Does Piracy Lead to Product Abandonment or Stimulate New Product Development?

Evidence from Mobile Platform-Based Developer Firms

Miric, Milan; Jeppesen, Lars Bo

Document Version
Accepted author manuscript

Published in:
Strategic Management Journal

DOI:
10.1002/smj.3208

Publication date:
2020

Does Piracy Lead to Product Abandonment or Stimulate New Product Development? Evidence from Mobile Platform-Based Developer Firms

Milan Miric and Lars Bo Jeppesen

Journal article (Accepted manuscript*)

Does Piracy Lead to Product Abandonment or Stimulate New Product Development?: Evidence from Mobile Platform-Based Developer Firms

Milan Miric[1*] & Lars Bo Jeppesen[2]

[1]University of Southern California,
Marshall School of Business (mmiric@marshall.usc.edu);
[2]Copenhagen Business School (lbj.si@cbs.dk);
[*]*Corresponding Author*

Abstract

Research Summary: With the growth of digital platforms, understanding the role of property rights on those platforms has become increasingly important. Digital piracy, the unauthorized copying and distribution of digital products, is therefore an important strategic issue, both because of lost revenues and because it is thought to decrease innovation. Yet, while the latter effect is often argued, empirical evidence is limited. We study whether piracy affects innovation and whether it leads firms to shift to different types of innovations. By studying a large piracy event in a mobile app marketplace, we find that piracy leads to a decrease in the release of incremental innovations, such as bug fixes, but does not decrease more substantial revisions. Additionally, it is associated with subsequent new product development.

Managerial Abstract: For many platform companies, a critical issue is understanding how piracy and imitation should be regulated, motivated in part by a common narrative that piracy will eliminate innovation on these platforms. The present paper suggests that these effects are slightly more nuanced. We find that piracy does lead to a decline in incremental innovations, such as bug fixes or appearance tweaks, but no discernible decline in more major innovations, such as feature updates or entirely new versions. This implies that piracy can shape the type of innovation, potentially leading to products that are less polished and refined but not affecting the overall level of innovation.

Keywords: *Digital Piracy, Digital Innovation, Platforms, Mobile Applications, New Product Development*

Running Header: *Piracy, Product Abandonment & New Product Development?*

1 Introduction

For many modern digital platform companies, their success depends on being able to foster innovation and attract and retain third parties to create complementary products (Eisenmann et al., 2011; Parker and Van Alstyne, 2017; Jacobides et al., 2018; McIntyre and Srinivasan, 2013). A number of recent studies have sought to understand the implications of various factors in shaping how these platform companies are able to attract third parties and foster innovation (Boudreau and Jeppesen, 2015; Wen and Zhu, 2019; Cennamo and Santalo, 2013; Kapoor and Lee, 2013). Property rights can play an important role in encouraging innovation by providing innovators with exclusive control over the technologies they create (Teece, 1986). This can, in turn, allow firms to have a unique position relative to their competitors and limit the threat of imitation by others in the marketplace, creating an incentive for innovation. Yet, within many digital industries, questions remain whether property rights are, in fact, effective at fostering innovation (Greenstein et al., 2013; Goldfarb et al., 2014). One well-known case relates to digital piracy, the unauthorized copying and distribution of digital products. While studies have established that digital piracy may affect the revenues that firms capture (Waldfogel, 2012b; Rob and Waldfogel, 2007), it remains an open question whether and how piracy influences subsequent innovation.

One prominent narrative for how piracy influences innovation argues that the economic losses from piracy diminish incentives for innovation. For instance, within the packaged software industry, where piracy is said to account for more than $63 Bn in lost revenues annually, leading firms have come out to warn against the economic consequences of piracy (BSA, 2016). Autodesk argued that piracy "*undermines the economy generally and software innovation in particular*." Similarly, the CEO of Adobe stated that Adobe may abandon support for Chinese language versions of their software as a consequence of their products being pirated in China. Even in the context of mobile platforms, where the platform owner has considerable discretion regarding their policy to govern and control third-party software, piracy is argued to have a huge negative effect on firm revenues. For instance, market research firms estimate that \$ 17Bn in revenues (approx. 14 Bn installs) have been lost as a result of software piracy on mobile app markets.[1] This is also the narrative put forth to motivate why legislators need to regulate piracy (SIIA, 2018). Yet, while academic research has provided evidence that piracy reduces the profits of firms (Novos and Waldman, 1984; Waldfogel, 2012a; Athey and Stern, 2015), there is limited evidence that aligns with the narrative, suggesting that piracy does not necessarily deter innovation or long-term innovation strategies. A number of studies have identified benefits of piracy, particularly surrounding the growth of consumer awareness that occurs as a result of piracy that, in turn, can lead to higher revenues (Peukert et al., 2017; Givon et al., 1995). Focusing on the context of "copycat" applications in the Apple App Store, Wang et al. (2015) found that copycat apps can lead to increased awareness for the original application and in turn lead to greater demand for the original product.

