The Supply and Demand for Data Privacy: Evidence from Mobile Apps*

Bo Bian[†] **Xinchen Ma**[‡] **Huan Tang**[§]

Abstract

This paper investigates how consumers and investors react to the standardized disclosure of data privacy practices. Since December 2020, Apple has required all apps to disclose their data collection practices by filling out privacy "nutrition" labels that are standardized and easy-to-read. We web-scrape these privacy labels and first document several stylized facts regarding the supply of privacy. Second, augmenting privacy labels with weekly app downloads and revenues, we examine how this disclosure affects consumer behavior. Exploiting the staggered release of privacy labels and use the non-exposed Android version of each app to construct the control group, we find that after privacy label release, an average iOS app experiences a 14% (15%) drop in weekly downloads (revenue) when compared to its Android counterpart. The effect is stronger for more privacy-invasive and substitutable apps. Moreover, we observe negative stock market reactions, especially among firms that harvest more data, corroborating the adverse impact on product markets. Our findings highlight data as a key asset for firms in the digital era.

Keywords: Privacy labels, App Transparency Tracking Policy, Data protection, Mobile apps, Product Markets, Event studies

JEL Classification: D12, D22, G14, G30, L15, L86, Q01

*We thank Pat Akey, Sylvain Catherine, Lorenzo Garlappi, Ben T. Leyden, Jorge Padilla, Eric Richert, Michela Verardo and seminar and conference participants at Bayes Business School, Peking University, London School of Economics, the USC Macro Finance Conference, the Rome Junior Finance Conference, WFA Early Career Women Workshop, Chicago Booth Conference on Data and Welfare in Household Finance, CEPR Household finance Seminar, Toulouse Yale-Regulating the Digital Economy workshop, Platform and Data Workshop at Bank of Canada, the UBC Winter Finance Conference, the 2022 Annual Conference in Digital Economics, and the Economics of Platform Online Seminar for their valuable comments. Bo Bian thanks the Social Science and Humanities Research Council (SSHRC) and the Centre for Innovative Data in Economics Research (CIDER) at UBC for financial support. First version: December 01, 2021. This version: October 20, 2022.

[†]Bo Bian: UBC Sauder School of Business; bo.bian@sauder.ubc.ca

[‡]Xinchen Ma: London School of Economics; x.ma25@lse.ac.uk

[§]Huan Tang: London School of Economics & CEPR; huan.ht.tang@gmail.com

1 Introduction

The past decades have witnessed a digital revolution, shifting economic activities from offline to online markets. Along with this sweeping change, personal data has become an essential element of business, fueling a $227 billion-a-year data industry.[1] The rise of social media and Big Tech showcases the potential of data monetization at scale (Tambe et al., 2020); however, the risk of privacy intrusion and data breach looms large at the same time – of which the Cambridge Analytica scandal is but one alarming example.

As a response to growing public concerns about data privacy issues (Goldfarb and Tucker, 2012b) and cybersecurity risks, bold regulatory moves have emerged in various jurisdictions.[2] Most recently, the Security and Exchange Comission (SEC) proposed making cybersecurity disclosures mandatory for US public firms.[3] Meanwhile, a nascent set of academic work starts to associate firm valuation and corporate policies with cybersecurity risks and data breaches (Akey et al., 2021; Florakis et al., 2020). These discrete events originate from firm's continuous data harvesting, and their impact on firms depends crucially on consumers' attitudes towards privacy. Despite major regulatory efforts and heated public discussions, there is limited large-scale evidence on the market for data privacy. How much privacy do firms supply? Can we consistently measure the scope and the purpose of data collection? How much do consumers demand privacy and does it translate into the valuation of firms that thrive on monetizing personal data?

This paper aims to shed light on the supply and demand for data privacy by exploiting Apple's privacy label policy. Since December 2020, Apple has required all apps to disclose their data collection practices by filling out privacy "nutrition" labels.[4] These privacy labels disclose the types of data being collected and how the data is used in a standardized and easily digestible format. Extracting information from privacy labels, we first quantitatively assess

[1] Source: https://www.strategy-business.com/article/Tomorrows-Data-Heroes. See also Goldfarb et al. (2015); Tucker et al. (2018); Goldfarb and Tucker (2019a,b) for discussions on the digital economy and the value of personal data in marketing and economics.

[2] For example, EU's General Data Protection Regulation (GDPR) in 2018 and California Consumer Privacy Act (CCPA) in 2020.

[3] Source: https://www.sec.gov/news/statement/gensler-cybersecurity-20220309.

[4] The global revenue from mobile apps is over 300 billion USD in 2020 and is expected to reach 600 billion by 2025. Source: https://www.statista.com/forecasts/1262892/mobile-app-revenue-worldwide-by-segment.

the supply of privacy across a large set of mobile apps developed by both private and public firms. More importantly, we examine how consumers and investors react to the disclosure of the (un)ethical collection and deployment of data. To the best of our knowledge, this paper is among the first to investigate the real and financial implications of privacy-related practices.

We first build a database of firms' privacy practices by web-scrapping the privacy labels of popular and active apps in the US as well as the global market, and provide descriptive statistics. We focus on the top 10,000 apps ranked by the total annual downloads across Apple's App Store and Google Play in 2020. These apps account for over 80% of the store-wide downloads and 90% of the store-wide revenue in 2020. For each app, we observe the types of data collected, the specific data items within each data type, and data use, from app functionality to third-party advertising. Apple's official categorization defines 14 data types and 32 specific data items, as well as six data uses. Furthermore, data types and specific data items are displayed in three categories depending on how widely they are shared with other parties: *Data Used to Track You*, *Data Linked to You*, *and Data Not Linked to You*.

Analyzing the web-scrapped privacy labels, we document the following key facts on firms' data collection practices. First, on average, an app in our sample collects 24 data items across 16 data types, with substantial variations across apps (standard deviation is 19 and 11, respectively). We find that 80% of data items collected are used for purposes unrelated to the functionality of an app. Data are most frequently collected for product personalization and developer's advertising or marketing. More importantly, 60% of apps collect data to track users (or their devices) and share user data across different apps, advertising networks, and companies. Worse still, sensitive information collected within this category could be sold to data brokers.

We then investigate what app characteristics are associated with more data collection. Apps that collect more data have a larger market share, a younger age, a higher rating, and more in-app purchases. They are also more likely to be developed by publicly-listed firms. These results hold after controlling for app categories fixed effects. Out of over 20 categories, we find that gaming apps gather the most data for third-party advertising, while shopping apps are the top data collector for multiple data uses including developers' advertising or

marketing, analytics, and product personalization. Apart from these two categories, news, food & drinks, and social networking apps are also heavy data collectors for purposes other than app functionality. These heterogeneities suggest that data from users of these services are easiest to monetize.

After providing insights into apps' data collection practice, we turn to consumers' behavior and investigate the product market effects of mandatory disclosure of privacy-related practices. We merge privacy labels with comprehensive app-level weekly download and revenue data from an app analytics platform. Our unique empirical setting allows us to draw causal inferences on the impact of privacy labels on digital consumption. Since the privacy label policy only applies to iOS apps and affects iPhone users, the corresponding Android apps serve as a natural control group. The iOS and Android versions of any app likely provide the same digital services and collect a similar amount of data from users. Comparing iOS and Android apps within the same period therefore allows us to control for any app-, category-, and market-level demand shocks. The only difference, from the point of view of users, is the disclosure of data collection practices, which remains unchanged for Android users but is altered for iOS users upon the release of privacy labels. Consistent with this assumption, we observe a parallel trend in the demand for the apps across the two platforms before the policy.

Exploiting the privacy label release, we estimate a difference-in-differences model, where iOS apps form the treatment group and corresponding Android apps the control. Our central finding is that, following the release of privacy labels, relative to its Android counterpart, the iOS version of a given app on average experiences a 14% decline in weekly downloads and a 15% decline in revenue from user subscriptions and in-app purchases. This result is robust to different combinations of fixed effects, clusters, and alternative sampling criteria.

If indeed consumers adjust their consumption for app services downwards because of privacy concerns, we should observe a larger decline in downloads and revenue for more privacy-invasive apps. To test this hypothesis, we measure apps' privacy protection along two dimensions: the amount of data collected and the intrusiveness of data uses. In particular, we include a triple interaction term with data collection intensity in the difference-in-differences model. We show that the decline in downloads and revenue increases in in data collection

intensity. For example, the treatment effect for apps that collect data to track consumers is 2.2 times that for apps that do not collect such data. Splitting apps into quartiles based on the number of data items or data types collected, we observe a sharp contrast in consumer reaction. Apps at the bottom quartile of data collection intensity do not experience a statistically significant drop in downloads while apps in the top quartile do. This is consistent with our prediction that app users decrease their consumption of services from less privacy-centric apps once privy to their data collection activities.

We also measure app-level privacy protection by the intrusiveness of data uses. We compute the number of data types or items collected for each of the six data use categories and rank them in the following order (from high to low privacy intrusiveness): third-party advertising, developer's advertising or marketing, analytics, product personalization, other purposes, and app functionality.[5] While the first four data collection purposes tend to be more privacy-invasive, the last two are rather neutral. Consistent with our previous finding, consumers react more negatively when apps collect more data for privacy-invasive purposes.

Admittedly, a consumer's choice to download and use the app depends on other factors besides its privacy protection feature. By revealed preference, a new user chooses to forgo an app only if the disutility from relinquishing access to personal data outweighs the utility gain from the digital services. Therefore, we hypothesize that consumers demand less of an app service only when it is easy to find a close substitute as an outside option. Presumably, apps that are more successful and mature are less substitutable than other apps. Consistent with this, we detect a weaker negative influence of privacy label release on app downloads among popular, mature apps that command a larger market share.

How do consumer preferences over data privacy vary across countries and what drives cross-country heterogeneities? To answer this question, we expand the sample to 95 countries and repeat our baseline difference-in-differences analysis to obtain country-specific DiD estimators. We can associate these coefficients with country-level factors including legal environment, general trust, and survey-based data privacy concerns. We first document the role of legal institutions. Countries with stronger legal protection of privacy and law enforcement react less negatively to privacy labels, presumably because consumers consider

[5] "Other purposes" refer to other purposes not listed in the specified five purposes of data use.

the current legal protection of their personal data adequate. Second, as proxies for general trust, confidence in the press and major companies negatively correlates with consumers' demand for privacy. In the end, consumer attitudes regarding data use and privacy also matter. Consumers in countries with more severe privacy concerns react more negatively to privacy label release.

The weaker demand for privacy-intrusive apps should naturally lead to deterioration in firm performance. Given that privacy labels were introduced only about a year ago, data is not yet available to study longer-term consequences. Instead, we rely on stock market returns for the 485 public firms that have active apps to document the financial implications of the privacy label policy. We show that the stock market responds negatively to the release of privacy labels. Using the release date of the firms' most popular app as the event date, we document that the 6-month cumulative abnormal returns (CAR), evaluated against CAPM and Fama-French factor models, is between -5.74% and -8.17%. We find similar results using the average release for each firm, the first day when the privacy policy became effective (Dec 14, 2020), or the app-level release date as the event date. Consistent with the heteregenous reaction on the product markets, we find a stronger negative effect for firms that collect more data and rely more on mobile users for sales. In particular, heavy data collectors experience a six-month CAR of -10.95%, in contrast to the -4.02% for the light data collectors. Firms in the retail and service industries are more negatively exposed than other firms, seeing a six-month CAR of -11.53%. Consistent with the negative response in the product and financial markets, we also show firms that collect more data experience a larger drop in earnings than those that collect less data.

Taken together, our results highlight the important role of data, as a key asset, for firms thriving on harvesting and monetizing data. We also provide an explanation for the so-called privacy paradox: a lack of transparency in firm data collection practices leads consumers to share excessive data.[6] When transparency improves, consumers and investors are allowed to discipline firms' behavior in the collection and deployment of digital user data through their consumption and investment decisions. As such, our paper makes the

[6]Prior work document the existence of the privacy paradox, using a combination of survey responses and behavioral data. See, for example, Chen et al. (2021) and Athey et al. (2017).

case that standardizing and mandating the reporting of data collection practices potentially fosters sustainable development of the digital economy.

Related literature We contribute to a few strands of literature. First, this paper joins the fast-growing literature that studies the implications of data protection regulations and policies (Goldfarb and Tucker, 2012a; Chiou and Tucker, 2017). Although there has been a lot of academic attention on GDPR,[7] this paper is the first to examine the impact of Apple App Store's privacy label policy, which represents privacy-focused efforts by the technology industry rather than legislative parties. Unlike comprehensive privacy laws enacted at different levels of governments that directly set rules for data handling, the privacy label policy deals with one specific aspect of data protection - the disclosure of data collection practices to consumers and the capital market. By making users of digital services more informed of firms' data collection activities, the privacy label policy relies on the "invisible hand" to discipline firms when it comes to personal data protection. Our results provide justifications for such a policy and encourage future efforts in improving the transparency and quality of firms' data collection policies.[8]

More importantly, our unique laboratory allows us to make several important deviations from the existing work. First, while prior work has examined the impact of privacy regulations on the ability of firms to collect consumer data (Aridor et al., 2020; Bessen et al., 2020; Peukert et al., 2021), hence the supply of privacy, little attention is paid to consumers' responses to privacy regulations. The only paper we are aware of is Schmitt et al. (2020) which documents a negative effect of GDPR on website visits. However, exploiting the adoption of GDPR does not easily distinguish the supply and demand effects since GDPR changes the supply of privacy. Also, due to complex regulatory spillovers (Peukert et al., 2021), it is challenging to define a clean control group when one utilizes place-based policies like GDPR for identification. In contrast, privacy label release is a platform-based shock to disclosure, which should shift the demand of iOS users only without altering the supply of privacy. Com-

[7] Existing work has linked GDPR to industry dynamics (Janssen et al., 2021), VC financing (Jia et al., 2021), and the ability of firms to collect consumer data (Aridor et al., 2020; Bessen et al., 2020; Peukert et al., 2021).

[8] Ebert et al. (2021) show, in an experiment, that concise privacy notices are indeed a promising approach to raise user awareness for privacy information.

bined with the iOS-Android pair-structured data, we are able to identify the causal effect of privacy-related disclosures on consumers' demand for digital services. Second, we focus on apps' downloads and revenue, which directly translate into firm-level sales and profit. We then naturally extend the analysis to stock market reactions. Third, the privacy label policy we study is applicable globally with identical standards and enforcement, allowing us to compare consumer reactions across countries using the same empirical setting.

Second, our findings reveal the relationship between the demand for privacy and firm value. Several recent papers investigate how firms and the financial market respond to cybersecurity events, and in particular, data breaches. Akey et al. (2021) find that data breaches lead to adverse stock market responses and burn reputational capital. Bana et al. (2021) document that firms react to breaches by investing in cybersecurity talent. Florakis et al. (2020) construct a firm-level cybersecurity risk measure based on firms' 10-K filings. Huang and Wang (2021) detect financial consequences of a reported data breach for bank loan terms. Complementing those studies, our paper develops new measures of privacy provision, furthering the understanding of firms' data collection, sharing, and use.

In particular, our new approach to capture the cross-section of privacy provision helps build the foundation for future work on privacy and firms. It is challenging to systematically assess firms' privacy policies. These documents are often unstructured, subject to idiosyncratic formatting, and full of legal terminologies. Pioneering work including Ramadorai et al. (2019) and Amos et al. (2021) crawl and assess the transparency of privacy policies of US public firms and popular websites. Closer to our methodology, some work has examined the privacy permissions on Google's Android market (Chia et al., 2012; Sarma et al., 2012). Compared with previous work, our measure is more comprehensive, transparent, and by construction more consistent across firms and jurisdictions due to the harmonized framework underlying Apple's privacy labels. Importantly, we measure firms' data collection intensity not only in terms of scope but also the *intrusiveness* of data uses. The multi-layer structure of the privacy labels then allows for flexible aggregation of privacy provision along the dimensions of use, data type (e.g., contract information, location, and browsing history), and more granular data item (e.g., name, phone number, email address, physical address).

Last, the paper adds to a burgeoning literature on consumer demand for privacy. A

few studies in the law and marketing literature have investigated the willingness to pay for privacy in the context of Internet-based marketing and offline transactions (Tsai et al., 2011; Jentzsch et al., 2012; Preibusch et al., 2013).[9] Acquisti et al. (2013) discuss the difference between the willingness to pay for a more privacy-protective offer and the willingness to accept a less privacy-protective offer. Tang (2019) uses a field experiment to quantify how much borrowers value their social network ID and employer contact in an online lending context. Al-Natour et al. (2020) show, using survey data, that consumer are less willing to pay for apps with higher privacy uncertainy. Lin (2021) separates consumers' privacy preferences into intrinsic and instrumental components to quantify the welfare implication of consumers' privacy preferences. Prince and Wallsten (2021) use discrete choice surveys to measure individuals' valuation of online privacy across countries and data types and discuss potential policy implications. Our paper differs from prior work in that we use large-scale field data and examines the heterogeneity in consumer demand for privacy among a broad set of countries.

2 Background

2.1 Mobile app market

Android and iOS are the two leading mobile operating systems. In January 2022, the worldwide market share of Android is close to 70% while iOS accounts for around 25%.[10]. In the US, iOS holds down a larger share than Android, claiming more than half of the mobile operating system's market. Consumers typically download and install mobile apps from online app stores for both iOS and Android. A mobile app has a page on each mobile store that provides detailed information about the app, such as its functions, category, update history, ratings, reviews, and price.

Compared with traditional digital medias such as web browsers from a desktop, mobile apps are taking up an increasingly amount of Internet usage time. According to Comscore (2019), smartphones account for 70 percent of the total digital media time in the US. Out of

[9] For a thorough review of this literature, see Acquisti et al. (2016).

[10] https://www.statista.com/statistics/272698

the time spent on smartphones, consumers in both developed and developing countries use mobile apps substantially more than mobile web browsers. For example, in the United States, the ratio of app usage time is 88% (eMarketer, 2020) and this ratio climbs up every year. The COVID-19 pandemic has driven the demand for mobile internet usage even higher.[11] Given the ultimate importance of the mobile app industry, studying consumer and investor behavior in this industry is essential to understanding the role of privacy-related disclosures in today's data-driven economy.

2.2 Apple's privacy label policy

While the size of the mobile app market has been on the rise for many years, disclosures about apps' data collection practices are very limited and tend to be obscure. For example, privacy policy documents are often complex, context-specific, unstructured and hard to quantify. There is also substantial heterogeneity in privacy disclosure across firms, making an objective comparison of data collection practices difficult.

On June 22, 2020, Apple announced its plan to switch to a new version of its iOS operating system, iOS 14, which was eventually released on September 16, 2020. In this biggest update ever to iOS, Apple introduced the privacy "nutrition" labels on App Store product pages of iOS apps to help users understand how apps handle customer data. Resembling the Nutrition Facts label on food packaging, privacy labels disclose an app's data collection practices, listing what data an app collects, why, and what the app developer does with it. The information is all presented in a standardized and easy-to-read format, allowing users to engage in a deep review of the privacy practices of apps they may install on their iOS devices.

More specifically, privacy labels inform app users of the details of data types the app collects from them, the purpose of collecting each data type, whether the data is used to track users or linked to them, and whether the data can be shared with third parties. After Apple's initial announcement of its intention to require privacy disclosures in iOS 14 in June, on November 5, 2020, Apple stated that developers would have to provide this information promptly or risk losing the ability to update their apps. On December 14, 2020, Apple

[11] https://www.weforum.org/agenda/2021/08/how-the-pandemic-sparked-a-data-boom/

officially launched privacy labels for all iOS device owners running the latest version of iOS 14. From that day onward, developers have been obliged to fill out the privacy labels when releasing a new version of their apps. Apple states that the labels have to be up to date and accurate every time a developer submits a new update.

In a nutshell, apple's privacy label policy greatly increases the level of reporting harmonization. Such harmonization can in turn allow investors and other stakeholders especially consumers to compare privacy practices across firms, leading to capital market consequences and real effects.

3 Data and Stylized Facts

3.1 Data

Sampling criteria We restrict our attention to popular and active apps. Out of the millions of different apps available for users to download, the top ranked ones account for the lion's share of downloads and revenue.[12] To obtain a comprehensive list of popular apps, we rely on app-level weekly downloads and revenue information provided by Sensor Tower, a company that consolidates data on millions of mobile apps and publishers from over 100 countries. Download and revenue statistics are estimated by Sensor Tower by combining actual data provided by their partner apps with an array of attributes available from app stores, including app rankings, category, and the number of reviews.

