Delaying Informed Consent: An Empirical Investigation of Mobile Apps' Upgrade Decisions*

Working Paper

Raveesh K Mayya
Robert H. Smith School of Business,
University of Maryland
College Park, MD 20742

Siva Viswanathan
Robert H. Smith School of Business,
University of Maryland
College Park, MD 20742

Abstract

We study apps' decisions to upgrade to Android 6.0, which restricts their ability to seek blanket permissions to sensitive user information at download, instead requiring them to request à la carte permissions at run-time. Mobile apps on Android had a choice of upgrading to Android 6.0 anytime over a three-year window instead of being forced to upgrade immediately. Given the choice of upgrading to version 6.0 that provides mobile apps with the latest platform features or staying with an earlier version that provides them with better access to user information, our study seeks to examine the upgrade decisions of apps in the Google Play Store. Analyzing a unique panel dataset of 13,599 most popular apps for 24 months, we find that apps that traditionally over-seek permissions (i.e., seek more permissions than those required for the app's functionality) strategically delay upgrading. We find that such upgrade delays are more likely when the permissions sought are non-essential for the app's operations. More importantly, we find that such strategic delaying of upgrades come at a cost to apps in terms of marketplace outcomes. We discuss the implications of our findings for app providers as well as platform operators.

Keywords: Information Privacy, strategic behavior, mobile platform, software upgrade

* This version of the working paper has been prepared for the 2019 Workshop on the Economics of Information Security (WEIS), 3-4 June 2019, Boston, MA, USA. Please contact the authors for the latest version of this paper. The authors wish to thank the Ed Snider Center for Enterprise and Markets at the University of Maryland for generous research support.

1. Introduction and Research Questions

The past decade has witnessed a marked rise in consumer sensitivity to online information collection and privacy. About 86% of internet users have taken steps to avoid surveillance by organizations during their online browsing sessions (Rainie et al. 2013). Recent cases of massive breach in privacy such as the Cambridge Analytica scandal (Cadwalladr and Graham-Harrison 2018) have only amplified such user concerns. With the rapid proliferation of mobile devices, such concerns have naturally extended to mobile devices. Smartphone mobile apps have traditionally obtained blanket permissions from users to access their sensitive information when they download the apps. IT security researchers have shown that over a third of apps seek more permissions than needed, which increases the risks of data misuse (Felt et al. 2011). Users have increasingly become sensitive to such practices and have been proactively taking measures to protect their privacy ranging from abandoning apps to abandoning the platform altogether (Pingitore et al. 2017).

In an attempt to respond to these concerns and provide users with better control over their mobile footprint, mobile platforms such as Android have released upgrades that provide consumers with fine-grained control over their information. In this study, we examine one such upgrade – the release of Android's version 6.0 in late 2015. Android 6.0 restricts apps' ability to seek blanket permissions to access sensitive user information at download, and instead requires them to seek standalone (à la carte) permissions during run-time. In earlier versions of Android, upon downloading the app, consumers automatically agreed to provide all information listed by the app such as access to user information (contacts, phone memory, phone log, etc.) or the user's hardware (camera, GPS, etc.). Android 6.0 allowed users to download apps without granting any permissions and then required the apps to seek standalone permissions from users during run-time (see Figure 1). Such a change in Android's privacy policy provided users with a choice to use the "watered-down"[1] version of the app by providing those permissions "as you go".

In earlier cases where platforms implemented such changes to their security and privacy policies such as in Apple iOS in 2012 and Microsoft's User Access Control for the OS kernel in Windows Vista OS in 2007, platforms used their market power to force all apps to adhere to such policy changes to be eligible to run in their latest version. Due to issues with hardware fragmentation, however, Android gave mobile apps a window of three years to upgrade, instead

[1] Users can grant partial permissions but still run the app, to the extent that such permissions are not required for the features.

of requiring them to upgrade immediately to the latest version. An app's choice of when to upgrade (anytime between 2015 and 2018) meant that only those apps that upgraded to the latest version were required to adhere to the new privacy policy. Apps could choose not to upgrade immediately by continuing to use the Android Software Development Kit (SDK) from an older version. These apps would continue to seek permissions at download (instead of at runtime) even when such apps ran on phones that were upgraded to the latest version. Such apps could also release feature updates or fix bugs using older SDKs. In other words, apps that delay upgrade would be "fully functional," as earlier. However, these apps that stayed with the older version of Android would forgo access to latest platform features, optimizations, and support.

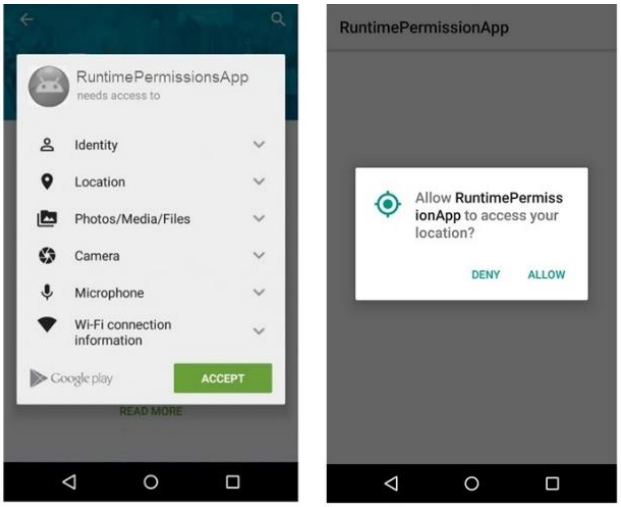


Figure 1: Change in permission-seeking from download to run-time (Photo Courtesy: GeneXus Community)

Given the tradeoff between upgrading early to benefit from the platform's latest features and staying with the older version to retain blanket access to user information, our study examines the choices made by different types of apps and their consequent outcomes. Specifically, we ask and answer the following questions: (i) *What are the characteristics of apps that delay upgrading to seeking run-time permissions?*, (ii) *Does the essentiality of permissions to an app's working impact its decision to delay upgrading to seeking run-time permissions?*, (iii) *How does upgrading to seeking run-time permissions affect the outcome for apps?*, (iv) *How does the timing of the upgrade to seeking run-time permissions affect the outcome for apps?* The first two questions seek to understand whether apps behave strategically in their upgrade decisions, while the last two questions identify the impacts of upgrade delays.

2. Relevant Research and Research Gaps

Our research draws from and contributes to three distinct streams of literature: responses to privacy policies, buyer privacy behavior and early movers.

2.1 Users' responses to Privacy Policies

Users' responses (or lack thereof) to privacy concerns in an online market have been extensively studied in the literature (see (Bélanger and Crossler 2011) for a survey of privacy research in IS). Consumers' concerns about information privacy may stem from their limited control on how their personal information may be used by the online seller (Acquisti and Grossklags 2005b; Dinev and Hart 2006; Pavlou et al. 2007). Such concerns are quite valid, especially when the secondary usage of their private information may come at a cost to them in their future transactions (Varian 2009). Concerns about asymmetric information also impact their risk perceptions about sellers that may use/distribute their private information without their knowledge. Research on Trust-Risk Frameworks have shown that privacy concerns can impact consumer beliefs of trust on, and risk from, the seller (Chellappa and Sin 2005; Malhotra et al. 2004; Pavlou et al. 2007).

Under such situations, users have expressed preference to take control over their information. For example, Phelps et al. (2000) surveyed 556 respondents and found that a significant percentage of respondents (84%) want to have more control over the use of personal data. Similar findings have been reported by Hoffman et al. (1999), where over 87% of consumers believe they should have complete control over their personal information captured by online portals. Acquisti and Grossklags (2003) have also documented a range of surveys where consumers have sought to gain control over secondary use of their sensitive information. In some cases, consumers have also walked out of transactions with sellers or have forgone potentially beneficial services instead of relinquishing control over their personal information (Pingitore et al. 2017). These studies have examined buyers' concerns and responses independent of sellers' actions.

Most papers in this stream have examined consumers' reactions to sellers' endogenous choice of adopting privacy practices. Online users are found to respond positively on websites that have guaranteed fair information practices to protect users' individual privacy. Users have also been found to reduce sharing concerns and risk perceptions on those websites (Culnan and Armstrong 1999; Dinev and Hart 2006; Hui et al. 2007). Consumers are also willing to pay a premium while buying from such sellers (Tsai et al. 2011). Nowak and Phelps (1995) have shown that users exhibit

fewer concerns about data collection when websites allow consumers to opt-in for data collection by explicitly giving permissions. When sellers invest efforts to be open about their information collection practices, consumers respond positively to personalized advertisements that use their data (Aguirre et al. 2015), or are more willing to make use of personal services that uses personal data (Chellappa and Sin 2005). When sellers demonstrate that the information gathered is relevant to transactions, consumers are more willing to provide such information (Zimmer et al. 2010). Tucker (2014) finds that consumers' clicks on personalized ads doubled after consumers were given control over their information. In these cases, consumers respond positively to sellers taking actions to win the buyers' trust over their data collection.

While there is a substantial body of research on users' responses to privacy changes, there are hardly any studies that have examined sellers' responses to exogenous changes in privacy policies. Recent studies have investigated firms' responses to adhering to the European Union's GDPR (General Data Protection Regulation) compliance (Garber 2018; Gradwohl 2018) but none have studied sellers' responses in the context of online or mobile platforms.

2.2 Users' Privacy Preferences

Despite an increase in the incidences of stated privacy concerns by buyers, studies have consistently documented a paradox in internalizing privacy choices. Researchers have observed that users that express concerns about privacy do not act to address such concerns when they have a choice (Acquisti and Grossklags 2003; Acquisti and Grossklags 2005b; Hann et al. 2007; Spiekermann et al. 2001). For example, in an online setting, users who are otherwise concerned with their privacy do not invest time to go through all these settings and/or read privacy policies (Hoofnagle et al. 2010; Jensen and Potts 2004). Furthermore, multiple studies have found that users that express their desire to protect their privacy are willing to exchange their sensitive information (such as disclosing their social security number) for short-term benefits (Acquisti and Grossklags 2005a; Hui et al. 2007). Further, despite stating their desire to pay a premium for privacy (Tsai et al 2011), users have demonstrated a lack of willingness to pay for privacy when such an option is made available to them (Brunk 2002; Rose 2005).

Such inconsistencies may be explained by the trade-off between the cost incurred by the users to protect the information versus the benefits derived from disclosing the information. Users that exhibit such "privacy paradox" seem to choose between potential losses from unauthorized use of private information and potential gains from a transaction that seeks private information (Acquisti

and Grossklags 2003). The inability to evaluate all the payoffs from protecting their privacy versus disclosing it in a transaction further adds to the variability in responses that each user displays to different types of sensitive information. For example, it may be extremely difficult for an Android user to evaluate why the most popular Battery Optimizer app would seek permission to access the downloader's contact list and the potential ways in which that app could make use of their contacts. A combination of cost-benefit calculus and bounded rationality dictates how user respond despite their stated preference of protecting their privacy at all times. Despite a large volume of existing literature on users' privacy behavior, not many studies have examined the changes in users' privacy behavior when the choice of privacy is made less costly. It is also unclear whether users have a fixed privacy preference or whether they value various sensitive information differently.

