

User-generated content, strategic moderation, and advertising

Leonardo Madio¹ Martin Quinn²

¹Toulouse School of Economics & CESifo ²Católica-Lisbon School of Business and Economics

> 8th Giorgio Rota Best Paper Award. The Conference, May 11, 2020

- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- In 2018, internet users spent 2.8 million years online.
- Social media accounts for 33% of the total time spent online.

- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- In 2018, internet users spent 2.8 million years online.
- Social media accounts for 33% of the total time spent online.
- UGC lack external validation (Allcot et al. 2016).
- UGC may entail negative externalities

- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- In 2018, internet users spent 2.8 million years online.
- Social media accounts for 33% of the total time spent online.
- UGC lack external validation (Allcot et al. 2016).
- UGC may entail negative externalities for
 - Advertisers (i.e., brand safety).

- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- In 2018, internet users spent 2.8 million years online.
- Social media accounts for 33% of the total time spent online.
- UGC lack external validation (Allcot et al. 2016).
- UGC may entail negative externalities for
 - Advertisers (i.e., brand safety).
 - Society (i.e., extremist propaganda, hoax news).

- Most platforms host user-generated content (UGC)
 - e.g., Youtube, Twitter, Facebook, Snapchat, IG



- In 2018, internet users spent 2.8 million years online.
- Social media accounts for 33% of the total time spent online.
- UGC lack external validation (Allcot et al. 2016).
- UGC may entail negative externalities for
 - Advertisers (i.e., brand safety).
 - Society (i.e., extremist propaganda, hoax news).

beginframe



Brand Safety Remains Top Advertising Concern as Incidents Rise 25 Percent

DoubleVerify Reports Latest Global Trends in Digital Media Quality



"Brand's exposure to inappropriate content. This could be an ad published next to, before, or within an unsafe environment" (Bannerflow).

On The One Hand...

TECH \ YOUTUBE \

Disney reportedly pulls ads from YouTube following child exploitation controversy

Following Nestlé and Epic Games

By Julia Alexander | Feb 20, 2019, 4:04pm EST



AdAge

FRIDAY WAKE-UP CALL: YOUTUBE HAS BRAND SAFETY ISSUES, AGAIN. AND 'FEARLESS GIRL' WILL GET MOVED

By Angela Doland. Published on April 20, 2018.

YouTube Inc (+ Add to myFT

List of advertisers to leave Google over YouTube backlash

Companies pull spending due to extremist content



Leaving Google: a growing list of prominent advertisers have pulled spending from the group's YouTube platform in a the backlas over extremist content © FT montage

"The Adpocalypse"

Madio & Quinn

Content Moderation as a Solution?

Timeline: Google's path to tackling YouTube brand safety problems

Addressing brand safety took on a new urgency after an advertiser revolt in the spring of 2017. Here's the rundown of what the company has done to address advertiser concerns.

Ginny Marvin on June 26, 2017 at 11:14 am



APRIL 8, 2019 by Kerry Flynn

			•	-	
M	ad	10	×.	J	uinr
				~	

On the Other Hand...

When Tumbir bans porn, who loses?

Verizon is leaving the engine of internet culture to sputter and die, and its communities to scramble for a new home.

By Kaitlyn Tiffany | @kait_tiffany | kaitlyn.tiffany@vox.com | Dec 4, 2018, 5:00pm EST





• The ban was adequately designed to keep "content that is not brand-safe away from ads"

Madio & Quinn

Main trade-off

• strong content moderation increases advertisers' willingness-to-pay.

Main trade-off

- strong content moderation increases advertisers' willingness-to-pay.
 - ... but may deter users, which lowers advertisers' willingness-to-pay.

Main trade-off

- strong content moderation increases advertisers' willingness-to-pay.
 - ... but may deter users, which lowers advertisers' willingness-to-pay.

Questions:

1 How does a platform choose content moderation and ad prices?

Main trade-off

- strong content moderation increases advertisers' willingness-to-pay.
 - ... but may deter users, which lowers advertisers' willingness-to-pay.