Based on the combined total downloads from Apple App Store and Google play in 2020, we select the top *10,000* apps in each of the following ten countries for which Sensor Tower reports to have wide data coverage: United States, United Kingdom, Russia, France, Canada, Japan, Australia, Germany, Korea, Italy. This list contains not only app names, but also unique identifiers specific to the different versions of each app in Apple App Store (iOS) or Google Play (Android). For example, the unique identifier for the Facebook app is *284882215* in the Apple App Store and *com.facebook.katana* on Google Play. The different versions are connected through a unified app identifier, which works as a pass-through between iOS

[12]For example, big techs like Google, Amazon, and Facebook are all active on the mobile app market and have considerable market power and often are at the center of antitrust discussions.

and Android versions. The iOS identifier for each app allows us to build the URL that directs us to the app's official page in Apple App Store. For Facebook, the URL is simply https://apps.apple.com/us/app/id284882215. We use the list of URLs of all iOS apps in our sample to collect information on their privacy labels through web-scraping. We then merge privacy labels with weekly downloads and revenue statistics of each app on both platforms to form our regression sample. Below we describe the steps we take in detail.

Privacy labels Our first step is to scrape each app's privacy labels from its official webpage on the Apple App Store. Figure A.1 provides a screenshot of Facebook's App Store page. The privacy label section comes right after the ratings and reviews. Every app's privacy label section follows the same structure as illustrated in Figure 1. There are four layers in the privacy labels. The first layer consists of three categories: *Data Used to Track You*, *Data Linked to You*, and *Data Not Linked to You*. If the app doesn't collect any data, there will be a tag showing "*Data Not Collected*". We provide Apple's official description of these three categories in Table B.1. The main difference between the *Data Used to Track You* category and the other two categories is that data items in the former category are shared across different apps, ad networks, and companies. Data items that are not shared should be reported in the latter categories. For *Data Linked to You*, and *Data Not Linked to You* categories, the second layer specifies six purposes of data use: app functionality, third-party advertising, developer's advertising or marketing, product personalization, analytics, and other purposes. Panel B of Table B.1 provides a definition of each data use category. *Data Used to Track You* does not have this second layer. The third layer includes 14 different data types that the app collects and uses, as self-reported by the app's developer. The same data type can appear under all data use purposes in the second layer and all three first-layer categories. We list these 14 data types in Figure 1. Last, the fourth layer further reports detailed data items within each data type in the third layer. One data item can only appear under one corresponding data type but can be collected for more than one data use. There are 32 data items in total. For example, five data items belong to the data type "contact information": name, phone number, email address, physical address, and other user contact information.

The three categories (1st layer) and data types (3rd layer) are displayed on the main page, while the data uses (2nd layer) and specific data items (4th layer) are reported in an expanded window that is displayed when one clicks anywhere in the App Privacy section in the mobile version or on the "see detail" tab at the right corner of the privacy label section as shown in Figure A.1.

For each app, we scrape the privacy labels and structure them in the four-layer structure described above. We started this process in July 2021 and have collected the privacy labels every month since then. It is possible that privacy labels evolve over time if app developers change their data collection practices or make corrections between the initial release and later releases. We make efforts on two fronts to mitigate this potential issue and we find little evidence of major privacy label changes, suggesting that rather a consistent supply of privacy over time.

First, we track the privacy label changes of apps on a quarterly basis from July 2021 onwards. We report in Panel A of Table B.2 that the fraction of apps with label updates ranges between 1% and 2.6% over the seven quarters following the policy. This fraction is even smaller (between 0.4% and 1.1%) among apps developed by public firms, as shown in Panel B. We then examine more granular, month-to-month privacy label changes in Figure A.3, both for the full sample and for each app category. On average, 0.57% (0.21%) of apps increased (decreased) the total number of data items collected, and 0.27% (0.09%) of apps turned on (off) the collection of *Data Used to Track You*. This confirms that adjustment of data collection practices is costly and infrequent. Second, we evaluate the changes between the initial release and the August-2021 version using Web archives from Wayback Machine. Among the 8,092 apps that have released privacy labels by August 2021, we find historical archives for 4,864 apps on Wayback Machine. Focusing on this subsample, we detect a near-perfect correlation (0.98) between the post-August version and the historical versions in the number of data items collected and the collection status of *Data Used to Track You*. Based on these observations, we treat the privacy labels as a time-invariant app feature and use the labels as of August 2021 to measure apps' data collection activities.

Since 2020/12/14, developers are obligated to publish privacy labels when they submit a version update to Apple App Store. We therefore determine the release date of privacy

labels primarily based on app version history. For most apps, the date of its first version update after 2020/12/14 became the release date of its privacy labels. To obtain information on app version history, we scrape the content stored under "Version History" in the "What's New" section of each app's Apple App Store page. A pop-up window contains the 25 most recent update records, also providing information on version release dates, and a description of major changes. Some developers update their apps so frequently that the first version update after 2020/12/14 is not included in the pop-up window. To avoid this potential truncation, we further scrape the full version update history for these apps from App Annie. In this way, we make sure that we always have the full set of version updates on or after 2020/12/14.

Although privacy label releases correspond with the first version update for the majority of apps, we notice that a small group of developers voluntarily release privacy labels before releasing a new version of their apps. To address this issue, we rely on historical screenshots of each app's official webpage in Apple App Store from the Wayback Machine. The first date on which the Wayback Machine captured a page with the privacy label section becomes the release date, if this date precedes the first version update on or after 2020/12/14.

We plot the heat map of the label release dates for the US top 10k app sample in Figure A.2. During the last weeks of December 2020, we observe a concentrated wave of privacy label releases. Around 300-500 apps released their privacy labels every day in the first two weeks after 2020/12/14. Starting from the beginning of 2021, the daily number of apps releasing privacy labels dropped to below 100 gradually. Even still, for almost every day in the first half of 2021, there have been apps disclosing their data collection practices. The variations in the timing of the release of privacy labels allow us to control for platform-specific shocks to digital demand.

App downloads and revenue We collect app-level characteristics and the time series of weekly revenue and downloads for each app from Sensor Tower. Sensor Tower estimates apps' weekly downloads and revenue based on publicly available app ranking history and their proprietary data sources. Sensor Tower also assembles most of the information displayed on an app's download page in Apple App and Google Play, such as app categories, reviews, and

ratings. A comprehensive list of app characteristics we consider in the empirical analysis is provided in Table 1.

Financials and stock price of public firms For the subset of app developers that are public firms, we collect their daily stock returns from the Center for Research in Security Prices (CRSP). We first use the developers' stock tickers, trading exchanges, and firm names (from Sensor Tower) to find the unique identifier, PERMCO, for each developer's company and the PERMNOs for securities they issued. Then, among the securities, we select common stocks that trade on NYSE, AMEX, and NASDAQ. We further require the stocks to be non-delisted as of 2020/12/14. After initial screening, our sample includes 501 public firms and 3616 apps developed by them with valid privacy label information. "No Details Provided" and non-active apps are dropped. To measure firm performance, we focus on quarterly earnings and daily stock returns. Abnormal returns are estimated using CAPM, Fama-French 3 factor model (FF3), Fama-French-Carhart 4 factor model (FFC4), and Fama-French 5 factors model (FF5), and factor returns are obtained from French's data library. The estimation window is 2019/12/01 to 2020/09/30, covering event windows and estimation periods of all label release events in our sample.

3.2 Measuring data collection intensity

We are first interested in the amount of data collected by apps. Concise and uniformly structured privacy labels allow us, for the first time, to measure data collection practices of different apps and firms in a consistent way. Based on the structure of the privacy labels described above, we construct the following measures for the data collection intensity.

First, we build an indicator variable for whether the app collects data used to track customers. This indicator equals one as long as there is a strictly positive number of data items appearing in the *Data Used to Track You* category. As mentioned above, data items in this category can be shared across different apps, ad networks, and companies. For example, an app may share users' device location data or email lists with a data broker. Another example is an app sharing users' list of emails, advertising IDs, or other IDs with a third-party advertising network that uses the information to re-target users in other developers'

apps. We regard apps with such data collection practice to be more privacy invasive.

Second, we construct two continuous variables for data collection intensity: the number of data types and data items collected by each app. Intuitively, an app that collects more data types or data items is more privacy invasive.

Last, we differentiate between different data uses. We count the number of data types and data items collected for each of the six data use categories because different uses may have different implications for consumer privacy. For example, data collection for improving the performance of the app or implementing security measures is less privacy-invasive than data collection for third-party advertising or marketing. Information on data use is only available for the *Data Linked to You* and *Data Not Linked to You* category. Hence, within each of these two categories, we calculate the number of data types and items collected for each data use.

3.3 Descriptive statistics

Our empirical analyses are based on three app samples. Our main sample consists of the top 10,000 apps based on the combined annual downloads across the two platforms in 2020. It is worth noting that not all top 10k apps have a presence in both the Apple App Store and Google Play. For our empirical design described in the next section, we further require the app to have both an iOS and an Android version. This leaves us with 7,692 apps on each of the platforms, or 7,692 iOS-Android pairs. We refer to this sample as the 'top10k US' sample. Although these apps constitute less than 0.3% of the universe of apps, they account for more than 80% of the store-wide downloads and over 90% of the store-wide revenue in 2020, as shown in Figure A.4, Panel A.[13]

In addition, we construct a sample of apps that are developed by publicly listed firms. We refer to this sample as the 'public firm sample'. These apps are not necessarily included in our top10k sample. This sample allows us to study investors' responses to apps' data collection practices. Panel B of Figure A.4 shows the share of downloads and revenue of those apps. Over the year 2020, apps of public firms account for around 35% (50%) of

[13]There are in total about 2.2 million apps on Apple's app store and 3.48 million on Google Play in 2021. Source: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/.

store-wide downloads (revenue).

Last, we are interested in the heterogeneity of supply and demand for data privacy protection across countries. In this regard, we construct a sample of top10k apps in the ten countries mentioned above. We refer to it as the 'international sample.'

Table 1 provides descriptive statistics on the data collection activities of top10k US apps. Out of 7,692 iOS apps with corresponding Android versions, 6,344 apps have provided their privacy labels by August 2021. The remaining 1,348 apps have either become inactive in 2021 or have not provided data by the end of our sample period. Among the 6,344 apps, 60% collect data to track users. The average number of data types collected by these apps is 15.6, with a median of 14 and a maximum of 80. The average number of data items collected is 24.1, with a median of 20 and a maximum of 167. Dividing data by their purposes, the category with the highest number of data items are app functionality and analytics, followed by product personalization and developer's advertising or marketing. However, large heterogeneities prevail among apps. For example, while the average number of data items collected for third-party advertising is only 1.6, the median is 0, and the maximum number is 24. A similar variation can be found for other data use categories.

We report other app characteristics in Panel A. About 70% of apps offer in-app purchases. The average rating of apps is 4.4, with a standard deviation of 1. The average age is 4.5 years old, and the oldest app in our sample is 13 years old.[14] Among the top10k apps, 20% of apps are developed by public firms. In terms of content rating, around 30% of apps are suitable for age 12 or above. Last, the average app is available in 80 countries and the maximum is 102.

In Panel B of Table 1, we present summary statistics of weekly downloads and revenue for the iOS and Android versions of our top10k apps separately. Within each platform, an app may have multiple versions, including a main mobile version, a tablet version, and even a lite version.[15] We therefore present summary statistics on the downloads and revenue of the main version and also all versions combined.

[14]Note that the average age for the universe of apps would be much smaller because we focus on the most popular and active apps that presumably also have a longer lifespan.

[15]For example, Facebook has a lite version on both iOS and Android. See https://apps.apple.com/ae/app/id1393838612 and https://play.google.com/store/apps/details?id=com.facebook.lite.

A few observations are worth noting. First, the iOS versions generally have a higher download volume and revenue. For example, combining all versions, the average iOS weekly downloads for a top10k app is 19.1 thousand, while it is 12.3 thousand for the Android version. The average weekly revenue is $48.2 thousand for the iOS versions and $29.7 thousand for the Android versions. This is not surprising as iOS has a larger market share in the US, and iOS users have a higher willingness to pay for digital services than Android users. Similar to data collection intensity, we observe substantial heterogeneities across apps in their weekly downloads and revenue. The most popular iOS app has more than 6 million weekly downloads and a weekly revenue of $15 million. In Section 4, we examine the dynamics of app downloads and revenue for the iOS and Android versions and provide evidence that they track closely with each other.

3.4 Determinants of data collection intensity

To better understand the determinants of apps' data collection practices, we investigate the association between app characteristics and data collection intensity. We start with regression analysis, where we use the three measures of data collection intensity as the outcome variable and regress it on app characteristics. Results are presented in Table 2. We run each regression with app category fixed effects (even columns) and without them (odd columns). There are in total 26 primary categories, including Books, Business, Developer Tools, Education, Entertainment, Finance, Food & Drink, Games, Graphics & Design, Health & Fitness, Lifestyle, Magazines & Newspapers, Medical, Music, Navigation, News, Photo & Video, Productivity, Reference, Shopping, Social Networking, Sports, Stickers, Travel, Utilities, and Weather.

Results are generally consistent across all three measures. First, apps with a larger market share within their app category collect more data. Because the majority of apps in our sample have a market share close to zero, we visualise this positive relation using the app's ranking in Figure 2. Panels A and B show the number of data types and items collected by apps double sorted by ranking and app categories, respectively. Both panels exhibit darker colors in the high-ranking region within each app category, indicating that highly ranked apps (i.e., those with higher market shares) collect more data than peer apps

with low rankings. This is consistent with Kesler et al. (2017), which also documents a robust positive relationship between market power and data collection.

In addition, Table 2 shows that apps with in-app purchases and a higher rating collect more data. One explanation is that apps with in-app purchases and higher ratings usually have a rich set of app features that require additional user data to function. Older apps tend to collect less information, presumably because these apps have already fine-tuned their data collection practices over time and therefore are not collecting unnecessary information. Apps that operate in more than 100 countries, which we refer to as global apps, collect less data. This can be driven by regulatory spillover – developers have to comply with the most strict data protection policy across the globe. Apps with content rated 12+ and 17+ collect the most data, compared to apps with other content ratings, suggesting that data from teenagers and adults are more valuable than data from kids.

Last, apps developed by publicly listed firms are more privacy invasive in terms of the number of data items and data types collected, but they are not more likely to collect data that track users. This could be driven by public apps having a higher market share, so we further sort apps by ranking in Panels B and C of Figure 2, and plot the number of data items and types collected by public firm apps versus private firm apps against their respective rankings. The red diamonds represent public apps and the blue dots represent private apps. It is visible that for any given ranking, apps developed by public firms still collect between 5-10 more data types and 10-20 more data items. One explanation could be that public firms are more sophisticated in harvesting and monetizing data, incentivizing them to collect more data. An alternative explanation is more cautious reporting from public firms since they are under intense public scrutiny.

Heterogeneity across app category Another important determinant of apps' data collection practices is app category. Figure 2 already shows some suggestive evidence. Based on the heatmap, apps in games, shopping, social networking, news, and travels are top data collectors, among all 23 categories.[16] We can further divide the number of data types and

[16]We combine a few categories together since there are too few apps in these categories. So the number of categories reduces from 26 to 23. More specifically, we reassign Magazines & Newspapers to News, Graphics & Design to Photo & Video, Stickers, and Developer Tools to Others.

items collected by apps into the six data use categories. Figure 3 shows the average number of data types collected by app category. Because data use is reported for both *Data Linked to You* and *Data Not Linked to You*, we report the descriptive statistics for these two first-layer categories in Panels A and B, respectively.

Based on Panel A of Figure 2, shopping apps surpass any other apps in the number of data types they collect in the *Data Linked to You* category. It is followed by travel, food & drink, news, and social networking. Panel B suggests, on the other hand, that apps in games and news categories collect substantially more data in the *Data Not Linked to You* category. Moving to different data uses, we find that these apps collect more data in each of the six categories.

To further examine whether top data collectors harvest data for more privacy-intrusive purposes or not, we regress the number of data types in each of the six data use categories on the indicator of app categories. We rank the intrusiveness of data uses from high to low in columns 1-6 in the following order: third-party advertising, developer's advertising or marketing, analytics, product personalization, other purposes, and app functionality. The first four categories are privacy intrusive while the last two are neutral. Results are reported in Table 3. The reference (omitted) group is the "others" category. Relative to the reference group, shopping apps are collecting more data types in all four privacy intrusive categories. Navigation, news, social networking, and sports apps collect more data in three out of the four privacy-intrusive categories. Moreover, within each data use category, gaming apps collect the most data for third-party advertising, and shopping apps are the top collector in the other three privacy-intrusive data use categories.

Heterogeneity across countries Having studied the US sample, we now move to the international sample to examine cross-country heterogeneities in apps' data collection practices. Figure A.5 provides a summary of the average number of data types and items collected for each country. Over 85% of these apps are available in more than one country, which implies that a simple average measure of data collection intensity for the most popular set of apps can be highly correlated across countries, preventing us from understanding the impact of local regulation or user preference on privacy protection. To construct a more informative,

country-specific measure for the supply of data privacy protection, we focus on two samples. Panels A and C include apps that are available in only one country which we label as local apps. Panels B and D include all apps in the international sample but assign each app to the country where it has the most downloads in 2020. As a result, global apps such as Facebook and Google are treated as US apps. The idea is that data collection practices of an app are more likely to be shaped by the regulation and consumer preference of its most important market.

For each country, we display the average number of data types and items collected in each of the three first-layer categories. Each bar represents the total number of data types/items collected in each country. As shown in Panel A and C, among local apps, apps available exclusively in the US collect considerably more data than apps targeting the audience of other countries, as measured by both the total number of data types/items collected, or those in the more privacy intrusive category (*Data Used to Track You* and *Data Linked to You*). Canada, Germany, UK, and France form the second-tier of countries in terms of data collection intensity. This pattern largely holds when we switch to global apps in Panels B and D. In the empirical analysis, we also provide quantitative evidence on how users react to the release of privacy labels in each country.

4 Consumers' Reaction to Privacy Label Release

Thus far, our analysis has focused on the data collection practices of apps, which primarily concern the supply of privacy. In this section, we delve into product market dynamics by examining how the release of privacy labels shapes consumer demand.

Two challenges have prevented researchers from estimating consumers' demand in response to data privacy disclosures. First, they lack high-quality, micro-level data about users' consumption of digital services at a relatively high frequency. The literature has so far relied on website visits and rankings, which not only suffer from inconsistent measurement but also provide only an incomplete picture of digital consumption given the growing importance of mobile Internet and apps. Second, even with granular data, researchers often grapple with identification issues. Disentangling the effects driven by the demand for pri-

vacy from other demand-side factors is difficult. Market-wide shocks like COVID lockdowns, category-specific shocks like age restrictions on violent gaming apps, and even app-specific shocks such as a major version update, can all shift the demand for digital services and therefore app downloads or revenue. A privacy-related shock like the introduction of GDPR is also prone to identification issues. To start, GDPR renders every app treated, either directly in cases where GDPR is applicable or indirectly through regulatory spillover (Peukert et al., 2021). Without a control group, it remains difficult to rule out alternative explanations based on time-varying factors like COVID. Worse still, policies like GDPR affect firms' data collection practices and hence the supply for privacy, so any changes in downloads or revenue could be driven by the supply of privacy rather than demand for privacy (Peukert et al., 2021; Johnson et al., 2021).

This paper addresses the above challenges. First, we focus on the increasingly important mobile app market and leverage a comprehensive dataset that covers new user downloads and revenue at a weekly frequency for the universe of apps. This offers us valuable insights into users' consumption of digital services at a granular level across different markets. More importantly, our unique empirical setting allows us to construct a counterfactual that is as close to ideal as possible. Since the privacy label policy only applies to iOS apps and affects iOS users only, the corresponding Android apps naturally serve as the control group. The iOS and Android versions of any app likely provide the same digital services and collect a similar amount of data from users. Comparing iOS and Android apps within the same period therefore allows us to control for any app-, category-, and market-level demand shocks. The only difference, from the point of view of users, is the disclosure of data collection practices, which remains unchanged for Android users but is altered for iOS users due to the release of privacy labels.

To validate our assumption that the demand for iOS and Android apps follows a common trend, we compare the aggregate downloads and revenue on the two platforms during 2020/01-2021/07 in Figure 4. The solid line represents iOS versions of all apps and the dashed line Android versions. The two lines follow a strikingly similar trajectory. For example, in March and April 2020, the downloads of both iOS and Android apps soared due to the increasing demand for apps during the first wave of COVID-19 lockdowns. This pattern

also holds when we divide apps by category, or plot the app-level monthly growth rate of iOS version against that of the Android version. We report these additional results in Figure A.6.

4.1 Baseline Effects

Using a staggered difference-in-difference approach, we first estimate the causal impact of privacy label release on the demand for digital services. Our main regression sample is the top10k US sample described in Section 3. For each app in each week, we have two observations, one for the iOS version, one for the Android version. The sample period is 2020/01-2021/08. We begin by investigating the importance of the iOS version of apps relative to the Android version over time. More specifically, Figure 5 plots the percentage share of iOS downloads over the sum of iOS and Android downloads over a 40-week window around the release of the privacy label. Panel A includes apps that collect data to track consumers. Panel B (C) includes apps in the top quartile of the number of data types (data items) collected. Two patterns emerge. First, in all three panels, there is a sudden fall in the share of iOS downloads over total downloads right after week 0, when a given app releases its privacy labels. Second, the share of iOS downloads follows a downward trend after privacy label release. Third, over the pre-event period, there seems to be some weak trends in the share of iOS downloads, although this trend is sometimes positive (Panel C) and sometimes negative (Panel A and Panel B). To take into account any potential pre-trends, we allow for platform-specific linear trends in the regression specification.