2.3 Temporal Choice of Response

In case of adherence to policy enactments, a key choice that firms make while responding to such policies is the timing of adherence. Literature on early movers finds that early movers enjoy advantages and that the market share of entrants can be causally linked to their time of entry into the market (Kalyanaram and Urban 1992; Mitchell 1991). Timing also determines how long the advantages of early movers can be sustained (Kerin et al. 1992). Social identity theories suggest that being early to the market is advantageous to some firms as long as such firms stay in the market memory as "originals" (Barnett et al. 2012). Almost all studies emphasize that early movers obtain and capitalize on asymmetries (Lieberman and Montgomery 1988). For example, studies have shown that only strategic pioneers gain advantages from early movement (Carow et al. 2004). In other words, a majority of studies on sellers' temporal choices study the positive outcomes of early movement choices.

A key difference exists between early movers to capitalize on rare information/asset and the first responders in an exogenous policy enactment setup. Most market entry decisions are often made to maximize profits and gain market share, while the timing of policy adoption decisions are often made to minimize potential losses due to policy enactments. There are hardly any studies that have examined firms' temporal choices under a policy enactment setup. Recent studies that examine firms' responses to governmental privacy policies have shown that well-informed market players may be able to utilize the timing of movement as a signal of trustworthiness. Garber (2018) comments that businesses that are early adopters of the European Union's GDPR (General Data

Protection Regulation) compliance may be able to use this as a differentiator from their competing firms and win consumer confidence. Gradwohl (2018) finds that, despite GDPR compliance being a requirement for only those firms doing business in Europe, North American firms that adopt such compliances may signal that they follow global best practices. But we are unaware of such investigations on an online platform setup. In essence, literature survey on early movers point towards a lack of rigorous investigations studying the timing of policy adherence on medium- to long-term revenue and reputational implications on platform.

2.4 Research Questions

We find several gaps in literature that reference understanding the phenomenon of information privacy and policy adoption on platforms. Why do some of the apps delay their upgrades while others upgrade on time? We are unaware of studies that examine whether delaying upgrades are demand driven, cost related or due to strategic behavior by apps. Hence our first research question asks and answers, *"What are the characteristics of apps that delay upgrading to seeking run-time permissions?"*. We consider multitudes of potential reasons for delay and employ econometric methods to identify the real reasons and to rule out alternative explanations. An app's decision to delay upgrade may be due to cost related to upgrading, the lack of demand for an upgrade from the existing users or due to strategic behavior by apps. The first reason why an app might delay upgrading would be the cost associated with such upgrades. Upgrading to a latest version requires that developers of the apps understand changes in the SDK, understand new features and modify app flow (if needed) based on best practices recommended by the latest SDK. We account for such variations by controlling for heterogeneity in developer characteristics (such as number of apps by each developer in a month) in models. Next, our analyses are run on most downloaded apps, hence this could further address cost-side concern as one would expect developers of popular apps to have enough bandwidth to upgrade and keep their apps relevant and current. Finally, including app characteristics such as app size and app users' activity (i.e., count of reviews left by app users) also controls for cost reasons in the supply side.

Another major reason why apps might delay upgrading may stem from user demand. One might argue that apps may have lesser incentive to upgrade if they cater to a majority of consumers that will not upgrade their phones to the latest Android version. While such a scenario (existing consumers not upgrading) may occur, there are good reasons why those apps should still upgrade.

Android maintains a quality score for listed apps in App Store[2] and explicitly states that quality score, and therefore the app's Play Store ranking, may be affected by delaying upgrade. A better Play Store rank aids in better discovery in the marketplace when new users search for a specific utility or when they observe "similar apps" list. Hence, any app that seeks to attract newer users would be better off upgrading. Furthermore, despite upgrading, apps would still seek download-time permissions to sensitive information from consumers who do not upgrade. Therefore, platform's design decisions help rule out demand-related reasons.

Another explanation for why certain apps may have no incentives to upgrade is when the latest platform version provides new features that do not benefit them or their consumers. To state alternatively, only apps that find new platform features more appealing are more likely to upgrade. A survey of all changes that the platform provided during the upgrade in late 2015 informs us that all new features and improvements would impact all apps equally (see Appendix Table A3). For example, Android made improvements to notification services, which would equally improve the consumer communication efficiency of all apps. Thus, choosing not to upgrade based on lack of benefits from upgrading does not seem to be the driving reason.

By ruling out demand side reasons and cost reasons in the supply side, we are left with the only dominant explanation: apps strategically delay upgrading to the latest version. Some apps build business models around utilizing a consumer's personal information in ways that consumers may not agree with (Cimpanu 2018). Apps, especially free apps, often monetize by displaying advertisements through advertisement networks. Apps choose from hundreds of advertisement networks by embedding advertisement libraries. Some advertisement libraries are known to follow non-standard policies such as uploading the consumer's sensitive information to remote servers or by triggering code from remote servers on the apps (Cimpanu 2018; Grace et al. 2012). Such practices pose serious threats to the security and privacy of app users. Such reckless practices by advertisement networks are not completely without the apps' knowledge (Grace et al. 2012). We argue that these apps also have incentives to over-seek permissions, thereby exposing platform users to security vulnerabilities. Hence, when platforms enact policies to restrict blanket permissions, such apps would choose to hold off upgrading to this version for as long as possible. This line of reasoning provides us with an easy and effective falsification test: investigating the upgrade propensity of the same apps to an earlier platform version where apps still sought blanket

[2] https://developer.android.com/docs/quality-guidelines/core-app-quality

permissions at download. Intuitively, apps that over-seek permissions should not potentially worry about upgrading to versions where the platform does not restrict blanket permissions. Our research setup provides for such a falsification test.

In our second research question, we test our conjecture about strategic delaying behavior of apps. Addressing this question helps address the gap in literature about consumers' privacy preferences under changing cost of privacy choice. If apps believe that consumers would indeed change their privacy behavior and exercise their control over private information, apps that seek more non-essential permissions should be delaying upgrade. On the other hand, if apps believe that consumer privacy behavior will not change, seeking non-essential permissions should not affect apps' decision to delay upgrade more than seeking essential permissions. Hence, we ask, *"Does the essentiality of permissions to an app's working impact its decision to delay upgrading to seeking run-time permissions?".* In the next section, we outline the methodology that we devise to determine the essentiality of each permission sought by the apps in our dataset.

In our third research question, we further address the research gap on consumers' privacy preferences by studying the impact of upgrading on apps' outcomes. The outcomes that we are interested in, are the number of essential and non-essential permissions that apps seek and the change in the rating in Play Store. These are the two measures that are directly linked to existing app users and their reactions. If the change in choice structure (from download-time to run-time) is indeed meaningful, consumers would not only reject unnecessary permissions but also start to complain/penalize for seeking unnecessary permissions. So, our third research question asks, *"How does upgrading to seeking run-time permissions affect the outcome for apps?".*

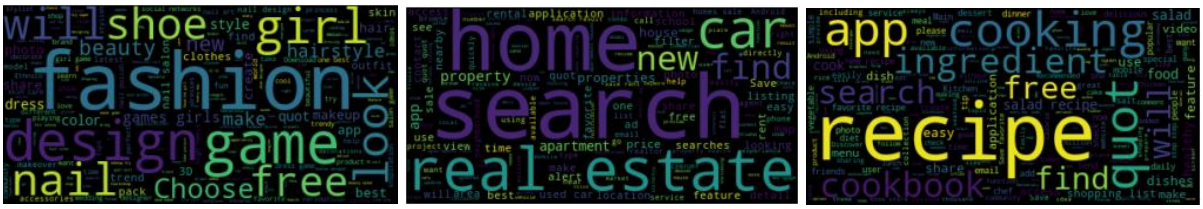
Our final research question pertains to the timing of mobile apps' decisions to upgrade to Android 6.0. Literature survey points to a lack of systematic study that measure outcomes for sellers that move early/late in a policy enactment setup. To this end, it is unclear whether being an early-mover in a policy adoption scenario is advantageous at all. One of the likely outcomes is that early movers among apps mostly face permission rejections from consumers who have just received fine-grained control over their sensitive information. The same consumer may become habituated to granting permissions by the time late-movers upgrade. An alternative outcome is that early movers among apps easily obtain all permissions because of consumers' initial fears of app malfunctioning if they deny permissions. Such consumers would slowly be habituated to rejecting permissions by the time later-movers upgrade. Hence, we ask, "*How does the timing of the upgrade*

to seeking run-time permissions affect the outcome for apps?". Specifically, we investigate how delaying an upgrade impacts seeking essential and non-essential permissions as well as an app's rating on Play Store. We describe our research context and data in the next section to demonstrate how our setup is ideal to address all the above-mentioned questions.

3. Research Context and Data

We examine the above research questions on Android, the world's largest smartphone platform with over 80% global market-share in mobile operating systems (Statista 2018). Apps need to follow a standard procedure to be listed in the Play Store, Android's official app marketplace. We assemble a unique panel data of apps between April 2016[3] and March 2018 to examine the voluntary decisions of apps to upgrade to Android 6.0 (and above) over the two years. We compile the dataset by installing around 13,600 top downloaded Android apps on emulators (over 1 Tera Bytes) and updating these apps on a monthly basis. We specifically develop an android app that scans the Android App Database of the host emulator and determine the Android version of all installed apps, along with a list of permissions sought by the apps from users. Installing apps is the only way to extract the Android version number of that app (targetAPI) since an installed file carries the signature (SDK version) of the targetAPI. We update all installed apps on a monthly basis to compile a 24-months panel. To distinguish permissions that are essential for app operations from permissions that aren't essential, we devise a methodology using a neural-network-based word-embedding technique, i.e., skip-gram Word2Vec technique (trained over 100 billion words from Google News Dataset) to sub-categorize apps such that each sub-group has apps that have similar utility (see Figure 2). Appendix B details the procedure followed above. Next, based on the methodology in Sarma et al. (2012), we statistically determine the essentiality of each permission sought by each of the apps. Specifically, we code those permissions requested by more than 75% of the apps in a sub-category as essential permissions. To build the intuition, let us consider navigation apps: permission to access GPS would be essential for any navigation app to function, hence at least 75% of such apps would seek access to GPS. On the contrary, access to the phone log does not seem essential to navigation apps, hence we would expect less than 75% of the apps to seek permissions to access the phone log.

[3] To be consistent with Technology adoption theories (Mahajan and Muller 1998), our panel starts when innovators (top 2.5% of adopters) and some of early adopters (2.5%) switch to Android 6.0. Android 6.0 installation base crossed 5% in April 2016.



2A: Fashion related 2B: Buy and Sell information 2C: Cooking related

Figure 2: Distribution of words in sample sub-categories from Android's category called "Lifestyle"

Finally, we also collect details of over 2 million Android apps from the Android Play Store website. We collect information such as app description, rating, download count bucket, count of reviews, categories, date of last update, number of screenshots uploaded, file size and developer ID of all apps. We collect this data for 24 months. Merging this dataset with the earlier dataset gives us an unbalanced panel of 13,599 apps for 24 months (April 2016 to March 2018), resulting in 278,955 app-month observations.

3.1 Variable Definitions

Table 1 provides a detailed description of the variables used in the major analyses. Appendix Table A1 provides the summary statistics and correlation matrix of the variables.

Variables of Interest: Our first three variables of interest are defined as the total number of dangerous[4] permissions sought, number of essential dangerous permissions sought, and number of non-essential permissions sought. In the following section, we elaborate on our methodology of coding essential and non-essential permissions. Briefly stated, these are dangerous permissions that are statistically determined as being essential (or not) for an app's functioning. Since the total normal permission count sought by any app highly correlates with the total dangerous permission count, we compute the ratio of dangerous permissions sought (three different ratios) to normal permissions sought and use them as dependent variables in logistic regression analyses. Variable $upgrade_{it}$ is coded as 1 if the given app performs an upgrade to the target version (6.0) or above. Finally, *rating* captures the overall rating that the app has received at the end of month *t*.

