# Competition and Privacy

Yifei Wang[1]

[1]Graduate School of Business, University of Pittsburgh

June 14, 2026

**Abstract**

The question of whether competition enhances or impedes a firm's commitment to protecting consumer privacy has been debated in competition policy circles. This paper studies how a shift in competition affects the nature and type of a firm's intrusion on consumers' privacy. I study this question in the context of Android app markets in China.I investigate a regulatory change in 2017 that reduced competition in censored app categories by prohibiting censorship-circumvention tools commonly used to access foreign apps banned by the government. This regulatory change made banned apps less accessible and reduced competition faced by permitted apps in censored categories, but did not affect apps in uncensored categories. I documented empirically that the regulation substantially reduced the availability of VPNs and banned apps in China while increased market concentration in censored categories.

I use a synthetic differences-in-differences approach and compare smartphone permissions requested by permitted apps in censored categories to apps in uncensored categories before and after the regulatory change. By analyzing monthly historical app installation packages, I show that reducing competition led to a significant increase in both the number of permissions requested by apps and the number of permissions with heightened privacy implications. The evidence indicates that the removal of competition increases intrusive privacy practices by altering firms' incentives to compete in the privacy dimension and by shifting firms toward capturing and monetizing the engagement of a more captive user base. I find no evidence that the increase in permission

requests reflects the introduction of new product functionalities, a strategic response to entry, or a major change in apps' business models.

# 1 Introduction

One of the frontier topics in competition policy is whether increased competition leads to more or less privacy for consumers. This has been argued in two landmark antitrust cases against two of the largest tech platforms – against Google and Meta. In both cases, the antitrust authority argued that their market power has led them to reduce privacy for consumers.[1] The U.S. Federal Trade Commission argued in its (amended) complaint against Facebook in 2021 that "without meaningful competition, Facebook has been able to provide lower levels of service quality on privacy and data protection than it would have to provide in a competitive market".[2]

These regulatory arguments reflect a broader intellectual view that the governance of digital technologies is shaped not only by legal rules but also by other institutional forces. In his influential book *Code and Other Laws of Cyberspace*, Lawrence Lessig proposes a widely cited framework in which behavior in cyberspace is regulated through four modalities: law, social norms, markets, and technological architecture (Lessig, 2009). While the roles of legal regulation, social norms, and technological design have received substantial attention in the privacy literature (e.g., Acquisti et al., 2016; Dubé et al., 2025; Goldfarb and Tucker, 2011, 2012; Johnson et al., 2020; Johnson, 2025; Ke and Sudhir, 2023), the role of markets in shaping firms' privacy practices – and the extent to which market competition disciplines privacy outcomes – remains relatively understudied.

Conceptually, the relationship between competition and privacy is ambiguous. On the one hand, stronger competition may discipline firms to offer better privacy protections in order to attract and retain users who value privacy. On the other hand, competition may lead firms to collect and use more consumer data – for example to improve personalization or because competition drives prices toward zero and shifts revenues toward data-based monetization such as targeted advertising. Direct evidence on these competing forces remains limited, with existing studies providing only suggestive evidence that larger firms

[1] https://www.justice.gov/opa/press-release/file/1328941/download

[2] https://www.ftc.gov/system/files/documents/cases/ecf_75-1_ftc_v_facebook_public_redacted_fac.pdf

may adopt more privacy-protective practices. For example, the marketing literature finds that larger firms tend to collect less sensitive data (Kummer and Schulte, 2019) and may be better positioned to implement procedures for informed consent (Campbell et al., 2015).

This paper attempts to provide some of the first empirical evidence to investigate whether a shift in competition affects firms' privacy practices on consumers. Empirically answering this question is challenging. The degree of competition in a market is often simultaneously influenced by factors that also affect how intrusively firms collect consumer data in that market, such as product features, market size, and policy regulations. Changes in competitive structures are infrequent in many markets and often coincide with other time-varying factors, making it challenging to isolate their causal effects.

To address these challenges, I study an unexpected shock to the competitive environment of Chinese app markets introduced by a major change in China's censorship circumvention policy. Due to internet censorship, many popular international apps, such as Google, Facebook, and Dropbox, were banned in mainland China. However, users could still access these apps through censorship circumvention tools, primarily virtual private networks (VPNs). Prior to 2017, using VPNs and similar tools for censorship circumvention was not formally regulated in China. However, on January 22nd, 2017, as part of its effort to strengthen Internet control, the Chinese government issued an official announcement restricting the provision and use of VPNs and similar cross-border Internet connection services. The new regulation targeted the use of such tools for bypassing censorship, prohibited those not approved by the government, and restricted their use to intra-firm communications only. The regulation was followed by an immediate enforcement campaign.

Although the regulation targeted censorship circumvention tools and was not intended to change the competitive environment of Chinese app markets, it nonetheless reduced competitive pressure on domestic apps by further restricting users' access to foreign apps that had previously been banned but remained accessible through VPNs. In censored markets such as news and messaging apps, where many foreign providers were banned, the regulation created an exogenous reduction in the set of available products. Chinese

apps in these markets (henceforth 'permitted apps') therefore experienced an exogenous removal of competition. On the other hand, uncensored markets, such as banking apps and weather apps, were unaffected by the regulation. These markets were not subjected to extensive censorship as of January 2017, and did not have major foreign providers that required circumvention tools to access. The difference in censorship across markets provides an opportunity to identify the causal effect of the competition change introduced by the regulation of censorship circumvention tools. I document that the regulatory change led to a surge in VPN exit and increased difficulty to access banned foreign apps. The average Herfindahl-Hirschman index in censored markets also increased substantially relative to uncensored markets, reflecting an increase in market concentration in censored markets following the regulation.

I use a synthetic differences-in-differences (SDID) approach (Arkhangelsky et al., 2021) to identify the causal effect of reduced competition on the degree of privacy intrusion by apps in censored markets. I analyze a unique dataset of over 300,000 historical app installation packages and examine the Android permissions requests made by permitted apps in censored markets ('treated apps') compared to those by apps in uncensored markets ('control apps') before and after the regulatory change. I find that, relative to the (synthetic) control group, treated apps request significantly more permissions, including more privacy-sensitive permissions, following the reduction in market competition. This difference is contributed by the increase in a wide range of permissions with serious privacy implications and is consistent across a battery of robustness checks. Collectively, the findings suggest that the removal of competition resulted in a substantial increase in privacy-intrusive behavior among affected apps.

I investigate the mechanisms underlying these findings. I show that reduced competition limits consumer choice and weakens firms' incentives to compete on privacy. The increase in permission requests is larger among treated apps that competed more closely with the banned foreign apps on privacy prior to the regulatory change, and in markets where domestic privacy competition was limited and thus less able to discipline firms once foreign

competition was removed. For more salient and sensitive permission requests, the increase is concentrated primarily among close competitor apps and in markets with low domestic privacy competition, whereas less salient permission requests increase broadly across apps and markets. This suggests that firms adopt more intrusive practices when the removal of foreign apps most effectively reduced consumers' outside options and constrained switching.

I further show that reduced competitive pressure shifts firms' strategies toward greater engagement optimization and monetization, which likely explains the rise in aggressive privacy practices. The increase in permission requests is concentrated among treated apps that monetize user engagement through advertising or in-app purchases and is insignificant for apps that do not directly monetize user attention. In addition, the increase in permission requests among treated apps is accompanied by greater adoption of software development tools related to analyzing, retaining, and engaging users. Taken together, the evidence suggests that as switching becomes less effective as a disciplining device, firms invest more heavily in data collection and intrusive technologies that help them extract value from a more captive user base.

## 2 Literature Review

The findings of this paper build on a small but growing literature that investigates the effect of market competition on consumer privacy protection. Much of this literature takes a legal perspective and examines conceptually whether insufficient competition hurts consumer privacy protection and whether antitrust laws should be used to address this concern (Ohlhausen and Okuliar, 2015; Wasastjerna, 2018; Day and Stemler, 2019; Economides and Lianos, 2019). Complementing this legal perspective, a small number of economics and managerial research investigates how market structure affects firms' strategic decisions that have privacy implications, such as whether to adopt data-based personalization and price discrimination (Acquisti and Varian, 2005; Chen et al., 2022). The majority of this literature so far has been conceptual or theoretical in nature. One exception is perhaps Marotta-Wurgler (2016), who examines the privacy policies of 249 firms in seven mar-

kets. They conclude that firms in more competitive markets follow stricter data security standards. Despite the importance of their findings, Marotta-Wurgler (2016) relies on cross-market comparisons and does not address the endogeneity of cross-sectional differences in competition. This paper adds to the existing literature by empirically investigating the *causal* effect of competition on firms' privacy practices. Unlike previous studies that are either theoretical or rely on cross-market comparisons, this paper offers one of the initial empirical studies that identify the causal effect of competition by examining an exogenous shock to the competitive environment introduced by a regulatory change.

This paper is related to, but different from, the literature on how privacy considerations and policies shape market competition. Theoretically, Casadesus-Masanell and Hervas-Drane (2015) show that consumers' privacy considerations can soften competition. Campbell et al. (2015) show that privacy regulations may intensify competition and are more likely to hurt small and new firms. Empirically, Johnson et al. (2023) and Peukert et al. (2022) find that the General Data Protection Regulation (GDPR) increased the concentration of the web-vendor market in favor of the big players. Janssen et al. (2022) find that the GDPR led to more exits and fewer entries in app markets. This paper also investigates the relationship between privacy and market competition, but from a different perspective. The key question of this paper is how competition changes affect privacy, instead of how privacy concerns or regulations affect the competitive environment.

Finally, this paper also adds to the broader literature on the non-price effects of changes in competition. Past research has investigated the effect of changes in competition on a wide range of non-price outcomes, such as product quality (Mussa and Rosen, 1978; Banker et al., 1998; Goolsbee and Petrin, 2004; Matsa, 2011), product variety (Berry and Waldfogel, 2001; Sweeting, 2010; Fan, 2013), investment (Pindyck, 2007), innovation (Igami and Uetake, 2020; Cunningham et al., 2021), and risk taking (Hellmann et al., 2000; Jiang et al., 2018). This paper adds to the existing literature by addressing a previously unexplored aspect: the provision of privacy. The findings of this paper contribute to the understanding of the various consequences of changes in market competition.

Privacy is sometimes thought of as a dimension of product quality, which is extensively studied in the literature on non-price effects of competition. However, privacy differs from typical dimensions of product quality, such as battery life, in at least three ways. First, privacy is often seen as a fundamental human right (Warren and Brandeis, 1989; Economides and Lianos, 2019; Lin, 2022), rather than as a quality dimension that is determined solely by a producer's profit considerations. Second, a firm's choice of privacy practices involves an inherent trade-off between protecting privacy and providing better personalization, both of which are valued by consumers. The costs associated with improving quality (such as extending battery life), on the other hand, are typically not directly recognized or valued by consumers. Third, while consumers can often observe and evaluate product quality, privacy is hard to evaluate (Acquisti et al., 2016; Goldfarb and Que, 2023). Consumers are often unaware of how their data is used and shared (Tsai et al., 2011; Kehr et al., 2015), and their privacy preferences are often uncertain and context-dependent (Spiekermann et al., 2001; Acquisti et al., 2013; Athey et al., 2017). These distinctions emphasize that privacy is an important outcome in its own right, instead of a straightforward extension of product quality.

# 3 Institutional Background

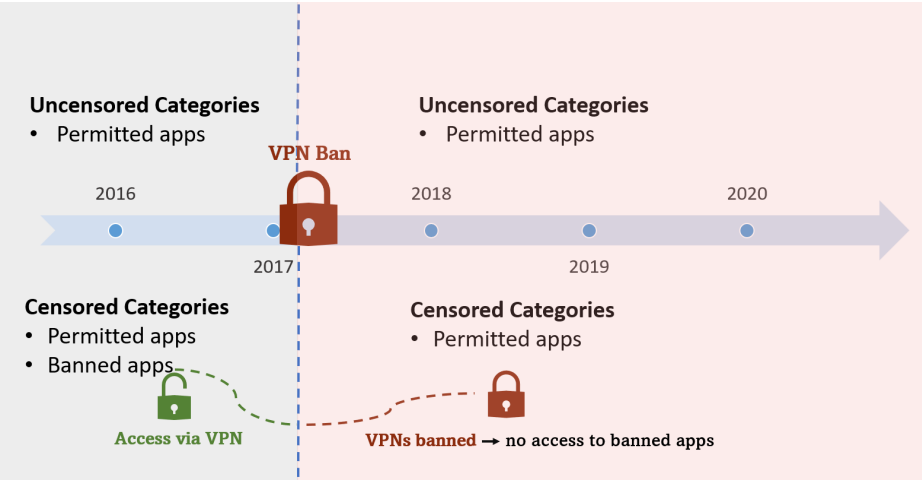
This section provides background information on the development of Internet censorship in China and the role of censorship circumvention tools in the digital ecosystem. Figure 1 summarizes how the ban on censorship-circumvention tools reshaped competition in Chinese mobile app markets. This regulatory shift provides the key source of exogenous variation for the identification strategy in Section 5.

## 3.1 Internet Censorship and Censorship Circumvention in China

In 2003, China launched the Golden Shield Project, a nation-wide project for network security and information infrastructure.[3] The purpose of the project is "to improve the

[3] http://www.gov.cn/zfjs/2006-11/17/content_445189.htm

Figure 1: Illustrating the Regulatory Impact on Competition




efficiency and crime investigation of the police system", and to "provide information sharing and support for public security services".[4] As part of this project, China introduced a set of legislative measures and information technologies, often referred to as the 'Great Firewall,' to block websites and content from foreign providers prohibited in mainland China (Clayton et al., 2006; Chen and Yang, 2019; Kerner, 2022). These restrictions cover many widely used global digital services, including Google, Facebook, YouTube, and Dropbox.

