

# When Big Brother Meets Big Profit

Yunyun Lv

October 7, 2022

## Abstract

The Chinese government (“Big Brother”) demands that Social Network Service (SNS) companies (which make “Big Profit”) comply with their order to censor sensitive content. The government threatens to punish SNS companies in case of disobedience. At the same time, the government also benefits from the tax paid by these companies. I document and explain two noticeable trends in China: (i) The level of censorship by SNS companies increases when their de facto tax rate decreases. (ii) Bigger SNS companies exhibit higher levels of compliance with the government than smaller SNS companies. The tax rate is determined by the government, serving as a commitment device: If the government sets a higher tax rate, it benefits more from the profit made by the company. Hence the government has weaker incentives to punish the company in case of disobedience. Therefore, the company also has weaker incentives to comply with the government. The higher censorship levels of bigger SNS platforms can be explained by stronger incentives for the government to stop circulating sensitive information on bigger platforms.

## 1 Introduction

The tragic death of an overworked content moderator at Bilibili, one of the largest video streaming platforms in China, during the Lunar New Year holiday in 2022 startled the public. People were angry for many reasons: Bilibili’s first reaction was to deny that the employee was overworked by deleting his employee profile and attendance record in the internal system. After the news was leaked, the

discussion of his death immediately became one of those “sensitive” topics that are prohibited both inside the company and on Bilibili.com. The provoked public indignation by the death of the overworked censor only made Bilibili censors even busier censoring one more topic. This unexpected incident also made the public aware of the pervasiveness of content moderation on Social Network Service (SNS) platforms in China. Bilibili finally released a public announcement, saying that 1,000 more content moderators would be hired in 2022 to ease the workload of the content moderation team. There were over 2,400 employees in this team, which was roughly 30% of the total employees at Bilibili. Yet Bilibili is not the biggest job provider for content moderators in China: By the end of 2020, ByteDance has more than 20,000 content moderators out of its 100,000 employees.

Why do SNS companies hire so many censors? Besides the need to scrub content associated with violence, pornography, fraud, etc., they also need to comply with orders of the Chinese government (the ‘Big Brother’) to censor contents that are deemed politically sensitive or offensive. The platforms do not have much to negotiate with the government when it comes to censorship of sensitive information. But they can choose how fast they respond to such orders: Empirical findings show that different platforms do have different delays of response after the same event/news shock happens (Liu 2020). Also, disobedience by social media companies is quite common: 16% of government directives are disobeyed by Weibo because of concerns about censoring more strictly than its competitors (Miller 2018). One reason SNS companies do not want to censor too much is that most censorship are done via friction, in other words, by increasing users’ cost to communicate (Roberts 2018); while internet users are extremely impatient: 57 percent of users will abandon a site if it takes more than three seconds for the webpage to load, of which 80 percent will not return (Nagy 2013). Although censorship via friction does not mean an increase in the loading time of a website, it often increases the time needed for users to successfully post a message, upload a photo or video, and to decipher encrypted messages posted by other users. [Common practices of censorship include, but is not limited to, suspending an account temporarily, deactivating an account permanently, hiding a post in the timeline with or without notifying the poster, disabling the reply/forward function for a post, adding a misinformation warning tag to a post.] (Should I

keep these examples?) Increased cost of communication drives active users away, possibly to a competitor’s platform. Another reason is that censorship is costly. Censorship by algorithm requires regular maintenance. The set of keywords that the government wants to block is changing everyday. It is expensive to hire machine learning experts who can write efficient programs and to hire a large group of censors who manually check suspicious contents.