Independent Variables: Our main independent variables are app characteristics that represent the app's demand, its appearance on the Play Store and the developer's activities on the platform. Specifically, we capture the total count of ratings that are provided by all users, the version of the platform that the app currently targets, number of days elapsed since the developer has pushed an

[4] https://developer.android.com/guide/topics/permissions/overview.html#normal-dangerous

update (features, bug fixes or upgrades) to the platform, total number of apps that the developer has published on Play Store in any given month, count of screenshots that the app has in its Play Store page, and the file size of apps. Playstore mandates that a minimum of 2 and a maximum of 8 screenshots[5] per target category (phone, tablet, TV and Wear OS) may be uploaded on Play Store. If the developer has uploaded a video, we count them in the same variable.

Table 1. Variable Explanation

Variable	Description
Key Variables	
$dangerous\_permissions_{it}$	Count of dangerous permissions sought by app *i* in month *t*
$essential\_dangerous\_permissions_{it}$	Count of essential dangerous permissions sought by app *i* in month *t*
$nonessential\_dangerous\_permissions_{it}$	Count of non-essential dangerous permissions sought by app *i* in month *t*
$normal\_permissions_{it}$	Count of normal permissions sought by app *i* in month *t*
$upgrade_{it}$	If the app *i* upgrades to target new API (of 23 and above) in the month *t.*
$rating_{it}$	The accumulative review rating of app *i* by the end of month *t.*
App Characteristics	
$rating\_count_{it}$	Total number of ratings given by raters for app *i* by the end of month *t.*
$dayssinceupdate_{it}$	Number of days since app *i* has been updated by the end of month *t*.
$screenshots_{it}$	Number of screenshots (including optional video) uploaded by app *i* by the end of month *t*
$developer\_appcount_{it}$	Number of apps in playstore by the developer of app i of month *t*
$filesize_{it}$	File size (in MB) of app *i* by the end of month *t*
Control Variables	
$category\_group_{i}$	App category group that app *i* belongs to.
$download\_bucket_{it}$	Download bucket (group) that app *i* belongs to at the end of month *t.*
$month\_dummy_{t}$	A vector of dummies that represent if month *t* is April-2016, May-2017, etc.

Note: Missing values in file sizes (such as "varies with devices" or months when file sizes were not displayed in Play Store) are handled by replacing them with (a) values carried forward from previous months or (b) mean value of file-sizes of all sub-category for the month in the order listed.

Control Variables: We use a set of control variables to account for unobserved app characteristics. We classify the app categories into 6 groups based on prior research (for example, (Ghose and Han 2014)) and on evolution in the category's overlaps. We classify the apps into 6 category groups: Online Content Consumption (Media and Entertainment), Learn and Explore, Personal (Social and Lifestyle), Mobile Specific Utilities (services that exist solely for the use of smart phones), Mobile Access Utilities (mobile apps provided by offline or internet utility firms), and Games. Appendix Table A2 provides brief explanation of the category groups. The variable $download\_bucket_{it}$ captures the range of app downloads that Play Store publishes.[6] Finally, $month\_dummy_{t}$ takes a single value for observations in a month-year pair.

[5] https://support.google.com/googleplay/android-developer/answer/1078870?hl=en

[6] The download buckets that Android publishes are: 1-5 downloads, 5-10, 10-50, 50-100, 100-500, 500-1k, 1k-5k, 5k-10k, 10k-50k, 50k-100k, 100k-500k, 500k-1M, 1M-5M, 5M-10M, 10M-50M, 50M-100M, 100M-500M, 500M-1B and 1B-5B downloads.

4. Analyses and Results

4.1 What are the characteristics of apps that delay upgrading to seeking run-time permissions?

We employ logistic regression models to address our first question: the likelihood that the app delays upgrading to the latest version of Android.

$$\textbf{Logit(upgrade}_{it}\textbf{)} = c_0 + c_1 * dangerous\_permissions\_ratio_{it} + c_2 * rating\_count_{it} + c_3 * rating_{it} + c_4 * rating^2_{it} + c_5 * dayssinceupdate_{it} + c_6 * screenshots_{it} + c_7 * filesize_{it} + c_8 * developer\_appcount_i + \Gamma * download\_bucket_{it} + Z * month\_dummy_t + \Lambda * category\_group_i + \varepsilon_{it} \quad (1)$$

where the subscript *i* indexes apps and subscript *t* indexes months. Using a logit model allows us to investigate which of the covariates increases or decreases the propensity of apps to delay upgrade (Gopal and Gao 2009; King et al. 2005). A positive (negative) sign for a covariate informs us that an increase in the value of the covariate decreases (increases) the propensity of the app to delay upgrade. In our investigation, we consider multitudes of demand-side and supply-side cost reasons and employ econometric methods to rule out alternative explanations. For example, we replace the independent variable $dangerous\_permissions\_ratio_{it}$ with a matrix of individual dangerous permissions to study which particular set of permissions predict a delay in upgrading. Controlling for the developer's total count of apps each month and restricting our analyses to popular apps (with over 100,000 downloads) addresses concerns regarding cost reasons for delaying the upgrade. Android's platform policies help address the demand side reasons for delaying. For example, Android's announcement of a Quality Score based on which non-upgraded apps are penalized, as well as Android's design choice that allow apps (upgraded or otherwise) to seek download-time permissions from app users that haven't upgraded their phones, would ensure that apps do not have demand-side reasons for delaying an upgrade. Finally, we also carefully examine Android 6.0 update to check if there were any changes (apart from run-time permission changes) or any new feature additions that may impact apps' decision to delay upgrade. Appendix Table A3 carries the list of changes and new features that Android 6.0 introduced. We can see that none of the changes made to existing features or addition of new features in Android 6.0 may negatively affect apps' functionality. Hence, we are confident that there are no other upgrade reasons that may impact apps' strategies of (delaying) upgrading.

Outcomes of model (1) have been presented in Table 2. We find that an app's likelihood of delaying upgrade to the latest platform version increases with an increase in the ratio of dangerous

permissions sought by the apps, indicating that apps prefer to retain control over access to the users' private information. The point estimate of $rating_{it}$ is positive and significant while that of the quadratic form of $rating_{it}$ is negative and significant indicating that apps with very high or very low ratings are more likely to delay upgrade[7]. Apps that have higher number of people providing reviews, an indicator of deeper engagement with consumers, have a higher propensity to delay the upgrade. The number of days since the latest update, an indicator of how actively the app has been maintained, negatively impacts the propensity to upgrade. App file-size and number of screenshots uploaded to the Play Store negatively impacts the propensity to delay upgrade. File-size and count of screenshots of an app may indicate its sophisticated nature, while the frequency of update may be correlated with the quality of the app (Ghose and Han 2014). Finally, the positive and significant estimate of $developer\_appcount_{it}$ suggests that apps from a relatively smaller developer are more likely to delay upgrade.

Table 2. Analysis of Upgrading to Latest Version of Android

Dependent variable	(1) Logit - $upgrade_{it}$	
$dangerous\_permissions\_ratio_{it}$	-0.012***	(0.001)
$rating\_count_{it}$	-0.000***	(0.000)
$rating_{it}$	1.902***	(0.235)
$rating_{it} * rating_i$	-0.279***	(0.030)
$games_i$	-0.158***	(0.023)
$personal\_apps_i$	-0.152***	(0.028)
$utility\_mobilespecific_i$	-0.247***	(0.025)
$utility\_mobileaccess_i$	0.241***	(0.031)
$learn\_explore_i$	0.062*	(0.028)
$category_i$=content_consumption$_i$ (baseline)		
$below\_1million_i$	-0.624***	(0.017)
$5million\_10million_i$	0.335***	(0.022)
$10million\_50million_i$	0.280***	(0.022)
$above\_50million_i$	0.202***	(0.049)
$download_i$=1million_5million$_i$ (baseline)		
$dayssinceupdate_{it}$	-0.005***	(0.000)
$screenshot_{it}$	0.013***	(0.001)
$filesize_{it}$	0.005***	(0.000)
$developer\_appcount_{it}$	0.000***	(0.000)
Constant	-2.276***	(0.455)
Month Dummies	YES	
# of Obs.	178693	
Pseudo R-Squared	0.195	

*** p<0.001, ** p<0.01, * p<0.05; Robust standard errors in parentheses.

Shifting our focus to app characteristics, we find that, compared to the "online content consumption" apps (such as on-demand video services), mobile specific utility apps (such as the

[7] We employ the Stata command 'utest' for this test. The inverted U test suggests that propensity to upgrade peaks at the rating of 3.41 and falls above this rating.

messaging, navigation, food delivery apps), game apps and personal apps (such as social networking, dating apps) are more likely to delay upgrade to the new version of the platform. This finding is consistent with reports that have found that apps of these categories are more likely to over-seek permissions (Maheshwari 2017; Stamm 2018). Finally, we find that apps that have less than 1 million downloads are significantly less likely to upgrade to the newer version while apps with more than 5 million downloads are more likely to upgrade to the latest version.

In further investigating individual dangerous permissions, we find that only apps that seek permissions that may be perceived as sneaky, such as (passively) accessing the users' information (call logs, network information, phone memory, SMS etc.) or their hardware (microphone, fine GPS location, etc.) are more likely to delay upgrade to Android 6.0. We present the analyses described above in Appendix Table A4. Our research setup, choice of time period, use of relevant controls, and Android's incentive structure to upgrade help us rule out alternative explanations (demand side reasons and cost reasons) and convincingly point to the strategic nature of apps in delaying upgrading to the latest version. We run a falsification test that examines the app's likelihood of upgrading to an earlier version of Android (Android 5.0 and 5.1) and find that apps that delay their upgrade to Android 6.0 did not delay when upgrading to an earlier version of Android that did not have such restrictions in accessing user information (Appendix Table A5). Upon carefully investigating the changes/new features that were introduced in Android 5.0 and 5.1[8] (2 versions of Android Lollypop), we find that Android platform had revoked restrictions (enforced in an earlier version) that prevented third party apps from accessing USB files that did not belong to each app's own home directory. Android also introduced screen capturing and sharing function, programmatically access camera devices (new camera API), added feature to access app usage history and introduced feature to allow apps to access battery usage log [9]. These features should incentivize apps that seek more (potentially non-essential) dangerous permissions to upgrade to targetAPI 21/22 but not targetAPI 23.

4.2 Does the essentiality of permissions to an app's working impact app's decision to delay upgrading to seeking run-time permissions?