Motivated by the descriptive evidence in Figure 5, we now formally estimate the following regression specification:

$$Y_{i,p,t} = \beta_1 iOS_{i,p} \times Post_{i,t} + \beta_2 Post_{i,t} + \alpha_i + \phi_{age,p} + \theta_t + \lambda iOS_p \times t + \varepsilon_{i,p,t} \qquad (1)$$

in which the subscript i, p, and t denote app, platform, and week respectively. The app-specific event indicator, $Post_{i,t}$ equals one for all the weeks after the respective app's release date and zero otherwise. The treatment indicator, iOS_i, equals one (zero) if the observation corresponds to the iOS (Android) version of an app. The outcome variable, $Y_{i,p,t}$, is logarithm of the weekly downloads or revenue of app i on platform p in week t. We scale the outcome variables using $log(1 + y)$ as the distribution of downloads and revenue numbers are highly

skewed as shown in Table 1. We add app fixed effects α_i so that the variation in the outcome variable comes from the difference between the iOS and Android versions of the *same* app. We include year-week fixed effects, θ_t, to control for seasonality and other common shocks to the consumption of all mobile apps. With $iOS_p \times t$ fixed effects and $Platform \times t$ as an additional regressor, our specification also allows for different time trends for iOS and Android apps over their life cycle and calendar years. We double-cluster standard errors by app developer and year-week. Our coefficient of interest, β_1, captures the effect of privacy label release on users' consumption of mobile apps. The key variable of interest is $Post_{i,t} \times iOS_{i,p}$, and its coefficient ($\beta_1$) captures the differential effect of privacy label release on app downloads and revenue from the iOS platform versus those from the Android platform. We expect β_1 to be negative if consumers are on average averse to data collection by firms.

The regression results are reported in Table 4. Column 1 shows that following the release of privacy labels, the average weekly downloads of the iOS version drop by 11.7%, relative to its Android counterpart. The effect stays quantitatively similar when we aggregate downloads over the various versions (e.g., iPad, Lite versions) of the same app within each platform, as shown in Column 2. In column 3, we include only apps that have updated their privacy label section on their apple store pages by August 2021. Around 10% of apps provide no privacy label section either because they haven't released a new version update and are therefore not obliged to report this information, or because they became inactive in the app market. For the remaining 90% of apps, we find a slightly stronger negative impact, reaching almost 14%. This is unsurprising, because the estimated treatment effect on the "never-treated" app pairs is effectively zero, lowering the overall point estimate. In columns 4-6, we use weekly revenue as the outcome variable and repeat the analysis. Results are consistent with weakened demand for apps: weekly revenue decrease by 13%-15%, depending on the sample used. We choose the regression sample used in columns 3 and 6 as our main sample for the rest of the analysis.

4.2 Heterogeneity

Data Collection Intensity The results from the baseline regression specification suggest that consumers indeed react negatively to the disclosure of data collection practices. Do consumers respond even more negatively when apps are more privacy-invasive? To answer this question, we further test whether the negative reaction from consumers depends on the intensity of data collection by constructing three measures that capture the heterogeneities across apps in their data collection. The first measure is an indicator variable for whether the app collects any data in the *Data Used to Track You* category. As explained in Section 3, data items in this category are collected to be shared with data brokers, which we consider most privacy invasive. The second and third measures are the number of data types and data items collected. These two variables reflect the breadth of data collection activities, and we transform them into logarithmic terms. We add a triple interaction term to Equation (1) and estimate the heterogeneous effects across apps with different levels of privacy intrusion using the following specification:

$$\begin{aligned} Y_{i,p,t} = \quad & \gamma_1 iOS_p \times Post_{i,t} \times Data\_Collection\_Intensity_i + \gamma_2 iOS_p \times Post_{i,t} + \gamma_3 Post_{i,t} \\ & + \alpha_i + \phi_{age,p} + \theta_t + \lambda iOS_p \times t + \varepsilon_{i,p,t} \end{aligned} \tag{2}$$

Regression results are reported in Table 5. We find that apps with a high data collection intensity suffer a larger decline in downloads and revenue. In fact, apps with relatively low data collection intensity in general do not experience a significant change in their downloads or revenue. According to column 1, apps that collect data items in the *Data Used to Track You* category see a 10.3% decline in downloads. In contrast, apps that do not collect such data are not affected, as suggested by the coefficient of $iOS \times Post$, which is negative but insignificant. Moving to columns 2-3, we show that the coefficient on downloads monotonically decreases in the number of data items or data types collected. In particular, a 10% increase in the number of data types (items) collected translates into a 0.78% (0.71%) higher drop in weekly downloads. The impact on revenue is similar. Apps that collect data to track consumers experience a 10.9% larger decline in revenue – 1.2 times larger (0.109/0.91) than the baseline effects on the apps that do not collect such data. Moreover, a 10% increase in

the number of data types (items) collected leads to a 0.42% (0.43%) higher drop in weekly revenue.

Dynamics To zoom into the dynamic effects, we replace the *Post* dummy with a series of time indicators and report the estimated coefficients of the interaction terms with the iOS indicator in Figure 6. For the period between week -5 and week +5, we assign an indicator for each week, and for the period before week -5 and after week +5, we assign one indicator to the whole period $\mathbb{1}(t < -5)$ (or $\mathbb{1}(t > -5)$).

Splitting apps by whether the app collects data in the *Data Used to Track You* section in Panels A and B, we first confirm the results in Table 5. Only those that collect such data experience a persistent decline in downloads. Moreover, before the label release, there is no discernible difference in downloads between the iOS and Android versions of the same app. This reassures that the inclusion of iOS time effects effectively controls for any potential platform-specific time trend. Starting from week 0, the DID coefficient becomes negative and continues to decrease till the end of our observation window. This persistent decline in downloads is consistent with consumers' increasing awareness of the privacy labels over time. Furthermore, we sort apps into quartiles based on the number of data types and items collected. Panel C (D) presents results for the apps in the *top* (*bottom*) quartile of the number of data types collected, while Panels E and F use samples split by the number of data items collected. Results are similar: the treatment effects are concentrated among top data collecting apps and the negative effect grows stronger over time.

Data collection purpose So far, we have treated data items equally despite the fact that firms collect data to serve different purposes. Some purposes are clearly more privacy-invasive than others, and we expect consumers to react more strongly to data collection for more privacy intrusive uses. To investigate heterogeneous responses to different purposes of data collection, we rank the six purpose categories based on the degree of privacy-intrusion (from high to low): third-party advertising, developer's advertising or marketing, analytics, product personalization, other purposes, and app functionality. The purpose category is only available for data items in the *Data Linked to You* and *Data Not Linked to You* sections, so

we calculate the total number of data types and data items in each purpose category across these two sections.

Using the same triple difference specification and the number of data types collected for each purpose, we test whether consumers react more negatively when apps collect a greater amount of data for more privacy-intrusive purposes. Results are reported in Table 6 and are consistent with our hypothesis. Panel A shows that collecting data for third-party advertising, developer's advertising or marketing, analytics, as well as product personalization adversely impacts the downloads. In contrast, when the data items are collected for other purposes or for app functionality, we observe no significant, negative influence. In Panel B, we use app revenue as the outcome variable and find that the data collection for the most intrusive purpose, third-party advertising, has the largest impact on revenue. Collection for other purposes also have negative effects on revenue, however, they are not statistically significant. Using the total number of data types collected only in the *Data Linked to You* section yields similar results, as reported in Table B.3.

Substitutability The above results highlight consumers' aversion to data collection and their privacy concerns. Of course, a consumer's choice to download and use the app depends on many other factors besides its privacy protection feature. By revealed preference, a new user will choose to download the app only if the utility gain from the digital services outweighs the disutility from relinquishing access to personal data. Therefore consumers will demand less of an app service only when they can find a close substitute that is acceptable as an outside option. We measure how substitutable an app is in various ways. We first use the download ranking of an app in 2020. The idea is that successful apps gain a large customer base and hence a better ranking because of high user viscosity. We divide apps into ten deciles based on their ranking and include this measure as a continuous variable in the regression. A better ranked app likely offers a higher value to price ratio and is less substitutable. Similarly, we also measure an app's popularity within its own app category. For each app, we calculate its market share as its total annual downloads over the total annual downloads of all apps in the same category. Then, we assign an indicator variable that equals one if the app has the top decile market share among its peers. Last, we use the

app's age as a proxy for its substitutability as only the most attractive apps survive.

In Table 7, we find evidence that unsubstitutable, popular apps are less exposed to the adverse impact of releasing their privacy labels. Column 1 shows that moving up in the store-wide ranking by one decile attenuates the negative impact of privacy label release by 1.6 percentage points. In column 2, we find that apps at the top decile of the market share distribution experience a 5.9 percentage points smaller decline in downloads, mitigating almost half of the baseline effect experienced by the remaining 90% of apps. According to column 3, older apps are subject to a significantly smaller drop in app downloads after privacy label release. In columns 4-6, when using revenue as the outcome variable, the point estimates for the triple interaction term have the same signs but are insignificant both statistically and economically. One explanation is that higher and lower ranked apps have vastly different pricing strategies, making the comparison in revenue patterns less credible and statistically more challenging than the comparison in download patterns.

4.3 Additional discussions on identifying assumptions

Besides the common trend assumption discussed above, other assumptions required for the identification of the treatment effect include the Stable Unit Treatment Value assumption (SUTVA) and no treatment effect on the population in the pre-treatment period (NEPT). Here we discuss the validity of these assumptions in our empirical context. One potential violation of SUTVA is cross-platform information acquisition. A very privacy-conscious Android user may learn about an app's data collection practice by making the effort to visit its iOS webpage via a web browser and reduce the demand for the corresponding Android app. While we can not directly rule out this behavior, it biases our estimate downwards. Another potential violation of SUTVA is that users may switch from iOS to Android, or vice versa, as a response to privacy label release. Our results are unlikely to be driven by this for two reasons. First, switching operating systems is costly for consumers. Second, based on the market share of operating systems on smartphones reported in Appendix Figure A.7, the market shares of iOS and Android remain stable during our sample period.

The NEPT assumption requires no behavioral changes of iOS users in anticipation of the privacy label release. For example, it is violated if iOS users delay the downloads of

new apps till their privacy labels are released. The lack of pre-trends in Figure 6 suggests little evidence of the anticipatory effect. Even if it exists, delaying downloads would point to an underestimation of the true treatment effect. Another way to mitigate this concern is to examine the demand for an iOS app relative to its Android counterpart around the *announcement* of the privacy label policy in June 2020. If the relative demand does not change around the policy announcement, biases caused by the anticipation effect is likely limited. Figure A.6 Panel B provides supporting evidence. Each dot in the subfigure represent a pair of monthly growth rates for the iOS and Android versions of the same app and we do so for all 12 months in 2020. The dot clouds center around the 45-degree line and there is no visible shifts around June.

It is worth noting that our DID estimators capture the average treatment effects on the treated (i.e., iOS users). To identify the average treatment effect on the population, one would need to assume that Android users, upon observing the same information, reduce their demand for Android apps by the same magnitude. In light of the demographic difference between iOS users and Android users (Bertrand and Kamenica, 2018), additional research is required to better understand the privacy attitudes of iOS and Android users.[17]

4.4 Robustness checks and Placebo tests

In this section, we show that our main results are robust to a wide range of alternative sample choices and regression specifications. In addition, using various placebo treatment dates generates insignificant results, which further supports our identification strategy.

We first report the results from robust checks with alternative regression specification and app samples. The estimated DID coefficients in Equation (1) associated with these tests are plotted in Panel A of Figure 7. Each row corresponds to an estimated coefficient with either an alternative sample or an alternative specification. For reference, the top row shows the baseline estimate of -0.117 with a t-statistic of 2.34 in column 1 of Table 4.

We start with alternative regression samples. We consider the top10k apps excluding the Christmas holiday period, apps that release their labels rather late (more than two

[17]The only study we are aware of that compares the privacy attitudes of iOS and Android users is Abrokwa et al. (2021). Using a survey with 494 US participants, the authors find no significant differences in privacy attitudes of different platform users.

months after 2020/12/14), top5k apps, apps ranked between 5k and 10k, apps developed by public firms among the top10k apps, as well as all apps developed by public firms.[18] In addition, we extend the sample to the beginning of 2019 to include more pre-treatment observations. Across all alternative samples, the estimated coefficients are negative and significant, as shown in Figure 7. In particular, excluding the apps that release their labels two months after the policy implementation mitigate concerns about strategic behaviors of app developers. For example, developers could strategically release the privacy label on days when customers are less attentive. In this case, our point estimate will be biased downwards. We do not find a significant difference in the point estimates between the two samples. Moreover, excluding holiday weeks, including 2019 in the pre-treatment period, or excluding apps that release their privacy labels rather late leads to an insignificant change in the DID coefficient. Moreover, apps developed by public firms are hit hard, potentially because those apps collect much more information than their peers, as shown in Figure 2. Last, consistent with our previous findings, top 5k apps experience a smaller drop in downloads compared to apps ranked between 5k and 10k, because consumers have less attractive outside options for more popular apps.

Our results are also robust to alternative regression specifications. We begin by using 2020/12/14 as a uniform treatment date for all apps. This allows us to test whether there are some anticipation effects for apps that do not release labels immediately after 2020/12/14. The argument is that once the first batch of privacy labels are available, consumers adjust their expectations about the data collection intensity of other apps, hence most of their reactions can materialize right after 2020/12/14. However, the point estimate changes little. We also add app *times* week fixed effects so that any app-level time-varying shocks can be differenced out. We further include category *times* platform fixed effects, which control for platform-specific time-invariant attributes for each category of apps. Adding these fixed effects has little impact on the estimated coefficient. Last, we show that clustering standard errors only by app developers does not affect the significance levels.

In Panel B, we present the results from various placebo tests, where we vary the treatment dates. We first advance the app-specific treatment dates to exactly one year or two

[18] Holiday weeks include the week starting from 2020/12/21 and 2020/12/28.

years before the actual ones. Then, we consider 2019/12/14 and 2018/12/14 as the placebo treatment dates to examine if seasonality can explain our results. Last, we take the dates of the apps' first version update following 2019/12/14 and 2018/12/14 as the placebo treatment dates. This helps us address the concern that the drop in downloads or revenue may be driven by the app's version update as opposed to the privacy label release. Importantly, in all placebo tests, we drop observations after 2020/12/14 to minimize contamination from the effect of the actual policy change. All placebo tests yield insignificant estimates and the point estimates are close to zero.

Overall, the results from the robustness checks and placebo tests show that our results are not driven by the choice of the regression sample or specification, and that the results cannot be explained by other factors such as apps' version updates or seasonality in digital consumption through apps.

5 Cross-country comparison

In this section, we show that our results are not specific to the US app market. We repeat our regression analysis for the top 10k apps in the 10 countries as described in Section 3. We primarily focus on these countries for two reasons. First, Sensor Tower report to have a wider data coverage on these mobile markets, in the sense that they have access to the actual performance data provided by more partner apps in these countries. The superior data coverage is likely driven by the larger and more developed mobile app markets in these countries. Second, when comparing the demand for privacy across countries, ideally one wants to fix the supply and examine apps commonly available in all countries. The tradeoff between the number of such apps and the number of countries leads us to choose this country set. Nonetheless, later in the section, we extend our analysis to 90 countries and analyze a restricted set of apps to show what country-level factors explain the heterogeneous consumer reactions.

5.1 Consumer's reaction to privacy label release in ten countries

We start by plotting the estimated DID coefficient for these 10 countries in Panel A of Figure 8. More specifically, for each country, we focus on the apps in the top and bottom quartile of data collection intensity, measured by the total number of data types and items collected. Panel A presents the results for data types and Panel B for data items. To facilitate comparison, for each country in each panel, we display the DID coefficient for the top quartile apps and bottom quartiles apps side by side.

We document the following. First, across all countries, apps in the bottom quartile of data collection intensity do not experience a significant change in downloads after privacy label release. In contrast, we observe a negative shock to downloads for apps in the top quartile in most countries. These negative coefficients are particularly significant for the US, Canada, and the UK. The effect is around -20% in the US and Canada, and -12% in the UK. In Austria and Japan, the effect is close to being significant at 95% level. In Germany, France, Italy, Korea, and Russia, we do not find that consumers react strongly to the disclosure of data collection activities through mobile apps.

To further examine the heterogeneity in consumer reactions to the disclosure of firms' data collection practices, we construct a sample of global apps that are available in all the 10 countries. In total, there are 1,284 such apps among the country-specific top10k apps. Then, we study the same the list of apps across countries and focus on the difference in consumer reactions. As reported in Panel B of Figure 8, two distinct patterns emerge. First, in all countries, we observe a positive effect on the light data collectors and a negative effect on heavy data collectors. This suggest that there is indeed positive spillover among globally popular apps. Second, users in France exhibit the most negative reaction, followed by users in the US and Canada. One caveat is the reduced statistical significance in this set of estimates because of the smaller sample size.

5.2 What explains country-level heterogeneity? Role of privacy protection, public trust, and privacy concerns

We continue by exploring what explains the heterogeneous responses to Apple's privacy label policy across countries. We expand the sample to all countries covered by SensorTower and repeat our baseline DID analysis in every country to obtain a country-specific coefficient (β_1 in Equation (1)). After applying top 10k sample filter, we are left with 90 countries with sufficient data to estimate the DiD coefficient.

Ideally one would use apps commonly available in all countries to hold the supply of privacy constant. However, each country's top 10k list does not necessarily overlap significantly with each other, and requiring each app to be present in every country's top 10k list would greatly limit the regression sample size.[19] To address this issue, we consider the universe of apps developed by publicly listed firms instead of being confined to the top 10k list. There are over 6,000 of such apps and they are typically available globally, hence the heterogeneity in country-specific responses is more likely to reflect the demand for privacy from app users in each country.

We show that consumers' responses to the privacy label release are associated with each country's data regulation and privacy protection, public trust, and consumer attitudes regarding data use and privacy. More specifically, we measure these factors using three sets of variables. The first set are proxies for each country's legal protection of data privacy, obtained from Comparitech. We consider both the legislative status of privacy regulations and the enforcement of these regulations. The higher the score, the stronger the privacy protection and enforcement. The second set of variables reflects general public trust in the press or major companies. We obtain these measures from the World Value Survey (2022). The lower the value, the higher the trust level. The last set of variables concerns consumer attitudes towards data privacy issues. The first one is directly linked to concerns regarding personal data use and is obtained from the Data Confidence Index by Datum Future. It is calculated as the average agreement score for statement "*I worry about how my personal*

[19] In the previous section, we show that requiring each app to be present in 10 countries' top 10k lists already reduces the sample to 1,284 apps.

data is being used by companies." The higher the value, the stronger the consumer concern. The second variable, willingness to pay for a security certificate, is obtained from the CIGI-Ipsos Global Survey on Internet Security and Trust. It is solicited by asking respondents the following question: "*For an equivalent product valued at $1,000, how much more would you be willing to pay to have a product with a security certification mark from the government?*" This measure is expressed in dollars.

Figure 9 depicts the relationship between the country-specific estimated coefficient and the three sets of measures, using population size as the regression weight. Panels A and B focus on the current state of privacy protection. The upward sloping line reveals that countries with stronger legal protection of privacy and law enforcement react less negatively to privacy labels, potentially because consumers consider the current legal protection of their data privacy adequate. In Panels C and D, we turn to confidence in the press and major companies as proxies for general trust. As the graphs made clear, the level of general trust, despite not directly related to data privacy issues, negatively correlates with the demand for privacy. In Panels E and F, we further explore the role of consumer attitudes regarding data use and privacy. With both measures, we find consistent evidence that consumers in countries with more severe privacy concerns react more negatively to privacy label release.

6 Privacy Label Release and Firm Valuation

In this section, we present early evidence on the impact of the policy on firm performance. As documented above, apps with high data collection intensity experience a greater decline in downloads and revenue after the release of privacy labels. The weaker demand for privacy-intrusive apps can dent firm performance in the short-run. Given the recency of the policy, we rely on stock market returns of public firms with an active app to document the policy's effect on firm performance. We show that firms that harvest more data indeed underperform in the stock market and see a drop in quarterly earnings.[20]

[20]In the long-run, firms may adjust business models and investment strategies to optimize the supply of privacy. However, given that privacy labels were introduced very recently, data is not yet available for longer-term consequences. Hence, we focus on the short-term impact of the policy on firm performance, measured by the stock market returns.

