When Big Brother Meets Big Profit

Yunyun Lv

October 10, 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 China's largest video streaming platforms, 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 had 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. Nevertheless, Bilibili is not the largest 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, and fraud, among other things, they also need to comply with orders of the Chinese government (the "Big Brother") to censor content that are deemed politically sensitive or offensive. The platforms do not have much to negotiate with the government regarding censorship of sensitive information. But they can choose how fast they respond to such orders: Empirical findings show that different platforms have different response delays after the same event/news shock (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). SNS companies do not want to censor too much because most censorship is 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 is not done by simply increasing 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 decipher encrypted messages posted by other users. The increased communication cost drives active users away, possibly toward a competitor's platform. Common practices of censorship on SNS platforms include, but are 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. 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 every day. It is expensive to hire machine learning experts who can write efficient programs and hire a large group of censors who manually check suspicious content.

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 platforms 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 punishments can trigger a greater loss. On Dec. 14th, 2021, Weibo stock hit multi-year lows after CAC fined Weibo three million yuan ($470,000), the maximum amount of such fines, for having violated cybersecurity laws and published illegal information repeatedly. Within one day, Weibo's stock price fell by 10% in Hong Kong and 4.8% in the U.S. stock market. The huge drop in market value reflects the reduced confidence level of the public about the company's future profitability.

These punishments hurt the "Big Profit" made by SNS companies and the government's tax revenue. Take the year 2019 as an example: China's total tax revenue was $2,300 billion. Weibo paid around $100 million in tax, while Tencent paid $2.1 billion. The top 500 private companies jointly contributed $200 billion in 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 across companies and time. Because the government provides some preferential tax treatments to some companies. For example, companies that the government approves as "High 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 the tax year 2017-2019 (which are recognized in the year 2018-2020). Weibo lost these benefits while Tencent managed to plow on for one more year in 2021. Figure 1 [1] shows the self-reported number of deactivated accounts due to politically sensitive or offensive content on the two platforms, Weibo and Tencent. Since the two platforms are of different sizes, the numbers are normalized by each platform's Monthly Active Users (MAU) so that the comparison is meaningful. The shaded areas indicate when preferential tax benefits are available. The data is collected from monthly and weekly reports posted by official accounts or on service centers of these platforms, including Weibo Administrator (微博管理员), which has 135 million followers, and Tencent Guard (腾讯卫士), which serves more than one billion WeChat users and 800 million QQ users. It represents the level of censorship these SNS companies want the public and, more importantly, the government to know.

The figure shows two noticeable trends: (i) The level of censorship by SNS companies is higher when their *de facto* tax rate is 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 *de facto* 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 because it also hurts its revenue. Therefore, the company also has weaker incentives to comply with the government to censor content. The difference in censorship levels on SNS platforms of different sizes can be explained by stronger incentives of the government to stop circulating sensitive information on bigger SNS platforms.

[1] Add another plot: Normalize everything using the average frequency of censorship of all companies.

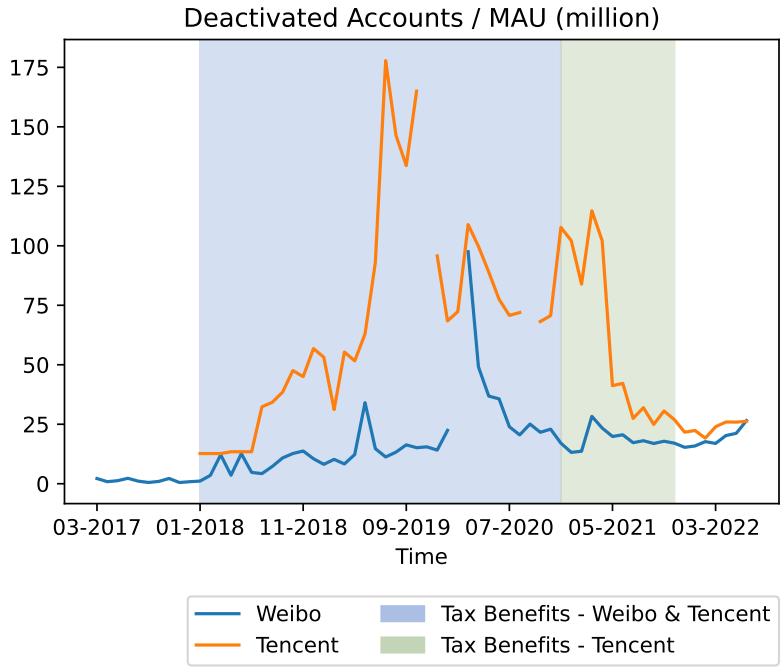

Figure 1: Number of Deactivated Accounts Due to Political Sensitive Content Every One Million MAU