The second research question pertains to consumers' permission-granting preferences. Extant literature treats such preferences as static, i.e., consumers treat all sensitive information in the same

[8] https://developer.android.com/about/versions/android-5.0

way and would have a similar granting action in all contexts. To address this research question, we replace the variable *dangerous_permissions_ratio*$_{it}$ in the model with variables that measure the essential and non-essential permissions respectively, to investigate whether the essentiality of a permission impacts the app's decision to upgrade:

$$\textbf{\textit{Logit(upgrade}}_{it}\textbf{\textit{)}} = c_0 + c_1 * essential\_dangerous\_permissions\_ratio_{it} + c_2 * rating\_count_{it} + c_3 * rating_{it} + c_4 * rating^2_{it} + c_5 * dayssinceupdate_{it} + c_6 * screenshots_{it} + c_7 * filesize_{it} + c_8 * developer\_appcount_{it} + \Gamma * download\_bucket_{it} + Z * month\_dummy_t + \Lambda * category\_group_i + \varepsilon_{it} \quad (2)$$

$$\textbf{\textit{Logit(upgrade}}_{it}\textbf{\textit{)}} = c_0 + c_1 * nonessential\_dangerous\_permissions\_ratio_{it} + c_2 * rating\_count_{it} + c_3 * rating_{it} + c_4 * rating^2_{it} + c_5 * dayssinceupdate_{it} + c_6 * screenshots_{it} + c_7 * filesize_{it} + c_8 * developer\_appcount_{it} + \Gamma * download\_bucket_{it} + Z * month\_dummy_t + \Lambda * category\_group_i + \varepsilon_{it} \quad (3)$$

where the subscript *i* indexes apps and the subscript *t* indexes months. If apps expect users to treat all dangerous permissions in a similar way, we should ideally see negative and significant values for the c_1 coefficient of both models above. A different sign and/or statistical significance of these two coefficients would mean that apps' propensity to delay upgrade would be different depending on whether the dangerous permissions sought are essential or non-essential.

Table 3. Analysis of Upgrading to Latest Version of Android

Dependent variable	(1) Logit - *upgrade*$_{it}$		(2) Logit - *upgrade*$_{it}$	
essential_dangerous_permissions_ratio$_{it}$	0.001	(0.001)		
nonessential_dangerous_permissions_ratio$_{it}$			-0.015***	(0.001)
rating_count$_{it}$	-0.000***	(0.000)	-0.000***	(0.000)
rating$_{it}$	1.792***	(0.233)	1.836***	(0.235)
rating$_{it}$ * *rating*$_i$	-0.264***	(0.030)	-0.270***	(0.030)
games$_i$	-0.100***	(0.023)	-0.147***	(0.023)
personal_apps$_i$	-0.141***	(0.028)	-0.155***	(0.028)
utility_mobilespecific$_i$	-0.272***	(0.025)	-0.261***	(0.025)
utility_mobileaccess$_i$	0.171***	(0.031)	0.220***	(0.031)
learn_explore$_i$	0.085**	(0.028)	0.059*	(0.028)
category$_i$*=content_consumption*$_i$ *(baseline)*				
below_1million$_i$	-0.615***	(0.017)	-0.623***	(0.017)
5million_10million$_i$	0.326***	(0.022)	0.339***	(0.022)
10million_50million$_i$	0.274***	(0.022)	0.287***	(0.022)
above_50million$_i$	0.194***	(0.049)	0.217***	(0.049)
download$_i$*=1million_5million*$_i$ *(baseline)*				
dayssinceupdate$_{it}$	-0.005***	(0.000)	-0.005***	(0.000)
screenshot$_{it}$	0.013***	(0.001)	0.013***	(0.001)
filesize$_{it}$	0.005***	(0.000)	0.005***	(0.000)
developer_appcount$_{it}$	0.000***	(0.000)	0.000***	(0.000)
Constant	-2.418***	(0.453)	-2.164***	(0.455)
Month Dummies	YES		YES	
# of Obs.	178693		178693	
Pseudo R-Squared	0.192		0.196	

*** p<0.001, ** p<0.01, * p<0.05; Robust standard errors in parentheses.

We present the results of models (2) and (3) in Table 3. We find that the likelihood of upgrade decreases only for apps that seek more non-essential permissions (column 2, Table 3), while essential permissions do not affect such decisions (column 1, Table 3). This finding points to strategic delaying by apps that seek permissions that consumers may deem as unnecessary for the apps' operations.

4.3 How does upgrading to seeking run-time permissions affect the outcome for apps?

We argue that if choices over privacy can be exercised in a costless manner, consumers will act consistent to their expressed concerns over privacy. Such a change in user behavior can be inferred by observing how apps alter their information-seeking habits, i.e., both the essential and non-essential dangerous permissions. We also study how such upgrades impact its outcome in the Play Store, as measured by ratings. Since the treated apps upgrade in different months, i.e., in a staggered manner, we set the month that treated app upgrades to time 0 and adjust pre- and post-months in reverse chronological order and sequential order, respectively. Such a normalization of time is performed based on prior research (Autor 2003; Fang et al. 2014; Gao and Zhang 2016). We use Propensity Score Matching (PSM) to dynamically match apps that upgrade to seeking run-time permission in our panel (treated apps) with those apps that don't upgrade (control apps). This matching is followed by a Difference in Differences (DiD) technique (Meyer 1995) to establish causality of the relationship of the treatment on the treated. Specifically, techniques combining PSM to pre-process the dataset followed by DiD have been used in research where the treatment may be affected by selection bias (Liu and Lynch 2011; Mayya et al. 2017; Smith and Todd 2005). We apply the nearest one-to-one neighbor matching and without replacement of the control samples. To account for the difference in time periods when treatments are introduced (app A upgrades in June 2016, app B upgrades in August 2016), we follow a dynamic matching technique where we match the treated apps with control apps, one month before the month of treatment. The covariates used for matching include app characteristics such as the app rating, count of rating, file size, screenshots uploaded, and date since the latest update. Also, we explicitly ensured that treated and control apps come from the same download bucket and belong to the same category group. Panel A of the Appendix Table A6 shows us that these covariates were well-balanced after matching. The PSM procedure resulted in 2628 treated and 2628 control mobile apps in the final sample.

Alternatively, we also use Look-Ahead Propensity Score Matching (LA-PSM) to pre-process the dataset (Bapna and Umyarov 2015) to ascertain that the outcomes are not driven by the matching technique. The LA-PSM uses a Look-Ahead approach in determining control samples, in that control samples are drawn from a pool of apps that will eventually upgrade to the new version sometime in the future but not during the month under consideration. Such a matching technique would not just match the treated samples with control samples based on the time-variant and invariant covariates, but also on the intent to switch. The LA-PSM procedure resulted in 1434 treated and 1434 control mobile apps in the final sample. Panel A of Appendix table A7 shows covariate balancing.

$$\textbf{essential\_dangerous\_permissions}_{it} = \alpha_{0a} + \alpha_{1a}*rating\_count_{it-1} + \alpha_{2a}*rating_{it} + \alpha_{3a}*post\_upgrade_t + \alpha_{4a}*upgrade\_group_i*post\_upgrade_t + \alpha_{5a}*dayssinceupdate_{it} + \alpha_{6a}*filesize_{it} + \alpha_{7a}*screenshots_{it} + \alpha_{8a}*developer\_appcount_{it} + \alpha_{9a}*download\_bucket_i + \alpha_{10a}*month\_dummy_t + w_i + \varepsilon_{ait} \quad (2a)$$

$$\textbf{nonessential\_dangerous\_permissions}_{it} = \alpha_{0b} + \alpha_{1b}*rating\_count_{it-1} + \alpha_{2b}*rating_{it} + \alpha_{3b}*post\_upgrade_t + \alpha_{4b}*upgrade\_group_i*post\_upgrade_t + \alpha_{5b}*dayssinceupdate_{it} + \alpha_{6b}*filesize_{it} + \alpha_{7b}*screenshots_{it} + \alpha_{8b}*developer\_appcount_{it} + \alpha_{9b}*download\_bucket_i + \alpha_{10b}*month\_dummy_t + w_i + \varepsilon_{bit} \quad (2b)$$

$$\textbf{log(rating}_{it}\textbf{)} = \beta_0 + \beta_1*rating\_count_{it-1} + \beta_2*dangerous\_permissions\_ratio_{it-1} + \beta_3*post\_upgrade_t + \beta_4*upgrade\_group_i*post\_upgrade_t + \beta_5*dayssinceupdate_{it} + \beta_6*filesize_{it} + \beta_7*screenshots_{it} + \beta_8*developer\_appcount_{it} + \beta_9*download\_bucket_i + \beta_{10}*month\_dummy_t + w_i + \eta_{it} \quad (3)$$

In the above models, $upgrade\_group_i$ is a binary variable that carries a value of 1 for apps in the treated group and a value of 0 for apps in the control group. For each treated-control pair, the variable $post\_upgrade_t$ carries a value of 0 before the treated app's upgrade and 1 after the upgrade to the latest version. Coefficients α_{4a}, α_{4b} and β_4 in the above models provide us with the required DiD estimates. Parallel trends assumption states that the control and treated apps should have a similar time trend before the treatment. We follow the recommendations in literature (Autor 2003) where we estimate models similar to (2a), (2b), and (3), but by modeling time trends of the dependent variables in the models. From Panel B of Appendix Tables A6 and A7, we see that the Parallel Trends Assumption is not violated by our matching techniques.

We present the outcomes of models (2) and (3) in Table 4. From column (1), we find that upgrading to a new platform version does not impact the app's count of essential dangerous permissions sought, as seen by coefficient α_{4a} (i.e., the interaction term), which is insignificant. On

the other hand, from column (2), we find that the apps significantly reduce their non-essential dangerous permissions by 6.01% (i.e., 0.163/2.71) as seen by coefficient α_{4b}. Figure 3 visualizes the impact of upgrading on contextually essential and non-essential permissions. Taken together, these findings suggest that apps find it harder to seek only those permissions that are not essential for their working, after upgrading to seeking run-time permissions.

Similarly, we find that consumers welcome such upgrades positively by giving a higher rating, as seen by the positive and significant result of coefficient β_4 in column (3) of Table 4. This coefficient (0.0010238) means that, on an average, about 620 app users (i.e., 0.0010238 * 4.04 mean rating * 150079.3 mean raters) who have previously rated the app, increase their rating by 1 star after the app upgraded to the latest version of the platform. This finding suggests that app users positively acknowledge the app's decision to upgrade and provide users with access to newer platform features. We check the robustness of our findings employing an alternative model specification, i.e., Conditional Fixed Effect Poisson count models (unreported). We also find qualitatively similar results with LA-PSM, further adding support to our analysis (Appendix Table A8).