Yet, there are many more nuanced factors through which piracy can increase innovation (Lieberman and Asaba, 2006). Since piracy allows customers to acquire a product either through an official or pirated version of the software, it effectively creates a close competitor (a perfect imitation) to the official version of the software (Belleflamme and Peitz, 2012). While this type of close imitation may lead to a loss of revenue, it may also lead to competitive adjustments by the imitated firm to attempt to re-capture lost revenues (Lieberman and Asaba, 2006; Ethiraj and Zhou, 2019; Posen and Martignoni, 2018; Wang and Shaver, 2014). In the absence of piracy, the software creator would have exclusive rights over a particular innovation and would maintain a unique position relative to other offerings in the marketplace (Teece, 1986). However, this may also leave the software creator with little incentive to innovate and create a new product, as doing so would cannibalize their existing revenues (Arrow, 1962). However, with piracy, the software creator can find their advantage entirely eroded, as a perfect copy of their product would be freely available in the marketplace. While on the one hand, this might lead to lower revenues, it could also push the firm to upgrade its products in order to "escape competition" and construct a superior new product

[1] Based on online report published by Koetsier (2018).

for which the updated version does not exist. This may also push the firm to alter the types of innovations that are created, as piracy may shift the firm to focus on creating larger product improvements such as major product revisions or developing new products, that consumers may want to purchase instead of using the freely available pirated version, rather than creating more minor bug fixes.

Our empirical analysis centers on exploiting a piracy attack that occurred within a prominent mobile application marketplace, the cydia marketplace for mobile apps, and led to the release of pirated versions of the entire population of apps in the marketplace simultaneously. The cydia marketplace was an important storefront for mobile apps for iOS (iPhone, iPad, iPod Touch, etc.) devices that allowed developers to upload applications that were not allowed by Apple to be released through the official App Store. Even though cydia was an unofficial marketplace, it was populated by more than fifty-thousand software titles created by more than ten-thousand individual developers. In 2015, a group of hackers copied the entire repository of this marketplace (all of the files) and posted them online for free. Unlike many piracy attacks, this was not directed at the software developers but at the platform. This provides us with the opportunity for an event study to understand how this piracy shock influenced innovative outcomes at the developer level. The outcomes we look at are whether, following the shock, firms continued to update their products and investing in further innovation or whether they abandoned their products, ceasing to invest in additional innovation. This research design provides us with several control groups that can serve as a counterfactual for our analysis. As a control group, we look both at free products that were unaffected by this piracy attack, paid (non-free) products that were previously hacked and therefore were not affected by this piracy attack, as well as products on a different platform (the Android platform). This research design allows us to overcome the problem associated with most studies of piracy, which is that the most popular and sought-after titles are those most likely to be pirated, biasing any analysis. The piracy attack studied here affected all firms simultaneously and therefore does not suffer from this bias.

We find that piracy led to a decline in innovation but that this is overwhelmingly driven by a decline in incremental innovations, such as bug fixes. Yet, we even find some evidence that piracy led to an increase in the creation of new products. Specifically, we find that this piracy attack led to a decline in innovations (0.06 per fewer updates per app per month, corresponding to a 28% reduction in product updates on average) but that this was caused overwhelmingly by minor bug fixes (the change in the release of feature updates was statistically indistinguishable from zero, while there was a 68% increase in new product releases relative to the counterfactual). These results are consistent with our predictions that piracy may decrease incremental innovations but lead to an increase in more substantial innovations. These results suggest that rather than piracy influencing innovation in general, it influences different types of innovations, diminishing sustaining innovations but potentially incentivizing new product development.