In case of disobedience the government punishes platforms by giving them a warning, collecting fines, or shutting them down, either temporarily or permanently. Toutiao (“Today’s Headlines”), one of the most popular news platform in China and a core product of ByteDance, was temporarily shut down for 24 hours in Dec. 2017 by the Cyberspace Administration of China (CAC). Four months later it permanently shut down its popular joke app Neihan Duanzi at the request of the State Administration of Radio and Television (SART). At the same time, its CEO Yiming Zhang openly apologized to regulators and the public for allowing the app to “lose its way” and announced an increase of content moderators from 6,000 to 10,000. In 2021 Weibo paid 44 fines totaling \$2.25mn (in USD) and Douban incurred 22 fines totaling \$1.42mn. These fines themselves are not huge relative to these companies’ revenues. But these public punishment can trigger a greater loss. On Dec. 14th, 2021 Weibo stock hits multi-year lows after CAC fines Weibo three million yuan (\$470,000), the maximum amount of such fines, for having repeatedly violated cybersecurity laws and published illegal information. Within one day, Weibo stock price fell by 10% in Hong Kong and 4.8% in U.S. stock market. The huge drop of market value reflects reduced confidence level of the public about the company’s future profitability.

These punishments do not only hurt the “Big Profit” made by SNS companies, but also the government’s tax revenue. Take year 2019 as an example: China’s total tax revenue was \$2,300 billion. Weibo paid around \$100 million tax while Tencent paid \$2.1 billion. The top 500 private companies jointly contributed \$200 billion of tax. The general Corporate Income Tax rate in Mainland China is fixed at 25%. But the *de facto* tax rate, which is the ratio of tax expense divided by pretax income, varies both across companies and across time. This is because the government provides some preferential tax treatments to some companies. For example, companies that are approve by the government as “High

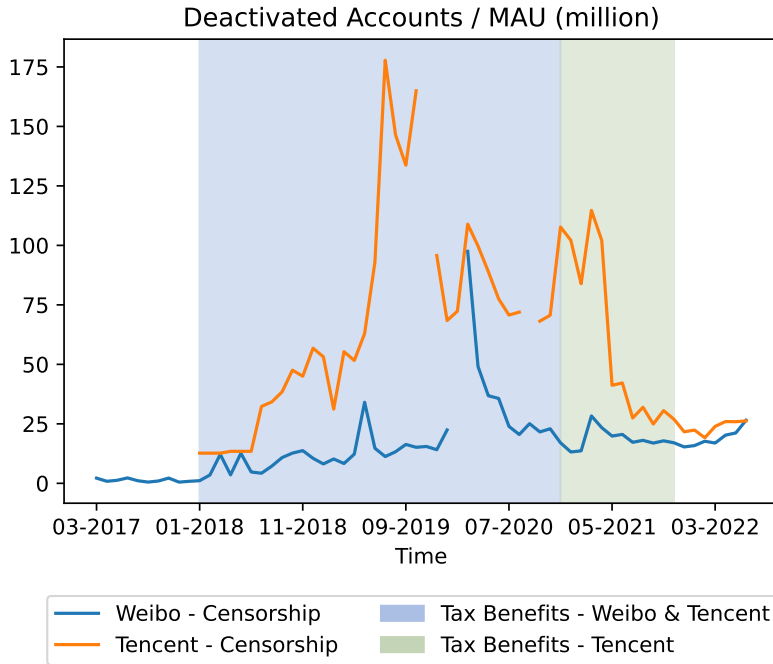


Figure 1: Ratio of Deactivated Accounts Related to Political Sensitive Contents to MAU

and New Technology Enterprise” or “Key Software Enterprise” are subject to a preferential corporate income tax rate of 15% and 10%, respectively. Both Weibo and Tencent enjoyed these preferential tax treatments for tax year 2017-2019 (which are recognized in year 2018-2020). Weibo lost these benefits while Tencent managed to plow on for one more year in 2021. The following figure<sup>1</sup> shows the self-reported number of deactivated accounts due to contents that are politically sensitive or offensive on the two platforms, Weibo and Tencent. Since the two platforms are of different sizes, the numbers are normalized by their own Monthly Active Users (MAU) so that the comparison is meaningful. The shaded areas represent years in which preferential tax benefits are enjoyed by Weibo and/or Tencent. From the figure two trends are noticeable: (i) The level of censorship by Social Network Service (SNS) companies are higher when their *de facto* tax rate is