Table 4. Analysis of Effects of Upgrading to Latest Version - PSM

Dependent Variable	(1) PSM DID - essential_ *dangerous_permissions*$_{it}$		(2) PSM DID - nonessential_ *dangerous_permissions*$_{it}$		(3) PSM DID – log(rating$_{it}$)	
dangerous_permissions_ratio$_{it}$					0.000	(0.000)
rating_count$_{it-1}$	-0.000*	(0.000)	0.000***	(0.000)	0.000***	(0.000)
rating$_{it}$	0.160***	(0.039)	-0.295**	(0.100)		
post_switch$_t$	0.019***	(0.004)	-0.010	(0.008)	-0.001***	(0.000)
upgrade_group$_i$**post_upgrade*$_t$	-0.012	(0.006)	-0.163***	(0.013)	0.001***	(0.001)
dayssinceupdate$_{it}$	0.000***	(0.000)	0.000***	(0.000)	-0.000	(0.000)
screenshot$_{it}$	0.000	(0.000)	0.001***	(0.000)	0.000	(0.000)
filesize$_{it}$	0.011***	(0.003)	-0.012	(0.007)	0.000*	(0.000)
developer_appcount$_{it}$	-0.000***	(0.000)	0.004***	(0.000)	0.000***	(0.000)
Constant	0.428**	(0.166)	3.757***	(0.424)	1.395***	(0.001)
Month Dummies	YES		YES		YES	
App Fixed Effect	YES		YES		YES	
Download Bucket Dummies	YES		YES		YES	
# of Obs.	106065		106065		106065	
R-Squared	0.140		0.025		0.108	

*** p<0.001, ** p<0.01, * p<0.05
Heteroskedastic Robust Standard Errors in parenthesis

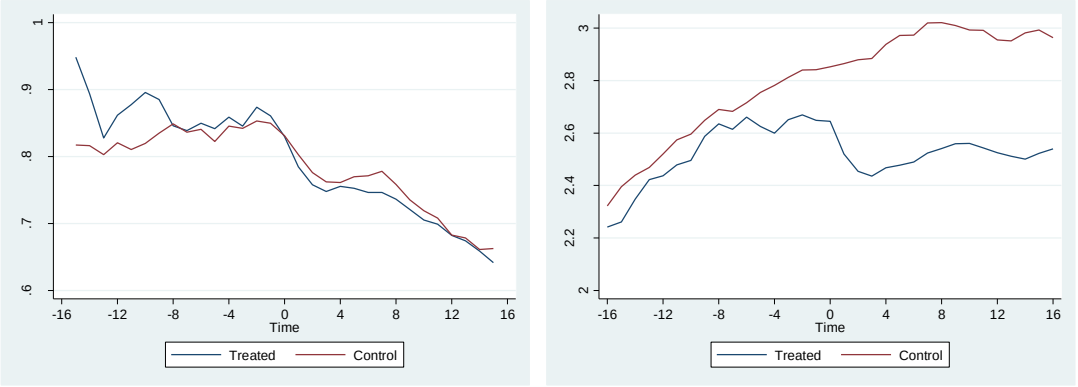


Figure 3: Line graph showing the parallel trends assumption for treated and control apps on essential and non-essential permissions sought for each Target API. Time 0 is the month of matching

4.4 How does timing of upgrade affect the outcomes of the apps?

Our final research question pertains to the timing of an upgrade in a policy adoption scenario. In our study, we distinguish early movers, i.e., apps that upgrade to Android 6.0 or higher before consumers start upgrading to Android 7.0 (1 year after the launch of Android 6.0), from late movers, and investigate how the timing of upgrade affects the outcome for apps. We employ a similar model as (2) and (3) with an additional interaction term. We include an interaction term between the dummy, representing late movers, and $post\_upgrade_t$ to identify the effect of delaying upgrading, over and above the effect of upgrading to the latest version. We present the outcome of this analysis in Table 5.

Table 5. Analysis of Effects of Delaying the Upgrading to Latest Version - PSM

Dependent Variable	(1) PSM DID - essential_ $dangerous\_permissions_{it}$		(2) PSM DID - nonessential_ $dangerous\_permissions_{it}$		(3) PSM DID – log(rating$_{it}$)	
$dangerous\_permissions\_ratio_{it}$					0.000	(0.000)
$rating\_count_{it}$	-0.000*	(0.000)	0.000***	(0.000)	0.000***	(0.000)
$rating_{it}$	0.160***	(0.039)	-0.299**	(0.100)		
$post\_switch_t$	0.018***	(0.004)	-0.007	(0.008)	-0.001***	(0.000)
$upgrade\_group_i$*$post\_upgrade_t$	-0.018	(0.012)	-0.104***	(0.023)	0.002***	(0.000)
$late\_upgrade\_group_i$*$post\_upgrade_t$	0.009	(0.012)	-0.080**	(0.025)	-0.001*	(0.000)
$dayssinceupdate_{it}$	0.000***	(0.000)	0.000***	(0.000)	-0.000	(0.000)
$screenshot_{it}$	0.000	(0.000)	0.001***	(0.000)	0.000	(0.000)
$filesize_{it}$	0.011***	(0.003)	-0.012	(0.007)	0.000*	(0.000)
$developer\_appcount_{it}$	-0.000***	(0.000)	0.004***	(0.000)	0.000***	(0.000)
Constant	0.426*	(0.166)	3.772***	(0.423)	1.395***	(0.001)
Month Dummies	YES		YES		YES	
App Fixed Effect	YES		YES		YES	
Download Bucket Dummies	YES		YES		YES	
# of Obs.	106065		106065		106065	
R-Squared	0.140		0.025		0.108	

*** p<0.001, ** p<0.01, * p<0.05
Heteroskedastic Robust Standard Errors in parenthesis

From columns (1) and (2) of Table 5, we see that the delaying of upgrade seems to affect contextually non-essential permissions (as seen from point estimate α_{5b}), while such a delay does not affect essential permissions. This result suggests that consumers indeed learn over time to reject non-essential permissions and the effect can be significantly felt by late upgraders among apps. From column (3), we find coefficient β_5 is negative and significant suggesting that delaying the upgrade has an adverse effect on the positive rating boost that apps receive post upgrade. In other words, delaying upgrade effectively diminishes the positive impact on rating that apps gain upon upgrading to the latest version of the platform. In essence, delaying upgrade to realize short-term profits (by retaining control over users' sensitive information) ends up affecting the app's outcomes on the platform. We find qualitatively similar results with LA-PSM technique, further providing strength to our analysis (Appendix Table A9).

5. Discussion and Conclusion

Our study is motivated by an increasing demand by consumers to take control over their sensitive information and a dearth in research examining the impact of platforms' mechanisms that transfer such controls to consumers. We investigate app's responses when Android, the largest smart phone ecosystem, released an upgrade that changed how consumers grant permissions to apps. We find that those apps that seek more dangerous permissions from consumers strategically delay upgrading to the latest version of Android. Within all dangerous permissions, propensity to delay upgrade increases when the apps seek permissions that are considered to be sneaky (i.e., running in the background). This finding is consistent with prior research which suggests that consumers' concerns of privacy stems from their uncertainty of how the information that are collected are used (Acquisti and Grossklags 2005a; Dinev and Hart 2006; Pavlou et al. 2007) . When users are more certain about how and when particular app collects sensitive information (for example, an app can make a phone call only when the consumer actively accesses such a feature from that app), they are more likely to be comfortable with the app and grant permissions. The likelihood falls when they are unsure when and how apps gather their information (for example, an app with an access to microphone can continuously monitor your conversations or other activities in the background). Similarly, in answering RQ2, we find that the propensity to delay upgrade increases among apps that seek irrelevant (to apps' working) permissions. This is consistent with prior research which suggests that consumers are more willing to provide their sensitive information if the information

gathered is relevant to transactions (Zimmer et al. 2010). Our finding in RQ2 points towards apps' belief that consumers will meaningfully exercise their control over privacy and grant only contextually relevant sensitive information to apps. Since apps that over-seek permissions usually build their revenue model around these sensitive permissions, they strategically delay upgrading to Android 6.0.

In answering RQ3, we find that apps that upgrade to Android 6.0 (or above) reduce seeking non-essential dangerous permissions while they continue to seek essential dangerous permissions. This finding is important because it suggests that shifting the control of private information towards consumers improves the information gathering behavior of apps. More specifically, this suggests that apps put an effort to reduce seeking unnecessary permissions that may raise alarms in users' mind. In investigating RQ4, we find that apps' ability to seek non-essential permissions fall further upon strategically delaying upgrade. This means that apps have a medium- to long-term price to pay if they delay upgrading, despite retaining control over users' information in short-time. Outcomes of RQ3 and RQ4 together help us further tease out the effect of the change in privacy policy. In RQ3, we found that consumers provide better rating when apps upgrade to seeking run-time permissions. Such an increase in rating may be attributed not just to privacy control improvements, but also additional feature improvements that are part of the Android 6.0 upgrade. However, a major difference between these two types of improvements is that privacy control improvements are applicable to all apps. Therefore, while app users don't have an expectation of what feature improvements they could get, they would definitely have expectations on privacy controls, especially once they see some apps providing them. Alternatively stated, while app users would always be happier if an app gives them feature improvements to features at any time, they are more likely to be unhappy if that app delays giving them privacy control. As expected, in RQ4, we find that such a gain in rating gets negated upon delaying the upgrade. Such a fall in rating due to delay strengthens our causal claim that apps pay a price specifically because of strategic delaying of upgrade.

Our study makes a number of significant contributions. Our study guides platform mechanism designs by showing that altering privacy policies to make choices less costly improves apps' information collection behaviors. Platforms' proactive measures to shift the control towards users may help address the privacy paradox that has been documented in privacy literature. By showing that even the popular apps systematically lower seeking non-essential dangerous permissions, we

are able to highlight the role of platform mechanisms under power asymmetry in safeguarding consumers' interest. We also contribute to the literature on privacy preferences by devising a methodology that statistically distinguishes contextually essential sensitive information from non-essential ones. By employing a neural-network based word embedding technique, we demonstrate that consumers have sophisticated and contextual preferences for information privacy. Consumers, despite expressing preferences to controlling sensitive information sharing, are willing to trade-off contextually essential sensitive information to derive utility from apps. Next, while extant research has investigated the buyers' perspective to information privacy, to the best of our knowledge, this paper is one of the first to investigate the sellers' responses to exogenous changes in information gathering practices. In doing so, we also add to the literature on early movers by showing that, even in case of exogenous policy enactment, early movers (among apps) stand to gain more. Our analyses show that upgrading to the latest version at the earliest is beneficial for an app's outcome in the marketplace, despite potentially disrupting its revenues in the short run. Alternatively, we also show that strategically delaying adhering to policies harm the sellers' outcomes on the platform. Finally, our research adds to the literature on temporal choices by suggesting that a fragmented platform such as Android should carefully design its upgrade window. Providing a longer or indefinite time-horizon may induce strategic behavior of delaying upgrades that may not be optimal for the platform, consumers or the apps themselves.

This line of research could be extended in many ways. A potential extension to our project would be to derive an importance measure for individual apps. Whether and how upgrade decisions vary with an app's relative importance to consumers (for example, a banking app vs. a game app) would provide additional insights. Another future extension is to study app interface designs, specifically the messaging design that may convey why granting certain permissions are necessary, thereby avoiding being rejected.

References

Acquisti, A., and Grossklags, J. 2003. "Losses, Gains, and Hyperbolic Discounting: An Experimental Approach to Information Security Attitudes and Behavior," *2nd Annual Workshop on Economics and Information Security*, pp. 1-27.

Acquisti, A., and Grossklags, J. 2005a. "Privacy and Rationality in Individual Decision Making," *IEEE Security & Privacy* (3:1), pp. 26-33.

Acquisti, A., and Grossklags, J. 2005b. "Uncertainty, Ambiguity and Privacy," *4th Annual Workshop on Economics and Information Security*.

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., and Wetzels, M. 2015. "Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness," *Journal of Retailing* (91:1), pp. 34-49.

Autor, D. H. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics* (21:1), pp. 1-42.

Bapna, R., and Umyarov, A. 2015. "Do Your Online Friends Make You Pay? A Randomized Field Experiment in an Online Music Social Network," *Management Science* (61:8), pp. 1902-1920.

Barnett, W. P., Feng, M., and Luo, X. 2012. "Social Identity, Market Memory, and First-Mover Advantage," *Industrial and Corporate Change* (22:3), pp. 585-615.

Bélanger, F., and Crossler, R. E. 2011. "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems," *MIS Quarterly* (35:4), pp. 1017-1042.

Brunk, B. 2002. "Understanding the Privacy Space," *First Monday* (7:10).

Cadwalladr, C., and Graham-Harrison, E. 2018. "Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach " in: *The Guardian*.

