For enforcement, this is likely the case because cases are driven by consumer complaints, and more established influencers have more followers who could complain. For example, when the FTC in the US sent warning letters to influencers in 2017, they only sent them to 90 mega-influencers and celebrities such as Naomi Campbell, rather than to a wider group of users.[11] These actions by the regulators are based on the idea that top influencers have the most eyeballs and that they may produce spillover effects on smaller influencers. However, it is not clear whether these spillovers fully materialised.[12] Moreover, it is plausible that the mass of all attention given to smaller influencers is actually larger than that given to the relatively few huge influencers that have been the focus of regulations thus far. Larger influencers may also

---

[11] David Ingram and Diane Bartz "FTC demands endorsement info from Instagram 'influencers'" (*Reuters,* 2017) <https:// www .reuters .com/ article/ us -usa -ftc -celebrities/ ftc -demands -endorsement -info -from -instagram -influencers -idUSKCN1BO2TE/> accessed 23 June 2024.

[12] There is no evidence that the effects of the warning letters expanded to the broader set of influencers, such as micro-influencers.

have inherent incentives to disclose sponsored content, as they are concerned about the reputational effects of non-disclosure and misleading customers.[13]

It is important to note that while the model in this chapter captures an important mechanism, it abstracts from many factors about the industry. This is a model where followers cannot be "deceived" about the content that they are consuming. In addition, this chapter only models the relationship between one influencer and one follower, without considering the choices followers make between different influencers. These aspects are discussed next.

3   INFLUENCER COMPETITION

The paper by Fainmesser and Galeotti[14] models interactions between multiple influencers and followers.[15] In this setting, the followers are identical, but influencers vary in terms of the "inherent" quality of content they post – some influencers post better (organic and sponsored) content than others that followers simply like more.[16] In this setup, followers should in principle all select the influencers with the highest inherent quality. However, the paper introduces friction where followers are not necessarily aware of all influencers in the market, and may not be able to find the highest-quality influencers. High inherent quality influencers are still going to have the most followers, but not all followers are aware of them, and so some lower-quality influencers are also going to have followers. In the model, this produces a power-law (hockey-stick shaped) distribution of follower counts that mimics the real-world distribution

---

[13]   A French survey of influencer disclosure by the ARPP (a self-regulatory advertising association) found that while top influencers with more than 1 million followers correctly disclosed approximately 70% of the commercial content they posted in 2020, only 40% of commercial content posted by influencers with approximately 10,000 followers was correctly disclosed. See <https://www.arpp .org/ influence -responsable/ observatoire -influence -responsable/ > for additional details.

[14]   Fainmesser and Andrea Galeotti (n 7).

[15]   Marketers are also included in the model, though they play a relatively small role and are excluded from this discussion.

[16]   The language in Fainmesser and Galeotti (n 7) refers to recommendations that the influencers make – for example, for particular products. These can be either organic (i.e., not compensated and based on the influencer's own experiences), or sponsored (i.e., compensated). The assumption in the paper is that influencers select better products organically than with sponsorship, leading to better outcomes for followers who receive more organic recommendations. For simplicity and consistency of language, content is simply referred to as sponsored or organic.

of nano-, micro-, macro-, and mega-influencers – i.e., there are a small number of high-quality influencers with a very large number of followers, and a large number of influencers with a small number of followers.

The key choice for each influencer, as in Mitchell's paper,[17] is the share of sponsored content that they post. Sponsored posts are worse for followers than organic posts. In the baseline model, followers cannot distinguish between organic and sponsored content, and they simply choose to follow influencers based on the overall well-being they receive from these influencers (the sum of an influencer's inherent quality, and the quality of the content they post). It is important to note that influencers are competing with one another in this setting. An influencer knows that if they increase the share of sponsored content they post, their quality will fall, and they will receive fewer followers. The extent of the sensitivity of the number of followers to the share of sponsored content depends on the friction in the market described above. Without any friction, the model predicts that the optimal choice for influencers is to post no sponsored content, since followers would only follow the highest overall-quality influencer, and any sponsored content reduces overall quality. With friction, influencers know that reducing their quality by posting additional sponsored content will not necessarily drive all of their followers away. As such, influencers with higher inherent quality know that they attract more followers, and so they have additional followers to lose by posting more sponsored content. Therefore, in equilibrium, the model predicts that larger influencers, who generally have higher inherent quality, will exploit their position by posting more sponsored content. Notably, this is both consistent with real-world outcomes and with Mitchell's predictions, though those come from a completely different model and a different setup. Unlike Mitchell's model, this is a "static" model – it is not concerned with the evolution of influencers' content, but rather with a "snapshot" of the market at one point in time.