<sup>1</sup>Is the title clear? I will figure out how to draw the graph using solid line and dashed line for the two companies later. Also maybe I should normalize everything using the average frequency of censorship of all companies as the ref. level?

lower. (ii) Bigger SNS companies censor more frequently than smaller ones. The intuition is that the tax rate serves as a commitment device of the government: If the government cancels some tax benefits for an SNS company, then the tax rate is higher, implying that the government cares more about the profit made by the company. Hence the government has weaker incentives to punish the company in case of disobedience. Therefore, the company also has weaker incentives to comply with the government to censor contents. The difference in censorship levels on SNS platforms of different sizes can be explained by stronger incentives of the government to stop circulating of sensitive information on bigger SNS platforms.

## 2 Benchmark Model

There are two players: The government,  $G$ , and the platform,  $P$ . In the first period,  $G$  chooses (and commits to) a tax rate  $\tau \in [0, 1]$ . Tax rate is public information. Non-strategic users spend a fixed amount of time on the platform. The platform's profit from active users is normalized as 1. The fraction of sensitive contents users spread, denoted by  $\alpha$ , follows a uniform distribution on  $[0, 1]$ . The distribution is public information but the realized  $\alpha$  is unknown to  $P$ . Only  $G$  observes  $\alpha$  in the final period. It captures the changing and private preference of the government on political topics. In the second period,  $P$  chooses their compliance level,  $\beta \in [0, 1]$ , measured as the fraction of sensitive contents  $P$  successfully catches. A higher  $\beta$  is more costly. In the final period,  $G$  decides whether to shut down  $P$ . Let  $d \in \{0, 1\}$  represent  $G$ 's decision, where  $d = 1$  means to shut down  $P$ , and  $d = 0$  means not to shut it down. If  $d = 1$ , both  $P$  and  $G$  get zero. If  $d = 0$ ,  $G$  taxes  $P$ 's revenue at the chosen tax rate  $\tau$ . The revenue for  $P$  comes from active users, but they also has to pay the cost of censorship. So  $P$ 's expected payoff is given by

$$\Pi(\beta) = (1 - d)(1 - \tau - \kappa_1\beta),$$

where  $\kappa_1$  is the coefficient of censorship cost. Notice that the censorship cost includes the cost of hiring content moderators, the reputation cost for censoring, and loss of active users for censoring more. The government's expected payoff is given by its tax revenue minus dis-utility from the existence of sensitive contents

on the platform:

$$V(\tau, d) = (1 - d) [\tau - \kappa_2 \alpha (1 - \beta)],$$

where  $\kappa_2$  represents how much  $G$  hates sensitive contents being spread on the platform.

The solution concept is Perfect Bayesian Equilibrium.

**Proposition 1.** *There are two possible cases of equilibrium:*

- *If  $(1 - \kappa_1)\kappa_2 < 1$ , then  $\tau^* = 1, \beta^* = 0$ , and  $\mathbb{E}(d^*) \geq 0$  at equilibrium.*
- *If  $(1 - \kappa_1)\kappa_2 < 1$ , then  $\tau^* = 1 - \kappa_1, \beta^* = 1 - \frac{1 - \kappa_1}{\kappa_2} > 0$ , and  $d^* = 0$  at equilibrium.*

When censorship is too costly while the government does not hate sensitive contents that much, the government taxes everything, and the platform never censors. Shut-down happens when  $\alpha$  is high. An example of this equilibrium is North Korea: Any platform on which people can communicate is completely under control of the government. In the opposite case, the platform censors frequent enough just to keep the government indifferent between shutting it down or not in the worst case (when  $\alpha = 1$ ). There is always censorship by the platform and shut-down never happens. When the cost of censorship,  $\kappa_1$ , increases, the government sets a lower tax rate to incentivize censorship and the equilibrium level of censorship increases. When the government's dis-utility from sensitive contents,  $\kappa_2$ , increases, the platform is forced to censor more so that the government is still indifferent between shutting it down or not in the worst case.