Carow, K., Heron, R., and Saxton, T. 2004. "Do Early Birds Get the Returns? An Empirical Investigation of Early-Mover Advantages in Acquisitions," *Strategic Management Journal* (25:6), pp. 563-585.

Chellappa, R. K., and Sin, R. G. 2005. "Personalization Versus Privacy: An Empirical Examination of the Online Consumer's Dilemma," *Information Technology and Management* (6:2-3), pp. 181-202.

Choe, E. K., Jung, J., Lee, B., and Fisher, K. 2013. "Nudging People Away from Privacy-Invasive Mobile Apps through Visual Framing," *IFIP Conference on Human-Computer Interaction*: Springer, pp. 74-91.

Cimpanu, C. 2018. "Google Restricts Which Android Apps Can Request Call Log and Sms Permissions." *ZDNet* Retrieved January 9, 2019, from https://www.zdnet.com/article/google-restricts-which-android-apps-can-request-call-log-and-sms-permissions/

Culnan, M. J., and Armstrong, P. K. 1999. "Information Privacy Concerns, Procedural Fairness and Impersonal Trust: An Empirical Investigation," *Organization Science* (10:1), pp. 104-115.

Dinev, T., and Hart, P. 2006. "An Extended Privacy Calculus Model for E-Commerce Transactions," *Information Systems Research* (17:1), pp. 61-80.

Fang, V. W., Tian, X., and Tice, S. 2014. "Does Stock Liquidity Enhance or Impede Firm Innovation?," *The Journal of Finance* (69:5), pp. 2085-2125.

Felt, A. P., Chin, E., Hanna, S., Song, D., and Wagner, D. 2011. "Android Permissions Demystified," *Proceedings of the 18th ACM Conference on Computer and Communications Security*: ACM, pp. 627-638.
Gao, H., and Zhang, W. 2016. "Employment Nondiscrimination Acts and Corporate Innovation," *Management Science* (63:9), pp. 2982-2999.
Garber, J. 2018. "GDPR–Compliance Nightmare or Business Opportunity?," *Computer Fraud & Security* (2018:6), pp. 14-15.
Ghose, A., and Han, S. P. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60:6), pp. 1470-1488.
Gopal, A., and Gao, G. 2009. "Certification in the Indian Offshore It Services Industry," *Manufacturing & Service Operations Management* (11:3), pp. 471-492.
Grace, M. C., Zhou, W., Jiang, X., and Sadeghi, A.-R. 2012. "Unsafe Exposure Analysis of Mobile in-App Advertisements," *Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks*: ACM, pp. 101-112.
Gradwohl, R. 2018. "Firms Quick to Adopt Eu Data Regs Will Have First-Mover Advantage," in: *The Hill*.
Hann, I.-H., Hui, K.-L., Lee, S.-Y. T., and Png, I. P. 2007. "Overcoming Online Information Privacy Concerns: An Information-Processing Theory Approach," *Journal of Management Information Systems* (24:2), pp. 13-42.
Harbach, M., Hettig, M., Weber, S., and Smith, M. 2014. "Using Personal Examples to Improve Risk Communication for Security and Privacy Decisions," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 2647-2656.
Hinneburg, A., and Keim, D. A. 1999. "Optimal Grid-Clustering: Towards Breaking the Curse of Dimensionality in High-Dimensional Clustering," *Proceedings of the 33rd International Conference on Very Large Data Bases*, pp. 506-517.
Hoffman, D. L., Novak, T. P., and Peralta, M. 1999. "Building Consumer Trust Online," *Communications of the ACM* (42:4), pp. 80-85.
Hoofnagle, C. J., King, J., Li, S., and Turow, J. 2010. "How Different Are Young Adults from Older Adults When It Comes to Information Privacy Attitudes and Policies?," *SSRN Working Paper*, https://ssrn.com/abstract=1589864.
Huberman, B. A., Adar, E., and Fine, L. R. 2005. "Valuating Privacy," *IEEE Security & Privacy* (3:5), pp. 22-25.
Hui, K.-L., Teo, H. H., and Lee, S.-Y. T. 2007. "The Value of Privacy Assurance: An Exploratory Field Experiment," *MIS Quarterly* (31:1), pp. 19-33.
Jensen, C., and Potts, C. 2004. "Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices," *Privacy Policies as Decision-Making Tools: An Evaluation of Online Privacy Notices*: ACM, pp. 471-478.
Kalyanaram, G., and Urban, G. L. 1992. "Dynamic Effects of the Order of Entry on Market Share, Trial Penetration, and Repeat Purchases for Frequently Purchased Consumer Goods," *Marketing Science* (11:3), pp. 235-250.
Kelley, P. G., Cranor, L. F., and Sadeh, N. 2013. "Privacy as Part of the App Decision-Making Process," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 3393-3402.
Kerin, R. A., Varadarajan, P. R., and Peterson, R. A. 1992. "First-Mover Advantage: A Synthesis, Conceptual Framework, and Research Propositions," *The Journal of Marketing* (56:4), pp. 33-52.

King, A. A., Lenox, M. J., and Terlaak, A. 2005. "The Strategic Use of Decentralized Institutions: Exploring Certification with Iso 14001 Management Standard," *Academy of Management Journal* (48:6), pp. 1091-1106.
Lieberman, M. B., and Montgomery, D. B. 1988. "First-Mover Advantages," *Strategic Management Journal* (9:S1), pp. 41-58.
Liu, B., Andersen, M. S., Schaub, F., Almuhimedi, H., Zhang, S. A., Sadeh, N., Agarwal, Y., and Acquisti, A. 2016. "Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions," *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*, pp. 27-41.
Liu, X., and Lynch, L. 2011. " Do Agricultural Land Preservation Programs Reduce Farmland Loss? Evidence from a Propensity Score Matching Estimator," *Land Economics* (87:2), pp. 183-201.
Mahajan, V., and Muller, E. 1998. "When Is It Worthwhile Targeting the Majority Instead of the Innovators in a New Product Launch?," *Journal of Marketing Research* (35:4), pp. 488-495.
Maheshwari, S. 2017. "That Game on Your Phone May Be Tracking What You're Watching on Tv," in: The New York Times.
Malhotra, N. K., Kim, S. S., and Agarwal, J. 2004. "Internet Users' Information Privacy Concerns (Iuipc): The Construct, the Scale, and a Causal Model," *Information Systems Research* (15:4), pp. 336-355.
Mayya, R., Ye, S., Viswanathan, S., and Agarwal, R. 2017. "Who Forgoes Screening in Online Markets and When? Evidence from Airbnb," *SSRN Working Paper*, https://ssrn.com/abstract=3082018.
Meyer, B. D. 1995. "Natural and Quasi-Experiments in Economics," *Journal of Business & Economic Statistics* (13:2), pp. 151-161.
Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. 2013. "Distributed Representations of Words and Phrases and Their Compositionality," *Advances in Neural Information Processing Systems*, pp. 3111-3119.
Mitchell, W. 1991. "Dual Clocks: Entry Order Influences on Incumbent and Newcomer Market Share and Survival When Specialized Assets Retain Their Value," *Strategic Management Journal* (12:2), pp. 85-100.
Mnih, A., and Hinton, G. E. 2009. "A Scalable Hierarchical Distributed Language Model," *Advances in Neural Information Processing Systems*, pp. 1081-1088.
Nowak, G. J., and Phelps, J. 1995. "Direct Marketing and the Use of Individual-Level Consumer Information: Determining How and When "Privacy" Matters," *Journal of Direct Marketing* (9:3), pp. 46-60.
Pavlou, P. A., Liang, H., and Xue, Y. 2007. "Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective," *MIS Quarterly* (31:1), pp. 105-136.
Phelps, J., Nowak, G., and Ferrell, E. 2000. "Privacy Concerns and Consumer Willingness to Provide Personal Information," *Journal of Public Policy & Marketing* (19:1), pp. 27-41.
Pingitore, G., Rao, V., Cavallaro, K., and Dwivedi, K. 2017. "To Share or Not to Share," Deloitte University Press.
Rainie, L., Kiesler, S., Kang, R., Madden, M., Duggan, M., Brown, S., and Dabbish, L. 2013. "Anonymity, Privacy, and Security Online," *Pew Research Center* (5).
Ramos, J. 2003. "Using Tf-Idf to Determine Word Relevance in Document Queries," *Proceedings of the First Instructional Conference on Machine Learning*, pp. 133-142.

Rose, E. 2005. "Data Users Versus Data Subjects: Are Consumers Willing to Pay for Property Rights to Personal Information?," *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*: IEEE, pp. 180c-180c.

Sarma, B. P., Li, N., Gates, C., Potharaju, R., Nita-Rotaru, C., and Molloy, I. 2012. "Android Permissions: A Perspective Combining Risks and Benefits," *Proceedings of the 17th ACM Symposium on Access Control Models and Technologies*: ACM, pp. 13-22.

Smith, J. A., and Todd, P. E. 2005. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?," *Journal of Econometrics* (125:1-2), pp. 305-353.

Spiekermann, S., Grossklags, J., and Berendt, B. 2001. "E-Privacy in 2nd Generation E-Commerce: Privacy Preferences Versus Actual Behavior," *Proceedings of the 3rd ACM Conference on Electronic Commerce*: ACM, pp. 38-47.

Stamm, S. M., Tripp; Kuronen, Jessica. 2018. "How Pizza Night Can Cost More in Data Than Dollars," in: The Wall Street Journal.

Statista. 2018. "Android Version Market Share Distribution among Smartphone Owners," in: *Statista*.

Tsai, J. Y., Egelman, S., Cranor, L., and Acquisti, A. 2011. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research* (22:2), pp. 254-268.

Tucker, C.E., 2014. "Social networks, personalized advertising, and privacy controls," *Journal of Marketing Research* (51:5), pp.546-562.

Varian, H. R. 2009. "Economic Aspects of Personal Privacy," in *Internet Policy and Economics*. Springer, pp. 101-109.

Wei, X., Gomez, L., Neamtiu, I., and Faloutsos, M. 2012. "Permission Evolution in the Android Ecosystem," *Proceedings of the 28th Annual Computer Security Applications Conference*: ACM, pp. 31-40.

Zimmer, J. C., Arsal, R. E., Al-Marzouq, M., and Grover, V. 2010. "Investigating Online Information Disclosure: Effects of Information Relevance, Trust and Risk," *Information & Management* (47:2), pp. 115-123.

Appendix A

Table A1.

Panel A: Summary Statistics

Variables	# of Obs.	Mean	Std. Dev.	Min	Max
Key Variables					
dangerous_permissions$_{it}$	278,955	3.46	2.65	0	22
essential_dangerous_permissions$_{it}$	278,955	0.74	0.77	0	10
nonessential_dangerous_permissions$_{it}$	278,955	2.71	2.36	0	20
normal_permissions$_{it}$	278,955	8.21	4.72	0	65
upgrade$_{it}$	278,955	0.48	0.50	0	1
rating$_{it}$	278,955	4.04	0.37	1.86	4.96
App Characteristics					
rating_count$_{it}$	278,955	150079.30	1108175.00	39	75718104
dayssinceupdate$_{it}$	278,955	176.84	188.52	0	1087
screenshots$_{it}$	278,955	12.50	5.55	2	33
developer_appcount$_{it}$	278,955	24.64	27.89	0.03	1800
filesize$_{it}$	278,955	29.16	86.82	1	1914

Note: These statistics are based on the sample before propensity score matching.