The degree of Internet censorship varies across product categories. While some categories such as communication and media apps were heavily censored with many major foreign providers blocked, other categories such as banking apps and weather apps were not subject to extensive Internet censorship. I refer to the former as 'censored markets' and the latter 'uncensored markets'. Throughout the paper, 'banned apps' refers to foreign apps in censored markets that were blocked by the Great Firewall. 'Permitted apps' refers to apps either in uncensored markets or in censored markets but not blocked.

Despite the Internet censorship, users in China can still access blocked content through circumvention tools, most commonly virtual private networks (VPNs). VPNs encrypt Internet traffic and allow users in mainland China to access the Internet as if located in another

[4] https://www.zgbk.com/ecph/words?SiteID=1&ID=43557&Type=bkzyb

country. Demand for banned websites and apps has made China one of the world's largest VPN markets. A 2014 report estimated that over 90 million Internet users in China had used VPNs to access restricted social media platforms,[5] and a 2017 Statista report found that 31% of surveyed Internet users in China had used a VPN in the previous month.[6]

### 3.2 The Regulation on Censorship Circumvention Tools

For many years, China did not formally regulate VPNs or other censorship-circumvention tools, allowing users to relatively easily bypass Internet censorship and access banned apps and websites. This changed on January 22, 2017, when the Ministry of Industry and Information Technology (MIIT) issued the "Notification on Regulating the Market of Internet Connection Services."[7] The notification marked a major shift in the regulation of Internet access. In particular, Section 2 Article 4 of the notification prohibited self-constructed or leased cross-border Internet connections without approval from China's telecommunications authorities, explicitly naming VPNs as the sole example of such services. It also required telecommunications service providers to keep VPN user records and inform users that these services were limited to intra-firm communication. MIIT simultaneously launched a campaign to "clean up" the Internet connection services market. [8]

The MIIT notification was China's first legal document to explicitly regulate the provision and usage of VPNs. Although it did not ban all VPNs, it substantially increased the difficulty of using VPNs for censorship circumvention. The regulation received widespread domestic and international media coverage and had a major impact on China's VPN market. For example, Bloomberg reported that the government had directed China's three main telecommunications providers to block individual VPN access by February 2018.[9] Apple

[5] https://blog.gwi.com/chart-of-the-day/90-million-vpn-users-in-china-have-accessed-restricted-social-networks/

[6] https://www.statista.com/statistics/301204/top-markets-vpn-proxy-usage/

[7] http://www.cac.gov.cn/2017-01/23/c_1120366809.htm, last accessed on January 22, 2023.

[8] The enforcement campaign was originally expected to complete by March 31st, 2018, but was later extended to March 31st. The extension was announced on May 10th, 2018. https://www.miit.gov.cn/zwgk/zcwj/wjfb/txy/art/2020/art_f53db3415f884c038a77bb2754fb9ea2.html

[9] https://www.bloomberg.com/news/articles/2017-07-10/china-is-said-to-order-carriers-to-bar-personal-vpns-by-february#xj4y7vzkg

also announced it had removed nearly 700 VPN apps from the Chinese iOS App Store following the MIIT notification.[10] Since then, several individuals and firms have reportedly been fined or jailed for selling VPN services.[11,12]

Importantly, although VPNs also encrypt traffic and protect online identity, individuals in China primarily use them to bypass censorship and access restricted content. Because the regulation targeted the circumvention usage of VPNs and because MIIT stated that VPNs remained permitted for secure communications, I interpret the 2017 regulation primarily as a restriction on censorship circumvention, which affected app markets by changing the competition between foreign and domestic providers.

### 3.3 Regulatory Impact on App Access and Competition

This section provides background evidence that the new regulation was enforced and meaningfully affected app market competition.

The analysis uses data on VPNs and banned apps from the Chinese iOS App Store, collected from an app analytics site with historical information on app entry, exit, and reviews. I use iOS data because the focal Android app store does not provide historical data on apps that exited the platform, such as many VPNs, or apps banned by the government.[13] Although iOS and Android app stores differ, both were subject to the same government regulations. Moreover, the analysis focuses on major VPNs and banned apps widely used across device types, making iOS availability a reasonable proxy for Android availability.

I first examine VPN accessibility using a sample of the most popular VPN products in the Chinese market priot to the regulation. Specifically, Figure 2a tracks 49 incumbent VPN apps that appeared at least once in the top 1000 China iOS app rankings in 2016. It

[10] https://cn.wsj.com/articles/CN-BGH-20180118112906

[11] https://www.theguardian.com/world/2017/dec/22/man-in-china-sentenced-to-five-years-jail-for-running-vpn

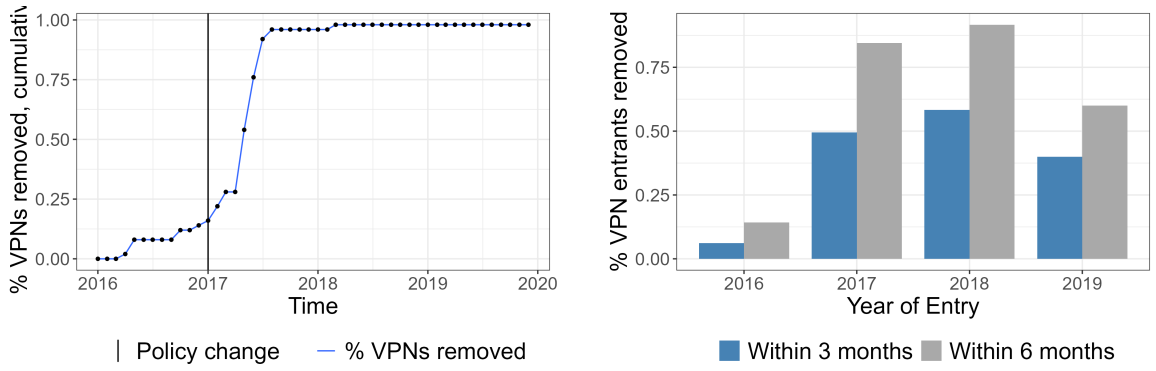
[12] http://www.jinmao.com.cn/cn/researchdetail.aspx?id=3406

[13] The Android app store I study, like most Chinese Android app stores, did not list apps banned in mainland China. Android users typically had to use VPNs to download these apps from official websites or non-Chinese app stores, making their access unobservable in Chinese Android app-store data. On iOS, banned apps remained listed and downloadable from the App Store but could not be used without a VPN or other circumvention tool.

plots the cumulative share of VPNs that became inaccessible after being removed by either developers or the app store. The figure shows a sharp, discontinuous increase in VPN exits beginning about five months after the regulatory announcement in January 2017. By March 2018, all apps in the sample had been removed from the App Store.

The removal of leading incumbent VPNs would not restrict access to banned apps if new providers replaced them. Figure 2b investigates this by plotting the share of new VPN entrants that exited within three or six months of entry.[14] In 2016, only 6% exited within three months and 14% within six months. After the regulatory change, the exit rates rose sharply to 50% and 85%, respectively, and remained high thereafter. This suggests that enforcement displaced not only incumbent VPNs but also new entrants. Consistent with this, Figure A1 shows that over 95% of major VPNs were inaccessible by September 2018, even after accounting for entry.

Figure 2: Removal of Major Incumbent and Entrant VPNs




(a) Removal of Incumbent VPNs

(b) Removal of Entrant VPNs

Notes: Panel (a) analyzes 49 leading VPN apps in China that appeared in the top 1,000 China iOS app rankings in 2016. It plots the cumulative number of these VPNs that exited or were removed from the iOS App Store over time. Panel (b) reports, by year, the share of VPN entrants that exited or were removed within three or six months of entry. The sample includes VPN apps that appeared at least once in the top 1,000 China iOS app rankings between 2016 and 2019. For 2016, only apps entering before October 1 are included to ensure at least three months of follow-up.

To understand how the restricted access to VPNs affected users' experiences with the

[14]The sample comprises VPN apps ranked in the top 1,000 in China's iOS App Store at least once between 2016 and 2019, representing the most popular services during the study period. For the 2016 cohort, only apps entering before October 1, 2016 are included to ensure at least three months of pre-regulation observation.

banned apps, I analyze the reviews and ranking of major banned apps in the Chinese iOS app store. A GPT-based textual analysis shows that the proportion of user reviews mentioning app inaccessibility rose substantially following the regulatory change, suggesting that the ban on VPNs created a significant barrier to accessing these apps. An analysis of within-category app ranking reveals a clear decline in the relative ranking performance of banned foreign apps compared to competitors within the same category following the regulatory change. I discuss the details of this analysis in Section A1.

Finally, I examine whether the regulatory change increased market concentration in censored categories. Because historical downloads or usage data were unavailable, I use monthly app reviews as a proxy for usage.[15] I collected reviews for apps ranked in the top 300 of each iOS category each month from 2016 to 2020,[16] and used them to calculate app-level market shares and category-level HHI. I classify categories as censored if major foreign apps were blocked by the Great Firewall, and uncensored otherwise, as detailed in Section 4.3. A difference-in-differences analysis shows that, after the regulatory change, average HHI in censored categories rose by 146.3 points relative to uncensored categories ($p = 0.016$). This approaches the 200-point threshold used by the U.S. DOJ and FTC to flag significant antitrust concerns, indicating a substantial rise in concentration in censored markets.

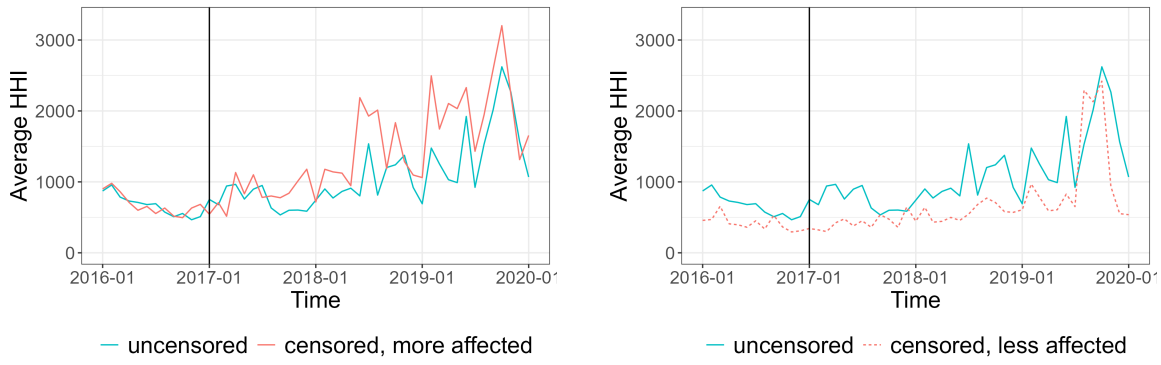
Figure 3 plots average HHI over time by exposure to the regulatory change. Panel (a) compares uncensored categories with more affected censored categories, where multiple banned foreign apps ranked in the top 300 before 2017. Panel (b) compares uncensored categories with less affected censored categories, where banned foreign apps existed but did not enter the top category rankings. Figure 3 shows that the average HHI of the more affected censored categories tracked that of the uncensored categories in 2016, but increased substantially after the regulatory change. This divergence did not occur for the less affected censored categories. This pattern suggests that the HHI increase in censored

[15] Ratings may better proxy usage, as they require less effort than reviews. However, rating dates are unavailable unless linked to reviews, so I use monthly reviews to estimate historical trends.

[16] Monthly Chinese iOS category rankings were collected on the first day of each month from January 2016 to January 2020.

categories reflects reduced competitive pressure from banned foreign competitors, especially in categories where they had been important players.

Figure 3: Policy Change Affected Market Concentration




(a) More Affected Categories

(b) Less Affected Categories

Notes. This figure shows the monthly Herfindal Hirschman Index (HHI) for each market (app category), calculated with the number of iOS user reviews as a proxy for app usage. A DID regression analysis shows that the average HHI for censored markets increased by 146.3 points ($p = 0.016$) following the regulatory change, compared to uncensored markets, suggesting that the regulatory restrictions on censorship circumvention tools led to a notable increase in market concentration in censored markets.

### 3.4 Summary and Discussion

The background evidence shows that the regulatory change substantially reduced access to censorship-circumvention tools and, in turn, to banned apps. As a result, it weakened competitive pressure on permitted apps in censored markets and increased market concentration. Two caveats, however, are important to note.

First, this evidence does not imply that Chinese consumers' access to banned apps was fully eliminated. While major VPNs became severely restricted, smaller providers may have faced less monitoring, and technically sophisticated users could develop less detectable circumvention tools. The regulatory change should therefore be viewed as substantially raising access barriers and reducing, rather than eliminating, competitive pressure from banned foreign apps. Second, the market concentration analysis provides suggestive evidence that

VPN regulation affected competition in censored categories. Because reviews are an imperfect proxy for usage, and the usage-review relationship is not observed or modeled, these results should be interpreted as qualitative rather than precise estimates of the regulation's impact.

# 4 Data and Measurement

## 4.1 Data

I obtained the app data from a major third-party Android app store in China. As Google Play was banned in China, third-party app stores are a major source for Android users to access and download apps. The app store I studied was widely recognized as one of the top ten third-party Android app stores in China in terms of market share. The apps listed on this app store are therefore a representative sample of the Android apps available to Chinese users. My data includes apps listed under each category defined by the app store. A unique feature of this app store is that it provides historical versions of an app in addition to the latest version. My data includes descriptive information for the most recent version of each app at the end of every month during my study period, including its name, developer, size, and release date.[17] The data contains 10,976 apps listed on the Android app store and their 306,756 monthly historical versions from January 2016 to January 2020. I collected permission used in each monthly version of each app in my sample by analyzing their Android Package Kit (APK) files. An Android Package Kit is a downloadable package that includes the code and assets needed to install and run an app on an Android device. In particular, it includes the AndroidManifest.xml file (refered to as the Manifest file hereafter), which is a mandatory xml file for Android apps to communicate information such as the permissions requested and the structure of basic app components to the Android operating system.[18]

[17] The dataset contains the latest version of a month if an app released multiple versions in that month.