Fainmesser and Galeotti simulate disclosure regulations in the model by giving followers the ability to distinguish between sponsored and organic content and ignore sponsored content. They also introduce heterogeneity to the followers – some followers are more likely to ignore sponsored content than others. Holding everything else constant, followers who ignore sponsored content receive higher benefits from following the same influencer as before.[18] This reflects the ability of many followers on social media to simply ignore or scroll through sponsored content in their feeds without paying a substantial

---

[17]   Mitchell (n 7).
[18]   The reason for this is that they still benefit from the influencer's inherent quality, but are now also able to get higher benefits from their content by only paying attention to organic posts and not to sponsored posts.

amount of attention to it, especially if it is clearly disclosed. If the amount of sponsored content in the market does not change, this should increase follower well-being. However, the influencers respond to these regulations by changing the amount of sponsored content they post. In particular, influencers know that some of their followers now no longer care about how much sponsored content they post. This means that, in a world with transparency regulations, if they increase their share of sponsored content, they will lose fewer followers as compared with a world without transparency regulations. This creates perverse incentives, effectively reducing competition between influencers and increasing the amount of sponsored content in the market, as each influencer attempts to earn more money out of their remaining followers who pay attention to the sponsored content. This change in the share of ads has no effect on the followers who do not pay attention to sponsored content. But it does substantially reduce the well-being of the other followers who still consume advertising and who are now substantially worse off. Compared with a world where there is no mandated transparency, these followers are now "stuck" with a higher ad load influencer.

All in all, this paper produces another channel through which the introduction of mandated transparency affects influencers' incentives in a way that affects the utility of social media users. In this case, competition between influencers is the main channel through which the effects occur. One policy suggestion from Fainmesser and Galeotti is that reducing platform frictions will intensify competition between influencers and reduce the overall ad load in the market. In practice, this means the collection of additional data on social media users so that better influencers can be recommended to them. Of course, this introduces further issues, as additional data collection by platforms can result in worse outcomes for social media users for a variety of other reasons.

4   POROUS REGULATIONS

Ershov and Mitchell's paper[19] focuses on another aspect of the influencer and follower relationship that is not directly accounted for by the other papers discussed above, but that plays an important role in any proposed regulation of social media markets: the language of posts, the beliefs of followers about the posts they see, and the presence of undisclosed advertising in markets, even after disclosure regulations. Similar to Mitchell's earlier paper,[20] Ershov and Mitchell also consider a single representative influencer interacting with a rep-

---

[19] Ershov and Mitchell (n 7).
[20] Mitchell (n 7).

resentative follower, and similar to Fainmesser and Galeotti,[21] this is a "static"/ single snapshot model. As in the previous papers discussed in this chapter, the key choice by the influencer is whether to post sponsored or organic content. Sponsored content is better for the influencer but worse for the follower, and organic content is the opposite. Unlike the other papers, Ershov and Mitchell consider the actual content of posts. When an influencer chooses to post a sponsored post, this post has to include certain "sponsored language" – words that connote some commercial intent, or that are dictated directly by the brands that they are advertising for.[22] When an influencer chooses to post an organic post, the post includes "organic language". Some words between the two types of posts will overlap, but words that are common in sponsored posts will be rare in organic posts and vice versa. For example, if influencers can only use the two words "love" and "sale", "love" would be more likely to appear in organic posts and "sale" would be more likely to appear in sponsored posts, although some sponsored posts may also include the word "love".