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 $\tau \in [0, 1]$ that applies to P in the next period. The tax rate is public information. Non-strategic users of the platform spend a fixed amount of time and hence generate a fixed amount of profit for the platform, which is normalized as 1. The fraction of sensitive content that the users spread, denoted by α, is a random variable that follows a uniform distribution on $[0, 1]$. The distribution is public information, but the realized α is unknown to P. Only G observes α in the final period. It captures the changing and private preferences of the government on political topics. In the second period, P chooses its compliance level, $\beta \in [0, 1]$, that measures the fraction of sensitive content that P successfully catches. Censorship is costly: choosing a higher β incurs a higher cost for P. In the final period, G decides whether to shut down P after observing α and β. Let $d \in \{0, 1\}$ represent G's decision. If $d = 1$ (G shuts down P), both P and G get zero. If $d = 0$ (G does not shut down P), G taxes P's revenue at the chosen tax rate τ. Notice that, in reality, the government punishes SNS platforms for disobedience in various ways. In this model, it is captured by assuming that G chooses whether to shut down the platform or not. The idea is that, no matter how the punishment is enforced, it harms the revenue of the SNS company and hence also the government's tax revenue.

Figure 2 summarizes the timeline of the game:

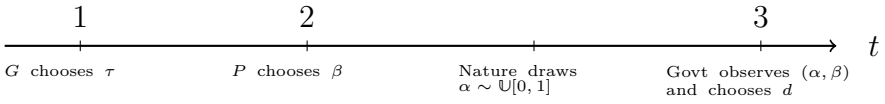

Figure 2: Timeline

If P is not shut down, the platform profits from active users and pays tax and the cost of censorship. P's payoff is given by

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

where κ_1 is the coefficient of censorship cost. Notice that the censorship cost includes the cost of hiring content moderators, the reputation cost of censoring, and the loss of active users for censoring more. The government's payoff when $d = 0$ is given by its tax revenue minus dis-utility from the existence of sensitive content on the platform:

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

where κ_2 represents how much G hates sensitive content being spread on the platform.

3 Results

The solution concept is Perfect Bayesian Equilibrium.

Proposition 1. *There exists a unique 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.*

If censorship is too costly (κ_1 is big) while the government does not hate sensitive content that much (κ_2 is small), the government taxes everything, and the platform never censors. A shutdown happens whenever α is high. An example of this equilibrium is North Korea: Any platform on which people can communicate is completely under the control of the government. In the opposite case, the platform censors frequently enough to keep the government indifferent between shutting it down or not in the worst case (when $\alpha = 1$). There is always censorship on the platform, and shutdown never happens.

Proof. In the final period, G's optimal strategy is to shut down the platform whenever α is too high:

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

Given this shutting-down strategy, P's expectation of shutting down probability is then

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

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 the cost of censorship, κ_1, and the tax rate, τ, are low enough. Notice that β^* is decreasing in τ: If the 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 ensure that the government will not shut them down in the worst case. This explains one of the trends observed above.

The government's expected payoff in the first period from choosing τ 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 τ is

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

□

4 Comparative Statics

Proposition 2. *The equilibrium tax rate τ^* decreases in κ_1 if $\kappa_1 < \tilde{\kappa}_1 = 1 - \frac{1}{\kappa_2}$, and jumps to 1 at the threshold $\tilde{\kappa}_1$. The equilibrium censorship level β^* increases in κ_1 and κ_2 if $\kappa_1 < \tilde{\kappa}_1$, and jumps to 0 at the threshold $\tilde{\kappa}_1$.*

If the cost of censorship, κ_1, increases, the government sets a lower tax rate to incentivize more censorship by the platform; the equilibrium level of censorship increases as a result. If the government's dis-utility from sensitive content, κ_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.

The difference in censorship levels on SNS platforms of different sizes can be explained by the positive correlation between β^* and κ_2: The same piece of sensitive information receives more attention on a bigger platform. Hence the government dislikes sensitive content on SNS platforms with more users. So κ_2 is higher if the SNS platform is bigger, and hence the equilibrium censorship level is higher.

In July 2022, the Russian internet regulator fined Google $370 million for promoting the dissemination of false content on YouTube. Just a week later, Russia fined Google $34 million under the pretext of breaching competition rules in the video hosting market. These punishments can be interpreted as the result of the sudden increase in α and κ_2 after the Russian invasion of Ukraine. The equilibrium censorship level β^* increases, whereas Google failed to meet the Russian government's expectations.

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.