In the final period,  $G$  shuts down the platform when  $\alpha$  is too high:

$$d^* = \mathbb{1}(\kappa_2 \alpha (1 - \beta) > \tau) = \mathbb{1}(\alpha > \frac{\tau}{\kappa_2(1 - \beta)}).$$

Given the government's optimal shutting down strategy,  $P$ 's expectation of shutting down probability is then

$$1 - \mathbb{E}(d^*) = \text{Prob}(\alpha < \frac{\tau}{\kappa_2(1 - \beta)}) = \min\{\frac{\tau}{\kappa_2(1 - \beta)}, 1\}.$$

Notice that if  $\tau \geq \kappa_2$ , then  $d^* = 0$  always and hence  $\beta^* = 0$  in equilibrium. The best response function of  $P$  is then

$$\beta^* = \begin{cases} \tilde{\beta}, & \tau < 1 - \kappa_1 \text{ and } \tau < \kappa_2 \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\tilde{\beta} = 1 - \frac{\tau}{\kappa_2}.$$

Observe that  $P$  censors only when cost of censorship,  $\kappa_1$ , and the tax rate,  $\tau$ , are low enough. Also notice that  $\beta^*$  is decreasing in  $\tau$ : As tax rate increases,  $P$  keeps a smaller share of its revenue while  $G$  gets a bigger share. Hence  $G$  has weaker incentives to shut down the platform, and  $P$  has weaker incentives to censor. The interior solution  $\tilde{\beta}$  makes the government indifferent between shutting down the platform or not in the worst case (when  $\alpha = 1$ ). The platform censors to make sure that they will not be shut down by the platform in the worst case.

The government's expected payoff in the first period from choosing  $\tau$  is

$$\mathbb{E}V(\tau, d^*) = \begin{cases} \int_0^1 [\tau - \alpha\tau] d\alpha = \frac{\tau}{2}, & \tau < 1 - \kappa_1 \text{ and } \tau < \kappa_2 \\ \int_0^{\frac{\tau}{\kappa_2}} [\tau - \kappa_2\alpha] d\alpha = \frac{\tau^2}{2\kappa_2}, & 1 - \kappa_1 < \tau < \kappa_2 \\ \int_0^1 [\tau - \kappa_2\alpha] d\alpha = \tau - \frac{\kappa_2}{2}, & \tau > \kappa_2. \end{cases}$$

Therefore the optimal  $\tau$  is

$$\tau^* = \begin{cases} 1 - \kappa_1, & \frac{1}{\kappa_2} < 1 - \kappa_1 < \kappa_2 \\ 1, & \text{otherwise} \end{cases}$$

## References:

- Jiang, Min (2012). “Internet companies in China: Dancing between the party line and the bottom line”. *Jiang, M.(2012). Internet companies in China: Dancing between the Party line and the bottom line. Asie Visions* 47.
- Jiang, Min and King-Wa Fu (2018). *Chinese social media and big data: big data, big brother, big profit?*
- King, Gary, Jennifer Pan, and Margaret E Roberts (2013). “How censorship in China allows government criticism but silences collective expression”. *American Political Science Review*, pp. 326–343.
- Liu, Z Jessie (2020). “Impact of Market Structure on Regulatory Compliance: Evidence from Online Censorship in China”. *Johns Hopkins Carey Business School Research Paper* 20-13.
- Miller, Blake Andrew Phillip (2018). “The limits of commercialized censorship in China”.
- Nagy, Zsolt (2013). “Improved speed on intelligent web sites”. *Recent Advances in Computer Science* 1.14, pp. 215–220.
- Pan, Jennifer (2017). “How market dynamics of domestic and foreign social media firms shape strategies of internet censorship”. *Problems of Post-Communism* 64.3-4, pp. 167–188.
- Roberts, Margaret E (2018). *Censored*. Princeton University Press.