6.1 Stock market returns

We first present strong evidence that the investors respond negatively to the release of privacy labels in the following six months. We measure stock market performance with cumulative abnormal returns. Using a 250-trading day estimation window that ends 10 trading days before label release, we first obtain expected returns from various benchmark models, such as the CAPM, Fama-French three-factor (FF3), Fama-French-Carhart four-factor (FFC4), and Fama-French five-factor (FF5) models. We further require that stocks have at least 126 trading days (6 months) or are delisted after the label release date. Note that we focus more on relatively longer-term stock market response because, unlike corporate announcements, individual events of privacy label releases are less likely to attract immediate public attention. Firms are not required to send a notification to consumers and investors. Furthermore, while market participants are familiar with firms' fundamentals or other well-documented determinants of stock prices, privacy-related information is fresh to most market participants, and more learning and analyses are required before trading takes place. It therefore may take longer for the market to incorporate privacy-related news into stock prices.

We carry out the analysis both at the firm level and at the app level. At the firm level, we first take the average privacy label release date among all apps developed by a public firm and consider this date as the firm's event date (Firm – avg.). Then we use the privacy label release date of a representative app – one that logs most downloads – as the firm's event date (Firm – rep.). In the last specification, we set the event date to be the official launch date of the Apple App Store's privacy label policy (2020/12/14). Our goal is to examine the potential impact of market anticipation towards privacy labels. At the app level, there are cases where multiple apps release privacy labels on the same date. In these cases, we count multiple releases by a firm as a single event. Panel A of Table 8 presents the six-month CAR after the event date for analyses at different levels. Using the CAPM model, we observe a CAR of -4.49% over the six months window following the privacy label release of a given firm's most popular app when compared with firms without any apps. This negative stock market response is significant and robust to alternative models. In fact, using FF3, FFC4, and FF5, we observe even more negative responses, with CARs reaching -7.22%, -6.10%, and

-6.10%, respectively. Importantly, similar results hold when we use alternative event dates, including firms' average release dates, 2020/12/14, and app-level release dates.

Consistent with the heterogenous reaction on the product markets, we also find a more negative effect for firms that collect more data and rely more on mobile users to generate revenue. We first split the sample by data collection intensity and repeat the event study for each subsample. We use the number of data types collected as the sorting variable and divide our sample into two groups: the above-median group (H) and the below-median group (L). We compare their CARs after privacy label release and show the results in Panel B of Table 8. Similar to results in Panel A, all CARs remain negative regardless of model specification. Using the CAPM model, firms that report above-median data collection intensity earn a six-month CAR of -13.30%, while their below-median counterparts earn a CAR of -4.76%. The difference is over 8% and significant, suggesting amplified market reactions to more privacy-intrusive apps. We plot the evolution of the CARs (estimated using CAPM, FF4, FFC4, and FF5) for the two groups of firms in the six-month window after each firm's privacy label release in Figure 10.The patterns in the two subsamples are consistent with the results reported in Table 8. It is worth noting that we do not observe an immediate drop in the CAR for either group. Instead, both groups experience the decline in stock market performance around 40 days after the privacy label release. This could reflect either investors' inattention to the event and/or the uncertainty about consumer demand for data privacy, consistent with the "privacy paradox". We present evidence on these explanations in Appendix C.

Besides the heterogeneity in the data collection intensities, we further document a more pronounced *H-L* difference among retail and service firms (SIC code range: 5200-5999 or 7000-8999). This result lends additional support to our interpretation that the underperformance was due to the adverse impact of the privacy label policy. Intuitively, firms that rely more on mobile users are more vulnerable to the lost of mobile app users and are therefore more likely to see a stronger negative stock market reaction following the release of privacy labels. Based on the last three rows of Table 8, when benchmarked agaist the CAPM model, firms in the above-median data collection group earn a CAR of -13.61%, while the below-median group earns a CAR of -4.12%. The difference is 9.49% and again significant. Similar results hold for FF3, FFC4, and FF5 models at similar or higher significance levels.

6.2 Quarterly earnings

We also show firms that collect more data experience a larger drop in earnings than those that collect less data. This is consistent with our findings on the product markets and financial markets. We focus on earnings per share (EPS) and estimate the following regression equation:

$$EPS_{i,t} = \gamma_1 Post_{i,t} \times \%(Data\ Used\ To\ Track\ You)_i + X_i'\beta + \alpha_i + \phi_t + \varepsilon_{i,t}, \quad (3)$$

where $\%(Data\,Used\,to\,Track\,You)_i$ is defined as the percentage share of firm i's apps that are rather privacy-invasive and collect data to track users. $Post_{i,t}$ equals one if the corresponding earnings quarter is 2021Q1 onward and hence subject to the impact of Apple's privacy policy. We expect γ to be negative. Firm-level controls, denoted by X', include quarter-end total assets (in logarithm), cash-to-asset ratio, tangible asset ratio, leverage ratio, and EBITA-to-asset ratios, all lagged by four quarters and winsorized at 1% at both tails. We also include firm and year-quarter fixed effects. To account for the importance of app business to a firm's earnings, we estimated a weighted linear regression, with the weight being each firm's total app downloads or revenue in 2020.

The regression results are reported in Table 9. Based on the point estimate in column 1, a one-standard-deviation increase (0.36 p.p.) in the share of privacy-intrusive apps decreases the firm's EPS by \$0.23 (0.36×0.64). Controlling for firm characteristics increases this effect to \$0.30 (0.36×0.83). In column 3, we report estimation results for firms in retail or services and find an even larger decline in EPS. In particular, a one-standard-deviation increase now translates into a \$0.52 (0.40×1.30) drop in the firm's EPS. In columns 4-6, we repeat the same set of regressions, but apply a different weight – firm's total app revenue. As a result, firms that do not charge for downloads or derive revenue from in-app purchase are automatically excluded from the sample. As expected, these firms' performance is more adversely impacted by the privacy label policy.

Our findings on the stock market returns and quarterly earnings suggest that privacy policies restricting the collection of data hurt firm performance. In the long-run, firms may adjust both price and data collection in response to the ever-tightening regulatory

environment. Kesler (2022) show that in reaction to a related privacy policy, the App Tracking Transparency framework implemented by Apple, more apps become paid apps and turn to in-app purchase as an alternative revenue source. Yet, it remains a question whether and how firms would adjust the provision of privacy in the long run.

7 Conclusion

Following Apple's privacy label policy, iOS app developers are required to report the collection and use of customer data in a transparent and digestible "nutrition label" format. We scrape privacy labels for the most popular iOS apps in ten countries to provide a valuable measure of data collection intensity. Supplementing this dataset with weekly downloads and revenues from Sensor Tower, we investigates how consumers react to the standardized disclosure of data privacy practices - a key element of corporate digital responsibility. We show that consumers are averse to data collection by apps, especially when their data is collected for privacy-invasive uses. Our findings highlight the lack of consumer awareness of firms' data collection practices as one important explanation for the privacy paradox – the discrepancy between an individual's intentions to protect their privacy and how they actually behave in the online marketplace. We also document negative stock market reactions, in particular among firms in the retail and service sector that harvest more user data. Overall, our findings suggest that data play a central role in firm valuations in today's digital economy.

References

Abrokwa, Desiree, Shruti Das, Omer Akgul, and Michelle L Mazurek, 2021, Comparing security and privacy attitudes between ios and android users in the us, in *SOUPS 2021: USENIX Symposium on Usable Privacy and Security*.

Acquisti, Alessandro, Leslie K. John, and George Loewenstein, 2013, What is privacy worth?, *The Journal of Legal Studies* 42, 249–274.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman, 2016, The economics of privacy, *Journal of Economic Literature* 54, 442–492.

Akey, Pat, Stefan Lewellen, Inessa Liskovich, and Christoph Schiller, 2021, Hacking corporate reputations, *Rotman School of Management Working Paper* .

Al-Natour, Sameh, Hasan Cavusoglu, Izak Benbasat, and Usman Aleem, 2020, An empirical investigation of the antecedents and consequences of privacy uncertainty in the context of mobile apps, *Information Systems Research* 31, 1037–1063.

Amos, Ryan, Gunes Acar, Elena Lucherini, Mihir Kshirsagar, Arvind Narayanan, and Jonathan Mayer, 2021, Privacy policies over time: Curation and analysis of a million-document dataset, in *Proceedings of the Web Conference 2021*, 2165–2176.

Aridor, Guy, Yeon-Koo Che, and Tobias Salz, 2020, The economic consequences of data privacy regulation: Empirical evidence from GDPR, NBER Working Paper 26900, Columbia University, Massachusetts Institute of Technology.

Athey, Susan, Christian Catalini, and Catherine Tucker, 2017, The digital privacy paradox: Small money, small costs, small talk, NBER Working Paper 23488, Stanford University, Massachusetts Institute of Technology.

Bana, Sarah, Erik Brynjolfsson, Wang Jin, Sebastian Steffen, and Xiupeng Wang, 2021, Cybersecurity hiring in response to data breaches, *Available at SSRN* .

Bertrand, Marianne, and Emir Kamenica, 2018, Coming apart? cultural distances in the united states over time, Technical report, National Bureau of Economic Research.

Bessen, James E., Stephen M. Impink, Lydia Reichensperger, and Robert Seamans, 2020, GDPR and the importance of data to AI startups, Working paper, New York University, Boston University.

Chen, Long, Yadong Huang, Shumiao Ouyang, and Wei Xiong, 2021, The data privacy paradox and digital demand, NBER Working Paper 28854, Luohan Academy, Princeton University.

Chia, Pern Hui, Yusuke Yamamoto, and Asokan N., 2012, Is this app safe? A large scale study on application permissions and risk signals, in *Proceedings of the 21st international conference on World Wide Web*, 311–320.

Chiou, Lesley, and Catherine Tucker, 2017, Search engines and data retention: Implications for privacy and antitrust, Technical report, National Bureau of Economic Research.

Comscore, 2019, Global state of mobile, Technical report.

Ebert, Nico, Kurt A. Ackermann, and Björn Scheppler, 2021, Bolder is better: Raising user awareness through salient and concise privacy notices, in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–12.

eMarketer, 2020, Us mobile time spent 2020, Technical report.

Florakis, Chris, Christodoulos Louca, Roni Michaely, and Michael Weber, 2020, Cybersecurity risk, Technical report, National Bureau of Economic Research.

Goldfarb, Avi, Shane M Greenstein, and Catherine E Tucker, 2015, *Economic analysis of the digital economy* (University of Chicago Press).

Goldfarb, Avi, and Catherine Tucker, 2012a, Privacy and innovation, *Innovation policy and the economy* 12, 65–90.

Goldfarb, Avi, and Catherine Tucker, 2012b, Shifts in privacy concerns, *American Economic Review* 102, 349–353.

Goldfarb, Avi, and Catherine Tucker, 2019a, Digital economics, *Journal of Economic Literature* 57, 3–43.

Goldfarb, Avi, and Catherine Tucker, 2019b, Digital marketing, in *Handbook of the Economics of Marketing*, volume 1, 259–290 (Elsevier).

Huang, Henry He, and Chong Wang, 2021, Do banks price firms' data breaches?, *The Accounting Review* 96, 261–286.

Janssen, Rebecca, Reinhold Kesler, Michael Kummer, and Joel Waldfogel, 2021, GDPR and the lost generation of innovative apps, NBER Working Paper 146409, University of Zurich, University of Minnesota, University of East Anglia, Georgia Institute of Technology.

Jentzsch, Nicola, Sören Preibusch, and Andreas Harasser, 2012, Study on monetising privacy: An economic model for pricing personal information, Report for the European Network and Information Security Agency (ENISA), Heraklion: European Network and Information Security Agency.

Jia, Jian, Ginger Zhe Jin, and Liad Wagman, 2021, The short-run effects of the general data protection regulation on technology venture investment, *Marketing Science* forthcoming.

Johnson, Garrett, Scott Shriver, and Samuel Goldberg, 2021, Privacy & market concentration: Intended & unintended consequences of the GDPR, Working paper, Boston University, University of Colorado, Northwestern University.

Kesler, Reinhold, 2022, The impact of apple's app tracking transparency on app monetization, *Available at SSRN 4090786* .

Kesler, Reinhold, Michael E. Kummer, and Patrick Schulte, 2017, Mobile applications and access to private data: The supply side of the android ecosystem, ZEW Discussion Paper 17-075, Unversity of Zurich, University of East Anglia.

Lin, Tesary, 2021, Valuing intrinsic and instrumental preferences for privacy, *Marketing Science* forthcoming.

Peukert, Christian, Stefan Bechtold, Michail Batikas, and Kretschmer Tobias, 2021, Regulatory spillovers and data governance: Evidence from the GDPR, *Marketing Science* forthcoming.

Preibusch, Sören, Dorothea Kübler, and Alastair R. Beresford, 2013, Price versus privacy: An experiment into the competitive advantage of collecting less personal information, *Electronic Commerce Research* 13, 423–455.

Prince, Jeffrey, and Scott Wallsten, 2021, How much is privacy worth around the world and across platforms?, Working paper, Indiana University, Technology Policy Institute.

Ramadorai, Tarun, Ansgar Walther, and Antoine Uettwiller, 2019, The market for data privacy, CEPR Discussion Paper DP13588, Imperial College London.

Sarma, Bhaskar P., Ninghui Li, Chris Gates, Rahul Potharaju, Cristina Nita-Rotaru, and Ian Molloy, 2012, Android permissions: A perspective combining risks and benefits, in *Proceedings of the 17th ACM symposium on Access Control Models and Technologies*, 13–22.

Schmitt, Julia, Klaus M. Miller, and Bernd Skiera, 2020, The impact of privacy laws on online user behavior, Working paper, Goethe University Frankfurt, HEC Paris.

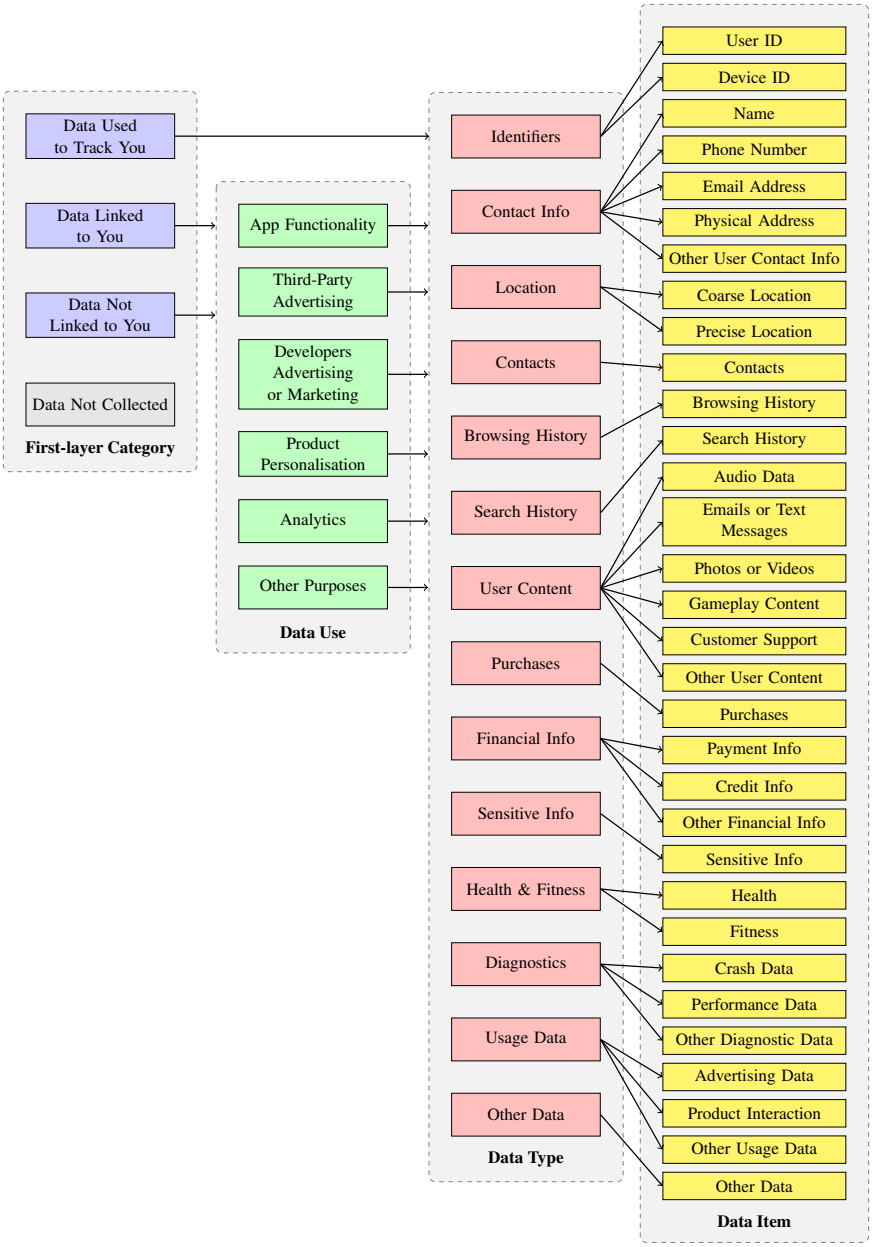
Tambe, Prasanna, Lorin Hitt, Daniel Rock, and Erik Brynjolfsson, 2020, Digital capital and superstar firms, Technical report, National Bureau of Economic Research.

Tang, Huan, 2019, The value of privacy: Evidence from online borrowers, Working paper, HEC Paris.

Tsai, Janice Y., Serge Egelman, Lorrie Cranor, and Alessandro Acquisti, 2011, The effect of online privacy information on purchasing behavior: An experimental study, *Information Systems Research* 22, 254–268.

Tucker, Catherine, A Agrawal, J Gans, and A Goldfarb, 2018, Privacy, algorithms, and artificial intelligence, *The economics of artificial intelligence: An agenda* 423–437.

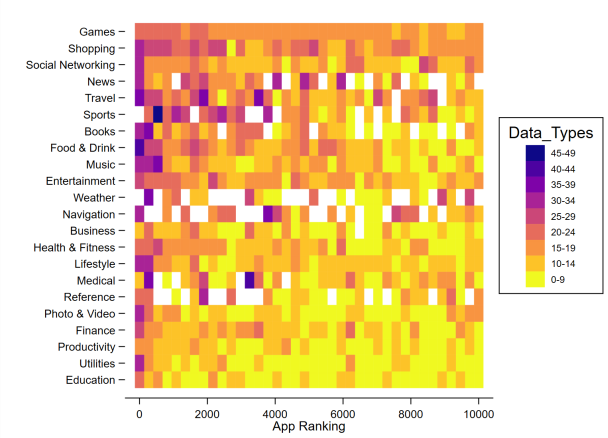
FIGURE 1
The Structure of Privacy Labels



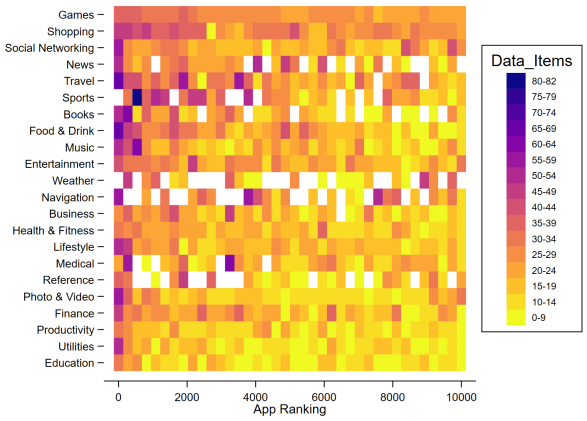
NOTE.— This figure shows the structure of apps' privacy labels adopted by all apps. There are four layers in the privacy labels. The first layer consists of three categories: *Data Used to Track You*, *Data Linked to You*, and *Data Not Linked to You*. If an app doesn't collect any data, it will only have *Data Not Collected* as the only layer in its privacy label. For the second layer, only *Data Linked to You* and *Data Not Linked to You* have this layer which shows 6 different purposes of data use. The third layer includes 14 different data types that the app collects, all data types can appear under each of the 6 purposes of data use in the second layer. The fourth layer reports 32 data items under the corresponding data type in the third layer. The first and the third layers are displayed on the main App Store page while the second and the fourth layers are only displayed in a pop-up window when one clicks on the "See Details" button at the upper right corner of the App Privacy section.