Questions:

- 1 How does a platform choose content moderation and ad prices?
- **2** Does competition generate more moderation?

Main trade-off

- strong content moderation increases advertisers' willingness-to-pay.
 - ... but may deter users, which lowers advertisers' willingness-to-pay.

Questions:

- 1 How does a platform choose content moderation and ad prices?
- **2** Does competition generate more moderation?
 - Competition often seen as a *panacea* in markets with network externalities.

Main trade-off

- strong content moderation increases advertisers' willingness-to-pay.
 - ... but may deter users, which lowers advertisers' willingness-to-pay.

Questions:

- 1 How does a platform choose content moderation and ad prices?
- **2** Does competition generate more moderation?
 - Competition often seen as a *panacea* in markets with network externalities.

• Main trade-off: Platforms engage in moderation *iff* they can increase ads.

- Main trade-off: Platforms engage in moderation *iff* they can increase ads.
 - brand safety effect > eyeball effect

- Main trade-off: Platforms engage in moderation *iff* they can increase ads.
 - brand safety effect > eyeball effect
 - with a monopoly, prices are non-monotonic in preference/aversion for content moderation.

- Main trade-off: Platforms engage in moderation *iff* they can increase ads.
 - brand safety effect > eyeball effect
 - with a monopoly, prices are non-monotonic in preference/aversion for content moderation.
- Would more competitive markets ensure more moderation?

- Main trade-off: Platforms engage in moderation *iff* they can increase ads.
 - brand safety effect > eyeball effect
 - with a monopoly, prices are non-monotonic in preference/aversion for content moderation.
- Would more competitive markets ensure more moderation?
 - If competition for eyeballs intensifies \Rightarrow moderation is reduced.
 - Ambiguous effect on the ad price.

The Model



- A volume 1 of safe content (i.e., pets, posts, holiday pics).
 - These entail positive externalities v for advertisers.

- A volume 1 of safe content (i.e., pets, posts, holiday pics).
 - These entail positive externalities v for advertisers.
- A volume *θ*(*m*) of **unsafe content** (i.e., nude, racist, extremist propaganda).

- A volume 1 of safe content (i.e., pets, posts, holiday pics).
 - These entail positive externalities v for advertisers.
- A volume *θ*(*m*) of **unsafe content** (i.e., nude, racist, extremist propaganda).
 - $m \in [0,1]$ is the content moderation effort s.t. $\theta(m=0) = 1$ and $\theta(m=1) = 0$.

- A volume 1 of safe content (i.e., pets, posts, holiday pics).
 - These entail positive externalities v for advertisers.
- A volume *θ*(*m*) of **unsafe content** (i.e., nude, racist, extremist propaganda).
 - $m \in [0,1]$ is the content moderation effort s.t. $\theta(m=0) = 1$ and $\theta(m=1) = 0$.
 - These entail negative externalities.

- A volume 1 of safe content (i.e., pets, posts, holiday pics).
 - These entail positive externalities v for advertisers.
- A volume *θ*(*m*) of **unsafe content** (i.e., nude, racist, extremist propaganda).
 - $m \in [0,1]$ is the content moderation effort s.t. $\theta(m=0) = 1$ and $\theta(m=1) = 0$.
 - These entail negative externalities.

We model brand safety concerns as follows:

$$\Omega = \mathbf{v} \cdot \mathbf{1} - \lambda \cdot \theta(\mathbf{m}),$$

Intuition: long-term net gain from being on a platform.

- A monopolist platform
 - maximizes profits Π = a × p C(m), with p the ad price, a ads, C(m) the moderation cost with C'(m) > 0, C''(m) > 0.

- A monopolist platform
 - maximizes profits Π = a × p C(m), with p the ad price, a ads, C(m) the moderation cost with C'(m) > 0, C''(m) > 0.
- A mass of advertisers
 - derive utility V = v × 1 − λ × θ(m) + r × n − p, i.e., with preference v for safe content and loss λ from unsafe content, r are marginal revenues from n users.