[18] See here for the official description for the AndroidManifest file: https://developer.android.com/guide/topics/manifest/manifest-intro

I also complemented the permissions data with information on the permission group and protection level associated with each Android system-defined permission, collected from Android's developer guide.[19] A permission group refers to a category of functionally-related permissions.[20] The protection level is a categorical variable that reflects the potential privacy or security risk a permission poses to the user.[21] The Android system defines four main protection levels for permissions: normal, dangerous, signature, and signatureOrSystem.[22] Of the four types, 'normal' permissions are the only group Android considers to have low privacy risk, though it is widely recognized among developers that even 'normal' permissions can sometimes involve significant privacy risk. After app installation, these permissions will be granted automatically to apps when requested, while permissions of other protection levels will only be granted to apps after certification or user authorization. In the analysis that follows, I adhere to the classification by Android and use 'normal' vs. 'sensitive' permissions as the primary way to classify permissions.

I collected data on banned foreign apps from the GreatFire Analyzer. Operated by GreatFire.org, the GreatFire Analyzer is a web tool that has been monitoring websites and keywords censored by the Great Firewall since 2011. In particular, it tracks the censorship status of the Alexa top 1000 domains in China.[23] I collected the list of Alexa top 1000 domains that were blocked by the Great Firewall prior to January 1st, 2017, as reported by the GreatFire Analyzer.[24] I matched the banned domains with the names and categories of their associated Android apps by searching for and matching the keyword of each banned domain name in Google Play. I also supplement this list with websites and apps reported

[19]Source: https://developer.android.com/reference/android/Manifest.permission, accessed on December 30th, 2022

[20] https://developer.android.com/guide/topics/manifest/permission-group-element

[21]The protection level of a permission also affects the procedure the Android system follows when an app requests the permission.

[22]See here for an explanation of each permission level: https://developer.android.com/guide/topics/manifest/permission-element

[23]Source: https://en.greatfire.org/search/alexa-top-1000-domains, accessed on 31st October, 2022.

[24]The GreatFire Analyzer tests each of the target domains on a regular basis and reports whether the domain was banned at the time of the test. Most domains were either completely banned or not banned in all the tests, but a small number of domains had inconsistent test results. For those domains, I classify them as being 'banned' during the sample period if the domain was banned in five consecutive tests after the first blockage, and 'not banned' if otherwise.

as being blocked in mainland China by various media sources.[25, 26, 27]

## 4.2 Measuring Privacy Intrusion

### 4.2.1 Definition and the Primary Measure

Following Warren and Brandeis (1989), I define privacy as "the right to be let alone". As discussed in Posner (1981), at least three key aspects are included within the general concept of privacy: concealment of information, such as when individuals object to their personal information being shared in unwanted way; freedom from intrusion or disturbance, such as when one complains about unsolicited phone calls or other disruptions violating their privacy; and freedom and autonomy, which extends to issues such as the right to make personal decisions. This paper considers all three dimensions collectively, without explicitly distinguishing among them.

I measure privacy intrusion with apps' permission requests. This measurement considers an increase in the number of permissions requested as the app becoming more privacy intrusive, conditional on its underlying functionality and baseline features (captured by app fixed effects). Permission requests are widely used in the literature as a proxy of the degree of privacy intrusion by Android apps (Mylonas et al., 2014; Krafft et al., 2017; Kesler et al., 2017; Kummer and Schulte, 2019). In the Android environment, apps request permissions from the operating system in order to access users' data or control device functions. The number of permissions therefore represents the level of data access and device control requested by an app.

Requesting access to data or device controls does not necessarily mean that such access is unwanted by consumers. However, it increases the risk of misuse, as firms may exploit these permissions in ways that are unobservable to consumers or inconsistent with their stated purposes. For example, an app might collect location data for navigation purposes

[25] https://zh.wikipedia.org/wiki/, accessed on January 29 2023

[26] https://www.vpnmentor.com/blog/the-complete-list-of-blocked-websites-in-china-how-to-access-them/, accessed on January 29 2023

[27] https://www.saporedicina.com/english/list-of-blocked-websites-in-china/, accessed on September 10, 2024

but also later use it for ad targeting. This risk is particularly salient in the Android system, where permission requests are one-off: during the period of my study, once a permission was granted, the app could access the associated data and resources as much as it wished, unless consumers manually revoke that permission.[28] In fact, recent reports revealed that well-known Chinese apps accessed users' location data as much as 75 times per hour after granted location permissions, even when the app wasn't actively in use.[29] Similar patterns have been observed and reported for other permissions (Reardon et al., 2019; Razaghpanah et al., 2018; Wijesekera et al., 2015).

Note that permission requests are technical commands submitted to the Android system and are distinct from obtaining explicit user consent through privacy agreements. While apps may access and use data in ways that are not fully disclosed in privacy agreements, they generally cannot access protected system resources without first obtaining the corresponding permissions.[30] Therefore, the alternative to requesting a permission is not undisclosed data use, but rather not accessing the associated data at all. Thus, requesting more permissions should be interpreted as an app accessing more types of data—thereby increasing privacy risks—rather than as an indication of greater transparency or respect for user privacy.

#### 4.2.2 Alternative Measures

I also presented two alternative measures for privacy intrusion. The first measure counts only the 'sensitive' permissions and excludes the 'normal' permissions labelled by Android. It serves as a robustness check to the main metric by measuring the changes in permission requests that impose substantial privacy risks to users. Notably, the 'sensitive' permissions cover only a small subset of Android permissions, as the Android permission classification system tends to be lenient on privacy concerns and often overlooks the potential privacy risks associated with 'normal' permissions, particularly when misused or exploited by malicious

[28] Consumers can only manually grant and revoke dangerous permissions but not other types.

[29] http://www.xinhuanet.com/2021-11/03/c_1128024150.htm

[30] Upon app installation, most permissions it requests are automatically granted by the Android system without additional user notification. Only a relatively small subset – those classified as "dangerous" permissions – require users to explicitly grant or deny access at runtime (in my sample, 41 out of 754 permissions requested more than 50 times fall into this category).

parties.

The second measure normalizes the number of permissions by a proxy for the number of functions an app provides. Although requesting more permissions generally imposes higher privacy risks for users regardless of functionality, the benefits from added features may outweigh these risks from the consumer's perspective, leaving the overall welfare implications ambiguous. The main metric captures only the level of potential privacy risk associated with permissions and does not account for the extent to which consumers may be willing to accept such risks when they are justified by functional improvements. In Section 7.3.1, I address this by presenting an alternative measure that captures the number of permissions requested *per function*. My findings are consistent across different measures for privacy intrusion.

### 4.3 Classifying Censored vs. Uncensored Markets

I match the Google Play category of each banned foreign app with the category tag(s) used by the focal Android app store. I classify an app category (as defined by the Android app store) as a censored market if any app in that category was banned by the Great Firewall according to the list identified in Section 4.1.[31,32] 22 app categories were classified as 'censored' (also referred to as the 'treated' apps) and 65 categories as 'uncensored' (the 'control' apps). I summarized the names and frequencies of the censored app categories alongside the 22 largest uncensored categories in Table A1. I present summary statistics for treated and control apps both before and after the regulatory change in Table 1. I address the possibility that treated categories varied in their exposure to the treatment due to differences in the number or prominence of banned apps in section 6.1.1. I check the robustness of the results to different classifications of the treatment and control group in

[31] A few banned apps can be matched to multiple category tags. For example, Facebook has the category tag 'social' on Google play, which has three matching categories on the app store I analyze: 'friend-making', 'communities', and 'chat'. I classified all three categories as in the treated group in this case.

[32] I excluded the category that contains VPN apps, the 'optimization' category, from the analysis. This is because many of the VPN apps were removed from app stores at a faster rate than other apps as a result of the policy (see the discussion in the identification section), making it harder to collect a representative sample of apps in this category during the period of this study (2016-2020). However, including it as a treated category doesn't affect the findings or conclusions of this paper qualitatively.

section 6.2.3.

Table 1: Summary Statistics

| Treatment | Variable | Mean | SD | N |
|---|---|---|---|---|
| Treated | N Permissions | 34.66 | 18.73 | 94658 |
| Treated | N Sensitive Permissions | 9.86 | 3.99 | 94658 |
| Treated | N Activities | 106.20 | 98.14 | 94658 |
| Treated | N Ad Activities | 5.94 | 20.66 | 94658 |
| Treated | N Payment Activities | 1.35 | 1.69 | 94658 |
| Control | N Permissions | 30.49 | 18.78 | 212098 |
| Control | N Sensitive Permissions | 8.93 | 4.07 | 212098 |
| Control | N Activities | 93.96 | 95.62 | 212098 |
| Control | N Ad Activities | 4.50 | 20.52 | 212098 |
| Control | N Payment Activities | 1.21 | 1.78 | 212098 |

This table presents summary statistics for apps in censored (treated) markets and uncensored (control)markets. Each observation represents a monthly snapshot of an individual app.

# 5 Identification

The key challenge to establish a causal interpretation between changes in competition and firms' privacy-intrusive behavior is that market competition is often determined endogenously by factors that may also influence firms' privacy practices. To address this challenge, I exploit variations in Internet censorship across markets prior to the regulatory change in 2017. These pre-existing differences mean that the regulatory change generated different degrees of competitive changes across markets. In censored app markets, the regulatory change reduced the number of available products by restricting access to foreign providers that had previously been accessible via VPNs or similar tools. In contrast, uncensored app markets, such as banking or weather apps, were largely affected because they were not subject to intensive Internet censorship and did not have important foreign providers that required VPNs to access. I identify the causal effect of regulation-induced changes in competition by comparing the privacy practices of permitted (i.e., domestic, non-banned) apps across censored and uncensored markets before and after the policy change, as shown in Figure 1.

I estimate treatment effects using the synthetic differences-in-differences (SDID) approach proposed by Arkhangelsky et al. (2021). SDID is well suited to this setting because censored and uncensored apps may differ systematically before treatment, making the parallel trends assumption less plausible for the unweighted control group. By using data-driven unit and time weights, SDID constructs a more comparable counterfactual trajectory for treated apps. This approach has been applied in recent empirical studies where counterfactual comparability is central to identification (Berman and Israeli, 2022; Lambrecht et al., 2023).

For a panel dataset consisting of $N$ units over $T$ time intervals, the SDID method estimates the average treatment effect on the treated (ATT) by solving the following equation :

$$
\begin{aligned}
&(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) \\
&= \underset{\tau,\mu,\alpha,\beta}{argmin} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \gamma_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \qquad (1)
\end{aligned}
$$

where $Y_{it}$ is the outcome variable for a *permitted* app $i$ at month $t$. $W_{it}$ is an indicator equal to one for *permitted* apps observed in censored markets after January 2017, when the new VPN regulation took effect, and equal to zero for *permitted* apps in uncensored markets, or for *permitted* apps in censored markets before the regulation (note that the analysis does not include banned apps). In my discussion that follows, I refer to permitted apps within censored categories as the 'treated group', and apps in uncensored categories (all of which were permitted) as the 'control group'. $\tau$ is the average treatment effect on the treated. The terms $\alpha_i$ and $\gamma_t$ are the units and time fixed effects, similar to the ones used in a conventional DID model. The terms $\hat{\omega}_i^{sdid}$ and $\hat{\lambda}_t^{sdid}$ denote the unit- and time-specific weights, respectively. The unit weights assign larger weights to control units that more closely approximate the treated units, while the time weights assign larger weights to pre-treatment periods that more closely approximate the post-treatment periods.

# 6 Empirical Results

Most of my empirical analysis focuses on the subset of apps that were listed on the Android platform *before* January 2016 (henceforth 'incumbent apps'), for which I have at least one year of pre-treatment observations. Using this subset of apps makes intuitive sense as they were in the market prior to the regulatory change, and therefore were directly affected by the removal of competition induced by the new regulation. Using incumbent apps also guarantees a balanced panel throughout the sample period I study, which is required by the SDID method. Throughout most of the paper, I analyze this subset of incumbent apps for the period from January 2016 to January 2020, except for in section 7.3.2 where I discuss entry.

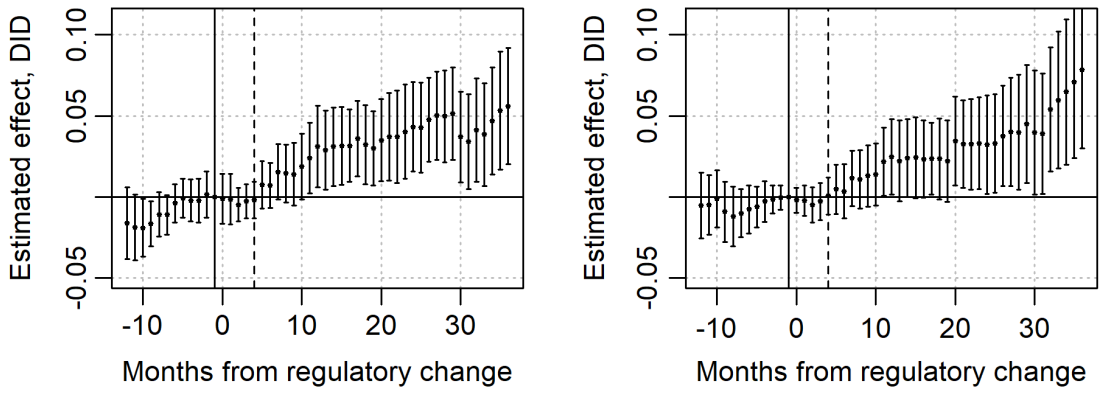
## 6.1 Main Effect

Figure 4 presents the event study plot from a simple difference-in-differences model. Here, I use a standard two-way fixed effects model to estimate the relative difference in permission requests between censored and uncensored categories each month, compared to their baseline difference in January 2017, conditional on app and year×month fixed effects. Panels (a) and (b) use the log of the total number of permissions and log of the number of sensitive permissions requested by an app as the outcome variable, respectively. The solid vertical line indicates the pre-treatment period $t = -1$. Figure 4 shows that apps in censored categories ('treated' apps) exhibited similar trends from those in uncensored categories prior to the regulatory change, but requested significantly more permissions after the change. While Figure 4 provides reassuring evidence that the parallel trends assumption is plausible in this setting, a closer look at Panel (a) also reveals that the difference in total permissions requested between the two groups changed slightly over time prior to the treatment. To provide further reassurance, I next use synthetic difference-in-differences to estimate the treatment effect and to account for any remaining time-varying differences between the two groups.