Panel B: Correlation Matrix

	Variables	# Obs.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1)	*dangerous_permissions$_{it}$*	278,955	1.00										
(2)	*essential_dangerous_permissions$_{it}$*	278,955	**0.50**	1.00									
(3)	*nonessential_dangerous_permissions$_{it}$*	278,955	**0.96**	**0.24**	1.00								
(4)	*normal_permissions$_{it}$*	278,955	**0.76**	**0.37**	**0.73**	1.00							
(5)	*upgrade$_{it}$*	278,955	**-0.04**	**-0.04**	**-0.04**	**0.01**	1.00						
(6)	*rating$_{it}$*	278,955	**0.10**	**0.08**	**0.09**	**0.19**	**0.11**	1.00					
(7)	*rating_count$_{it}$*	278,955	**0.12**	**0.05**	**0.12**	**0.18**	**0.03**	**0.11**	1.00				
(8)	*dayssinceupdate$_{it}$*	278,955	**-0.19**	**-0.18**	**-0.15**	**-0.22**	**-0.23**	**-0.17**	**-0.07**	1.00			
(9)	*screenshots$_{it}$*	278,955	**-0.11**	**-0.07**	**-0.11**	**-0.08**	**0.05**	**0.05**	**0.02**	**0.01**	1.00		
(10)	*developer_appcount$_{it}$*	278,955	**-0.08**	**-0.01**	**-0.09**	**-0.03**	**-0.04**	**-0.02**	**-0.02**	**0.05**	**0.00**	1.00	
(11)	*filesize$_{it}$*	278,955	0.00	**0.03**	**-0.01**	**0.02**	**0.06**	**0.04**	**0.07**	**0.03**	**0.30**	**0.01**	1.00

Table A2. App Categories and Category Groups

Category Group	Group Description	Categories in the Group
Content Consumption	Apps that allow consumption of content on a mobile platform	Comics Entertainment Media and Video Music and Audio Sports Video Players
Learn and Explore	Apps that allow users to learn and explore content on the mobile platform	Art and Design Books and References Education Medical News and Magazines Travel and Local Weather Parenting Libraries and Demo
Personal	Apps that are personal to users	Beauty Dating Health and Fitness Lifestyle Personalization Social
Utility – Mobile Access	These are the utility apps that have a major offline/web presence and provides online access through the platform	Business Auto and Vehicles Events Finance House and Home Shopping Transportation
Utility – Mobile Specific	These are utility apps that are present mainly on the mobile platform. Its existence depends on features of smart phone	Maps and Navigation Food and Drink Communication Tools Photography Productivity
Games	These are apps that are categorized as games by Google Playstore	Action Game Adventure Game Arcade Game Board Game Casino Game Casual Game Educational Game Music Game Puzzle Game Racing Game Role Playing Game Simulation Game Sports Game Strategy Game Trivia Game Word Game

Table A3. Android 6.0 Features

Panel A: Changes made to existing Android Features in Android 6.0

Change Type	Brief Description of change
Runtime Permissions	Changed how apps seek permission to access users' information
Doze and App Standby	Power saving optimization for idle devices and apps
Apache HTTP Client Removal	Apps can instead use an efficient class, *HttpURLConnection*
BoringSSL	Android moved from *OpenSSL* to *BoringSSL* library
Access to Hardware Identifier	Apps can use *access fine location* permission instead
Notifications	Improvements to how notifications can be updated
AudioManager	Use alternative methods to setting volumes and muting
Text Selection	Makes cut, copy, paste easier
Browser Bookmark	Apps are required to store bookmarks data internally (instead of global)
Android Keystore	Security enhancement of keystore provider
Wifi and Networking	Cannot alter Wifi Configuration created by other apps. Provide new APIs for network binding
Camera Service	Improvement to Camera API
USB Connection	Default connections through USB are in charge-only mode
Android for Work	Changes in Android for Work (device policies etc.)
APK Validation	Stricter validation of APK files to ensure there are no corruptions

List retrieved from https://developer.android.com/about/versions/marshmallow/android-6.0-changes

Panel B: New Features introduced in Android 6.0

New Feature	Brief Description of the New feature
Fingerprint Authentication	Allows apps to use fingerprints on supported phones
Confirm Credential	Authenticate users based on how recently they last unlocked phone
App Linking	Allows apps to associate with a web domain they own
Auto Backup for Apps	Allows apps for automatic full data backup
Direct Share	Allow apps to use APIs to share content with Android users
Voice Interactions	Improves conversational voice experience
Assist API	Allows apps to engage with users through an assistant
Adoptable Storage Devices	Users can adopt external storage devices
Notifications	Improves notification
Bluetooth Stylus Support	Added support for Bluetooth Stylus
Bluetooth low energy scanning	Improves power efficiency
Hotspot Improvement	Allows Hotspot 2.0
4K Display Mode	Allows compatible hardware to have 4K display resolution
Themeable Colors	Supports Theme attributes
Audio Features	A few improvements on Audio API
Video Features	A few improvements in Video API
Camera Features	A few new features such as Flashlight, reprocessing
Android for Work Features	A few new APIs for Android for Work

List retrieved from https://developer.android.com/about/versions/marshmallow/android-6.0

Table A4. Analysis of Upgrading to Latest Version of Android

Dependent variable	(1) Logit - $upgrade_{it}$	
$location\_coarse_{it}$	0.270***	(0.022)
$location\_precise_{it}$	-0.110***	(0.023)
$access\_bodysensors_{it}$	0.194	(0.179)
$make\_phonecall_{it}$	0.139***	(0.034)
$access\_camera_{it}$	0.329***	(0.019)
$find\_accounts_{it}$	-0.222***	(0.017)
$reroute\_outgoingcalls_{it}$	0.066	(0.059)
$read\_phonestatus\_and\_identity_{it}$	-0.669***	(0.015)
$read\_sms_{it}$	-0.211***	(0.053)
$receive\_mms_{it}$	0.129**	(0.042)
$record\_audio_{it}$	-0.325***	(0.024)
$send\_sms_{it}$	0.103*	(0.048)
$make\_sip\_call_{it}$	1.466***	(0.151)
$add\_voicemail_{it}$	-0.356	(0.371)
$read\_write\_usb_{it}$	-0.194***	(0.018)
$read\_write\_calendar_{it}$	-0.138**	(0.047)
$read\_write\_calllog_{it}$	-0.897***	(0.053)
$read\_write\_contact_{it}$	0.139***	(0.027)
$rating\_count_{it}$	-0.000***	(0.000)
$rating_{it}$	1.085***	(0.234)
$rating_{it} * rating_{it}$	-0.155***	(0.030)
$games_i$	-0.155***	(0.024)
$personal\_apps_i$	-0.191***	(0.028)
$utility\_mobilespecific_i$	-0.309***	(0.026)
$utility\_mobileaccess_i$	0.146***	(0.034)
$learn\_explore_i$	-0.033	(0.029)
$category_i$=content_consumption$_i$ (baseline)		
$below\_1million_i$	-0.616***	(0.017)
$5million\_10million_i$	0.377***	(0.022)
$10million\_50million_i$	0.359***	(0.023)
$above\_50million_i$	0.265***	(0.049)
$download_i$=1million_5million$_i$ (baseline)		
$dayssinceupdate_{it}$	-0.006***	(0.000)
$screenshot_{it}$	0.011***	(0.001)
$filesize_{it}$	0.007***	(0.000)
$developer\_appcount_{it}$	0.001***	(0.000)
Constant	-0.879	(0.454)
Month Dummies	YES	
# of Obs.	178,693	
Pseudo R-Squared	0.216	

*** p<0.001, ** p<0.01, * p<0.05; Robust standard errors in parentheses.

Table A5. Analysis of Upgrading to Earlier Version of Android

Dependent variable	(1) Logit – $upgrade\_api21_{it}$		(2) Logit – $upgrade\_api22_{it}$	
$dangerous\_permissions\_ratio_{it}$	0.004*	(0.002)	0.040***	(0.001)
$rating\_count_{it}$	-0.000*	(0.000)	-0.000**	(0.000)
$rating_{it}$	9.346***	(1.373)	5.155***	(0.609)
$rating_{it} * rating_i$	-1.205***	(0.173)	-0.703***	(0.078)
$games_i$	-0.132	(0.087)	0.166**	(0.054)
$personal\_apps_i$	0.168	(0.098)	0.171**	(0.063)
$utility\_mobilespecific_i$	-0.049	(0.090)	0.253***	(0.057)
$utility\_mobileaccess_i$	-0.370**	(0.135)	-0.418***	(0.090)
$learn\_explore_i$	-0.442***	(0.125)	0.246***	(0.069)
$category_i$=content_consumption$_i$ (baseline)				
$below\_1million_i$	-0.492***	(0.067)	-0.263***	(0.036)
$5million\_10million_i$	0.612***	(0.070)	-0.447***	(0.060)
$10million\_50million_i$	0.339***	(0.079)	0.114*	(0.051)
$above\_50million_i$	-0.254	(0.199)	0.275**	(0.100)
$download_i$=1million_5million$_i$ (baseline)				
$dayssinceupdate_{it}$	-0.006***	(0.000)	-0.005***	(0.000)
$screenshot_{it}$	0.008	(0.005)	0.025***	(0.003)
$filesize_{it}$	-0.003*	(0.001)	0.001	(0.001)
$developer\_appcount_{it}$	-0.000*	(0.000)	-0.000	(0.000)
Constant	-19.950***	(2.724)	-11.996***	(1.190)
Month Dummies	YES		YES	
# of Obs.	79,563		108,449	
Pseudo R-Squared	0.129		0.141	

*** p<0.001, ** p<0.01, * p<0.05; Robust standard errors in parentheses.

Table A6. Validity of Propensity Score Matching Procedure

Panel A: Covariate Balance - PSM

Covariates	Pre-Match (Overall)	Pre-Match (Treated)	Pre-Match (Control)	Difference	t-test	Post-Match (Treated)	Post-Match (Control)	Difference	t-test
Rating	4.01	4.03	4.00	-0.03	-17.53	4.00	4.01	0.01	1.20
Rating Count	123188	153233	106024	-47209	-12.06	73624	75736	2112	0.25
File Size	24.00	25.28	23.30	-1.98	-17.19	23.79	24.45	0.66	1.22
Screenshots	12.26	12.55	12.09	-0.46	-17.86	12.73	12.51	-0.22	-1.45
dayssinceupdate	193.88	122.57	234.63	112.06	122.40	170.68	179.30	8.63	1.74