A follower in this model does not know whether each post is organic or sponsored. The follower inspects each post briefly, and forms expectations/ beliefs about whether or not it is sponsored based on the language it includes. For example, if a post includes the word "sale", the follower infers that the post is more likely to be sponsored. Then, the follower pays attention to the post based on how sponsored they believe it is. Posts that followers believe to be more sponsored receive less attention. As before, the influencer's key choice is the share of sponsored posts. In equilibrium, the optimal number of sponsored posts is based on the average amount of attention they expect followers will pay to sponsored posts compared with organic posts. Put another way, an influencer will not post additional sponsored content if they know that followers are going to reduce their attention. As in the previous papers, although this is an abstract model, it represents essential features of influencers' content selection decisions, and their desire to maximise overall follower attention and engagement.

Disclosure regulations in this model reveal a portion of sponsored posts as sponsored. This is done by including disclosure words (i.e., "#AD") that do not exist in organic posts and that perfectly inform followers that the post is not organic. For simplicity, the paper assumes that a random portion of sponsored content is revealed in such a manner. This is meant to capture the porousness of real-world regulations, where (i) regulations often have inherent uncertainty

---

21   Fainmesser and Galeotti (n 7).
22   See Goanta and Wildhaber (n 2) for examples of such contractual relationships.

about what content should be disclosed and are updated regularly,[23] and (ii) different brands contracting with influencers have different expectations about disclosure.

This form of disclosure regulations has two effects on followers' perceptions and behaviour in the model. On the one hand, followers now know that the disclosed sponsored posts are actually sponsored and as a result will choose to not pay any attention to those. On the other hand, their perception of posts without disclosure tags also changes. On average, followers are going to have higher "trust" in undisclosed sponsored posts, because they know that a portion of sponsored posts have been caught by the filter. As a result, they will pay more attention to all undisclosed posts, including undisclosed sponsored posts that "escape" disclosure. *Ex ante*, it is not clear which effect will dominate the other. Therefore, the effect of disclosure regulations on the share of sponsored content in the market, including on the share of undisclosed sponsored content, is inherently ambiguous, since influencers make choices based on total follower attention to sponsored content. If the lack of attention to disclosed posts dominates, influencers will start posting fewer sponsored posts compared with the baseline setting without disclosure regulations. If, however, the additional attention provided to non-disclosed sponsored posts dominates, influencers will post more sponsored content, with the expectation that some of these additional posts will escape disclosure and end up as undisclosed posts that capture substantial consumer attention.

5   EMPIRICAL EVIDENCE

There is a variety of existing empirical literature on the value of disclosure and whether disclosure negatively or positively affects engagement.[24] The results are often mixed, partially because of the difficulty of effectively establishing causality in real-world data, and the challenge of accurately simulating social media networks in lab settings. In addition, context matters for disclosure

---

[23]   See more below on German regulations. In the US, the FTC released corrections and updates clarifying the existing rules regarding disclosure in 2017, 2019, 2020 and 2023.

[24]   See Hughes, Swaminathan and Brooks (n 6); Navdeep S. Sahni and Harikesh S. Nair, 'Sponsorship disclosure and consumer deception: Experimental evidence from native advertising in mobile search' (2020) Marketing Science 5; Zeynep Karagür, Jean-Michel Becker, Kristina Klein and Alexander Edeling, 'How, why, and when disclosure type matters for influencer marketing' (2022) International Journal of Research in Marketing 313; Fine F. Leung, Flora F. Gu, Yiwei Li, Jonathan Z. Zhang and Robert W. Palmatier, 'Influencer marketing effectiveness' (2022) Journal of Marketing 93; and Cao and Belo (n 6).

effects. While an "average" lifestyle influencer does experience negative engagement effects from posting disclosed ads, niche influencers who post about very specialised topics (i.e., ski equipment) and have dedicated audiences do not. Nonetheless, as described below, the average engagement for disclosed sponsored content is lower than undisclosed content by the same influencer, and the perception in the industry is that disclosure is bad for engagement, leading to low disclosure rates.[25]

In addition to their theoretical model, Ershov and Mitchell empirically evaluate the effects of changes in advertising disclosure regulations on the behaviour of influencers and on the interactions between influencers and followers. They evaluate the effects of changes in German influencer regulations. Germany has been a relatively early adopter of stricter regulations on influencers. Starting in early 2016, there were several initiatives, including by the German parliament, to impose stricter requirements for the labelling of advertising content on social media. In November 2016, the German Die Medienanstalten, a consortium of 14 state regulators, provided a set of new guidelines that mandate the disclosure of ad content.