I present graphical evidence for the synthetic differences-in-differences estimates in Fig-

ure 5. To better assess the resemblance of pre-treatment trends between groups after reweighting, I use the 'overlay' algorithm from 'synthdid' R package, which adjusts the trajectory of the synthetic control group by subtracting its (average) pre-treatment difference from the treatment group. The solid vertical line indicates the pre-treatment period $t = -1$, and the dashed line indicates five months after the initial regulatory announcement, which was when the removal of VPNs substantially increased in Figure 2. Figure 5 provides reassuring evidence that the SDID method generates a synthetic control group that matches well with the treatment group on pre-treatment trend. The divergence between treated and synthetic control apps emerged around five to six months after the initial policy announcement, which aligned with the timing of the large-scale regulatory enforcement of VPN removal shown in Figure 2. The difference in permission requests across groups increased gradually over time until early 2019 and remained relatively stable afterwards, although the difference in sensitive permissions continued to expand.

Figure 4: Regulatory Change Increased Permission Requests, Simple Difference-in-Difference

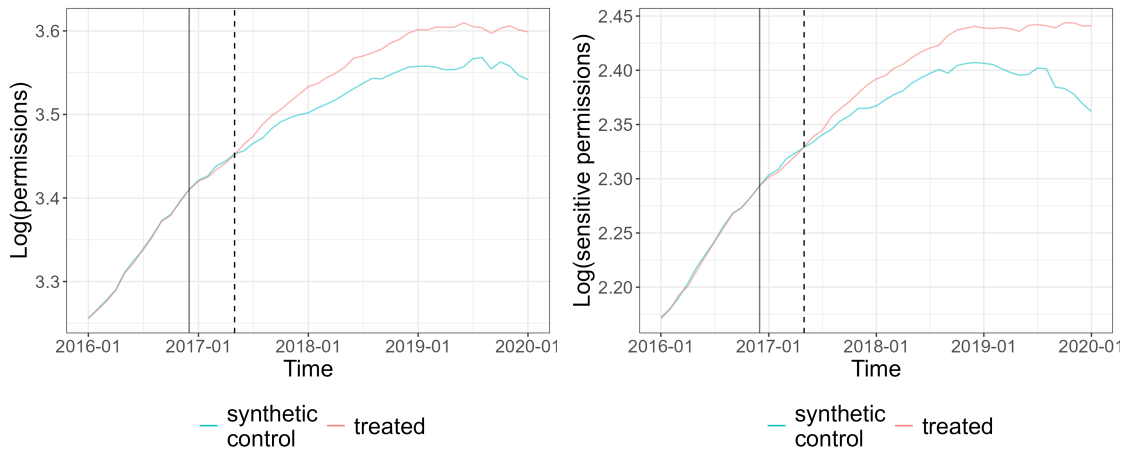



(a) All Permissions

(b) Sensitive Permissions

Notes: This figure presents the event study plot from a simple DID model that estimates the relative difference between censored and uncensored categories at each month, compared to their baseline difference in January 2017, controlling for app-specific and year-month fixed effects. Panel (a) plots the average number of permissions requested by an app. Panel (b) plots the average number of 'sensitive' permissions requested by an app.

Figure 5: Number of Permissions and Sensitive Permissions, SDID




Notes: This figure shows the permission requests from the treated and the synthetic control groups between January 2016 and January 2020. I use the 'overlay' algorithm from 'synthdid' R package to adjust the trajectory of the synthetic control group by subtracting its difference from the pre-treatment average of the treated group. Panel (a) plots the average number of permissions requested by an app. Panel (b) plots the average number of 'sensitive' permissions requested by an app.

Table 2 presents the numerical estimates. Columns (1) of Table 2 take the log of the number of permission requests as the outcome variable and estimates a conventional DID model, which resembles Equation (1) but does not have the unit and time weights $\hat{\omega}_i^{sdid}$ and $\hat{\lambda}_t^{sdid}$. The results show that the restrictions on censorship circumvention tools significantly increased permissions requested by treated apps. I then estimate a synthetic diff-in-diff model in Column (3) using the exact specification of Equation (1). The results shows that an average treated app requested 3.05% ($exp(0.030) - 1 = 0.0305$) additional permissions than the synthetic control group after the regulatory change. This suggests the decrease in competitive pressure in censored markets increased the degree of privacy intrusion of treated apps. I report the jackknife standard errors for the SDID estimator, as suggested by Arkhangelsky et al. (2021). Notably, the SDID estimate is qualitatively similar, but quantitatively slightly smaller than the conventional DID estimate. This suggests that the DID estimator likely overestimates the treatment effect when the raw control data is only a noisy match to the treated units.

The previous analysis uses the number of permissions requested as a proxy for the

degree of privacy intrusion by an app. While the total number of permissions indicates the extent of data access and device control requested by an app, permissions vary in their degree of privacy risks to users. As discussed in Section 4.2, I check the robustness of the findings using the set of permissions that are explicitly labeled as having higher than average privacy risks by Android in Columns (2) and (4) of Table 2. The outcome variable for this analysis is defined as the number of Android-designated 'dangerous' permissions, as well as custom permissions whose protection levels exceed 'normal'.[33] The later are developed by apps and are typically used to facilitate the sharing of sensitive data or functionality across apps. I refer to permissions with higher privacy risks, as captured by this outcome variable, as 'sensitive permissions' throughout the paper. The estimated effects on requests of sensitive permission are qualitatively and numerically similar to those on all permission requests. In particular, the SDID estimate indicates an average of 2.84% ($exp(0.028) - 1 = 0.0284$) additional sensitive permissions requested by treated apps as a result of reduced competition.

Table 2: The Removal of Competition Led to Increased Permission Requests

| | *DID* | | *SDID* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Log(permissions) | Log(sensitive permissions) | Log(permissions) | Log(sensitive permissions) |
| Treatment Effect | 0.038** | 0.033** | 0.030*** | 0.028*** |
| | (0.014) | (0.013) | (0.011) | (0.008) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 137,494 | 137,494 | 137,494 | 137,494 |

Notes: The unit of observation is an app × month. Cluster robust standard errors at the market and year × month level are reported for DID estimates (Columns (1) - (2)). Jackknife standard errors are reported for SDID estimates (Columns (3) - (4)). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 6.1.1 Effect is Due to Shift in Competition

I first provide evidence that the regulatory shock affected permission requests through competition rather than through alternative channels. In Table 3, I stratify treated categories

[33]This includes custom permissions whose protection levels are 'dangerous' or 'signature'. Notably, I did not include the Android-defined signature permissions because these permissions are designed for "apps that are signed with the same digital signature as the Android operating system or are included as part of the system firmware" based on the Android developer manual.

based on how much the competition in the category was affected by the regulatory change. Although all treated categories had at least one internationally popular app (that appeared in the Alexa Top 1000 ranking) blocked by the Great Firewall, banned foreign apps had larger market shares among Chinese users in some categories than in others, making these markets more affected by the regulatory change. To capture this difference in treatment exposure, I classify treated categories into three groups. A category is classified as 'less affected' by the regulatory change if none of the banned foreign apps in that category ranked among the top 300 apps in the category-specific rankings of the Chinese iOS App Store prior to the regulatory change.[34] A category is classified as 'more affected' if at least one banned foreign app ranked among the top 300 apps, and as 'most affected' if at least two such apps appeared, suggesting that banned foreign apps were significant competitors in these markets.

Table 3 reports the results for each stratum. The effect of the regulatory change is larger and statistically significant for treated apps in 'more affected' or 'most affected' categories. The effect is smaller and insignificant for treated apps in 'less affected' categories, where none of the banned foreign apps were among the top choices of Chinese users. Compared to the synthetic control group, treated apps in categories where competition was more affected requested 3.46% more permissions and 2.94% more sensitive permissions; treated apps in categories where competition was most substantially affected requested 5.44% more permissions and 4.39% more sensitive permissions. These findings align with expectations and confirm that the regulatory change on circumvention tools affected apps' behavior through the shift in competition.

### 6.2 Robustness Checks

I did a battery of additional analysis to check the robustness of the main finding. I discuss the results of these analysis below.

[34] My data collection covers the top 300 apps in each iOS category on the first day of each month in 2016. I map the iOS categories to the categories of the focal Android app store analyzed in this paper. I use the Chinese iOS category ranking in this analysis because there was no widely used app ranking that covered all major Chinese domestic Android app stores during this period to the best of my knowledge.

Table 3: Treatment Effect Is Larger for Categories where Competition Was More Affected

| | *Competition Less Affected* | | *Competition More Affected* | | *Competition Most Affected* | |
|---|---|---|---|---|---|---|
| | (1) All | (2) Sensitive | (3) All | (4) Sensitive | (5) All | (6) Sensitive |
| Treatment Effect | 0.018 | 0.023 | 0.034*** | 0.029*** | 0.053*** | 0.043*** |
| | (0.018) | (0.014) | (0.012) | (0.009) | (0.014) | (0.010) |
| App FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 101,381 | 101,381 | 127,449 | 127,449 | 116,277 | 116,277 |

Notes: Synthetic differences-in-differences estimates reported. Outcome variable is the log transformation of the number permissions and sensitive permissions. The unit of observation is an app × month. Jackknife standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 6.2.1 Regulatory Change May Reflect Differential Governmental Control

One identification concern is that the regulatory change may also signal greater overall government scrutiny and control in censored app categories, so treated apps' behavior could reflect differential governmental oversight rather than competition. Table A3 addresses this concern using variation within the news industry. While all news apps faced similar state regulation and oversight, national/global newspapers competed more directly with banned international newspapers than regional newspapers did. Comparing these apps helps isolate changes in competition while holding the industry-level government oversight constant.

In Table A3, I estimate the regulatory effect separately for national/global and regional news apps. If permission requests respond to reduced competition from major international newspapers, the effect should be larger for national/global news apps and smaller for regional ones. If the effect was instead due to differences in governmental control between treated and control categories, the effect should be similar across all news apps as they faced the same industry-level regulation. The results show a much larger effect for national/global apps, especially for sensitive permissions: their requests increased by 7.57%, compared with an insignificant 1.92% increase for regional apps. This pattern supports the competition mechanism rather than differential government influence.

### 6.2.2 Alternative Dependent Variables

I next check the robustness of the results to alternative specifications of the dependent variable. In Columns (1) and (2) of Table A4, I replicate the analysis in Columns (3) and (4) of Table 2 using the number of permission requests and sensitive permission requests (without the log transformation) as the outcome variables in the SDID regressions. The findings are robust to whether the outcome variables are log-transformed.

In Columns (3) and (4) of Table A4, I test whether treated apps requested more permissions simply because they added more new functions after the regulatory change. As discussed in Section 4.2, I divide the outcomes in Table 2 by the number of app activities. This captures the permissions requested per app function and takes into account that apps with more functions might naturally request more permissions. The results show that treated apps increased permission requests even on a per-function basis, suggesting that the effect cannot be explained by added functionality.

### 6.2.3 Alternative Treatment Classification

In Table A5, I test robustness to alternative treatment classifications. While most categories fall clearly into one group or the other, the classification of map and travel apps is less clear. I previously classify them as treated because Google Maps was blocked in China. However, anecdotal evidence suggests limited offline use of Google Maps remained possible for users who downloaded maps data abroad.[35] In Columns (1) and (2) of Table A5, I re-estimate the models without map and travel apps. The results are robust to the inclusion or exclusion of these categories.

A second challenge concerns the classification of lottery apps. GreatFire.org reported no blockages of major non-Chinese lottery services, though it did test blockages of sports-betting sites. In China, betting and lottery are legally distinct: betting is prohibited in

[35] https://www.reddit.com/r/chinalife/comments/1bsf4go/first_timer_is_it_ok_to_use_google_maps_with_a/, accessed September 6, 2025. Google Maps also appeared to have outdated geographic information in China after Google's exit, raising doubts about whether it remained a meaningful competitor even via VPNs.

mainland China,[36] while lottery is permitted and often serves public welfare functions for the state.[37] Given this distinction, I classify lottery apps as part of the control group in the main analysis. However, consumers may still view lottery and betting products as substitutes because both involve uncertain rewards. To assess how the classification of lottery apps affects the findings, I rerun the main analysis in Columns (3) and (4) of Table A5 excluding lottery apps. The results remain robust.

#### 6.2.4 Alternative Window of Analysis

The main analysis period spans the enactment of the EU's General Data Protection Regulation (GDPR) in 2018, which is a major privacy regulation protecting the personal data of individuals in the EU. Although the GDPR does not directly regulate China, it may indirectly affect Chinese apps by changing the behavior of foreign competitors or increasing privacy awareness among consumers and businesses. To address the concern that treated and control apps may differ in their exposure to the GDPR, I re-estimate the model using only pre-GDPR data in Columns (1) and (2) of Table A6. The results show a significant treatment effect relative to the synthetic control group even before the GDPR.

#### 6.2.5 VPN-related Permissions

Because VPNs replace users' real IP addresses with those of VPN servers, and IP addresses can be used to infer location, some users may use VPNs to avoid IP-based location inferences.[38] As the VPN ban makes hiding IP addresses much harder, it might increase consumers' tolerance of location data collection, which may encourage more requests of such permission from firms. In Columns (3) and (4) of Table A6, I re-estimate the model excluding all location-related permissions. The results remain robust, suggesting that this alternative mechanism is unlikely to drive the findings.