Panel B: Analysis of Effects of Upgrading to Latest Version - PSM

Dependent Variable	(1) essential_ *dangerous_permissions_ratio$_{it}$*		(2) nonessential_ *dangerous_permissions_ratio$_{it}$*		(3) log(rating$_{it}$)	
dangerous_permissions_ratio$_{it}$					-0.000	(0.000)
rating_count$_{it-1}$	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
rating$_{it}$	0.090	(0.097)	-0.744*	(0.347)		
*upgrade_group$_i$*time$_{t-7}$*	0.001	(0.011)	0.039	(0.020)	0.000	(0.000)
*upgrade_group$_i$*time$_{t-6}$*	0.012	(0.009)	-0.011	(0.014)	0.000	(0.000)
*upgrade_group$_i$*time$_{t-5}$*	0.001	(0.008)	0.004	(0.014)	-0.000	(0.000)
*upgrade_group$_i$*time$_{t-4}$*	-0.012	(0.010)	0.034*	(0.016)	-0.000	(0.000)
*upgrade_group$_i$*time$_{t-3}$*	0.002	(0.010)	-0.004	(0.020)	0.000	(0.000)
*upgrade_group$_i$*time$_{t-2}$*	0.007	(0.011)	0.002	(0.020)	0.000	(0.000)
*upgrade_group$_i$*time$_{t-1}$*	-0.012	(0.011)	-0.019	(0.022)	-0.000	(0.000)
*upgrade_group$_i$*time$_{t0}$*			(Omitted Base Case)			
*upgrade_group$_i$*time$_{t+1}$*	0.005	(0.008)	-0.056**	(0.020)	-0.000*	(0.000)
*upgrade_group$_i$*time$_{t+2}$*	0.004	(0.008)	-0.041*	(0.016)	0.000*	(0.000)
dayssinceupdate$_{it}$	0.000***	(0.000)	0.001***	(0.000)	-0.000	(0.000)
screenshot$_{it}$	0.000	(0.000)	-0.000	(0.001)	-0.000	(0.000)
filesize$_{it}$	0.022*	(0.009)	-0.017	(0.011)	0.000	(0.000)
developer_appcount$_{it}$	-0.001	(0.000)	0.003***	(0.001)	0.000*	(0.000)
Constant	0.430	(0.419)	5.378***	(1.415)	1.392***	(0.003)
Month Dummies		YES		YES		YES
App Fixed Effect		YES		YES		YES
Download Bucket Dummies		YES		YES		YES
# of Obs.		53672		53672		53672
R-Squared		0.082		0.031		0.089

*** p<0.001, ** p<0.01, * p<0.05

Heteroskedastic Robust Standard Errors in parenthesis

Table A7. Validity of Look Ahead Propensity Score Matching Procedure

Panel A: Covariate Balance – LA-PSM

Covariates	Pre-Match (Overall)	Pre-Match (Treated)	Pre-Match (Control)	Difference	t-test	Post-Match (Treated)	Post-Match (Control)	Difference	t-test
Rating	4.01	4.03	4.00	-0.03	-17.53	4.00	4.01	0.01	0.54
Rating Count	123188	153233	106024	-47209	-12.06	91155	91796	641	0.04
File Size	24.00	25.28	23.30	-1.98	-17.19	24.58	24.22	-0.36	-0.41
Screenshots	12.26	12.55	12.09	-0.46	-17.86	12.82	12.53	-0.29	-1.40
dayssinceupdate	193.88	122.57	234.63	112.06	122.40	135.66	132.32	-3.34	-0.69

Panel B. Analysis of Effects of Upgrading to Latest Version - LAPSM

Dependent Variable	(1) essential_ $dangerous\_permissions\_ratio_{it}$		(2) nonessential_ $dangerous\_permissions\_ratio_{it}$		(3) $\log(rating_{it})$	
$dangerous\_permissions\_ratio_{it}$					-0.001	(0.002)
$rating\_count_{it-1}$	-0.000*	(0.000)	0.000**	(0.000)	0.000*	(0.000)
$rating_{it}$	0.217	(0.114)	-0.916**	(0.337)		
$upgrade\_group_i * time_{t-7}$	0.026	(0.028)	-0.004	(0.047)	-0.000	(0.001)
$upgrade\_group_i * time_{t-6}$	0.037	(0.026)	-0.044	(0.030)	-0.001	(0.000)
$upgrade\_group_i * time_{t-5}$	0.002	(0.018)	0.023	(0.027)	-0.000	(0.000)
$upgrade\_group_i * time_{t-4}$	-0.007	(0.018)	0.055	(0.031)	-0.000	(0.000)
$upgrade\_group_i * time_{t-3}$	-0.003	(0.016)	0.015	(0.043)	0.000	(0.000)
$upgrade\_group_i * time_{t-2}$	0.012	(0.018)	0.005	(0.031)	0.000	(0.000)
$upgrade\_group_i * time_{t-1}$	-0.011	(0.018)	-0.022	(0.034)	-0.000	(0.000)
$upgrade\_group_i * time_{t0}$			(Omitted	Base Case)		
$upgrade\_group_i * time_{t+1}$	0.001	(0.012)	-0.028	(0.027)	-0.000*	(0.000)
$upgrade\_group_i * time_{t+2}$	-0.004	(0.010)	-0.017	(0.022)	0.000	(0.000)
$dayssinceupdate_{it}$	0.000**	(0.000)	0.001***	(0.000)	-0.000	(0.000)
$screenshot_{it}$	0.000	(0.000)	-0.001	(0.001)	-0.000	(0.000)
$filesize_{it}$	-0.007	(0.008)	-0.013	(0.019)	-0.000	(0.000)
$developer\_appcount_{it}$	-0.002**	(0.001)	0.003	(0.001)	0.000**	(0.000)
Constant	0.426	(0.486)	5.893***	(1.404)	1.394***	(0.003)
Month Dummies		YES		YES		YES
App Fixed Effect		YES		YES		YES
Download Bucket Dummies		YES		YES		YES
# of Obs.		30045		30045		30045
R-Squared		0.087		0.025		0.092

*** p<0.001, ** p<0.01, * p<0.05

Heteroskedastic Robust Standard Errors in parenthesis

Table A8. Analysis of Effects of Upgrading to Latest Version - LAPSM

Dependent Variable	(1) LAPSM DID - essential_ *dangerous_permissions*$_{it}$		(2) LAPSM DID - nonessential_ *dangerous_permissions*$_{it}$		(1) LAPSM DID – log(rating$_{it}$)	
dangerous_permissions_ratio$_{it}$					0.001	(0.001)
rating_count$_{it-1}$	-0.000***	(0.000)	0.000***	(0.000)	0.000***	(0.000)
rating$_{it}$	0.150***	(0.045)	-0.389**	(0.121)		
post_switch$_{t}$	0.024***	(0.007)	-0.084***	(0.016)	-0.001***	(0.000)
upgrade_group$_{i}$**post_upgrade*$_{t}$	-0.002	(0.009)	-0.043*	(0.019)	0.001*	(0.000)
dayssinceupdate$_{it}$	0.000***	(0.000)	0.000***	(0.000)	-0.000	(0.000)
screenshot$_{it}$	0.000	(0.000)	0.001*	(0.000)	0.000**	(0.000)
filesize$_{it}$	-0.005	(0.004)	-0.030**	(0.009)	0.000	(0.000)
developer_appcount$_{it}$	-0.002***	(0.000)	0.003***	(0.001)	0.000***	(0.000)
Constant	0.764***	(0.208)	4.253***	(0.524)	1.396***	(0.002)
Month Dummies	YES		YES		YES	
App Fixed Effect	YES		YES		YES	
Download Bucket Dummies	YES		YES		YES	
# of Obs.	58791		58791		58791	
R-Squared	0.145		0.017		0.114	

*** p<0.001, ** p<0.01, * p<0.05
Heteroskedastic Robust Standard Errors in parenthesis

Table A9. Analysis of Effects of Delaying the Upgrading to Latest Version - LAPSM

Dependent Variable	(1) LAPSM DID - essential_ *dangerous_permissions*$_{it}$		(2) LAPSM DID - nonessential_ *dangerous_permissions*$_{it}$		(3) LAPSM DID – log(rating$_{it}$)	
dangerous_permissions_ratio$_{it}$					0.001	(0.001)
rating_count$_{it}$	-0.000***	(0.000)	0.000***	(0.000)	0.000***	(0.000)
rating$_{it}$	0.150***	(0.045)	-0.393**	(0.121)		
post_switch$_{t}$	0.023**	(0.007)	-0.075***	(0.016)	-0.001**	(0.000)
upgrade_group$_{i}$**post_upgrade*$_{t}$	-0.006	(0.014)	0.018	(0.027)	0.001*	(0.000)
late_upgrade_group$_{i}$**post_upgrade*$_{t}$	0.008	(0.015)	-0.111***	(0.031)	-0.001^{+}	(0.000)
dayssinceupdate$_{it}$	0.000***	(0.000)	0.000***	(0.000)	-0.000	(0.000)
screenshot$_{it}$	0.000	(0.000)	0.001*	(0.000)	0.000*	(0.000)
filesize$_{it}$	-0.004	(0.004)	-0.030**	(0.009)	0.000	(0.000)
developer_appcount$_{it}$	-0.002***	(0.000)	0.003***	(0.001)	0.000***	(0.000)
Constant	0.761***	(0.208)	4.296***	(0.524)	1.396***	(0.002)
Month Dummies	YES		YES		YES	
App Fixed Effect	YES		YES		YES	
Download Bucket Dummies	YES		YES		YES	
# of Obs.	58791		58791		58791	
R-Squared	0.145		0.017		0.114	

*** p<0.001, ** p<0.01, * p<0.05, $^{+}$ p<0.1
Heteroskedastic Robust Standard Errors in parenthesis

Appendix B: Essential and non-essential Dangerous Permissions

A key task in our project is to determine which of the permissions sought are essential for the app's working. Our methodology, which uses app sub-categories to statistically determine the essentiality of a permission, based on a method proposed by Sarma et al (2012). We employ a skip-gram Word2Vec vector representation of words, a modern and highly effective text-mining technique. Such distributed vector representation of words that learns the locational similarity and context of words in a statement, has been proven to be more accurate than the traditional bag of words or n-gram vectors (Mnih and Hinton 2009). We use this technique over Term Frequency Inverse Document Frequency (TF-IDF), because TF-IDF techniques create a vector for each word in the corpus and does not consider the meanings and contexts of the words as well as grammatical nuances such as plurals (Ramos 2003). Furthermore, since we employ a popular clustering algorithm called k-means clustering to create subgroups of apps based on app description, TF-IDF technique may by less effective due to the curse of dimensionality (Hinneburg and Keim 1999). The approach by Mikolov et al. (2013) resolves these problems by employing a skip-gram model and creating a 300-dimensional vector representation for each word. The accuracy of such models improve with the size of the training set. Hence, we follow their approach to utilize the model trained on 100 billion words from a dataset derived from Google News.

First, we tag the Play Store text description of all apps with category label as determined by Android. These categories are pre-determined and maintained by the platform; thus, we are confident about the general similarity between apps under the same category. Next, we divide apps within each category into sub-groups of apps that have similar functionality and utility.

Codifying our process to determine the sub-category of apps, we:

I. obtain a 300-dimensional vector for each app based on app description collected from Play Store by employing skip-gram Word2Vec.

II. employ k-means Clustering Algorithm to determine the optimal sub-groups within a category by following the following sub-steps

    a. Determine the range of optimal clusters for each app sub-category *c* using the Elbow technique, AIC and BIC

b. Determine the smallest cluster size k_c among the optimal set of cluster sizes wherein each cluster has at least 5 apps.
c. Divide apps under each category c into k_c clusters (i.e., k_c sub-categories of each of *c* categories)

We also manually check a random sub-set of all algorithmically generated sub-categories to ensure that the sub-categorization has worked well.

Next, similar to the technique proposed by Sarma et al. (2012), we use these sub-categories to statistically determine which of the dangerous permissions sought are essential to each of the apps in those sub-categories. We code those permissions requested by more than 75% of the apps in a sub-category as essential permissions. The intuition is that, if a permission is essential for a given sub-category of apps, most apps in that sub-category would seek that permission. For example, most navigation apps would seek permissions to access the user's location. Conversely, permissions that are sought by a smaller number of apps in the sub-category are likely non-essential. We check the sensitivity of this sub-categorization by varying the threshold between 70% and 80% and the results are qualitatively similar