The regulatory changes were succeeded by legal actions taken against several influencers following complaints from consumer protection groups, and resulting in fines. Among other cases, in 2017 a sports YouTube influencer was fined for non-disclosure of advertising.[26] Interestingly, legal activity in Germany differed from other countries that also introduced additional regulations on influencer activity. In the US and France, for example, regulators (e.g., the FTC) also introduced new guidelines about the disclosure of advertising content. However, only the largest influencers were ever targeted. The FTC sent warning letters to the most prominent mega-influencers (primarily celebrities with millions of followers).[27] By comparison, the German influencers that received fines were small, even in comparison with other existing German influencers. For example, one prominent early non-disclosure case was against an influencer with only 50,000 followers.[28] This is small compared

---

[25]   See, for example, Arunesh Mathur, Arvind Narayanan and Marshini Chetty, 'Endorsements on Social Media: An Empirical Study of Affiliate Marketing Disclosures on YouTube and Pinterest' (2018) Proceedings of the ACM on Human-Computer Interaction 1, for evidence of very low disclosure rates for affiliate marketing on YouTube and Instagram.

[26]   Dirk Spacek 'Newsletter no. 119: Digital Influencer Marketing – Worldwide Legal Developments' (*Walderwyss Newsletter,* October 2017) <https:// www .walderwyss.com/user_assets/publications/2153.pdf> accessed 23 June 2024.

[27]   Ingram and Bartz (n 12).

[28]   Rossana Ducato, "One hashtag to rule them all? Mandated disclosures and design duties in influencer marketing practices. In Mandated disclosures

with some of the larger influencers in Germany, who have millions of followers. The end result of this is an environment where there is widespread concern by many influencers regarding potential legal action if they do not disclose the commercial content they post.

It is important to note that there appears to have been substantial legal uncertainty in Germany regarding the extent of regulations and enforcements. Different regional courts produced judgments that varied in severity, ranging from judges claiming that influencers must disclose every single one of their posts as ads (including non-explicitly commercial posts), to judges dismissing disclosure as unimportant. In that sense, disclosure rules are not "airtight" and there is no complete information about what should and should not be disclosed as sponsored. This means that, as in the theory model of Ershov and Mitchell's paper,[29] some sponsored posts could "escape" disclosure even under the stricter regulatory regime.

To look at the effects of changes in disclosure regulations on influencer behaviour and follower engagement, Ershov and Mitchell use Instagram data from CrowdTangle.com.[30] They look at a random sample of 6,000 German influencers and their posts from 2014 to 2020. To benchmark the behaviour of these influencers, they look at a sample of 6,000 Spanish influencers for the same period of time. Unlike Germany, Spain's regulations on commercial social media activity have not changed during the 2010s. As such, it is possible to think about this setting as a "natural experiment", where Spanish influencers serve as a "control" group for the German influencers. Comparing the behaviour of the two before and after the regulatory environment changed in Germany should allow for identifying the effects of changes in regulations on the market.

Part of the challenge of this empirical exercise is that even if disclosure regulations fully disclose all sponsored content, there is a substantial period of time in Germany without regulations where sponsored content is hidden. Sponsored content is also hidden in Spain for the duration of the sample period. To uncover sponsored posts, Ershov and Mitchell use a machine learning procedure. They trained a classification model on a sample of disclosed-sponsored and undisclosed German posts. The goal of the model is to use the text of

---

and design duties in influencer marketing practices." Catalina Goanta and Sofia Ranchordás (eds), *The Regulation of Social Media Influencers* (Edward Elgar, 2020).

29 Ershov and Mitchell (n 7).

30 CrowdTangle is a company owned by Meta that provides API-like access to public Facebook and Instagram data to researchers.

a post to predict whether it is a disclosed-sponsored post.[31] The training of the model is done to maximise the number of disclosed-sponsored posts correctly predicted, rather than the number of undisclosed posts correctly predicted. As a result, there are a number of posts that the model predicts are disclosed sponsored posts because of their text, even though they are not actually disclosed. These are the undisclosed sponsored posts. After training the model on a sample, the authors apply its predictions to the remaining German and Spanish data.