[36] The special administrative regions of Macau and Hong Kong have distinct legal systems. In particular, gambling is legal in Macau and is a major source of revenue for the region.

[37] https://legalpilot.com/country/china/ , accessed September 6, 2025

[38] Note that VPNs mask IP-based location but not GPS-based location on mobile devices.

# 7 Mechanism

Why does the shift in competitive pressure change firms' privacy practice? One explanation is that lower competitive pressure means consumers have fewer choices and are therefore less able to discipline firms' intrusive behavior through switching. When limited consumer choice constrains switching, firms have stronger incentives to aggressively collect data and request device access in order to better engage and monetize consumers. In this section, I first document that increases in privacy-intrusive practice are driven by incentives and practices related to capturing and monetizing user engagement. I then provide evidence that intrusive practices are systematically associated with lower firm incentives to compete on privacy due to limited (privacy-friendly) outside options for consumers. Finally, I discuss alternative explanations in section 7.3.

## 7.1 Regulatory Change Increased Intrusive Practices to Capture and Monetize Engagement

I examine whether expanded access to user data and device control reflects firms' attempts to increase and monetize user engagement in response to the shift in competitive pressure.

To better understand firms' incentives, I collect additional data on app monetization models by examining activities documented in their Android Manifest files prior to the regulatory change. I analyze the names of app activities and identify those associated with two primary methods of monetizing user engagement: displaying ads and processing payments for in-app purchases. Based on the presence of such activities, I classify apps into two groups: those that directly monetize usage and engagement – through advertising, in-app purchases, or both – and those that do not directly monetize engagement. In Table 4, I compare the treatment effect between 'high incentive' apps, which employed at least one of these monetization methods prior to the treatment, and 'low incentive' apps, which did not monetize engagement through either advertising or in-app purchases. The findings suggest that the effect of reduced competition on permission requests is concentrated among apps that directly monetize user engagement. Figure 6 further illustrates these patterns by

comparing changes in sensitive permission requests across monetization models. Treated apps that monetize via advertising or in-app purchases experienced significant increases in sensitive permission requests, with the largest percentage increases observed among apps relying on advertising-based monetization. The evidence in Table 4 and Figure 6 is in line with the hypothesis that firms increase privacy-intrusive practices after the treatment to better capture and monetize engagement.

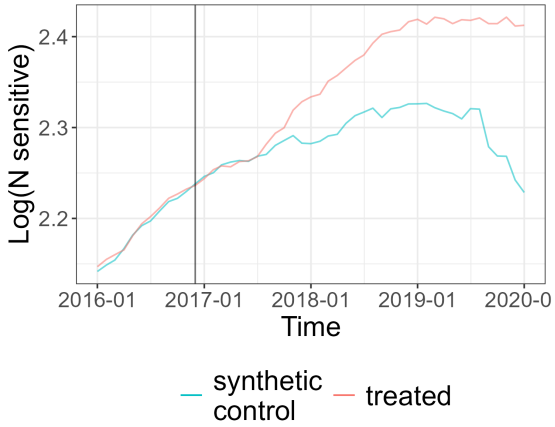
Table 4: Stratification by Monetization Models

| | *Apps with Monetization* | | *Apps w/o Monetization* | |
|---|---|---|---|---|
| | (1) All | (2) Sensitive | (3) All | (4) Sensitive |
| Treatment Effect | 0.029*** | 0.032*** | 0.038 | 0.005 |
| | (0.011) | (0.008) | (0.034) | (0.026) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 114,758 | 114,758 | 22,736 | 22,736 |

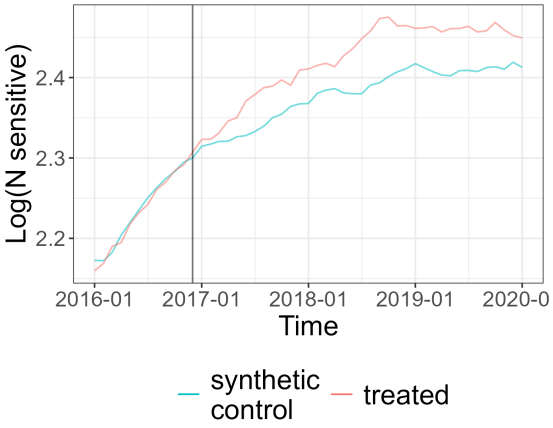
Notes: Synthetic differences-in-differences estimates reported. Outcome variable is the log transformation of the number permissions and sensitive permissions. The unit of observation is an app × month. Jackknife standard errors are reported for SDID estimates. Columns (1)–(2) analyze the subsample of apps that monetized through ads or in-app purchases (or both) prior to the treatment. Columns (3)–(4) analyze the subsample of apps that did not use either ads or in-app purchases prior to the treatment. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

If apps increased permission requests primarily to improve and monetize engagement, we should also observe increases in other concurrent efforts to capture and engage users among treated apps. I next examine whether such patterns exist. In particular, I analyze whether the increase in permission requests is accompanied by a greater adoption of app development tools designed to capture, analyze, and engage users among treated apps. I focus on the use of third party software development kits (SDKs), which are external software packages that developers integrate into their apps to provide specific functionalities, such as tracking user behavior, delivering targeted advertisements, managing authentication, or optimizing engagement. These libraries typically operate in the background and can access device permissions granted to the host app. I obtain information on the SDKs used by each app from its Android manifest file, which allows me to infer the external libraries and

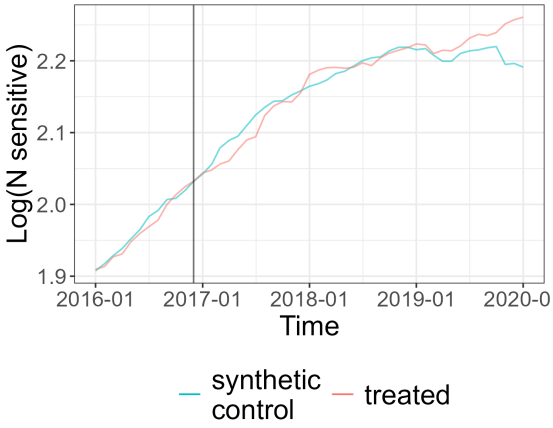
Figure 6: Number of Sensitive Permissions by Monetization Model, SDID




(a) Ads




(b) In-app purchases




(c) No monetization

This figure compares estimated treatment effects on sensitive permissions between apps that directly monetized user attention via advertising or in-app purchases, and apps that did not monetize prior to treatment. Panels (a), (b), and (c) present synthetic differences-in-differences estimates for treated apps using ad-based revenue models, in-app-purchase-based revenue models, and apps using neither monetization method before treatment, respectively. The estimated treatment effects for the three groups are 0.068 (SE = 0.014), 0.049 (SE = 0.018), and 0.010 (SE = 0.024).

components bundled with the app. I then identify SDKs related to user engagement and retention functionalities by pattern matching. Specifically, each SDK name was evaluated against a predefined set of patterns corresponding to functionality categories relevant to user retention/analytics and user engagement. These patterns capture both generic functional

descriptors (e.g., analytics, retention, tracking, metrics, push) and the names of widely used SDK providers associated with those functionalities. As a comparison group, I also identify SDKs related to user identity and authentication. While these SDKs also involve the use of personal information, they are not primarily designed to drive user engagement or retention, making them a useful benchmark for comparison. I provide additional details on the identification of SDK usage and functionality in Appendix A5.1.

I check the correlations between the number of SDKs used and the number of permissions requested in Table A8 using observations prior to the regulatory change. The use of all three types of SDKs positively correlates with the number of permissions requested, even after controlling for app functionality, suggesting that efforts to identify, monitor, and engage users are generally associated with greater data collection and more privacy-intrusive practices.

Table 5 reports estimated effects of the regulatory change on the usage of SDKs by functionality categories. The estimates show significant increases in the use of user analytics, retention, and engagement related SDKs among treated apps following the treatment. The effects are consistent across both the level and log specifications. In contrast, I find no significant effect on the use of SDKs related to user identity and authentication. The coefficients on identity-related SDK adoption are small and statistically insignificant, indicating that the treatment does not broadly increase all forms of data infrastructure. Instead, the pattern of results points to an expansion of app activities aimed to monitor user behavior, improve retention, and stimulate engagement.

Note that my data do not allow me to distinguish whether permission requests originate from first-party code or integrated third-party SDKs within the app. Therefore, the results in Table 5 should not be interpreted as evidence that newly added SDKs directly account for the increase in permission requests. Nevertheless, the findings indicate a broader shift toward engagement-maximizing product design, suggesting that developers respond to the treatment by expanding user monitoring and engagement capabilities, which likely contributes to the observed increase in permission requests (either through first-party im-

plementation or through the integration of third-party tools).

Table 5: Effects on SDK Usage

| | *Analytics & Retention* | | *User Engagement* | | *User Identity* | |
|---|---|---|---|---|---|---|
| | (1) N SDKs | (2) Log(N SDKs) | (3) N SDKs | (4) Log(N SDKs) | (5) N SDKs | (6) Log(N SDKs) |
| Treatment Effect | 0.081** | 0.032** | 0.160*** | 0.064*** | 0.009 | 0.001 |
| | (0.033) | (0.015) | (0.033) | (0.015) | (0.039) | (0.011) |
| App FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 137,494 | 137,494 | 137,494 | 137,494 | 137,494 | 137,494 |

Notes: Synthetic differences-in-differences estimates reported. The unit of observation is an app × month. Jackknife standard errors are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## 7.2 Regulatory Change Reshaped Firms' Incentives to Compete on Privacy

This section examines whether the removal of access to banned foreign apps constrained consumers' privacy choices and thereby weakened domestic firms' incentives to compete on privacy.

I first document that banned foreign apps generally position in the privacy-protective side of the market relative to their Chinese counterparts. To facilitate this analysis, I collected additional permissions data from a representative sample of major banned foreign apps that appeared in the top iOS China app rankings prior to the regulatory change.[39]. I define privacy distance as the difference between the number of permissions requested by a focal treated app and the average number requested by banned foreign apps in the same market, measured in the last observed month prior to treatment. Positive values indicate that the app is more privacy-intrusive than the banned apps, while negative values

[39]Because the focal platform no longer listed banned apps, I obtained historical permissions data for those apps from APKMirror using their most recent APK release prior to the regulatory change: https://www.apkmirror.com/

indicate that it is more privacy-protective.[40] The right panel of Figure 7 displays the distribution of privacy distance among treated apps. The zero line lies toward the left tail of the distribution, confirming that banned apps generally occupy the relatively privacy-protective end of their market, although a substantive mass of treated apps remains at or below this level. Their removal therefore reduced the set of privacy-protective options available to consumers and weakened competitive pressure on privacy in treated markets.

Figure 7: Distribution of Privacy Distance from Banned Apps

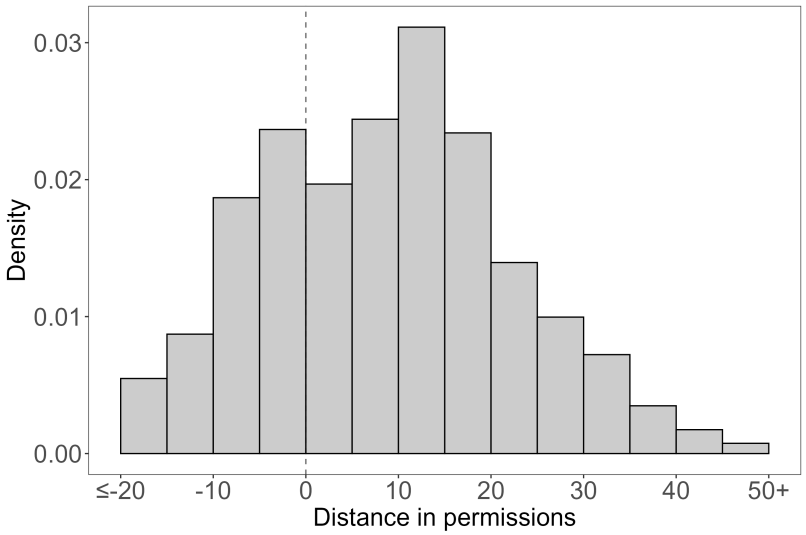



Figure 8: Distribution of privacy distance

Notes. This plot shows the distribution of privacy distance from banned apps within the same category.

Because foreign apps were closer substitutes for relatively privacy-protective Chinese apps than for more intrusive ones, their removal should have larger effects on the former since they competed more directly on privacy. In Table 6, I classify treated Chinese apps as more or less privacy intrusive using a within-market median split based on the number of sensitive permissions requested before the regulatory change. Columns (1)–(2) support this

[40]Because I only have permission data for major banned foreign apps that appeared in the top China iOS rankings, I adopt a broader market definition that maps treated apps to banned foreign apps by genre (the platform's broad category classification) in this analysis. This approach increases the number of treated apps for which the privacy distance measure can be meaningfully constructed.

prediction: permission requests increase substantially more among initially less intrusive, more privacy-protective Chinese apps than among more intrusive apps in the same market. Columns (3)–(4) further show that this heterogeneity is especially pronounced for sensitive permissions, which are likely to be more salient to consumers. Together, these results suggest that firms that had previously differentiated themselves by offering stronger privacy protection responded to the removal of foreign competitors by increasing more intrusive and salient permission requests. This pattern is consistent with standard models of horizontal differentiation, in which policy-induced changes have larger effects when competitive constraints are more binding.