This paper contributes to the literature that has looked at the implications of different platform strategies on innovation (Boudreau and Jeppesen, 2015; Wen and Zhu, 2019; Cennamo and Santalo, 2013). Recent papers have called for a greater understanding of the dynamic processes and policies through which innovation is created on such platforms (McIntyre and Subramaniam, 2009; Jacobides et al., 2018; Eisenmann et al., 2011). Platforms have considerable discretion regarding how they govern issues such as piracy and to what extent they allow "copycat" behavior on their platforms. Yet, there has been a lack of research regarding the property rights and property rights violations that occur as a consequence of piracy. This paper provides evidence regarding the nuances of how piracy influences innovations, which in turn informs optimal policies around property rights. In understanding the implications of piracy on innovation, this paper also contributes to the broader literature that has looked at property rights and the competitive imitation that occurs in the absence of those property rights (Teece, 1986; Nagaraj, 2018; Galasso and Schankerman, 2015; Luo and Mortimer, 2017). Piracy is an particularly important form of copyright violation, that affects many digital industries (Rob and Waldfogel, 2006; Oberholzer-Gee and Strumpf, 2007;

Athey and Stern, 2015). Understanding the implications of these piracy violations informs the broader literature on imitation, which has sought to "understand why imitation occurs and when it may have harmful implications" (Lieberman and Asaba, 2006, p. 336). While recent studies have documented how imitation can influence competitive adjustments (Wang and Shaver, 2014; Ethiraj and Zhou, 2019), they have not considered perfect imitation that occurs through piracy, as this paper does.

2 Literature Review

Here, we provide a background on studies that have looked at issues related to piracy and innovation.

2.1 Relevance of Piracy for Platform Strategy

Piracy is often discussed from the perspective of policymaking by studies that attempt to quantify piracy or identify optimal enforcement mechanisms. However, with the growth of digitization, digital platforms have begun to account for a larger share of economic activity. For these companies, their success depends on their ability to manage the balance between providing value for customers and complementers (Parker and Van Alstyne, 2017; Parker et al., 2017). This has spawned a growth of literature on the topic of platform strategy that tries to understand the unique set of strategic decisions in governing these platforms (Eisenmann et al., 2011; Parker and Van Alstyne, 2017; Jacobides et al., 2018; McIntyre and Srinivasan, 2013; Afuah, 2013; Wen and Zhu, 2019). This includes ensuring that firms developing complementary products (or complementers) have an incentive to develop for the platform and profit from their innovations (Boudreau, 2012; Wen and Zhu, 2019). This has spurred interest in platform strategy to understand the policies that platform companies can take in order to foster innovation, including policies on property rights (Huang et al., 2012; Miric et al., 2019). While piracy is an important issue anecdotally for platform companies , there has been little empirical work documenting the implications of piracy for platforms. One exception is Ishihara and Muller (2020), who study piracy in the case of console gaming platforms, and find that piracy of console games can lead to a shift in the strategy of the platform to focus less on developing exclusive products themselves and more on attracting more third-party complementers.

However, existing studies have documented that platform policies in limiting or encouraging complementers to enter and imitate one another may have considerable implications for innovation. Boudreau and Hagiu (2008) document how poor policies on the part of Atari led to widespread imitation and in turn a decline in the platform. Boudreau (2012) document how entry onto a platform might affect the innovative output of complementers. However, existing studies have not considered the specific but important case of piracy, and how it may shape the innovative output of complementers on a platform.

2.2 Property Rights and Innovation

Establishing, protecting, and enforcing property rights around innovations is a fundamental issue in firm strategy because property rights can allow firms to own and control particular innovations (Teece, 1986). Specifically, property rights ensure that the creator of an innovation has exclusive control over the innovation, excluding competitors, and in turn is able to profit from that innovation. However, whether property rights actually incentivize innovation is hotly debated. For instance, there is evidence from historical policy changes that property rights incentivize innovation (Moser, 2005, 2013; MacGarvie and Moser, 2015), but there is also evidence that these effects are somewhat more modest than the theoretical characterization of property rights (Lerner, 2009; Sakakibara and Branstetter, 2001). Similarly, there have been are numerous instances in which innovation was widespread, yet property rights such as patents and copyrights are only weakly enforceable (Cockburn and MacGarvie, 2009, 2011; Cohen and Lemley, 2001; Graham et al., 2009).

In the context of digital settings, such as software development on top of a digital platform, evidence on the impact of these property rights is similarly ambiguous. On the one hand, while patents and copyrights can be applied to protect digital innovations from imitators, they are not frequently used (Hall and Ziedonis, 2001; Miric et al., 2019; Graham et al., 2009). Similarly, copyright that is intended to protect against the unauthorized replication and distribution of goods is threatened in digital settings, where the costs of copying and distributing are close to zero (Goldfarb et al., 2015; Shapiro and Varian, 1998). This suggests that digital property rights are not entirely effective at protecting digital innovations but that, given the proliferation of digital goods and services, they may not be critical to facilitating innovation.