FIGURE 2
Data Collection Intensity

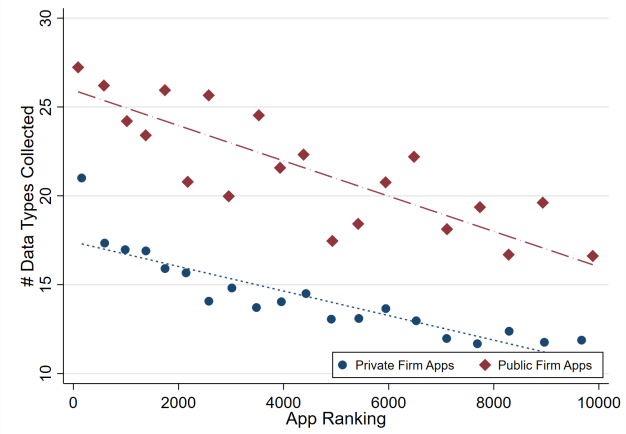
Panel A. # Data Types Collected by Category



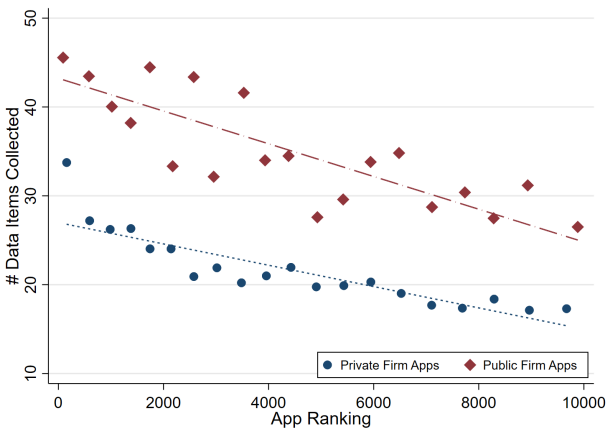
Panel B. # Data Items Collected by Category



Panel C. # Data Types Collected by Developer Status



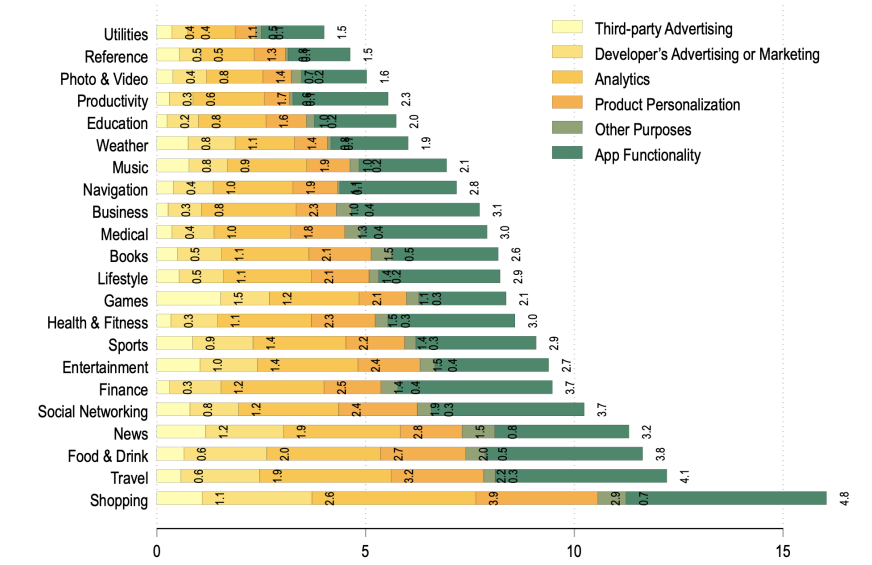
Panel D. # Data Items Collected by Developer Status



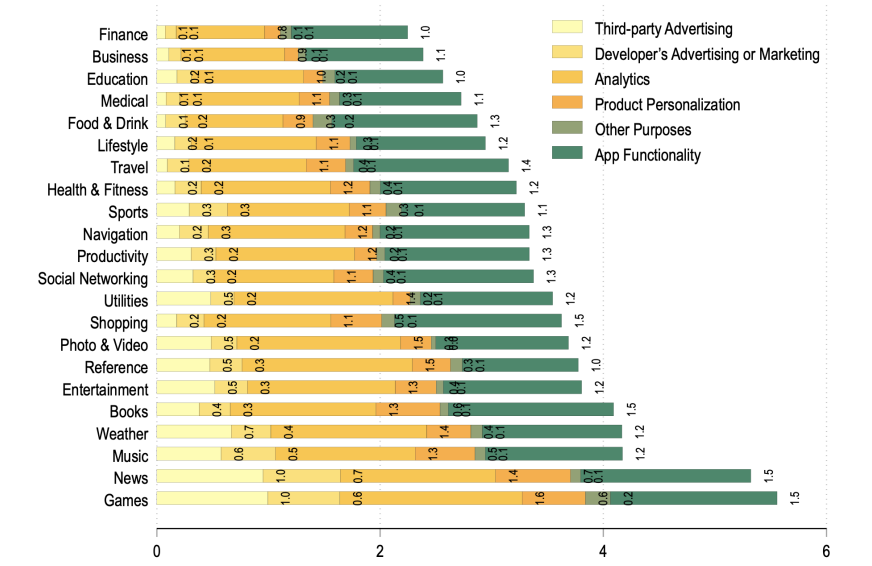
NOTE.— This figure visualizes the positive relationship between data collection intensity and app ranking. Panels A and B show the number of data types and items collected by apps double sorted by app ranking and app categories, respectively. Panel C and D plot a binned scatter of the number of data items and types collected by public apps sorted by app ranking.

FIGURE 3
Data Collection Intensity and Data Use

Panel A. Data Linked to You: # Data Types Collected



Panel B. Data Not Linked to You: # Data Items Collected

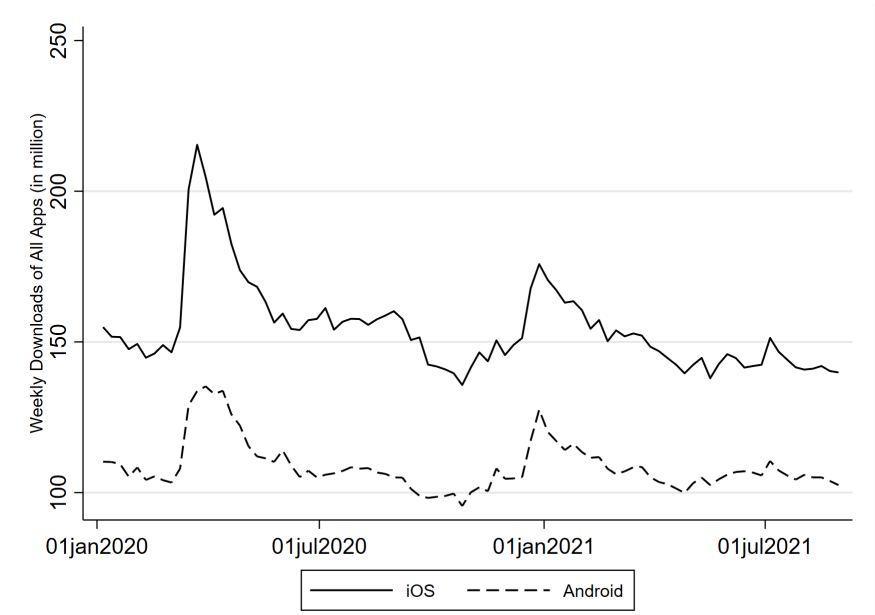


NOTE.— This figure presents the average number of data types collected under 6 data use purposes for each app category. Panels A shows the result of *Data Linked to You*, and Panel B shows the result of *Data Not Linked to You*

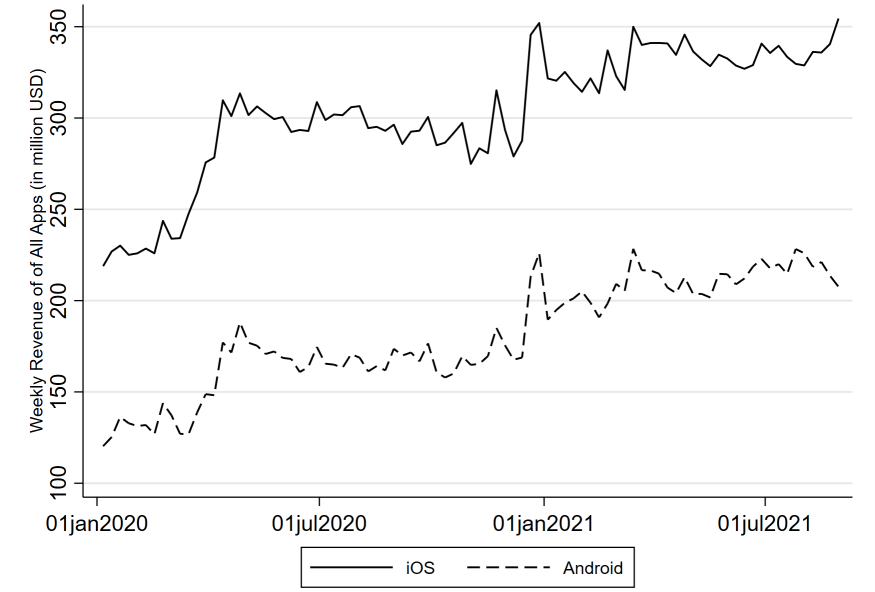
.

FIGURE 4
Download and Revenue of All Apps by Platform

Panel A. Total Weekly Downloads by Platform



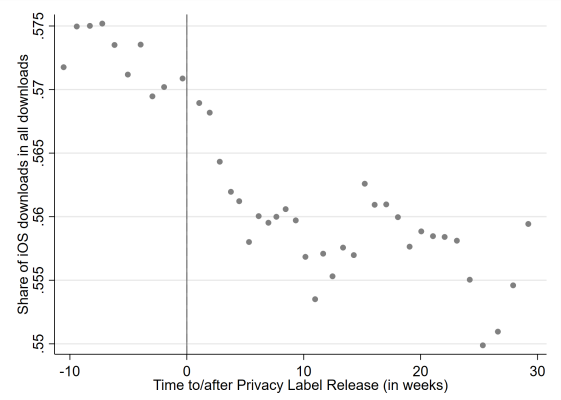
Panel B. Total Weekly Revenue by Platform



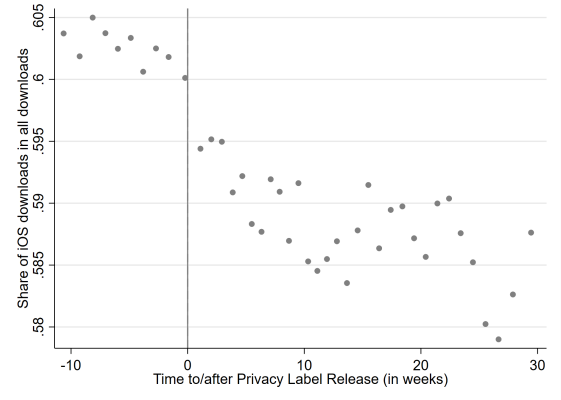
NOTE.— This figure shows the aggregate weekly downloads and revenue by iOS and Android platforms during 2020/01-2021/07. Panel A focuses on weekly downloads and panel B focuses on weekly revenue. The solid line represents iOS versions of all apps and the dashed line Android versions.

FIGURE 5
Share of iOS downloads over total downloads
around Privacy Label Release

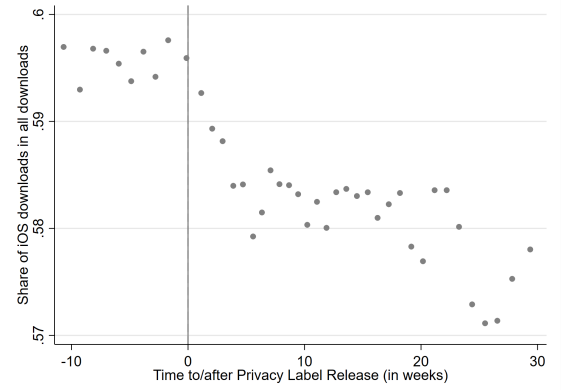
Panel A. Collect Data Used to Track You: Yes



Panel B. # Data Types Collected - High



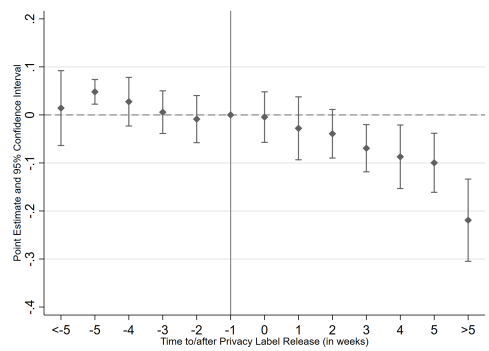
Panel C. # Data Items Collected - High



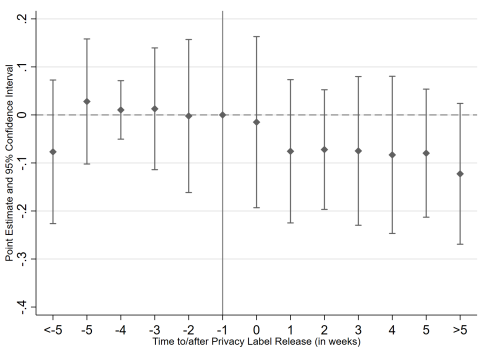
NOTE.— This figure shows the percentage share of iOS downloads over the sum of iOS and Android downloads in a 40-week window around the release of the privacy labels. Panel A includes apps that collect data to track consumers. Panels B and C include apps in the top quartile of the number of data types and items collected, respectively.

FIGURE 6
Impact of Privacy Label Release on Downloads

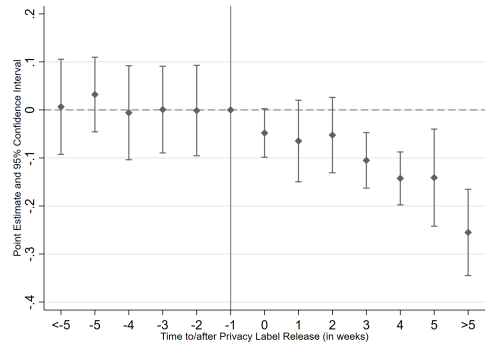
Panel A. Collect Data Used to Track You: Yes



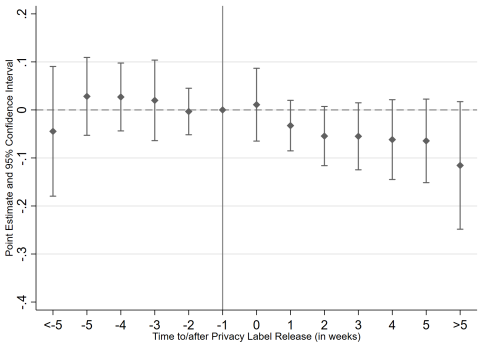
Panel B. Collect Data Used to Track You: No



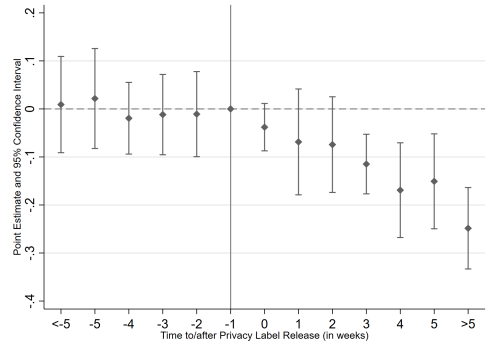
Panel C. # Data Types Collected - High



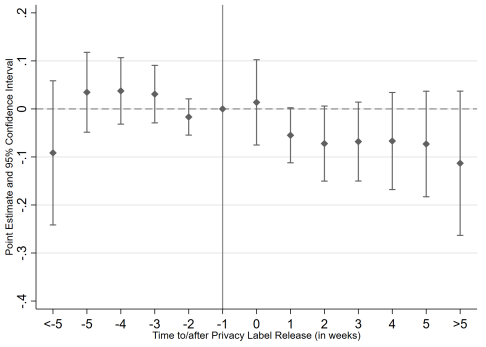
Panel D. # Data Types Collected - Low



Panel E. # Data Items Collected - High



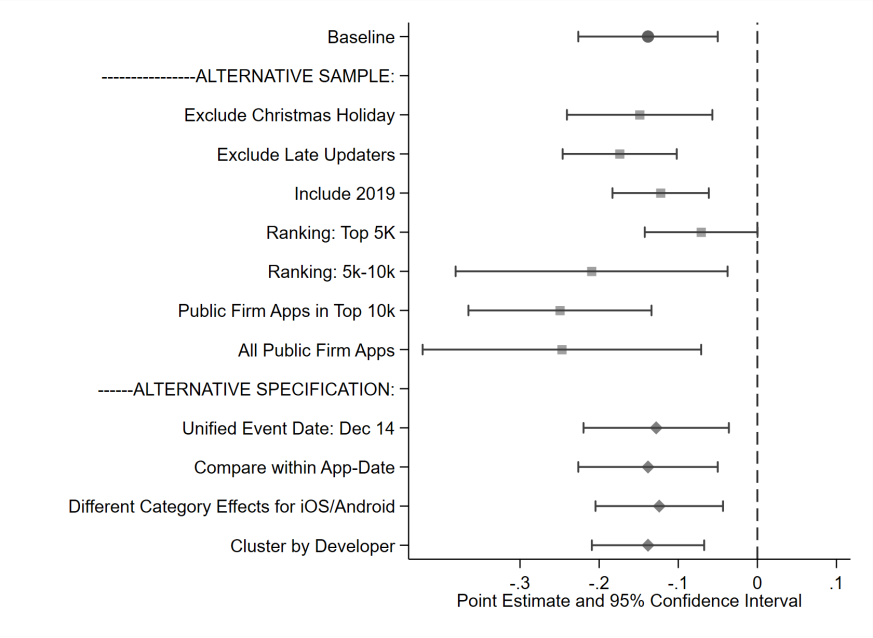
Panel F. # Data Items Collected - Low



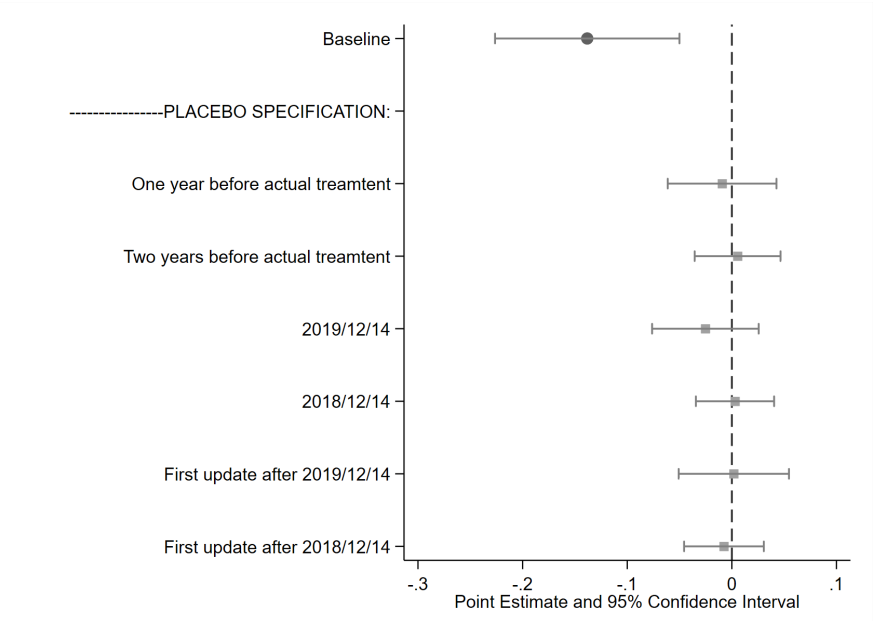
NOTE.— This figure plots estimated coefficients of the interaction term in Equation (1) where the *Post* dummy is replaced with a series of week indicators before and after privacy label releases. For the period between week -5 and week +5, we assign one indicator per week, and for the period earlier than 6 weeks before (after) the label release, we assign one indicator to the whole period. Panels A and B split the sample by whether the app collects data in the *Data Used to Track You* category. Panels C and D show the results for apps in the top and bottom quartile of the number of data types collected, respectively. Panels E and F show the results for apps in the top and bottom quartile of the number of data items collected, respectively.