Table 6: Stratification by the Pre-treatment App Intrusiveness

| | Intrusive Apps | Less Intrusive Apps | Intrusive Apps | Less Intrusive Apps |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | All | All | Sensitive | Sensitive |
| Treatment Effect | 0.016 | 0.040*** | 0.004 | 0.050*** |
| | (0.013) | (0.015) | (0.009) | (0.012) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 66,836 | 70,609 | 66,836 | 70,609 |

Notes: This table reports SDID estimates of the regulatory change on app permission requests, stratified by apps' pre-treatment privacy intrusiveness within their market. Apps are classified based on their percentile rank in permission requests within the market as of December 2016. Apps at or above the median are classified as more intrusive, while those below the median are classified as less intrusive. Columns (1)–(2) report effects on log total permissions, and columns (3)–(4) on log sensitive permissions. The unit of observation is an app × month. Jackknife standard errors are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I present additional evidence on shifts in firms' incentives to compete on privacy by exploring variations in the degrees of domestic differentiation on privacy across markets. In markets where domestic firms differentiate substantially on privacy practices, consumers still retain a range of privacy options even when foreign competitors are removed. Constrained by domestic competition, firms in these markets are less likely to substantially

increase privacy-intrusive practices. In contrast, in markets where domestic firms do not differentiate much on privacy, the removal of foreign competitors, who are generally more privacy-friendly, substantially reduces consumers' privacy options and further weakens domestic firms' incentives to protect privacy. Therefore, we should expect larger treatment effects in markets with lower baseline domestic differentiation in privacy, where the removal of foreign competitors leads to a significant reduction in the competitive pressure on privacy among Chinese firms.

In Table 7, I stratify markets by the extent of privacy differentiation among domestic firms. I measure privacy differentiation using the pre-regulation standard deviation in the number of dangerous permissions requested across Chinese apps within the same market. Markets with above-median standard deviations are classified as high-differentiation markets, while those with below-median standard deviations are classified as low-differentiation markets. Table 7 confirms that the effect of the regulatory change is indeed larger in markets with low differentiation on privacy among domestic firms. Notably, while the increase in total permissions differs only modestly between high- and low-differentiation markets (Columns (1) and (2)), the increase in high-salience sensitive permissions is concentrated in markets with low domestic privacy differentiation (Columns (3) and (4)). This suggests that firms respond more strongly to reduced competitive pressure by adopting more intrusive practices when consumers have fewer outside options. I note though that the evidence in Table 7 should be interpreted as suggestive rather than causal, since differentiation in privacy practices is itself endogenously determined and may be correlated with other underlying features of the market.

### 7.3 Ruling out Alternative Explanations

I now discuss alternative explanations for the effect that are ultimately not supported by the empirical evidence.

Table 7: Stratification by the Degree of Market Differentiation

| | High Privacy Differentiation | Low Privacy Differentiation | High Privacy Differentiation | Low Privacy Differentiation |
|---|---|---|---|---|
| | (1) All | (2) All | (3) Sensitive | (4) Sensitive |
| Treatment Effect | 0.026* | 0.036** | 0.011 | 0.052*** |
| | (0.014) | (0.016) | (0.011) | (0.012) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 79,184 | 58,310 | 79,184 | 58,310 |

Notes: This table reports SDID estimates of the regulatory change on app permission requests, stratified by market-level privacy differentiation. Differentiation is measured as the pre-treatment (through December 2016) standard deviation of high salience (dangerous) permissions across apps within a market. Markets with above-median dispersion are classified as high differentiation and those below the median as low differentiation. Columns (1)–(2) report effects on log total permissions, columns (3)–(4) on log low-salience (install-time) permissions, and columns (5)–(6) on log high-salience (dangerous) permissions. The unit of observation is an app × month. Jackknife standard errors are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 7.3.1 App Functionality

One possible explanation is that the increase in permission requests reflects changes in app functionality. As competitive pressure declines, treated apps may become more profitable and invest in new features. If the newly requested permissions support these features, the welfare implications may differ, especially if consumers value the added functionality more than the associated privacy risks.

I test this hypothesis by combining the previous analysis with data on app functionality. In Android terminology, an activity is a user-facing window, such as 'select photo' or 'send email'. The number of activities therefore provides a proxy for the breadth of an app's functionality. In Column (1) of Table A9, I estimate an SDID model testing whether the regulatory change affected the number of activities in treated apps relative to control apps. The results show no statistically significant difference. If anything, treated apps added slightly fewer functionalities on average, consistent with theories that reduced competition lowers quality and innovation.

### 7.3.2 Entry and Strategic Differentiation from Entrants

The analysis so far focuses on incumbent apps available before the 2017 regulatory change. Another potential explanation is that treated apps became more privacy-intrusive in response to new entrants to the market. I examine this possibility in Columns (2)–(4) of Table A9. Here, entry refers to the appearance of a new app in the relevant app-store category. Because of data limitations, I observe entrants only if they survived until my data collection in August 2022; apps that entered and exited earlier are unobserved. The entry results should therefore be interpreted as conditional on survival.

Columns (2) - (4) show no significant difference in the number of entrants, or the permissions requested by entrants between treated and control markets. This suggests that the removal of competition affected treated markets mainly through incumbent behavior rather than increased entry. This is plausible because the main competitors of banned foreign apps were likely established incumbent Chinese apps. The demand affected by the removal of banned apps is more likely to be captured by these incumbent firms, rather than by new entrants.

### 7.3.3 Changes in Monetization Models

Finally, I examine whether the regulatory change caused shifts in monetization strategies that could explain changes in data collection and permission requests. In Columns (5)–(8) of Table A9, I test whether treatment affected the adoption or intensity of ad-based and in-app-purchase-based monetization. Columns (5) and (7) show no statistically significant effect on whether treated apps use advertising or in-app purchases, while Columns (6) and (8) show no effect on monetization intensity, measured by the number of ad or in-app-purchase activities. Together with the findings in Section 7.1, these results suggest that the increase in permission requests is not driven by changes in monetization models or intensity, but instead reflects more privacy-intrusive efforts to engage and retain users among apps already monetizing their users.

# 8 Conclusion

This paper provides one of the first empirical studies of the relationship between competition and firms' privacy practices. I study this question in the context of Chinese Android app markets. In particular, I examine a regulation that prohibited censorship-circumvention tools commonly used by consumers to access apps banned by the government in censored app markets. This regulation reduced the competition faced by permitted apps in censored markets, but did not affect apps in uncensored markets. A synthetic differences-in-differences analysis comparing permitted apps in censored markets to those in uncensored markets reveals that the removal of competition resulted in a significant increase in permission requests by the former compared to the latter. This effect is driven by an increase in requests for a broad range of privacy-intrusive permissions.

I present evidence that the removal of competition affected firms' privacy practices by changing their incentives to compete for consumers on the privacy dimension. I find that reduced competitive pressure shifts firms' strategies toward greater engagement optimization and monetization. As switching becomes less effective as a disciplining device, firms invest more heavily in data collection and intrusive technologies that help them extract value from a more captive user base. I show that the effect cannot be explained by treated apps increasing their app functionality, engaging in strategic differentiation from new entrants, or substantially changing their business models.

The findings of this paper have important implications. The findings highlight that competition is an important factor that affects firms' privacy practices. The results show the removal of competition led to substantial increases in privacy-intrusive behavior across different markets and categories. The impact of insufficient competition on consumer privacy is likely to be especially pronounced in data-rich setting, such as audio/video and image-sharing apps, and when a product profits from monetizing consumer engagement. These findings provide an important perspective for understanding the potential harms of insufficient competition. They suggest that regulators should carefully consider the implications for consumer privacy when evaluating business practices that may reduce competition,

such as data-driven mergers. The findings also highlight the importance of coordinating regulatory efforts to protect privacy and promote market competition.

## References


Acquisti, A., L. K. John, and G. Loewenstein (2013). What is privacy worth? *The Journal of Legal Studies 42*(2), 249–274.

Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of economic Literature 54*(2), 442–492.

Acquisti, A. and H. R. Varian (2005). Conditioning prices on purchase history. *Marketing Science 24*(3), 367–381.

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021). Synthetic difference-in-differences. *American Economic Review 111*(12), 4088–4118.

Athey, S., C. Catalini, and C. Tucker (2017). The digital privacy paradox: Small money, small costs, small talk. Technical report, National Bureau of Economic Research.

Banker, R. D., I. Khosla, and K. K. Sinha (1998). Quality and competition. *Management science 44*(9), 1179–1192.

Berman, R. and A. Israeli (2022). The value of descriptive analytics: Evidence from online retailers. *Marketing Science 41*(6), 1074–1096.

Berry, S. T. and J. Waldfogel (2001). Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics 116*(3), 1009–1025.

Campbell, J., A. Goldfarb, and C. Tucker (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy 24*(1), 47–73.

Casadesus-Masanell, R. and A. Hervas-Drane (2015). Competing with privacy. *Management Science 61*(1), 229–246.

Chen, Y. and D. Y. Yang (2019). The impact of media censorship: 1984 or brave new world? *American Economic Review 109*(6), 2294–2332.

Chen, Z., C. Choe, J. Cong, and N. Matsushima (2022). Data-driven mergers and personalization. *The RAND Journal of Economics 53*(1), 3–31.

Clayton, R., S. J. Murdoch, and R. N. Watson (2006). Ignoring the great firewall of china. In *International workshop on privacy enhancing technologies*, pp. 20–35. Springer.

Cunningham, C., F. Ederer, and S. Ma (2021). Killer acquisitions. *Journal of Political Economy 129*(3), 649–702.

Day, G. and A. Stemler (2019). Infracompetitive privacy. *Iowa L. Rev. 105*, 61.

Dubé, J.-P., J. G. Lynch, D. Bergemann, M. Demirer, A. Goldfarb, G. Johnson, A. Lambrecht, T. Lin, A. Tuchman, and C. E. Tucker (2025). The intended and unintended consequences of privacy regulation for consumer marketing.

Economides, N. and I. Lianos (2019). Restrictions on privacy and exploitation in the digital economy: a competition law perspective. *Forthcoming, Journal of Competition Law and Economics, CLES Research Paper Series 5*(2019), 21–02.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review 103*(5), 1598–1628.

Goldfarb, A. and V. F. Que (2023). The economics of digital privacy. *Annual Review of Economics 15*.

Goldfarb, A. and C. Tucker (2012). Shifts in privacy concerns. *American Economic Review 102*(3), 349–353.

Goldfarb, A. and C. E. Tucker (2011). Privacy regulation and online advertising. *Management science 57*(1), 57–71.

Goolsbee, A. and A. Petrin (2004). The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica 72*(2), 351–381.

Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American economic review 91*(1), 147–165.

Igami, M. and K. Uetake (2020). Mergers, innovation, and entry-exit dynamics: Consolidation of the hard disk drive industry, 1996–2016. *The Review of Economic Studies 87*(6), 2672–2702.

Janssen, R., R. Kesler, M. E. Kummer, and J. Waldfogel (2022). Gdpr and the lost generation of innovative apps. Technical report, National Bureau of Economic Research.

Jiang, L., R. Levine, and C. Lin (2018). Does competition affect bank risk? *Journal of Money, Credit and Banking*.

Johnson, G. (2025). Unearthing privacy-enhancing ad technologies (peat): The adoption of google's privacy sandbox. *Available at SSRN 4983927*.

Johnson, G. A., S. K. Shriver, and S. Du (2020). Consumer privacy choice in online advertising: Who opts out and at what cost to industry? *Marketing Science 39*(1), 33–51.

Johnson, G. A., S. K. Shriver, and S. G. Goldberg (2023). Privacy and market concentration: intended and unintended consequences of the gdpr. *Management Science*.

Ke, T. T. and K. Sudhir (2023). Privacy rights and data security: Gdpr and personal data markets. *Management Science 69*(8), 4389–4412.

Kehr, F., T. Kowatsch, D. Wentzel, and E. Fleisch (2015). Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal 25*(6), 607–635.

Kerner, S. M. (2022). Great firewall of china. https://www.techtarget.com/whatis/definition/Great-Firewall-of-China. Accessed: 2023-01-20.

Kesler, R., M. E. Kummer, and P. Schulte (2017). Mobile applications and access to private data: The supply side of the android ecosystem. *ZEW-Centre for European Economic Research Discussion Paper* (17-075).

Krafft, M., C. M. Arden, and P. C. Verhoef (2017). Permission marketing and privacy concerns—why do customers (not) grant permissions? *Journal of interactive marketing 39*(1), 39–54.

Kummer, M. and P. Schulte (2019). When private information settles the bill: Money and privacy in google's market for smartphone applications. *Management Science 65*(8), 3470–3494.

Lambrecht, A., C. Tucker, and X. Zhang (2023). Tv advertising and online sales: a case study of intertemporal subsitution effects for an online travel platform. *Journal of Marketing Research*.

Lessig, L. (2009). *Code: And other laws of cyberspace*. ReadHowYouWant. com.

Lin, T. (2022). Valuing intrinsic and instrumental preferences for privacy. *Marketing Science 41*(4), 663–681.

Marotta-Wurgler, F. (2016). Self-regulation and competition in privacy policies. *The Journal of Legal Studies 45*(S2), S13–S39.

Matsa, D. A. (2011). Competition and product quality in the supermarket industry. *The Quarterly Journal of Economics 126*(3), 1539–1591.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic theory 18*(2), 301–317.

Mylonas, A., M. Theoharidou, and D. Gritzalis (2014). Assessing privacy risks in android: A user-centric approach. In *Risk Assessment and Risk-Driven Testing: First International Workshop, RISK 2013, Held in Conjunction with ICTSS 2013, Istanbul, Turkey, November 12, 2013. Revised Selected Papers 1*, pp. 21–37. Springer.

Ohlhausen, M. K. and A. P. Okuliar (2015). Competition, consumer protection, and the right [approach] to privacy. *Antitrust LJ 80*, 121.

Peukert, C., S. Bechtold, M. Batikas, and T. Kretschmer (2022). Regulatory spillovers and data governance: Evidence from the gdpr. *Marketing Science 41*(4), 746–768.

Pindyck, R. S. (2007). Mandatory unbundling and irreversible investment in telecom networks. *Review of Network Economics 6*(3).

Posner, R. A. (1981). The economics of privacy. *The American economic review 71*(2), 405–409.