It is important to mention that despite the overwhelming number of examples in the literature of digital copyright being violated (Godinho de Matos et al., 2017; Givon et al., 1995), there have been instances where property rights have been successfully upheld. For instance, Luo and Mortimer (2017) study the case of a stock photography platform that enforces the copyright of photographs and consider optimal strategies through which this can be enforced. Nagaraj (2018) studies copyright protection for digital books and finds that copyright has led to a reduction in the use of copyrighted images from Google Books. Therefore, digital property rights have been found to protect digital innovations, which implies that they may be important for incentivizing innovation. However, evidence for this remains scarce.

The debate surrounding the use of property rights is part of a broader debate surrounding the impact of competitive imitation on the innovative output of companies. Being able to prevent competitors from imitating a company's products is a key determinant of how firms sustain their unique position (Teece, 1986). There are various strategies that firms may use to prevent competitors from imitating their products (Ethiraj et al., 2008; Cohen et al., 2000) or devise competitive adjustments in response to imitation (Ethiraj and Zhou, 2019; Wang and Shaver, 2014). Often, the combination of product complexity, legal protections, and internal processes help firms limit imitation and provide them with differentiated or imperfect competitors. In cases of piracy, however, which often violates legal channels, the imitation which occurs is often a perfect (or close to perfect) copy of the company's own products. The implications of such close imitation are typically less studied, as they do not occur as part of the natural cycles of competitive imitation and adjustment, but rather as a consequence of piracy or other legal decisions, such as granting Paragraph IV protections for pharmaceutical drugs (Buccafuso and Heald, 2013) or the expiration of copyright terms (Buccafuso and Heald, 2013; Heald, 2014). Despite the limited existing research on such imitation, it remains an important issue, particularly in digital settings where piracy is widespread and question exist regarding its implications for innovation.

2.3 Digital Piracy and its Economic Consequences

Perhaps the most widely discussed form of property right violation in digital settings is related to piracy, the violation of digital copyright. Copyright was designed to protect authors of creative works (originally book authors) from having their creations duplicated and sold by others. This was originally instituted to protect writers and composers from having their books and music, copied (or re-written by hand) and then sold by others (Giorcelli and Moser, 2017; MacGarvie and Moser, 2015). Within digital industries, copyright infringement has become an important issue due to the ease with which digital products are reproduced and shared. This means that consumers can often acquire a pirated version of the product at a lower cost (often zero) than the un-pirated version. The availibility of a pirated version may lead to a loss of revenues as consumers who would have purchased the product shift towards the pirated version. This is often the narrative of industry associations, which argue that piracy results in vast sums of lost revenues.

A set of empirical studies have sought to document the magnitude and impact of piracy (Adermon and Liang, 2014; Athey and Stern, 2015; Bai and Waldfogel, 2012; Leung, 2013; Rob and Waldfogel, 2006; Smith and Telang, 2019). For instance, Danaher and Smith (2014) studied how the shutdown of music sharing sites influenced the consumption of movies in theaters. Similarly, Oberholzer-Gee and Strumpf (2007) study how music piracy influences sales of records and found that the effect of piracy is statistically indistinguishable from zero. One argument for why piracy may have a limited observable impact on revenues is that while it leads to a pirated version being offered to consumers, it also increases the overall level of consumption and consumer awareness of the product. For instance, Peukert et al. (2017), who study the same mega-upload shutdown as Danaher and Smith (2014), find that while it did lead to a decrease in awareness for less popular products, it was associated with an increase for more niche products that benefited from the increased consumer awareness that resulted from piracy. Lee (2018) performed a similar study on the sharing of pirated music over the BitTorrent network and found that piracy did lead to losses for top artists but also helped more unknown artists gain recognition and attract fans. Kretschmer and Peukert (2020) find evidence of a similar diffusion effect associated with online music streaming, where having content available to stream increased music sales by 10%. The potential benefits of diffusion may be sufficient for firms to deliberately choose to release a version of their product for free in order to generate greater product diffusion (Nan et al., 2018; Sundararajan, 2004). Within the software industry itself, there is evidence that piracy may increase diffusion and enhance user consumption. For instance, Givon et al. (1995) found that 60% of consumers tried pirated software products but that these same consumers were then responsible for 80% of subsequent software purchases. There is evidence of this effect as far afield as fashion piracy (knockoffs) which can results in greater awareness for the product that is copied (Appel et al., 2018).