They find that German regulations increased disclosure in Germany dramatically relative to Spain. This is – in and of itself – a notable finding which suggests that influencer disclosure decisions are responsive to regulatory changes, if these changes are strict enough.[32] At the same time, they also show that the amount of sponsored content increases in Germany relative to Spain after the regulatory environment in Germany strengthens. The magnitudes are substantial. The main results in the paper suggest that the share of sponsored content increases by at least 12%, relative to the baseline. Interestingly, although disclosure increases, because of the increase in sponsorship, the share of non-disclosed posts that are sponsored does not fall, and in fact increases. This means that consumers are exposed to more undisclosed sponsored content after regulations. These results are consistent with the mechanism outlined in the theory model above – as the regulatory environment becomes stricter but remains porous and uncertain, influencers may want to increase the amount of sponsored content they post, since some of it is bound to escape regulatory attention and detection.

Additional analysis examines engagement. Engagement for German influencers falls after regulations. Both likes and comments decrease for the average influencer, as compared with their Spanish counterparts. This may be simply because of the increase in sponsored content, or because of the increase in disclosed sponsored content, which followers may not be happy with. The study shows that engagement also falls when only looking at undisclosed content

---

[31]   The authors also convert the text of posts from German or Spanish into multilingual embedding space, representing each post with a 300-dimensional vector, which captures its place in linguistic-meaning space. This helps deal with translation challenges, as well as other issues coming from variations in language over time.

[32]   Popular press discussions often dismiss the effectiveness of disclosure regulations at changing influencer behaviour, suggesting they would not even affect disclosure. For example, see: Amelia Tait, 'Forcing social-media influencers to be clear about #ads? Good luck with that' (*The Guardian*, 25 January 2019) <https://www.theguardian.com/commentisfree/2019/jan/25/social-media-influencers-clear-ads-celebrities-authorities> accessed 23 June 2024.

(sponsored and non-sponsored), suggesting that the former story is the case here. Although there is no direct analogy between engagement and consumer welfare, it does suggest that followers are not better off with the additional ads. Additional evidence from forthcoming research suggests that there are notable heterogeneities in both disclosure rates and the effects of the regulatory change on influencers of different sizes and in different "industries". For example, larger influencers were disclosing more content prior to regulations and are less affected by the changes.

Further analysis shows that the ratio in engagement between undisclosed sponsored and organic content falls after regulations in Germany, moving to close to one in the treated period (i.e., one organic like for one undisclosed sponsored like). This suggests that, as predicted by the theoretical model, the disclosure of a subset of sponsored posts increases the beliefs of followers that the remaining posts are not sponsored, and so increases their engagement relative to organic posts. In the theory model, such an outcome would be interpreted as bad for consumers, who are in a sense "deceived" by the undisclosed content into liking it more than they would have if it was disclosed.

6   DISCUSSION

The three analytic theory models described above present distinct channels through which disclosure regulations may backfire, resulting in higher ad loads for many social media users, potentially more undisclosed sponsored content, and lower social media user well-being. All three of these can occur simultaneously. When regulations come in, influencers building their reputations may have an incentive to increase the amount of sponsored content they post due to concerns about the profitability of disclosed sponsored posts,[33] due to a change in the composition of their followers,[34] and due to the imperfection of regulations and to changes in their followers' beliefs.[35] All of these effects may reinforce one another.

There are also additional effects and channels that are not considered in these studies. For example, the choice of business model by influencers, and how these choices would be affected by changing regulations, is an important channel for consumer well-being,[36] but it is not studied by any of the papers discussed above. The closest paper to studying this examines the interaction

---

[33]   As in Mitchell (n 7).
[34]   As in Fainmesser and Galeotti (n 7).
[35]   As in Ershov and Mitchell (n 7).
[36]   See Catalina Goanta, 'Emerging Business Models and the Crowdfunding Regulation: Income Crowdfunding on Social Media by Content Creators' (2021)

between influencers and brands, and the information embedded in advertising content by different types of influencers.[37]

The empirical evidence presented in Ershov and Mitchell's paper[38] regarding the increasing ad load German Instagram users face after the tightening of disclosure regulations is also consistent with all three models presented above. They show the theoretical channel from their paper holds, but since they only have a small set of influencers and have no data on non-influencer users (e.g., followers), they are not able to test for the sorting of different followers across influencers and the re-allocation of followers after the regulatory regime is strengthened (as in Fainmesser and Galeotti's paper[39]). Testing the predictions of Mitchell's paper[40] is also challenging since they relate to relatively complex dynamic incentives. That said, related work shows that the share of sponsored content that influencers post increases over their "lifetime" as the number of followers they have also increases.[41] This confirms some of Mitchell's theoretical results regarding the cycle of influencer content posting.