Razaghpanah, A., R. Nithyanand, N. Vallina-Rodriguez, S. Sundaresan, M. Allman, C. Kreibich, and P. Gill (2018). Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*.

Reardon, J., Feal, P. Wijesekera, A. Elazari Bar On, N. Vallina-Rodriguez, and S. Egelman (2019). 50 ways to leak your data: An exploration of apps' circumvention of the android permissions system. In *Proceedings of the 28th USENIX Security Symposium (USENIX Security 19)*, pp. 603–620.

Spiekermann, S., J. Grossklags, and B. Berendt (2001). E-privacy in 2nd generation e-commerce: privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pp. 38–47.

Sweeting, A. (2010). The effects of mergers on product positioning: evidence from the music radio industry. *The RAND Journal of Economics 41*(2), 372–397.

Tsai, J. Y., S. Egelman, L. Cranor, and A. Acquisti (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information systems research 22*(2), 254–268.

Warren, S. and L. Brandeis (1989). The right to privacy. In *Killing the Messenger*, pp. 1–21. Columbia University Press.

Wasastjerna, M. C. (2018). The role of big data and digital privacy in merger review. *European Competition Journal 14*(2-3), 417–444.

Wijesekera, P., A. Baokar, Y. Hosseini, S. Egelman, D. Wagner, and K. Beznosov (2015). Appops: Fine-grained control of android permissions. In *Proceedings of the 24th USENIX Security Symposium (USENIX Security 15)*, pp. 821–836.

## A1 Additional Evidence for Changes in Competitive Pressure

In Figure A2a, I analyze user reviews of a sample of 16 banned apps that ranked among the top 1000 on the China iOS app store in 2016.[41] These apps represent the most popular banned apps among Chinese users prior to the regulatory change, and were ranked highly despite being banned in mainland China. I use gpt-3.5-turbo to analyze whether a review indicates that the user had difficulty accessing or using the app. Figure A2a shows that the average percentage of user reviews mentioning inaccessibility rose substantially following the regulatory change, suggesting that the ban on VPNs created a significant barrier to accessing these apps. In Figure A2b, I further analyze the monthly rankings of banned apps within their category.[42] Figure A2b plots, for the set of major banned apps that appeared at least once in the top 300 of their category during the study period, the share that ranked in the top 300 of their category each month before and after the regulatory change. The figure reveals a clear decline in their relative ranking performance compared to competitors within the same category following the regulatory change. The pattern is also robust to alternative cutoffs, such as the top 100 or top 200 in the respective category.

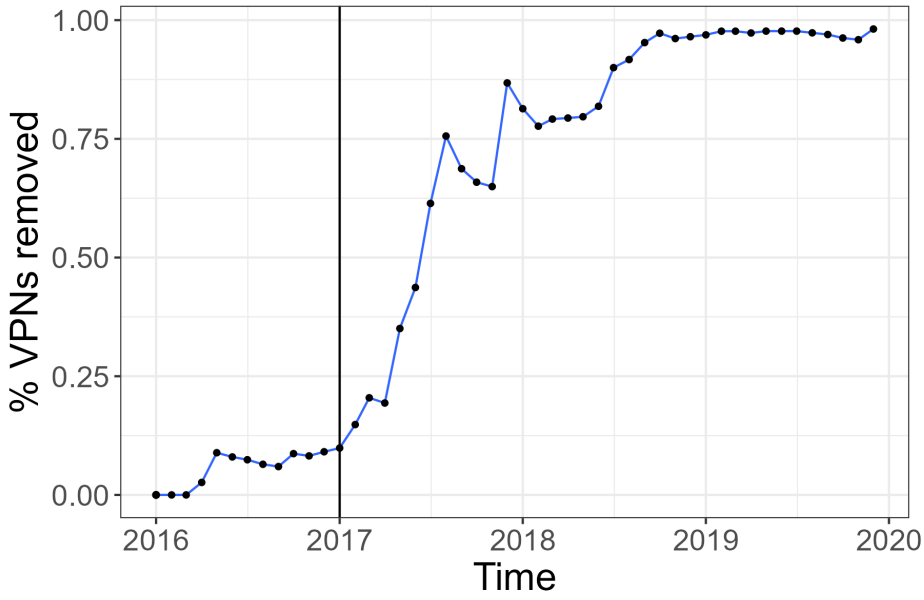
I also investigate whether Chinese users experienced difficulty using banned foreign apps when VPNs became less accessible. I analyze a sample of 16 banned apps that ranked among the top 1000 on the China iOS app store in 2016 in this analysis. These apps represent the most popular banned apps among Chinese users prior to the regulatory change, and were ranked highly despite being banned by the government. To understand uses' experience with these apps, I randomly sampled 200 user reviews[43] of each app and used the gpt-3.5-turbo model with few-shot prompting to determine whether a review indicated that the

[41] Note that users did not need a VPN to review a banned app in the iOS App Store. To leave a review, users had to have previously downloaded the app, but downloading a banned app from the Chinese iOS app store also did not require a VPN. A VPN was only necessary when actually using the app after it had been downloaded.

[42] This analysis includes banned apps that appeared at least once in the top 300 of their category during the study period, representing major players in their market.

[43] Notably, in the iOS app store, writing reviews for banned apps does not require VPNs. Users can post reviews for an app directly via its product page in the iOS app store.

Figure A1: Removal of All VPNs




Notes: This figure analyzes 269 VPN apps that ranked in the top 1,000 of China's iOS App Store at least once between 2016 and 2019, including both incumbents and post-2016 entrants. It thus captures the most popular VPNs among Chinese users over this period. Each month, the figure plots the share of all VPNs which had ever entered the market by that month that were inaccessible. By September 2018, over 95% of these leading VPNs were inaccessible - either removed by their developers or by the platform.

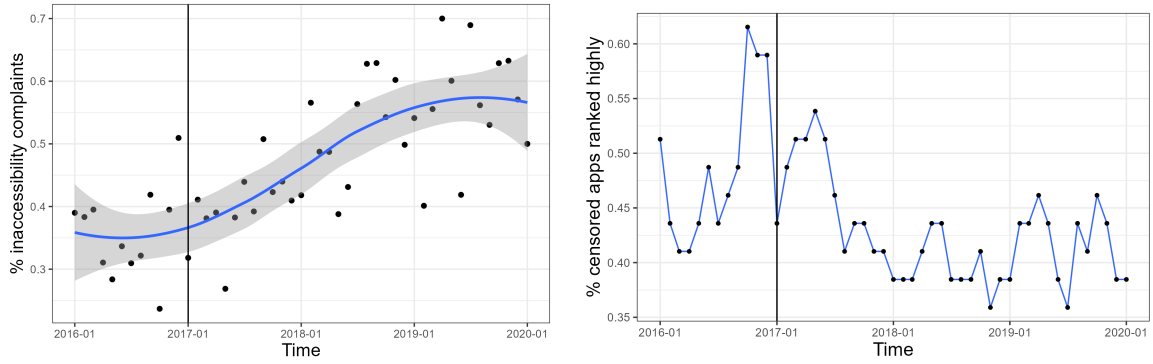
user had difficulty accessing or using the app. Figure A2a plots the average percentage of reviews complaining about inaccessibility over time. The share of user complaints about the inaccessibility of banned apps rose substantially following the regulatory change, indicating that the censorship introduced a significant barrier to accessing these apps.

To better understand the performance of banned apps relative to permitted ones after the regulatory change, I collected additional data on monthly category-specific ranks for Chinese iOS apps. I then analyzed the rankings of banned apps within the same category.[44] Figure A2b plots the share of banned apps that ranked in the top 300 of their category over time. The figure reveals a clear decline in their relative ranking performance compared

[44]The banned apps I analyze are those that appeared in the top 300 of their category at least once during the study period, representing major players in the market.

to competitors within the same category following the regulatory change. This pattern is robust to alternative cutoffs, such as the top 100 or top 200.

Figure A2: Changes in Accessibility of Censored Apps




(a) Inaccessibility Complaints of Censored Apps

(b) Category Rank of Censored Apps

Notes: Panel (a) analyzes a random subset of 200 user reviews for each of the 16 banned apps that ranked among the top 1000 in the iOS app rank in 2016. These banned apps represent the most popular banned foreign products among Chinese users prior to the regulatory change. This figure shows the average percentage of user reviews complaining about these apps being inaccessible and shows an increase in the percentage of inaccessibility complaints after the regulatory change. Panel (b) plots the share of banned apps that ranked in the top 300 of their category over time. The figure reveals a clear decline in their relative ranking performance compared to competitors within the same category following the regulatory change. This pattern is robust to alternative cutoffs, such as the top 100 or top 200.

## A2 Additional Definitions and Statistics

Table A1: Distribution of App Categories by Treatment Status

| Group | Category | Percent of Sample |
|---|---|---|
| Treated | Productivity | 4.88% |
| Treated | E-shopping | 4.31% |
| Treated | Information | 3.60% |
| Treated | News | 3.46% |
| Treated | Communities | 3.46% |
| Treated | Videos | 2.32% |
| Treated | Novel | 1.32% |
| Treated | Maps | 1.28% |
| Treated | Chat | 1.25% |
| Treated | Music | 1.03% |
| Treated | Friending | 0.96% |
| Treated | Travel | 0.96% |
| Treated | Livestream | 0.89% |
| Treated | Audio-book | 0.78% |
| Treated | Telecommunications | 0.64% |
| Treated | E-book | 0.53% |
| Treated | Photo book | 0.46% |
| Treated | File management | 0.39% |
| Treated | Short videos | 0.32% |
| Treated | Cloud storage | 0.25% |
| Treated | Browsers | 0.25% |
| Treated | Emails | 0.21% |
| Control | Learning | 5.49% |
| Control | Tools | 5.38% |
| Control | Medical | 4.10% |
| Control | Early childhood education | 3.24% |
| Control | Stocks | 2.78% |
| Control | Childcare | 2.71% |
| Control | Cars | 2.35% |
| Control | Examination | 2.35% |
| Control | Discounts | 2.14% |
| Control | Housing | 2.00% |
| Control | Career | 1.89% |
| Control | Office software | 1.78% |
| Control | Diet and exercise | 1.67% |
| Control | Games (children) | 1.67% |
| Control | Car renting | 1.57% |
| Control | English | 1.57% |
| Control | Door-to-door service | 1.50% |
| Control | Investment | 1.50% |
| Control | Food | 1.32% |
| Control | Banking | 1.25% |
| Control | Ticketing | 1.21% |
| Control | Finance and budgeting | 0.96% |

Notes: This table lists the treated markets (22 app categories) and the top 22 control markets in terms of the number of observations. The percentages are the percentage of observations from each category within the synthetic DID analysis sample.

Table A2: Android-defined Dangerous Permissions

| Permission Type | Permission Name | Permission Description |
|---|---|---|
| Phone | Accept_Handover | Allows accepting ongoing handover calls |
| Phone | Answer_Phone_Calls | Grants the ability to answer incoming calls |
| Phone | Call_Phone | Enables initiating phone calls |
| Phone | Read_Phone_Numbers | Access to read phone numbers |
| Phone | Read_Phone_State | Read current phone state |
| Phone | Use_SIP | Use Session Initiation Protocol for VoIP calls |
| Phone | Add_Voicemail | Add voicemail messages |
| Storage | Access_Media_Location | Access media files with location information |
| Storage | Read_External_Storage | Read external storage contents |
| Storage | Read_Media_Audio | Read audio files |
| Storage | Read_Media_Images | Read image files |
| Storage | Read_Media_Video | Read video files |
| Storage | Write_External_Storage | Write to external storage |
| Microphone | Record_Audio | Record audio using microphone |
| Camera | Camera | Access the camera |
| Sensors | Body_Sensors | Access data from body sensors |
| Sensors | Body_Sensors_Background | Access body sensor data in the background |
| Calendar | Read_Calendar | Read calendar events and details |
| Calendar | Write_Calendar | Create, modify, or delete calendar events |
| Location | Access_Background_Location | Access device location in the background |
| Location | Access_Coarse_Location | Access approximate device location |
| Location | Access_Fine_Location | Access precise device location |
| Contacts | Get_Accounts | Access device accounts |
| Contacts | Read_Contacts | Read contacts |
| Contacts | Write_Contacts | Modify or delete contacts |
| Call log | Process_Outgoing_Calls | Monitor, modify, or abort outgoing calls |
| Call log | Read_Call_Log | Read call history/log |
| Call log | Write_Call_Log | Modify or delete call history/log |
| SMS | Read_SMS | Read SMS messages |
| SMS | Receive_MMS | Receive and process MMS messages |
| SMS | Receive_SMS | Receive and process SMS messages |
| SMS | Receive_WAP_Push | Receive and process WAP push messages |
| SMS | Send_SMS | Send SMS messages |
| SMS | Read_Cell_Broadcasts | Read cell broadcast messages |

Notes: This table summarizes the permission types (groups) of Android system-defined dangerous permissions obtained from the Android developer manual on December 30th, 2022. At the time of our data collection, three additional permission types, 'Activity recognition', 'Nearby devices' and 'Notification' also contain dangerous permissions, but all the corresponding dangerous permissions were introduced after the time of the treatment (January 2017). I therefore exclude these permission types from the analysis in Figure **??**.

