

Information Control on YouTube During Russia's Invasion of Ukraine

Keywords: Censorship, YouTube data, War, Ukraine, Russia

Extended Abstract

The war between Russia and Ukraine takes place on physical battlefields as well as in the information space. This became apparent during Russia's invasion of Ukraine in 2014 and continues to be relevant during the recent invasion on the 24th of February 2022. Since the beginning of the ongoing war, numerous scholars across fields have pointed out that the so-called "information war" (Thornton, 2015; Darczewska, 2014) plays an important political and military role. While the scholarly community as well as the general public has largely focused on the *production* and *dissemination* of content, an important part of the informational struggle also takes place through information *control* both in Russia, Ukraine and even the EU (Golovchenko, 2022; European Commission, 2022).

We focus on online user activity on YouTube, one of the most popular social media platforms in Russia, Ukraine, and the world. While the platform has been criticized for allowing hate speech and misinformation to be spread, it also plays an important part in disseminating information and allowing regime-critical opinions to be shared in autocracies like Russia (Gainous et al., 2018; Reuter and Szakonyi, 2015). YouTube's double-edged sword nature makes it an important platform in the struggle for "truth" about the war.

The first weeks of the full-scale invasion were marked by attempts from Kremlin to restrict and slow down information flows not adhering to the regime's official narrative about the war. Having already tightened control over most of the established media outlets in Russia as well as Russian-owned social media. After February 24, 2022, it restricted and blocked some of the last mainstream alternatives to state-controlled media and social media: Facebook, Instagram, and Twitter. YouTube is so far one of the few major western social media that has not yet been restricted or blocked. Nevertheless, both Russian content creators on YouTube, as well as users from Russia, are under legal pressure from the Kremlin authorities. Perhaps most controversially, Russians now risk several years of imprisonment for "discrediting" Russian forces or spreading "disinformation" (i.e. information that does not align with official Russian sources) both online and offline. In practice, it is now illegal to discuss Russian atrocities or even to describe the "special military operation" (as it was officially named by Kremlin) as a "war" (Troianovski, 2022). Not long after the implementation of this law, it became evident that the regime not only targeted protesters and activists but also went after regular citizens posting critical messages online (Dixon, 2022). The Kremlin's drastic escalation of information control highlights how important it is to understand how political actors control information about Russia's invasion of Ukraine.

Pro-Kremlin disinformation about the war in Ukraine as well as Kremlin's strategic control over Russian media outlets has also been met with great concern in the West. On the 2nd of March 2022, the European Union responded by banning access to Sputnik and Russia Today

(RT) (European Commission, 2022). Finally, on March 11th, YouTube went a step further by announcing that they were blocking access to Russian state media as a whole in all countries (Reuters, 2022). In turn, interventions to the informational sphere were enacted both by Russia and YouTube within a short time period - the first aimed at restricting citizens' access to and interaction with information, while the latter aimed at preventing the spread of disinformation about the war.

Using publicly available data from YouTube's own API, this research sets out to 1) measure how effective YouTube's own ban is at reducing online activity on Russian state media and 2) examine whether the Russian censorship laws in March were followed by a decrease in user activity among Russian speaking YouTube channels. We do this by investigating the number of comments posted by users on regime-critical and pro-regime content. The data consists of 12,315,588 YouTube comments tied to 13,950 videos from 40 channels (selected based on domain knowledge) in the time period of 40 days before and 40 after the implementation of the Russian censorship law and YouTube's own ban on Kremlin-affiliated channels. We identified 10 pro-Kremlin media outlets that were banned by YouTube, 10 pro-Kremlin non-banned channels, 10 regime-critical channels as well as 10 entertainment channels used as a baseline. We approach our analysis descriptively (see main results in Figure 1)).