FIGURE 7
Robustness Checks and Placebo Tests

Panel A. Robustness Checks



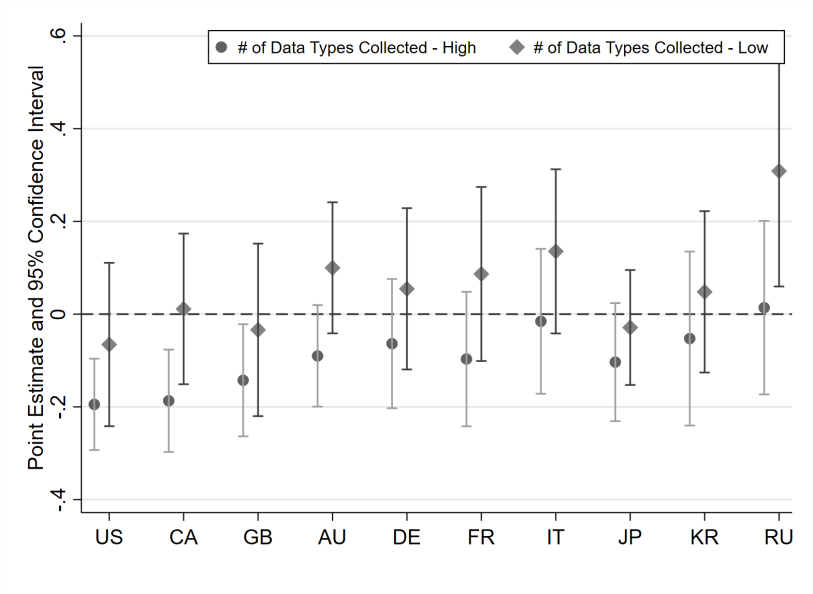
Panel B. Placebo Tests



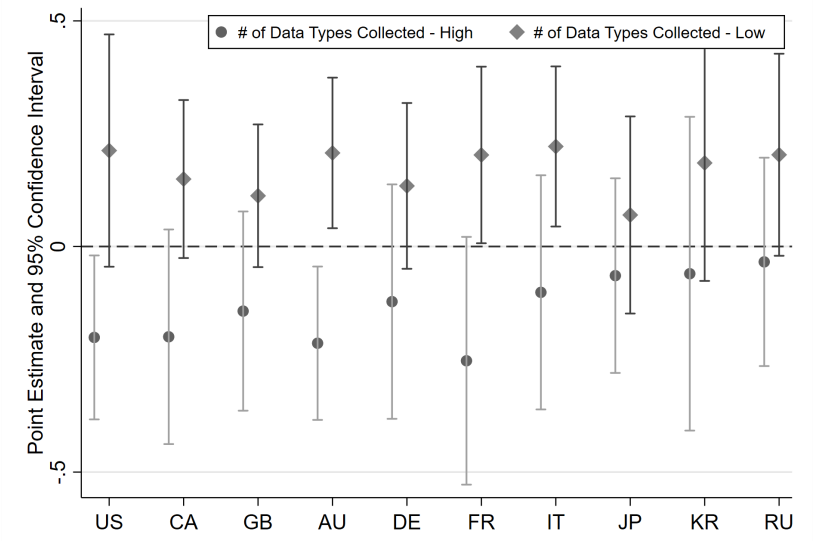
NOTE.— This figure plots estimated coefficients of the interaction term in Equation (1) for alternative regression sample and specifications (Panel A), or placebo treatment dates (Panel B). Each row corresponds to the estimated coefficient of one alternative sample, different fixed effects or clustering, or placebo treatment dates. The top row of each panel shows the baseline estimate in column 1 of Table 4.

FIGURE 8
Impact of Privacy Label Release on Downloads: International

Panel A. Country-specific Top 10k Apps



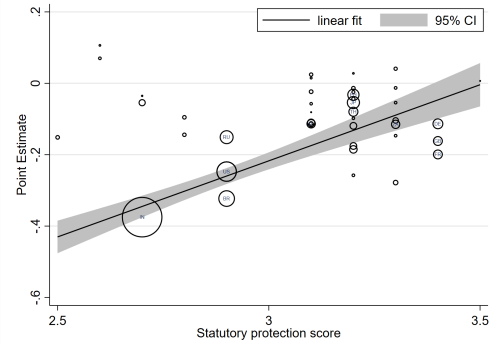
Panel B. Apps available in all 10 countries



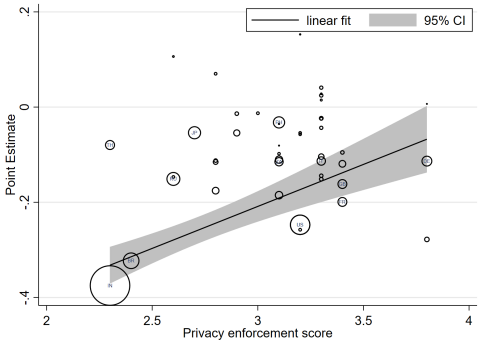
NOTE.— This figure plots estimated coefficients of the interaction term in Equation (1) for international app samples of 10 countries mentioned in Section 3. Both panels shows the estimated coefficients in the top and bottom quartile of the total number of data types collected. Panel A include the top10k apps in each country and Panel B only includes the global apps that are commonly available in the 10 countries.

FIGURE 9
What Explains Country-level Heterogeneity?
The Role of Regulation, Public Trust, and Privacy Concerns

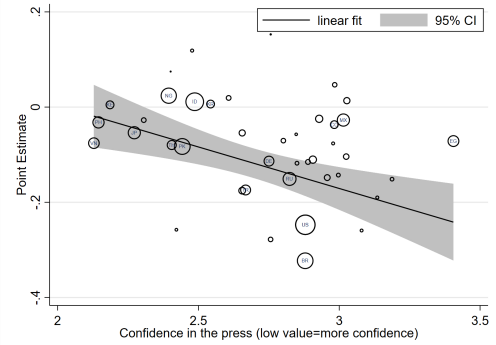
Panel A. Privacy protection regulation



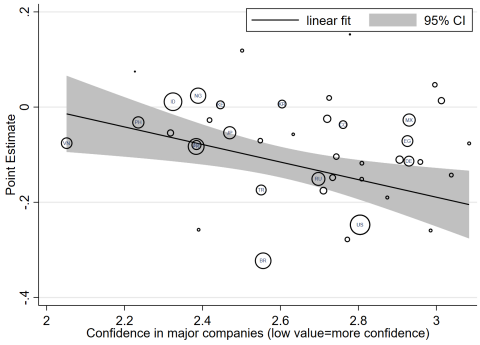
Panel B. Privacy regulation enforcement



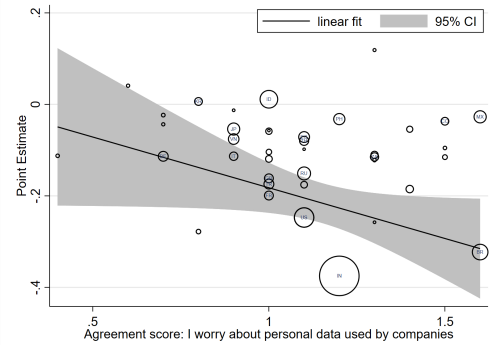
Panel C. Trust in the press



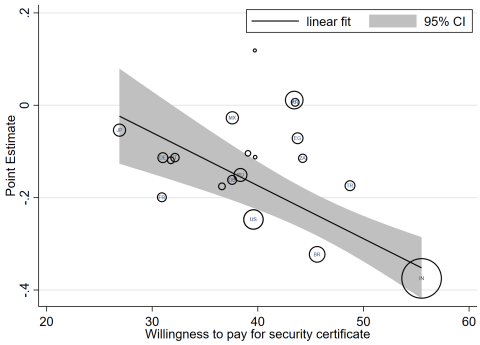
Panel D. Trust in major companies



Panel E. Concern over company's data use



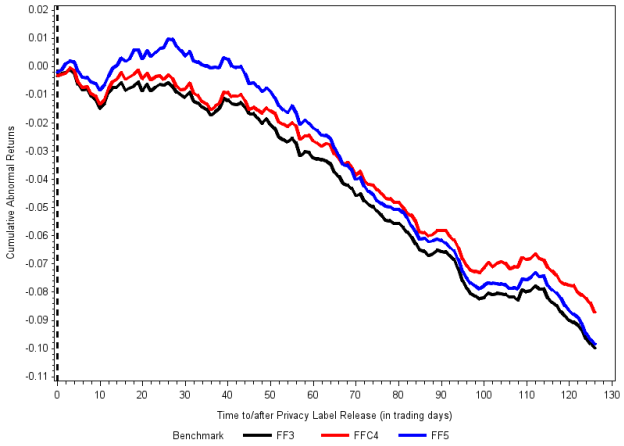
Panel F. WTP for security certificate



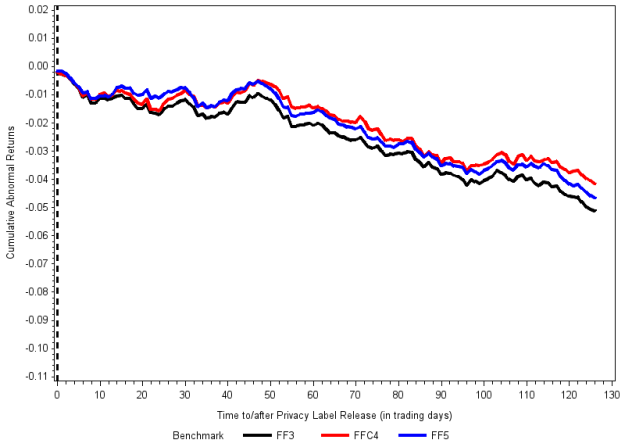
NOTE.— This figure shows the relationship between the impact of privacy label release on downloads and various index that capture the status of privacy protection, public trust, and privacy concerns. Only apps owned by public firms (which can be listed in any country) are included in this analysis. The fitted line is obtained by regressing the country-specific reaction to privacy label release on the respective index. Observations are weighed by country population, as represented also by the size of the circle. Panels A and B focus on the state of privacy protection obtained from Comparitech. The higher the score, the stronger the privacy protection and law enforcement. In Panels C and D, we examine the role of general trust in the press and major companies, respectively. Both measures are obtained from the World Value Survey (2022). A lower value corresponds to a higher confidence level. In Panels E and F, we explore two measures for attitudes towards data privacy issues. The first one, shown in Panel E, is obtained from the Data Confidence Index by Datum Future. It is the average agreement score for statement "*I worry about how my personal data is being used by companies.*" The higher the value, the stronger the concern. The second measure, willingness to pay for security certificate, obtained from the CIGI-Ipsos Global Survey on Internet Security and Trust, is solicited by asking respondents the following question: "*For an equivalent product valued at $1,000, how much more would you be willing to pay to have a product with a (Internet) security certification mark from the government?*" This measure is expressed in dollar term.

FIGURE 10
6-month CARs after the Release of Privacy Labels

Panel A. # Data Types Collected - High



Panel B. # Data Types Collected - Low



NOTE.— This graph presents the evolution of the average cumulative abnormal returns (CARs) after the release of firms' privac labels. For firms with multiple release date, we use the average date. CARs are either computed using CAPM and Fama-french factor models. Panels A and B split the sample by whether the app collects above-sample-median number of data types.

TABLE 1
Summary Statistics

Panel A. App Characteristics

	min	*mean*	*p*50	*sd*	*max*	*count*
Privacy Labels						
𝟙(Data used to track you)	0	0.6	1	0	1	6,344
# Data Types Collected	0	15.6	14	11	80	6,344
# Data Items Collected	0	24.1	20	19	167	6,344
# Data Items Collected - *By Purpose Category*						
Third-party Ad	0	1.6	0	3	24	6,344
Product Personalization	0	2.0	0	3	25	6,344
Developer's Adv or Mkting	0	2.0	0	3	24	6,344
Analytics	0	3.6	2	4	30	6,344
Other Purposes	0	0.6	0	2	28	6,344
App Functionality	0	4.5	3	5	32	6,344
Other App Characteristics						
Within-category Market Share	0	0.1	0	0	11	7,692
In-app Purchase	0	0.7	1	0	1	7,692
Ratings	1	4.4	5	1	5	7,692
App Age	1	4.5	4	3	13	7,692
Public Developers	0	0.2	0	0	1	7,692
Content Rating Age: 4+	0	0.6	1	0	1	7,692
Content Rating Age: 9+	0	0.1	0	0	1	7,692
Content Rating Age: 12+	0	0.2	0	0	1	7,692
Content Rating Age: 17+	0	0.1	0	0	1	7,692
#Countries	1	79.5	101	40	102	7,692

Panel B. Downloads and Revnues

	min	*mean*	*p*50	*sd*	*max*	*count*
iOS Weekly Downloads(k)						
Main Version	0.00	18.99	5.90	55.18	6170.35	435,735
All Versions	0.00	19.07	5.95	55.22	6170.35	435,735
iOS Weekly Revenues(k)						
Main Version	0.00	48.06	0.17	271.38	12763.41	435,735
All Versions	0.00	48.19	0.18	271.41	12763.41	435,735
Android Weekly Downloads(k)						
Main Version	0.00	12.21	4.11	29.71	3249.13	435,735
All Versions	0.00	12.29	4.13	30.07	3249.13	435,735
Android Weekly Revenues(k)						
Main Version	0.00	29.63	0.01	215.52	15087.69	435,735
All Versions	0.00	29.69	0.01	215.56	15087.70	435,735

NOTE.—This table provides descriptive statistics on the data collection activities of top 10k US apps and shows a comprehensive list of app characteristics we consider in the empirical analysis. Panel A shows summary statistics of app characteristics. Panel B presents summary statistics of weekly downloads and revenue for the iOS and Android versions of our top 10k apps separately.

TABLE 2
Determinants of Data Collection Intensity

	𝟙(Data used to track you)		**# Data Types Collected**		**# Data Items Collected**	
	(1)	(2)	(3)	(4)	(5)	(6)
Within-category Market Share	−0.030**	−0.002	2.553***	3.148***	4.491***	5.746***
	(0.01)	(0.01)	(0.33)	(0.34)	(0.57)	(0.59)
In-app Purchase	0.239***	0.162***	1.187***	1.404***	0.699	0.916
	(0.01)	(0.01)	(0.28)	(0.33)	(0.50)	(0.57)
App Age	−0.065***	−0.034***	−0.430***	−0.127	−1.406***	−0.807***
	(0.01)	(0.01)	(0.16)	(0.15)	(0.27)	(0.27)
App Age^2	0.004***	0.002***	0.048***	0.031**	0.115***	0.083***
	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)
Ratings	0.114***	0.077***	3.025***	2.304***	4.690***	3.500***
	(0.01)	(0.01)	(0.24)	(0.24)	(0.42)	(0.42)
Public Developers	−0.002	−0.016	8.052***	8.004***	14.859***	14.664***
	(0.01)	(0.01)	(0.34)	(0.34)	(0.60)	(0.60)
Global App	−0.082***	0.045***	−3.299***	−1.403***	−5.412***	−1.852***
	(0.01)	(0.01)	(0.28)	(0.31)	(0.50)	(0.55)
Content Rating Age: 9+	0.131***	0.009	0.648	−0.701	0.802	−1.601
	(0.02)	(0.02)	(0.56)	(0.56)	(0.99)	(0.99)
Content Rating Age: 12+	0.160***	0.094***	3.925***	3.168***	6.769***	5.663***
	(0.01)	(0.01)	(0.32)	(0.33)	(0.56)	(0.57)
Content Rating Age: 17+	0.116***	0.101***	2.189***	1.943***	2.721***	2.492***
	(0.02)	(0.02)	(0.42)	(0.42)	(0.73)	(0.74)
Constant	0.133***	0.236***	0.099	1.889*	2.489	5.012**
	(0.05)	(0.05)	(1.15)	(1.13)	(2.01)	(2.00)
Category FE		Y		Y		Y
Observations	6,344	6,344	6,344	6,344	6,344	6,344
R-sq	0.195	0.295	0.193	0.238	0.189	0.230

NOTE.—This table shows the association between app characteristics and their data collection intensity. The regression uses three measures of data collection intensity as the outcome variable and app characteristics as regressors. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 3
Data Collection Intensity by Use

	(1) Third-party Ad	(2) Developer's Ad or Mkting	(3) Analytics	(4) Product Personalization	(5) Other Purposes	(6) App Functionality
Books	0.377	0.729*	0.277	0.807**	0.327	0.003
	(0.31)	(0.37)	(0.31)	(0.39)	(0.26)	(0.37)
Business	0.112	0.394	0.199	0.360	0.195	−0.105
	(0.31)	(0.36)	(0.31)	(0.38)	(0.26)	(0.36)
Education	0.283	0.386	0.029	0.303	0.144	−0.433
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Entertainment	0.815***	0.789**	0.409	0.616	0.192	−0.169
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Finance	0.166	0.604*	0.231	0.467	0.217	0.045
	(0.31)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Food & Drink	0.340	0.914**	0.248	0.756**	0.276	0.125
	(0.31)	(0.36)	(0.31)	(0.38)	(0.26)	(0.36)
Games	1.209***	0.848**	0.426	0.606	0.286	−0.255
	(0.30)	(0.36)	(0.30)	(0.37)	(0.25)	(0.35)
Health & Fitness	0.341	0.644*	0.287	0.633*	0.181	−0.024
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Lifestyle	0.366	0.566	0.230	0.547	0.162	−0.098
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Medical	0.239	0.538	0.145	0.608	0.223	−0.013
	(0.31)	(0.37)	(0.31)	(0.38)	(0.26)	(0.36)
Music	0.777**	0.673*	0.252	0.541	0.185	−0.282
	(0.31)	(0.37)	(0.31)	(0.38)	(0.26)	(0.36)
Navigation	0.379	0.849**	0.545*	0.725*	0.067	0.312
	(0.32)	(0.38)	(0.32)	(0.39)	(0.27)	(0.37)
News	0.963***	0.982***	0.403	0.853**	0.226	−0.009
	(0.31)	(0.37)	(0.31)	(0.39)	(0.26)	(0.37)
Photo & Video	0.503*	0.528	0.225	0.306	0.137	−0.402
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Productivity	0.343	0.425	0.254	0.272	0.096	−0.164
	(0.31)	(0.36)	(0.31)	(0.38)	(0.26)	(0.36)
Reference	0.589*	0.420	0.264	0.422	0.089	−0.349
	(0.32)	(0.38)	(0.32)	(0.39)	(0.27)	(0.37)
Shopping	0.505*	1.083***	0.608**	1.099***	0.326	0.298
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Social Networking	0.568*	0.659*	0.340	0.725*	0.201	0.118
	(0.31)	(0.36)	(0.31)	(0.38)	(0.26)	(0.36)
Sports	0.775**	1.070***	0.458	0.889**	0.345	0.133
	(0.31)	(0.37)	(0.31)	(0.39)	(0.26)	(0.37)
Travel	0.302	0.931**	0.497	0.966**	0.169	0.278
	(0.31)	(0.37)	(0.31)	(0.38)	(0.26)	(0.36)
Utilities	0.437	0.285	0.080	0.134	0.143	−0.512
	(0.30)	(0.36)	(0.30)	(0.38)	(0.26)	(0.36)
Weather	0.857***	0.752*	0.242	0.451	0.176	−0.209
	(0.33)	(0.39)	(0.32)	(0.40)	(0.27)	(0.38)
Constant	0.000	−0.000	1.069***	0.173	0.000	1.573***
	(0.30)	(0.36)	(0.30)	(0.37)	(0.25)	(0.35)
Observations	6,344	6,344	6,344	6,344	6,344	6,344
R-sq	0.315	0.074	0.048	0.067	0.017	0.072

NOTE.—This table shows the heterogeneities across app categories in data uses. The outcome variable in columns 1 to 6 corresponds to the number of data types collected for the six data use categories, respectively. The reference (omitted) group is the "others" category.