- A monopolist platform
 - maximizes profits Π = a × p − C(m), with p the ad price, a ads, C(m) the moderation cost with C'(m) > 0, C''(m) > 0.
- A mass of advertisers
 - derive utility V = v × 1 − λ × θ(m) + r × n − p, i.e., with preference v for safe content and loss λ from unsafe content, r are marginal revenues from n users.
- A mass of Internet users
 - derive utility U = u × 1 + φ × θ(m) − γ × a, i.e., with preference u for safe content and φ for unsafe one. Ads create disutility γ.

- A monopolist platform
 - maximizes profits Π = a × p C(m), with p the ad price, a ads, C(m) the moderation cost with C'(m) > 0, C''(m) > 0.
- A mass of advertisers
 - derive utility V = v × 1 − λ × θ(m) + r × n − p, i.e., with preference v for safe content and loss λ from unsafe content, r are marginal revenues from n users.
- A mass of Internet users
 - derive utility U = u × 1 + φ × θ(m) − γ × a, i.e., with preference u for safe content and φ for unsafe one. Ads create disutility γ.



The Main Trade-off

Let $\Upsilon(m^*, p^*)$ be the marginal gain from moderation such that

$$\Upsilon(m^*,p^*) = -rac{d^a(m^*,p^*)\cdot \Psi}{rac{\partial D^a}{\partial p}}.$$

The Main Trade-off

Let $\Upsilon(m^*, p^*)$ be the marginal gain from moderation such that

$$\Upsilon(m^*,p^*) = -rac{d^a(m^*,p^*)\cdot\Psi}{rac{\partial D^a}{\partial p}}$$

where Ψ is the elasticity of profit with respect to moderation:



The Main Trade-off

Let $\Upsilon(m^*, p^*)$ be the marginal gain from moderation such that

$$\Upsilon(m^*,p^*) = -rac{d^a(m^*,p^*)\cdot\Psi}{rac{\partial D^a}{\partial p}}.$$

where $\boldsymbol{\Psi}$ is the elasticity of profit with respect to moderation:



Then, there is an incentive to engage in moderation $m^* \in (0,1]$ iff

the brand safety effect outweighs the eyeball effect

Example

- An Internet user is defined by a duple $(u, \phi) \sim U[0, \overline{u}] \times U[0, \overline{\phi}]$
- An advertiser is defined by a duple $(v, \lambda) \sim U[0, \overline{v}] \times U[0, \overline{\lambda}]$,



• **Competition**: Does competition create the right incentives for more moderation?

- **Competition**: Does competition create the right incentives for more moderation?
 - Relevant discussion in policy circles on allowing data portability, lowering barriers to entry ⇒ More competition.
- What do we do?

- **Competition**: Does competition create the right incentives for more moderation?
 - Relevant discussion in policy circles on allowing data portability, lowering barriers to entry ⇒ More competition.
- What do we do?
 - Two differentiated platforms in a *Hotelling* setting, with multi-homing advertisers and single-homing users.

- **Competition**: Does competition create the right incentives for more moderation?
 - Relevant discussion in policy circles on allowing data portability, lowering barriers to entry ⇒ More competition.
- What do we do?
 - Two differentiated platforms in a *Hotelling* setting, with multi-homing advertisers and single-homing users.
- What happens if **competition intensifies**?
 - Requires higher effort to attract consumers!

 $\Rightarrow This translate into a dual strategy for the platform :$

it can relax content moderation,

2 it can lower advertising nuisance (= increase ad prices!)

• In equilibrium, lower moderation and higher (lower) ad prices.

Takeaway

- Content moderation increases if the *brand safety effect* outweighs the *eyeball effect*.
 - Non-monotonic relationship between price and moderation with a monopoly;
- Market failures are around the corner. More competition lowers incentives to increase content moderation (below social optimum?)



• Absent advertisers, moderation levels would be much lower. But advertisers might be shortsighted.

Thank you!

leonardo.madio@tse-fr.eu
martinquinn@ucp.pt



Economics for the Common Good Manufacture des Tabacs 21, Allée de Brienne 31015 Toulouse Cedex 6 France www.tse-fr.eu

T. +33(0)5 61 63 36 90 F. +33(0)5 61 63 35 86