We find that YouTube's own ban on Kremlin-affiliated channels worked, and abruptly cut off users' interaction with videos from the banned channels as soon as the ban was announced (Lower right plot Figure 1). However, the data indicate a corresponding increase in user engagement with non-blocked pro-Kremlin channels (Upper left plot Figure 1). Closer inspections of content posted by these channels point to more extreme pro-war and anti-Ukraine sentiments than those exhibited by the official state-backed channels that were blocked by YouTube. This is an important implication to keep in mind, as this shows that while strategies of information control in democratic settings might be successful in curbing the consumption of harmful content, it might also push some users towards more problematic content. More research is needed to establish whether the data show a substitution effect for users previously following now-blocked state-backed channels, or whether the increase is caused by other users.

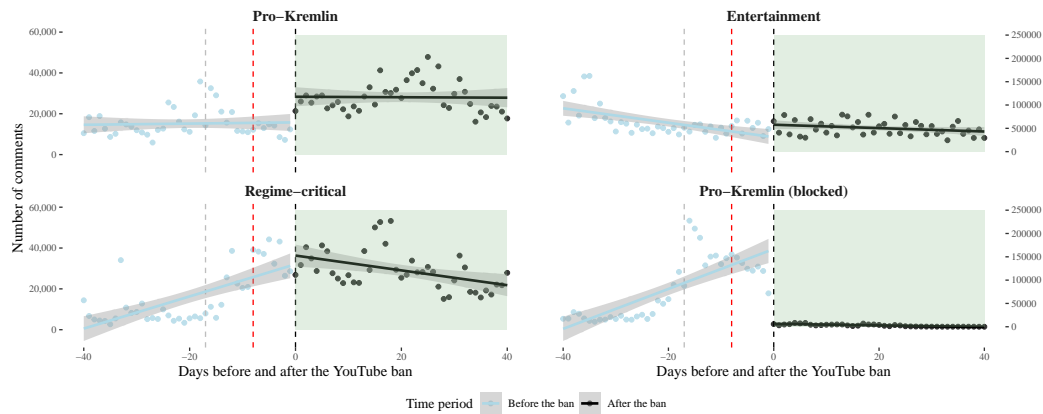
Regarding the Russian implementation of the drastic censorship law, we find no substantial change in the frequency of users' engagement with regime-critical content (Lower left plot Figure 1). This goes against our theoretical expectations and common wisdom in the literature - our main hypothesis was that the added cost risk connected to speaking up would decrease users' willingness to interact with content deemed as 'bad' by the regime. These findings do not rule out that users changed their behavior in response to the law. We continue our research into whether users instead change the way in which they speak when interacting with regime-critical content.

This paper provides unique evidence of user behavior on YouTube in the context of information control carried out during a war. The findings contribute to our understanding of people's online behavior in response to information control carried out by authoritarian regimes and social media companies respectively.

References

- Darczewska, J. (2014). The anatomy of Russian information warfare. The Crimean operation, a case study.
- Dixon, R. (2022). A Russian musician mounts a modest antiwar protest and pays the price. *The Washington Post*. Accessed February 21, 2022.
- European Commission (2022). Ukraine: Sanctions on Kremlin-backed outlets Russia Today and Sputnik. Accessed May 21, 2022.
- Gainous, J., K. M. Wagner, and C. E. Ziegler (2018). Digital media and political opposition in authoritarian systems: Russia's 2011 and 2016 дума elections. *Democratization* 25(2), 209–226.
- Golovchenko, Y. (2022). Fighting propaganda with censorship: A study of the ukrainian ban on russian social media. *The Journal of Politics*.
- Reuter, O. J. and D. Szakonyi (2015). Online social media and political awareness in authoritarian regimes. *British Journal of Political Science* 45(1), 29–51.
- Reuters (2022). YouTube announces an immediate block on Russian state-funded media channels globally. *Euronews.next*. Accessed May 21, 2022.
- Thornton, R. (2015). The Changing Nature of Modern Warfare: Responding to Russian Information Warfare. *The RUSI Journal* 160(4), 40–48.
- Troianovski, Anton, V. S. (2022). Russia Takes Censorship to New Extremes, Stifling War Coverage. *New York Times*. Accessed May 21, 2022.

Figure 1: Change in the number of comments for Pro-Kremlin media (band and non-banned), regime-critical and entertainment channels.



Note: 24th of February marked with a grey line, the 4th of March marked as a red and YouTube ban marked as a black line