TABLE 4
Impact of Privacy Label Release on Downloads and revenue: Baseline

	Downloads			**Revenues**		
	(1) Main version	(2) All versions	(3) All versions	(4) Main version	(5) All versions	(6) All versions
Post	0.240***	0.238***	0.203***	0.255***	0.284***	0.207***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
iOS × Post	−0.117**	−0.117**	−0.138***	−0.135***	−0.195***	−0.151***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Sample	Full	Full	Updated apps	Full	Full	Updated apps
Observations	966,216	966,216	871,470	966,216	966,216	871,470
R-sq	0.557	0.558	0.569	0.917	0.917	0.920

NOTE.—This table shows the regression results of Equation (1). The outcome variables are downloads and revenue. Columns 1 and 4 use all available versions of apps; Columns 2 and 5 aggregate the downloads over various versions of the same app. Columns 3 and 6 only include the apps that eventually updated the privacy label section on its apple store page. Year-week, app, and platform-age fix effects are included. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 5
Impact of Privacy Label Release on Downloads and revenue: Intensive Margin

	Downloads (all versions)			**Revenues (all versions)**		
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.250***	0.189**	0.200**	0.227***	0.267***	0.283***
	(0.04)	(0.08)	(0.08)	(0.04)	(0.07)	(0.07)
iOS × Post	−0.063	0.064	0.070	−0.091**	−0.049	−0.029
	(0.05)	(0.09)	(0.09)	(0.04)	(0.07)	(0.06)
iOS × Post × 𝟙(Data used to track you)	−0.103***			−0.109***		
	(0.03)			(0.04)		
iOS × Post × # Data Types Collected		−0.078***			−0.042*	
		(0.02)			(0.02)	
iOS × Post × # Data Items Collected			−0.071***			−0.043**
			(0.02)			(0.02)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.570	0.569	0.569	0.920	0.920	0.920

NOTE.—This table shows the regression results of Equation (2) for downloads and revenue. In the triple DID term, Columns 1 and 4 add an indicator variable for whether the app collects any data in the *Data Used to Track You* category, Columns 2 and 5 include the number of data types collected, and Columns 3 and 6 include the number of data items collected. Year-week, app, and platform-age fix effects are included. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 6
Heterogeneity in Data Use Category

Panel A. Downloads

	Downloads (all versions)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.281***	0.231***	0.180***	0.215***	0.213***	0.143**
	(0.05)	(0.04)	(0.06)	(0.05)	(0.04)	(0.06)
iOS × Post	−0.096**	−0.107**	−0.015	−0.110**	−0.131***	−0.091
	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.08)
iOS × Post × Third-party Ad	−0.049**					
	(0.02)					
iOS × Post × Developer Ad or Mkt		−0.044*				
		(0.02)				
iOS × Post × Analytics			−0.085***			
			(0.03)			
iOS × Post × Product Personalization				−0.040*		
				(0.02)		
iOS × Post × Other Purposes					−0.033	
					(0.03)	
iOS × Post × App Functionality						−0.039
						(0.03)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.571	0.570	0.569	0.569	0.569	0.570

Panel B. Revenue

	Revenue (all versions)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.237***	0.232***	0.221***	0.228***	0.220***	0.219***
	(0.04)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)
iOS × Post	−0.079**	−0.129***	−0.100*	−0.138***	−0.136***	−0.127**
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.06)
iOS × Post × Third-party Ad	−0.107***					
	(0.03)					
iOS × Post × Developer Ad or Mkt		−0.043				
		(0.03)				
iOS × Post × Analytics			−0.038			
			(0.03)			
iOS × Post × Product Personalization				−0.018		
				(0.02)		
iOS × Post × Other Purposes					−0.077	
					(0.05)	
iOS × Post × App Functionality						−0.011
						(0.03)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.920	0.920	0.920	0.920	0.920	0.920

NOTE.—This table shows the regression results of Euqation (2) for downloads and revenue. In the triple DID term, Columns 1 to 6 use the total number of data types and data items in each of the 6 purposes under *Data Linked to You* and *Data Not Linked to You* to interact with *iOS* × *Post*, respectively. Panel A shows the result for downloads and Panel B for revenue. Year-week, app, and platform-age fix effects are included. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 7
Heterogeneity in App Market Power

	Downloads (all versions)			**Revenues (all versions)**		
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.078*	0.200***	0.166***	0.193***	0.191***	0.248***
	(0.04)	(0.04)	(0.06)	(0.04)	(0.03)	(0.06)
iOS × Post	−0.046	−0.143***	−0.225***	−0.160***	−0.153***	−0.195***
	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)	(0.07)
iOS × Post × Platform-wide ranking decile	−0.016**			−0.001		
	(0.01)			(0.01)		
iOS × Post × 𝟙(Market share above 90^{th} pct.)		0.059*			0.013	
		(0.03)			(0.06)	
iOS × Post × Age			0.015**			0.008
			(0.01)			(0.01)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.571	0.569	0.569	0.920	0.920	0.920

NOTE.—This table shows the regression results of Equation (2) for downloads and revenue. In the triple DID term, we use three measures of the substitutability of an app to interact with $iOS \times Post$, respectively. Columns 1 and 4 use the download ranking of the app in the iOS store in 2020. Columns 2 and 5 use an indicator variable that equals one if the app's market share is above the 90^{th} percentile among its category. Columns 3 and 6 use app's age. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

TABLE 8
Abnormal Returns around the Release of Privacy Labels

Panel A. 6-month CAR after Label Release

Level	EventDate	Obs	CAPM	FF3	FFC4	FF5
App	Staggered	2,188	-7.25	-5.57	-4.80	-5.16
			(-16.55)	(-15.31)	(-12.80)	(-14.67)
Firm - avg.	Staggered	467	-10.95	-7.51	-6.54	-6.63
			(-9.89)	(-7.91)	(-6.71)	(-7.31)
Firm - rep.	Staggered	459	-8.17	-7.02	-6.14	-5.74
			(-7.66)	(-7.95)	(-6.71)	(-6.81)
Firm	14/12/2020	485	-3.38	-9.14	-7.60	-7.33
			(-3.27)	(-10.25)	(-8.08)	(-8.87)

Panel B. Subsample Analysis - # Data Types Collected

Firms	Group	Obs	CAPM	FF3	FFC4	FF5
All	H	226	-14.53	-10.95	-9.72	-9.88
			(-8.58)	(-7.67)	(-6.47)	(-7.31)
All	L	229	-7.70	-4.02	-3.41	-3.41
			(-5.50)	(-3.32)	(-2.82)	(-2.93)
All	H-L	455	-6.83	-6.93	-6.31	-6.47
			(-3.11)	(-3.70)	(-3.27)	(-3.62)
Retail & Service	H	82	-15.50	-11.53	-10.29	-10.31
			(-4.76)	(-4.13)	(-3.45)	(-4.06)
Retail & Service	L	82	-6.61	-3.07	-1.14	-2.39
			(-2.89)	(-1.47)	(-0.57)	(-1.20)
Retail & Service	H-L	164	-8.88	-8.46	-9.15	-7.92
			(-2.23)	(-2.43)	(-2.55)	(-2.45)

NOTE.—Panel A reports 6-month CAR after controlling for CAPM, FF3, FFC4, FF5 factors for both staggered and non-staggered samples, at App and firm level. Panel B reports results of firm-level subsample analysis using the number of data types collected as the sorting variable to divide the sample into halves. Subsample with above (below) median collection intensity is denoted by "H" ("L"). Staggered event dates are used in subsample analysis. T-statistics are reported in parentheses.

TABLE 9
Impact of Privacy Label Release on Earnings

Sample	All	All	Retail & Service	All	All	Retail & Service
	(1)	(2)	(3)	(4)	(5)	(6)
Post × %(Data Used to Track You)	−0.641**	−0.833**	−1.300**	−1.407***	−1.568**	−2.252**
	(0.28)	(0.40)	(0.50)	(0.37)	(0.60)	(0.86)
Size_{q-4}		−0.607	−1.198		−1.556	−1.083
		(0.63)	(0.73)		(0.94)	(0.81)
Cash_{q-4}		1.883	2.354		1.098	−1.919
		(2.68)	(3.53)		(2.18)	(2.13)
Tangible_{q-4}		−1.696	−4.844		−5.462	−18.908
		(5.40)	(6.85)		(7.03)	(12.76)
Leverage_{q-4}		−1.216	−1.851		−3.173**	−3.641**
		(1.54)	(1.81)		(1.45)	(1.72)
EBITDA_{q-4}		0.915	0.287		2.033	4.117
		(1.94)	(2.31)		(6.03)	(8.75)
Constant	2.100***	16.880	32.029*	1.798***	40.103*	31.138
	(0.04)	(16.26)	(19.07)	(0.09)	(23.13)	(21.30)
Weighted by	Downloads	Downloads	Downloads	Revenue	Revenue	Revenue
Firm	Y	Y	Y	Y	Y	Y
Quarter FE	Y	Y	Y	Y	Y	Y
Observation	3,260	3,095	1,364	521	505	277
R-sq	0.72	0.73	0.73	0.86	0.87	0.88

NOTE.—This table shows the impact of privacy label release on firm earnings as estimated in Equation (3). In columns 1-3, we weight observations by the annual downloads in 2020, and in columns 4-5, by the 2020 annual revenue. Columns 1-2 and 4-5 use the full set of public firms, while columns 3 and 6 only include firms in the retail and service industry. %(*Data Used to Track You*) is defined as the percentage share of a firm's apps that collect data to track users. Firm and year-quarter fixed effects are included. Standard errors are clustered at firm level and reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

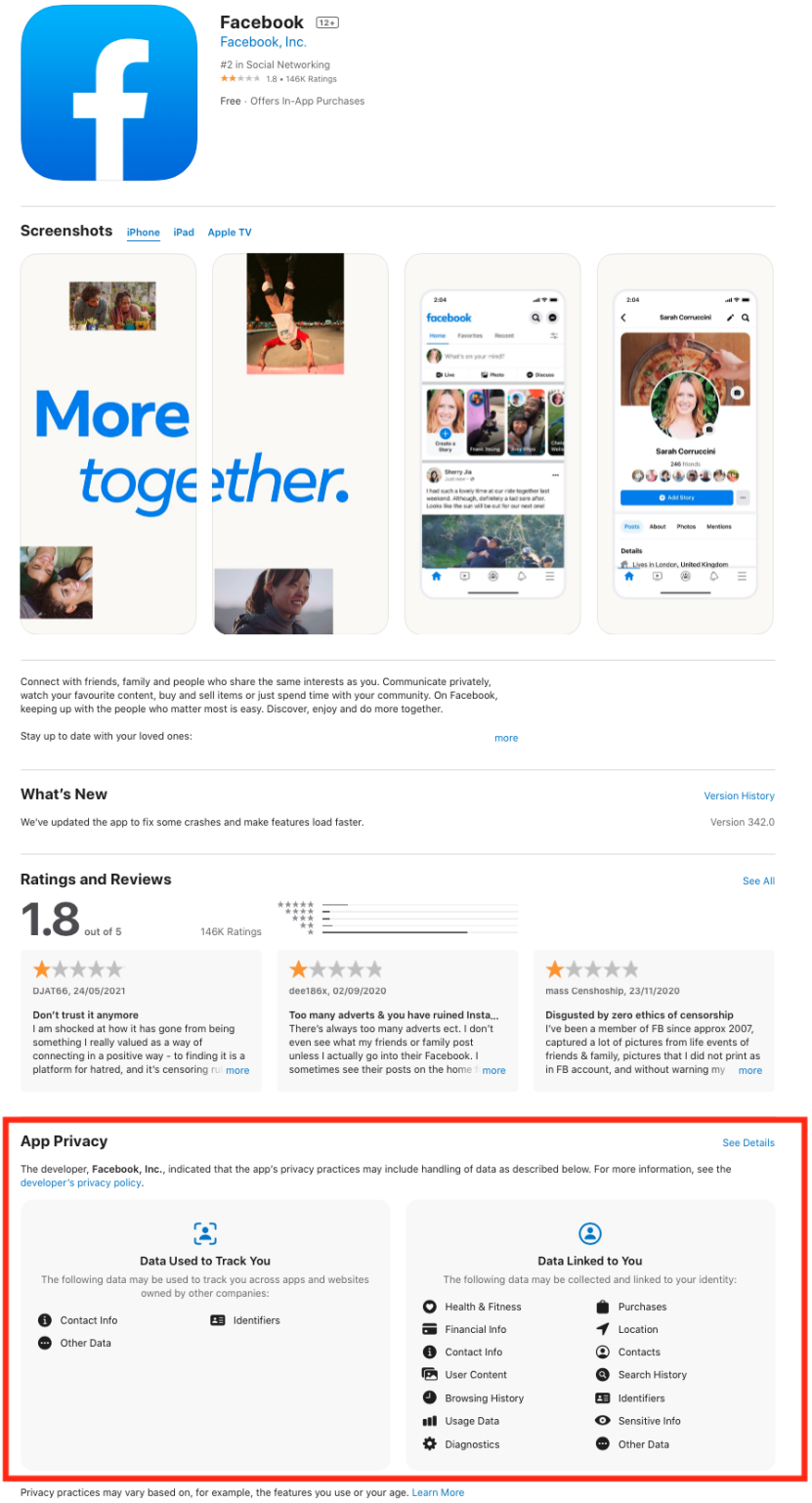
The Supply and Demand for Data Privacy: Evidence from Apple's Privacy Labels

Online Appendices

Bo Bian Xinchen Ma Huan Tang

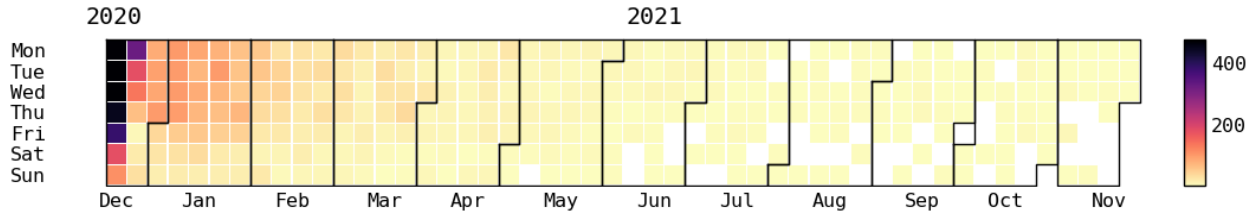
A Additional Figures

FIGURE A.1
An Example of Privacy Labels



NOTE.— This figure provides a screenshot of Facebook's App Store page at https://apps.apple.com/us/app/facebook/id284882215. The App Privacy section comes right after the section of Ratings and Reviews.

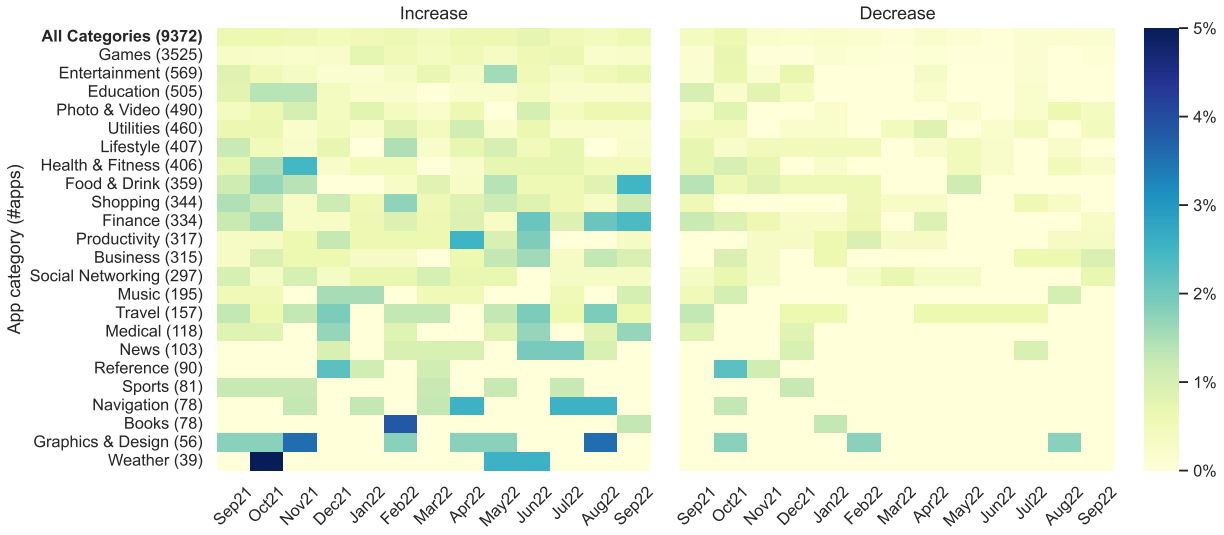
FIGURE A.2
Release Dates of Privacy Labels



NOTE.— This figure plots the privacy label release dates of Top 10k apps in the US based on downloads. The label release date is defined as the earlier of the following two dates: the first release date of a new version of the app on or after 2020/12/14, and the first date on which the Wayback Machine captured available screenshots of the privacy labels.

FIGURE A.3
Changes in the Data Collection Intensity Over time

Panel A. Total Number of Data Types Collected

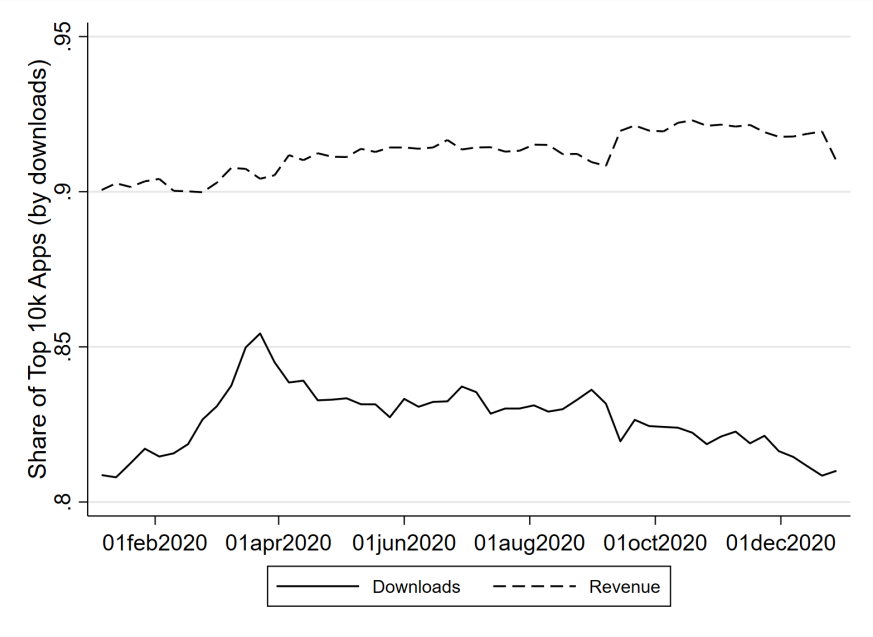


Panel B. Collection of Data Used to Track You

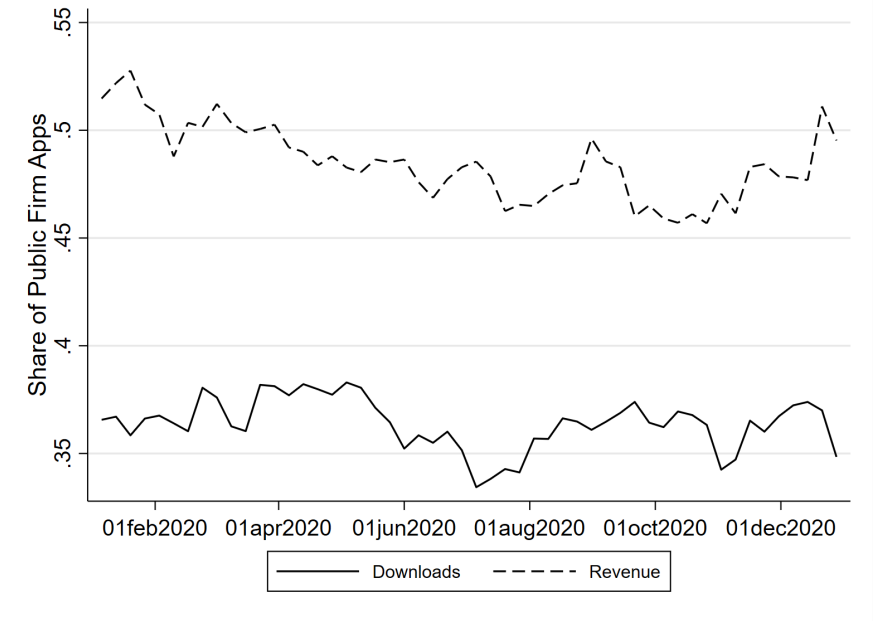
NOTE.—The heatmaps show the monthly fraction of apps that increase or decrease data collection intensity between September, 2021 and September, 2022. The fractions in Panel A are calculated based on the total number of data items collected and that in Panel B are based on whether the app changes its collection status of Data Used to Track You. The fractions are presented for both the full sample of Top 10k apps in the US and for each app category.

FIGURE A.4
Market Share of Top 10k Apps based on Downloads and Revenues

Panel A. Market Share of Top 10k Apps



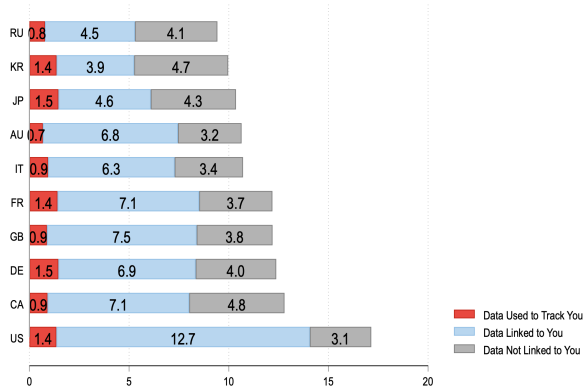
Panel B. Market Share of Apps Developed by Public Firms



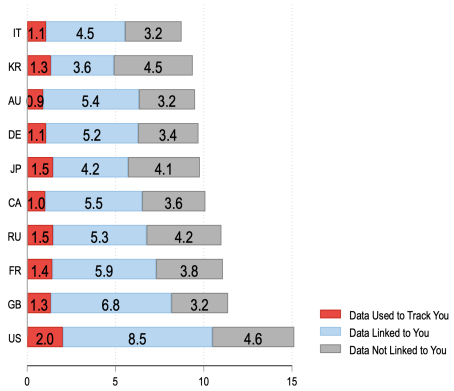
NOTE.— This figure plots the market share of apps based on Download and Revenue over time. Panel A focuses on the market share of Top 10k apps. Panel B focuses on the market share of apps developed by public firms.