## A3 Additional Robustness Checks

Table A3: Treatment Effect Is Larger for National News Apps Compared to Regional News Apps

| | *National & Global News* | | *Regional News* | |
|---|---|---|---|---|
| | (1) All | (2) Sensitive | (3) All | (4) Sensitive |
| Treatment Effect | 0.061 | 0.073** | −0.006 | 0.019 |
| | (0.054) | (0.028) | (0.032) | (0.024) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 92,855 | 92,855 | 94,570 | 94,570 |

Notes: Synthetic differences-in-differences estimates reported. Outcome variable is the log transformation of the number permissions and sensitive permissions. The unit of observation is an app × month. Jackknife standard errors are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Robustness Checks against Alternative Dependent Variables

| | *Alternative DV 1* | | *Alternative DV 2* | |
|---|---|---|---|---|
| | (1) N permissions | (2) N sensitive permissions | (3) Permissions per activity | (4) Sensitive per activity |
| Treatment Effect | 1.340*** | 0.270*** | 0.093*** | 0.040*** |
| | (0.471) | (0.076) | (0.028) | (0.009) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 137,494 | 137,494 | 137,494 | 137,494 |

Notes: Synthetic differences-in-differences estimates reported. The unit of observation is an app × month. Jackknife standard errors are reported for SDID estimates. Columns (1)–(2) use the raw number of total and sensitive permissions as outcomes. Columns (3)–(4) scale these outcomes by the number of Android activities, a proxy for app functions. The results show that treated apps requested more total and sensitive permissions after the regulatory change, including on a per-function basis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness Checks against Alternative Treatment Classification

| | *Alternative Treatment* | | *Alternative Control* | |
|---|---|---|---|---|
| | (1) Log(permissions) | (2) Log(sensitive perm.) | (3) Log(permissions) | (4) Log(sensitive perm.) |
| Treatment Effect | 0.035*** | 0.033*** | 0.029*** | 0.028*** |
| | (0.011) | (0.008) | (0.011) | (0.008) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 134,407 | 134,407 | 137,151 | 137,151 |

Notes: This table reports robustness checks using alternative treatment and control classifications. Columns (1)–(2) exclude map and travel apps, whose treatment status may be ambiguous due to limited access to Google Maps in China. Columns (3)–(4) exclude lottery apps, whose control status may be ambiguous because lottery apps and banned betting apps may be viewed as substitutes. The results remain robust across both alternative samples. Synthetic differences-in-differences estimates reported. The unit of observation is an app × month. Jackknife standard errors are reported for SDID estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Robustness Checks against Alternative Window and Excluding Location-based Permissions

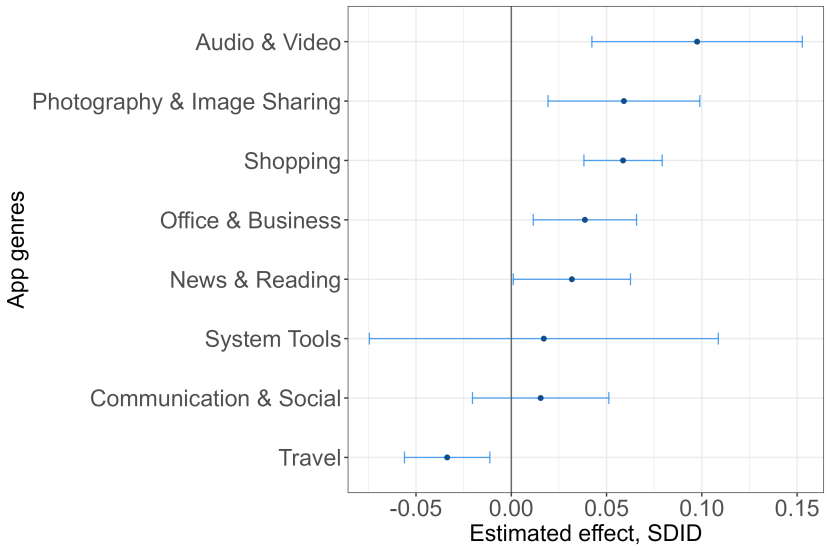
| | *Alternative Window* | | *Excluding Location* | |
|---|---|---|---|---|
| | (1) Log(permissions) | (2) Log(sensitive perm.) | (3) Log(permissions) | (4) Log(sensitive perm.) |
| Treatment Effect | 0.015* | 0.011* | 0.033*** | 0.038*** |
| | (0.008) | (0.006) | (0.011) | (0.008) |
| App FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 81,374 | 81,374 | 137,494 | 137,494 |

Notes: This table reports robustness checks using alternative analysis windows and permission types. Columns (1)–(2) restrict the sample to the pre-GDPR period to address potential differential effects of the GDPR. Columns (3)–(4) exclude location-related permissions to test whether potential changes in tolerance for location data collection drive the results. Synthetic differences-in-differences estimates reported. The unit of observation is an app × month. Jackknife standard errors are reported for SDID estimates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

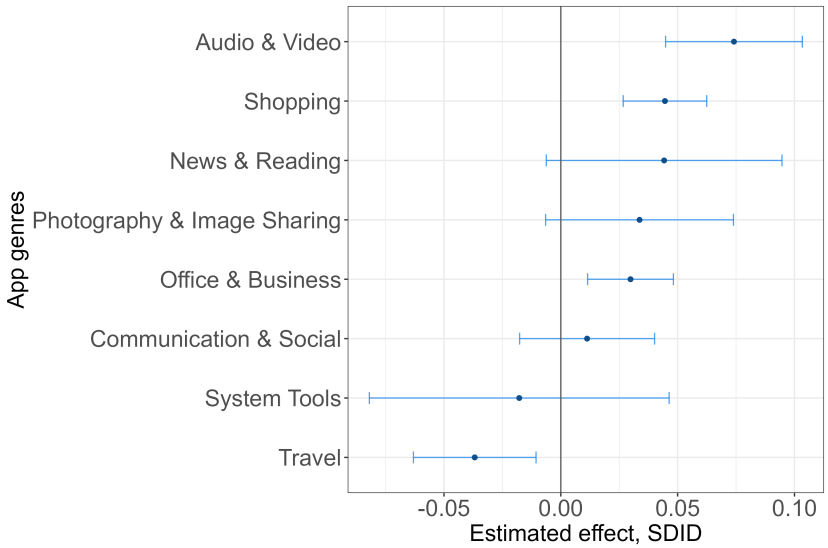
## A4 Additional Evidence for Heterogeneity Analysis

Figure A3: Treatment Effect by App Genre

(a) Estimated effect on the total number of permissions



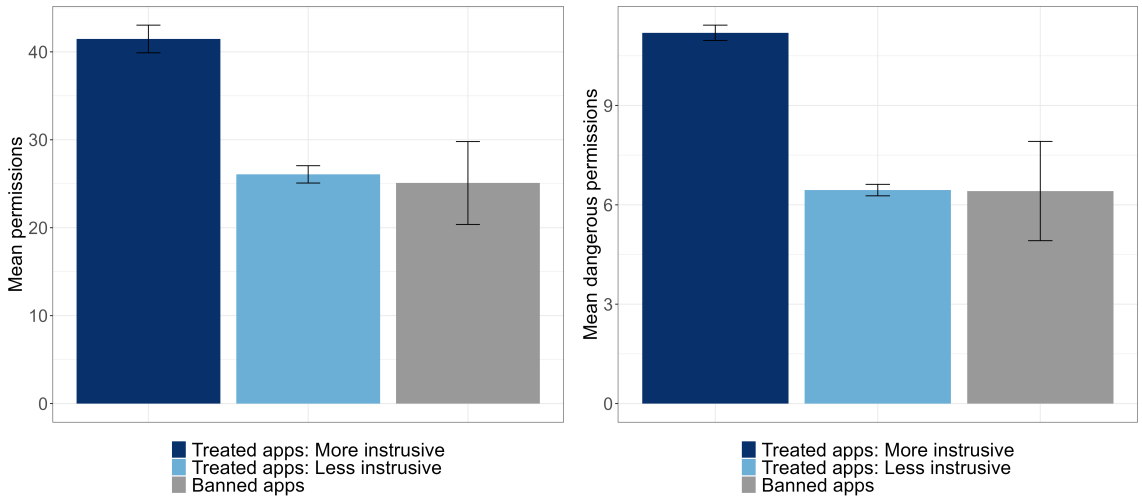

(b) Estimated effect on the number of sensitive permissions




Notes: Panel (a) presents SDID estimates of the effect on the number of permission requests for each app genre. Panel (b) presents the SDID estimates of effect on the number of sensitive permissions requests for each app genre. Genres are defined by the app store I study and contain several app categories (markets) that are functionally related. If a genre has both treated and control categories, I use all control apps, including those in the same genre and those in different genres, to construct the synthetic control group for treated apps in this genre.

# A5 Additional Evidence for Mechanism

Figure A4: Pre-treatment Permission Requests across Apps




This figure reports the mean number of permissions requested per app across app categories. Bars represent sample means, and vertical whiskers indicate 95 percent confidence intervals. The left panel uses the full set of permissions, while the right panel restricts attention to dangerous permissions only. Treated apps are classified as more intrusive or less intrusive depending on whether their number of dangerous permission requests lies above or below the median of apps in the same category.

## A5.1 Additional Details on Identifying SDK Adoption and Functionality

SDK adoption was identified from each app's `AndroidManifest.xml` file by extracting third-party namespaces. In particular, I developed an algorithm that parses the manifest file as an XML object and scans multiple component declarations for fully qualified identifiers in the Android namespace ({`http://schemas.android.com/apk/res/android`}), including `activity`, `service`, `receiver`, `provider`, `meta-data`, and `action` elements. The resulting strings are normalized to a package "root" by retaining the first three dot-delimited segments (e.g., `com.tencent.tauth`). This truncation reflects the common structure of Android identifiers (package + subpackage + class/authority) and yields a stable root-level namespace that can be mapped to SDK vendors.

To separate third-party from first-party identifiers, the algorithm reads the app's de-

clared package name from the manifest root (`package=...`) and constructs a first-party vendor prefix using the first two dot-delimited segments. Identifiers that match the app's own vendor prefix are excluded to avoid counting first-party components, and identifiers beginning with `com.android` are excluded to remove Android system components. To avoid spuriously flagging developer-defined permissions as third-party SDK signals, the algorithm collects all self-declared permissions and, when scanning `<uses-permission>`, removes any entries that match this set. Finally, all extracted roots are deduplicated and returned as a list of unique third-party package roots, which are interpreted as third-party SDKs/libraries adopted by the app.

SDK functionalities were identified using a rule-based classification approach based on regular-expression pattern matching applied to SDK package names. Specifically, each SDK was evaluated against a predefined set of regex patterns corresponding to three functionality categories relevant to user engagement and retention: analytics/retention, push notifications, and identity/authentication. The patterns were designed to capture both generic functional descriptors (e.g., analytics, tracking, metrics, push, notification, auth, login, token) and the names of widely used SDK providers associated with those functionalities (e.g., Firebase Analytics, AppsFlyer, Adjust for analytics; Umeng Push or Meizu Push for messaging; and WeChat or QQ for authentication). The functionality classification was then implemented through a function that systematically searched each SDK name for these patterns and assigned the SDK to the matching category. Table A7 lists the regular expressions used to identify SDK functionality categories.

### A5.2 Additional Evidence for Ruling out Alternative Explanations

Table A7: Regular expression patterns used to classify SDK functionalities.

| **Category** | **Regular Expression Pattern** |
|---|---|
| Analytics & Retention | `(analytics\|stat(s)?\|tracker\|tracking\|metrics\|measure\|event(s)?\|profile\|retention\|cohort\|funnel\|engage(ment)?\|abtest\|experiment\|crm\|growth\|umeng\|talkingdata\|growingio\|zhuge\|sensorsdata\|appsflyer\|adjust\|kochava\|branch\|singular\|firebaseanalytics\|googleanalytics)` |
| Engagement | `(push\|notification(s)?\|notify\|message\|jpush\|getui\|umengpush\|firebasemessaging\|fcm\|huaweipush\|xiaomipush\|mipush\|oppopush\|vivopush\|meizupush)` |
| Identity & Authentication | `(auth\|authentication\|authorize\|authorization\|login\|sign[\s_-]?in\|signin\|signup\|register\|account\|passport\|token\|jwt\|session\|credential\|oauth\|openid\|unionid\|wechat\|weixin\|qq\|qzone\|weibo\|alipay\|huaweiaccount\|miaccount\|oppoaccount\|vivoaccount\|firebaseauth\|comcmicssso\|commobileauth\|comaliauth\|comchuanglanshanyan(_sdk)?)` |

Table A8: Pre-treament Correlations between SDK Usage and Permission Requests

| | (1) Log(Permissions) | (2) Log(Sensitive Permissions) |
|---|---|---|
| Log(Analytics & Retention SDKs) | 0.064*** | 0.026*** |
| | (0.004) | (0.003) |
| Log(Engagement SDKs) | 0.106*** | 0.096*** |
| | (0.005) | (0.004) |
| Log(Identity SDKs) | 0.097*** | 0.051*** |
| | (0.005) | (0.004) |
| Log(App Activities) | 0.208*** | 0.162*** |
| | (0.004) | (0.003) |
| App FE | Yes | Yes |
| Month FE | Yes | Yes |
| Observations | 33,672 | 33,672 |
| $R^2$ | 0.94096 | 0.94025 |

Notes: Sample includes observations prior to January 2017. Standard errors clustered at the app level. The unit of observation is an app × month. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Ruling out Alternative Explanations

| | *Functionality* | *Entry* | | | *Monetization Models* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) N Activities | (2) N Entrants | (3) All Perms | (4) Sensitive Perms | (5) Any Ads | (6) N Ads | (7) Any Payment | (8) N Payment |
| Treatment Effect | −0.558 | −0.125 | 0.728 | 0.309 | −0.013 | −0.169 | −0.003 | −0.054 |
| | (2.137) | (0.327) | (1.846) | (0.484) | (0.013) | (0.539) | (0.012) | (0.058) |
| App FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 137,494 | 4,312 | 4,312 | 4,312 | 137,494 | 137,494 | 137,494 | 137,494 |

Notes: Synthetic diff-in-diff estimates are reported. Jackknife standard errors are in parentheses. The unit of observation for columns (1) and columns (5) - (7) is an app version × month. The unit of observation for columns (2) - (4) is a market (category) × month. The outcome variables for columns (2) - (4) are: the average number of new apps that entered a market at a given month, the average number of permissions requested by entrant apps at a given month (measured at the first month of entrance), and the average number of sensitive permission requested by entrant apps at a given month (measured at the first month of entrance). $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.