There are several implications for regulatory design coming out of this stream of research. First, it is crucial for the regulatory environment to be as clear as possible to influencers with respect to what content needs to be disclosed and how this disclosure must take place. Put another way, disclosure regulations must be as "airtight" as possible, to avoid any attempts by influencers to game the regulations and post a large volume of sponsored content with the hope that some of it escapes disclosure requirements. As suggested by Mitchell's paper, there is also reason to suggest that disclosure regulations should be more lax for larger influencers, who are likely to disclose in any case. That environment, which would be stricter for smaller influencers, would encourage influencers to post more organic content over the course of their career. Moreover, more popular brands and influencers already have substantially higher inherent incentives to disclose their sponsored content compared with smaller brands and influencers, as they have established reputations to uphold.

It should be noted that it would be difficult for any regulatory agency to substantially increase the scrutiny of content produced by a large group of social media users without the cooperation and assistance of the social media

---

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3885581> accessed 23 June 2024.

[37]   Pei and Mayzlin (n 6).
[38]   Ershov and Mitchell (n 7).
[39]   Fainmesser and Galeotti (n 7).
[40]   Mitchell (n 7).
[41]   Daniel Ershov and Matthew Mitchell, 'The Effects of Influencer Advertising Disclosure Regulations: Evidence From Instagram' (2020) Proceedings of the 21st ACM Conference on Economics and Computation (EC '20) 73.

platforms themselves. This, in turn, is unlikely to occur without additional regulatory intervention. In the US, for example, the main social media platforms bear no responsibility for the content that influencers post, including for any misleading or inappropriately labelled commercial content. In fact, given the current evidence that disclosed sponsored content is less engaging than undisclosed sponsored content (or organic content), it makes sense that platforms have little incentive to tackle it. Their goal, after all, is to maximise their own engagement and advertising revenues.

One potential avenue for future change is the implementation of the new Digital Services Act (DSA).[42] The goals of this act explicitly include consumer protection and the minimisation of misleading commercial content. As part of the Act, the EU Commission will be given powers to access data from the main EU-operating platforms, including TikTok, Instagram and Facebook. With increased scrutiny and accountability, it may be possible to incentivise the platforms to seriously consider to better monitor disclosure of commercial content, with potential reductions in the amount of undisclosed sponsored content.

That said, it is important to consider the role and incentives of platforms in this market more carefully. Over time, the role of platforms in mediating the relationships between influencers and advertisers changed. In the early days of influencer marketing, platforms were essentially entirely uninvolved in commercial transactions, which were completed using third-party marketing agencies. However, in recent years, likely because of the growth of the influencer market, platforms have become increasingly involved as intermediaries, pushing influencers and advertisers to connect through them. This potentially has positive implications for disclosure, as discussed above. However, it may also generate additional effects – with the weakening of third-party marketing agencies, platforms may gain too much market power, resulting in adverse welfare consequences for influencers, marketers and social media users more generally.[43] More research on the role of platforms in influencer marketing is needed.

---

[42]   Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC [2022] OJ L 277 (Digital Services Act).

[43]   One potential recent example of platforms exerting their power relates to their experiments with removing the number of likes posts receive from public view. These are allegedly done to improve users' mental health, but they also negatively affect third-party marketing agencies, who rely on scraping the platforms and using the number of likes for analytics. For more, see Paige Leskin, "Influencers are fighting for attention as Instagram tests removing likes from its platform: 'There's no audience applause at the end of a performance'" (*Business Insider,* 5 September 2019) <https://www.businessinsider.com/instagram-influencers-removing-likes-impact-2019-9?r=US&IR=T> accessed 23 June 2024.