FIGURE A.5
Data Collection Intensity: International Sample

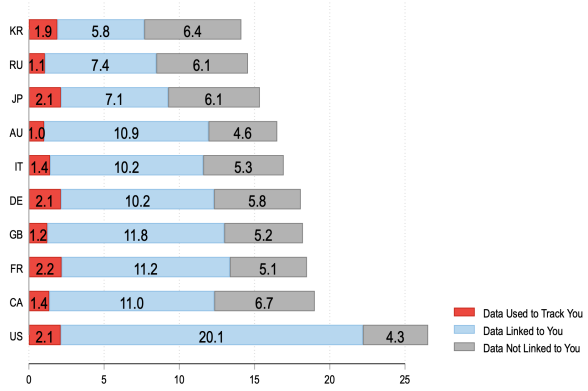
Panel A. # Data Types Collected: Local Apps



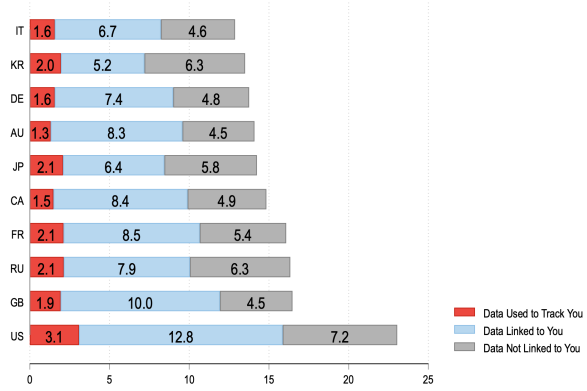
Panel B. # Data Types Collected: Main country



Panel C. # Data Items Collected: Local Apps



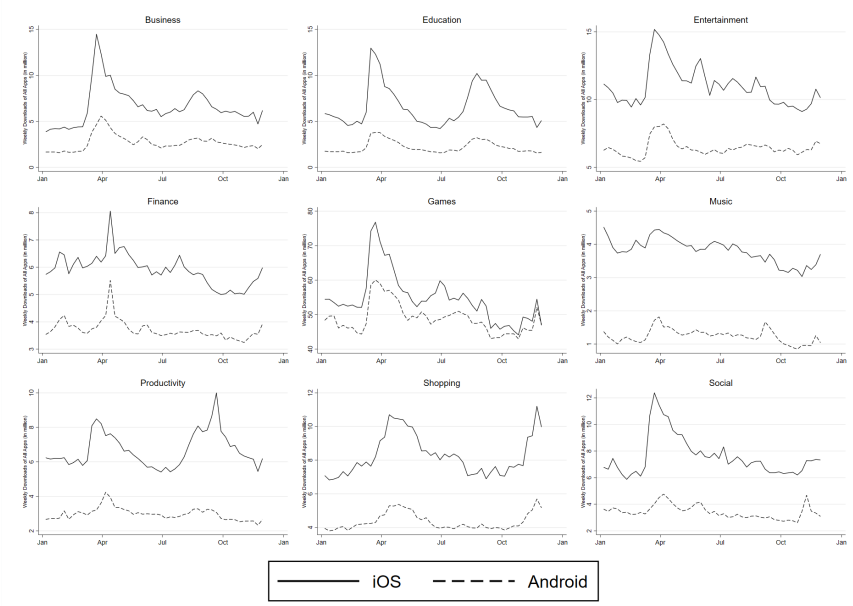
Panel D. # Data Items Collected: Main country



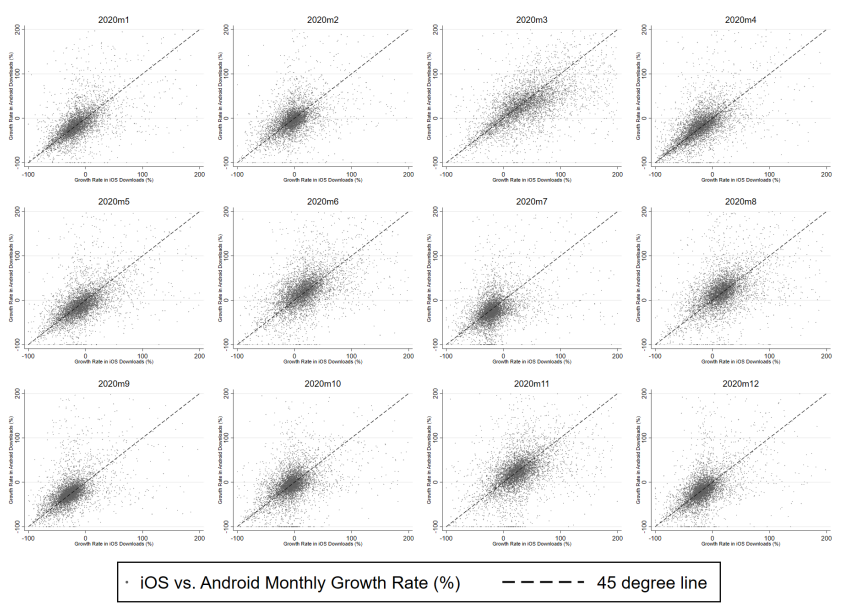
NOTE.— This figure shows the average number of data types (Panels A and B) and items (Panels C and D) collected for each country. Panels A and C include apps that are available in only one country (labeled as local apps). Panels B and D include all apps in the international sample, but assign each app to the country where it has the most downloads in 2020.

FIGURE A.6
Downloads: iOS vs. Android

Panel A. Download Growth by App Category: iOS vs. Android

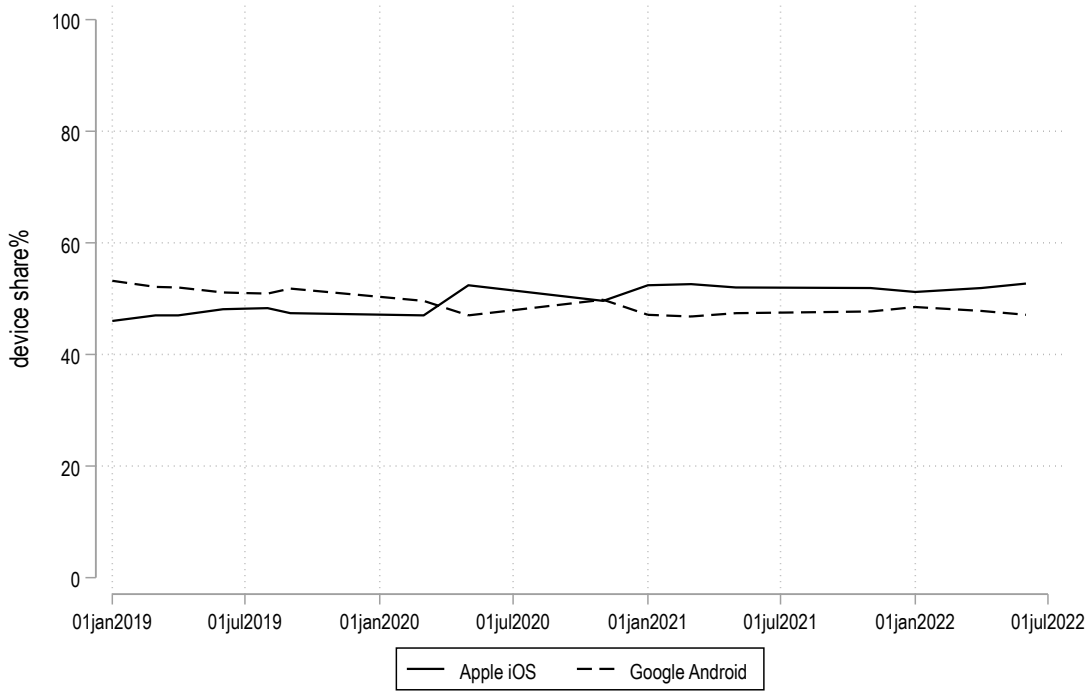


Panel B. App-level Download Growth by Month: iOS vs. Android



NOTE.— This figure shows that the iOS and Android downloads patterns evolve in parallel over time. Subfigure a plots the aggregate weekly downloads by app category for apps on the two platforms. Subfigure b plots the monthly growth rate of iOS apps against Android apps, in each of the 12 months in 2020. The dashed lines are the 45-degree line.

FIGURE A.7
Market Share of Apple OS vs. Google Android



NOTE.— This figure shows the market share of Apple OS and Google Android among all smartphone operating systems between 2019-2022. Source data is from Comscore: https://www.statista.com/statistics/266572/market-share-held-by-smartphone-platforms-in-the-united-states/.

B Additional Tables

TABLE B.1
Data Category Explained

Panel A. First-layer Category

Data Used to Track you	Data used to track you (or your device) and *shared* across different apps, ad networks, and companies
Data Linked to You	Data linked to you (and your real identity) that is collected by the app and company but *not shared*
Data Not Linked to You	Data not linked to you that the company generally aggregates into larger statistics

Panel B. Second-layer: Data Use Category

Third-Party Advertising	Such as displaying third-party ads in your app, or sharing data with entities who display third-party ads
Developer's Advertising or Marketing	Such as displaying first-party ads in your app, sending marketing communications directly to your users, or sharing data with entities who will display your ads
Analytics	Using data to evaluate user behavior, including to understand the effectiveness of existing product features, plan new features, or measure audience size or characteristics
Product Personalization	Customizing what the user sees, such as a list of recommended products, posts, or suggestions
App Functionality	Such as to authenticate the user, enable features, prevent fraud, implement security measures, ensure server up-time, minimize app crashes, improve scalability and performance, or perform customer support
Other Purposes	Any other purposes not listed

Note.—The definition of these categories is from https://apps.apple.com/story/id1539235847

TABLE B.2
Updates in Privacy Labels

Panel A. Number of Apps with Label Updates

				Total Number of Data Types		Data Used to Track You	
Month	Total apps	Valid label	Any change	Increase	Decrease	Turned on	Turned off
Sep21	9327	7312	97	57	36	28	20
Dec21	9327	7366	243	150	86	92	33
Mar22	9327	7367	181	140	37	59	16
Jun22	9327	7405	212	176	26	57	11
Sep22	9327	7377	182	134	42	72	15

Panel B. Number of Apps with Label Updates - Public Firm

				Total Number of Data Types		Data Used to Track You	
Month	Total apps	Valid label	Any change	Increase	Decrease	Turned on	Turned off
Sep21	6461	3880	29	13	15	6	10
Dec21	6461	3939	62	44	15	21	11
Mar22	6461	3926	25	21	4	9	5
Jun22	6461	3900	75	45	24	12	8
Sep22	6461	3893	55	40	13	11	5

NOTE.—This table reports quarter-to-quarter label updates for the US top10k sample and the public firm sample. We do not exclude apps that only have an iOS version but not an Android version. "Valid label" shows the number of active apps with released privacy labels. "Any change" reports the number of apps that have any updates in their privacy labels compared to the last quarter. "Increase" ("Decrease") represents the number of apps that report an increased (decreased) number of data items collected. "Turned on" shows the number of apps that turned on the data tracking by collecting data under Data Used to Track You category. "Turned off" shows the number of apps that turned off data tracking. The first rows of both panels report the updates from our first version of privacy labels in September 2021.

TABLE B.3
Heterogenous Effects on Downloads and Revenue: By Purpose Category

Panel A. Downloads

	Downloads (all versions)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.223***	0.197***	0.146***	0.197***	0.201***	0.134***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
iOS × Post	−0.116**	−0.119**	−0.087	−0.125**	−0.132***	−0.107*
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
iOS × Post × Third-party Ad	−0.044*					
	(0.02)					
iOS × Post × Developer Ad or Mkt		−0.041*				
		(0.02)				
iOS × Post × Analytics			−0.056**			
			(0.02)			
iOS × Post × Product Personalization				−0.025		
				(0.02)		
iOS × Post × Other Purposes					−0.039	
					(0.03)	
iOS × Post × App Functionality						−0.035
						(0.02)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.570	0.570	0.569	0.569	0.569	0.570

Panel B. Revenues

	Revenue (all versions)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post	0.214***	0.211***	0.190***	0.213***	0.213***	0.198***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.05)
iOS × Post	−0.108***	−0.134***	−0.139***	−0.139***	−0.145***	−0.163***
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
iOS × Post × Third-party Ad	−0.092***					
	(0.03)					
iOS × Post × Developer Ad or Mkt		−0.034				
		(0.02)				
iOS × Post × Analytics			−0.011			
			(0.02)			
iOS × Post × Product Personalization				−0.016		
				(0.02)		
iOS × Post × Other Purposes					−0.057	
					(0.05)	
iOS × Post × App Functionality						0.019
						(0.02)
Linear Trend	Y	Y	Y	Y	Y	Y
Year-week FE	Y	Y	Y	Y	Y	Y
App FE	Y	Y	Y	Y	Y	Y
Platform-Age FE	Y	Y	Y	Y	Y	Y
Observations	871,470	871,470	871,470	871,470	871,470	871,470
R-sq	0.920	0.920	0.920	0.920	0.920	0.920

NOTE.—This table shows the regression results of Equation (2) for downloads and revenues. In the triple DID term, Columns 1 to 6 use the total number of data types and data items in each of the 6 purposes under *Data Linked to You* to interact with *iOS* × *Post*, respectively. Panel A shows the result for downloads and Panel B for revenues. Year-week, app, and platform-age fix effects are included. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

C Investor Underreaction

Despite the negative stock market reaction over the six-month horizon, Figure 10 reveals that this effect does not kick in immediately. In this section, we provide evidence consistent with investors' underreaction to the release of privacy labels. There are several explanations for investor underreaction. First, Apple's privacy label requirement constitutes the first policy that specifically targets transparent disclosures of firms' data collection activities. In contrast, the scope of previous privacy protection regulations such as GDPR or CCPA is much wider, and often shift supply and demand simultaneously. Therefore, investors may not be able to fully understand and anticipate the policy's implications for data collection. In addition, given the ongoing debates on privacy paradox, it remains unclear whether consumers attach a significant monetary value to their personal data, and whether they would take actions to protect themselves by not sharing data. It is therefore ambiguous ex-ante whether firms with high data collection intensity would experience a sharp drop in the demand for their apps. Moreover, one can also argue that the information presented in privacy labels is not sufficiently salient and many consumers and investors may fail to notice these privacy labels.

One implication of above arguments is that once the uncertainty about the impact of mandatory disclosure on firms' data collection practice is resolved, and/or when the information on firms' data collection becomes more salient, investors should react more quickly. Consistent with this idea, we first show that following firms' first earnings release after the privacy label policy, there is a -3% announcement return over a 30-day window. Furthermore, we exploit a subsequent shock that improves the saliency of firm's data collection practice to consumers and show that investors react almost immediately to this shock.

C.1 Stock market returns around earnings announcement

In Section 6, we show that firms earnings drop following the policy. The earnings release effectively reduces uncertainty about the policy's effect. Thus, such earnings information should allow investors to better understand the policy's implications on firm fundamentals and incorporate it in their investment decisions.

Do investors indeed react strongly and promptly when key post-policy earnings numbers are disclosed? To answer this question, we implement an event study around firms' first earnings announcements post policy implementation. We require the earnings announcement to occur at least one month after the firm publishes its first privacy labels. For example, if the firm has multiple apps and releases its first privacy label in December 2020, we consider the 2021Q1 earnings announcement as the first event. Figure C.8 plots the CARs from 10 days before the earnings announcement to 30 days after the announcement. We compute CARs using the FFC4 model. The figure shows no evidence of abnormal returns in the ten days prior to the earnings announcement. In contrast, a negative abnormal return emerge right after the earnings announcement. The post earnings announcement drift suggests that investors fail to anticipate the deterioration in firm performance. This is consistent with our conjecture that market participants are initially uncertain about consumers' reactions, possibly due to the privacy paradox.

C.2 Stock market returns around the App Transparency Tracking policy

Another factor underlying investor underreaction concerns the saliency of the information on firms' data collection practices. One could argue that the information on privacy labels only draws attention from a limited subset of consumers and investors intially, preventing privacy-related information from being incorporated into stock price. If this argument bears any merit, we should observe a stronger stock market response when information on key aspects of firms' data collection practices becomes more salient. In this section, we provide evidence consistent with this argument by exploiting Apple's App Transparency Tracking (ATT) policy.

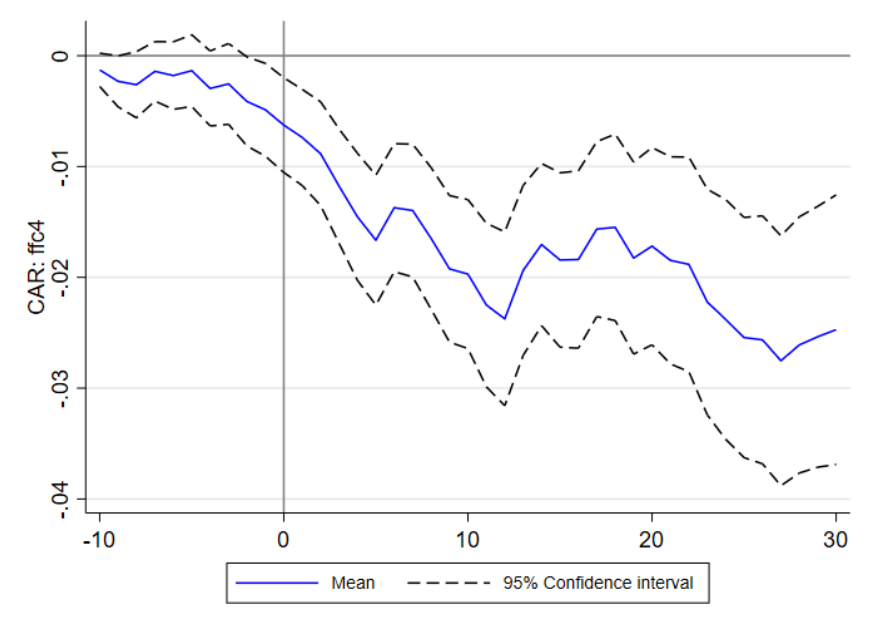
On April 26th, 2021, Apple adds a new feature to its operating system (iOS 14.5), which requires apps to seek permission from users before tracking their activity across other companies? apps and websites. By default, a user is opted-out of tracking when the app launches in the iOS 14.5 environment and The ATT request works on a per app basis, that is, users can decide to opt-in for specific apps as they choose. As of Feburary 2022, 18% of app users allowed tracking among those tho have been asked for permissions.[1]

[1]Source:https://www.flurry.com/blog/att-opt-in-rate-monthly-updates/

Importantly, the ATT policy is implemented in a more salient way than the privacy label policy, although both policies convey important information related to *Data Used to Track You*. While ATT requires the app to ask for permission in a pop-up window when users launch an app, privacy labels are merely displayed on the app's download page. To this end, the ATT policy can be viewed as a saliency shock.

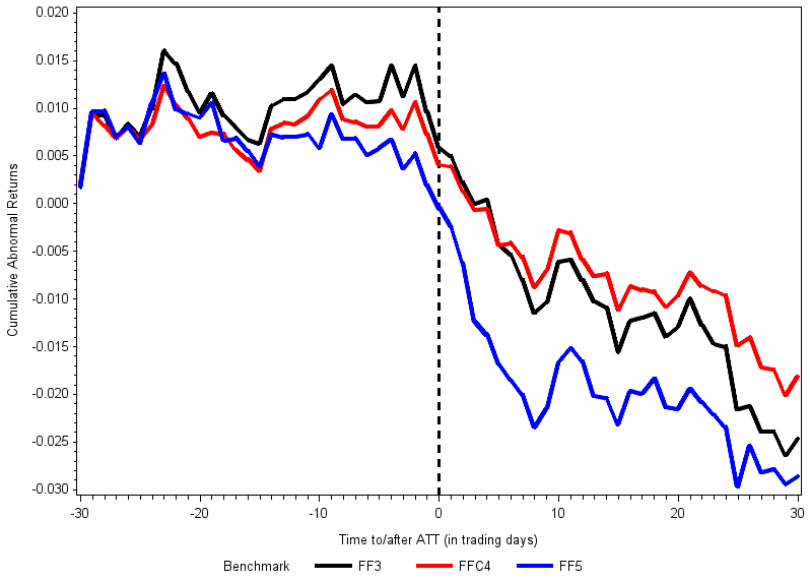
Based on the above arguments, we expect investors to act more swiftly once the ATT policy becomes effective. To verify this, we repeat the event study around the ATT implementation date and plot the CARs from 10 days before to 30 days after the event in Figure C.9. The event study reveals an immediate and sharp decline in the CARs right after the ATT implementation date, while the CARs are consistently around zero before the event. These patterns are robust to alternative benchmark models we use to estimate CARs and the 30-day CAR is around 2-3%.

FIGURE C.8
Event Study on the First Earnings Annoucements Post Label Release



NOTE.— This figure plots the average cumulative abnormal returns (CARs) around the first earnings release after firms' privacy label release. The event window starts ten days before and ends thirty days after the earnings release. We focus on the first earnings release that is at least one month after the firm publishes its first privacy label. For example, if the firm releases its first privacy label in December 2020, we consider the 2021Q1 earnings release as the first one. CARs are computed using the Fama-French four factor model.

FIGURE C.9
Event Study on the Implementation of the App Transparency Tracking Policy



NOTE.— This figure plots the average cumulative abnormal returns (CARs) around the implementation of the App Transparency Tracking Policy on April 26, 2021. The event window includes 30 days before and after the implementation date in the event window. CARs are computed using the Fama-french factor models.