Numerous theoretical papers have formulated models that explicate optimal firm policy to benefit from these different effects. For instance, many models have looked at how firms may optimize their product offerings in order to prevent piracy (Nan et al., 2018; Sundararajan, 2004; Tunca and Wu, 2013; Sinha et al., 2010). These studies often consider offering a service bundled with the product to encourage consumers to buy rather than to use the pirated version, while others documented how firms can offer a free tier of their products to avoid piracy altogether.

Many studies have looked explicitly at different forms in which piracy was regulated and the impact of these policies on the consumption of pirated goods (Aguiar et al., 2018; Batikas et al., 2019; Reimers, 2016; Luo and Mortimer, 2017). These studies document that while there may be opportunities to regulate piracy through either technological or policy interventions, their impacts on consumption are mixed, suggesting that piracy may continue to affect digital innovators even following these interventions.

2.4 Piracy and Innovation

Existing studies and market analysis have overwhelmingly focused on the economic losses caused by piracy, either in terms of aggregate firm revenues or paid consumption of products such as movies, music or software applications.[2] The most salient or concerning element surrounding the impact of piracy is that it results in lower incentives to develop new products and services.[3] The underlying rationale is that piracy results in a loss of revenues that leads firms to invest less in new product development, which in turn creates other economic losses, such as job losses. However, evidence surrounding this phenomenon is not clear cut, particularly due to the empirical challenges of studying how piracy influences innovation. For instance, Danaher and Smith (2017) find that while piracy did not lead to fewer movies being created, a nuanced shift in the types of movies being created, suggesting a shift from riskier titles to those that are more mainstream. Alternatively, Waldfogel and Aguiar (2018) find that music piracy and music sharing did not impact the supply of new music, suggesting that piracy did not have a measurable impact on innovation. Telang and Waldfogel (2018) find that the growth of piracy in Bollywood was associated with a decrease in the creation of new titles, suggesting that piracy led to a reduction in incentives for product development. In a recent paper, Bradley and Kolev (2019) examined the release of pirated versions of computer software through BitTorrent and find that this led to an increase in R&D spending by pirated firms, but a decline and delay in new product releases. The majority of existing studies suggest that piracy influences the demand for a particular product and, therefore, incentives for innovation. However, piracy also removes the exclusive rights of a software

[2]Even if this results in lower firm revenues, it may lead to enhanced consumer welfare as more consumers purchase these products and the monopoly power of the firm over its products decreases.

[3] For example, the Recording Industry Association of America (RIAA) argued that in 2007, the piracy of music and movies resulted in the loss of approximately 200,000 US jobs.

creator to distribute their products, which might create incentives for innovation in different ways. We discuss this potential mechanism in subsequent sections.

3 Theory and Hypotheses

Here, we develop theoretical arguments for how piracy influences incentives for innovation, which in turn translates into firms either updating their products or abandoning them. We focus on two separate sets of mechanisms. First, we consider the arguments put forth by previous literature regarding the impact of piracy on expected revenues and how it may influence innovation. Second, we introduce a new mechanism that has not been considered in previous studies relating to the fact that piracy may create an incentive for firms to innovate in order to escape competition, thus creating an incentive for innovation.

3.1 Piracy and Expected Revenues

As described thus far, software piracy can be thought of as the distribution of software products in ways that are unauthorized by the software creator. Therefore, a market with piracy can be considered a setting where consumers may choose between purchasing a product through an authorized channel or using the unauthorized version of the product, potentially priced at zero (Conner and Rumelt, 1991; Sundararajan, 2004). Intuitively, giving consumers the choice to acquire an unofficial version of the product, often for free, may dissuade some consumers from purchasing the product. This implies that piracy leads to a loss of company revenues, and that firms will expect lower revenues in future periods and therefore will be less likely to innovate. This is a prediction of much of the theoretical literature on piracy (Belleflamme and Peitz, 2012).

One exception where the availability of a pirated software application does not lead to a loss of revenues is when the availability of a pirated product can increase consumer awareness. This is an important mechanism that previous studies have identified in the case of movies (Peukert et al., 2017) and music (Lee, 2018). In fact, this mechanism can be sufficiently strong that firms may choose to create a second "free" version of their products in order to increase awareness and attract more consumers. This is the rationale behind many free-mium products (Sundararajan, 2004; Nan et al., 2018). Yet, this mechanism might not benefit all firms, given the fact that piracy can also reduce revenues. In particular, as Peukert et al. (2017) show, that increased consumer awareness may only benefit those products which are less popular and therefore gain customers from increased awareness, while highly popular "blockbusters" that are well-known are unlikely to gain consumer awareness and therefore likely to only lose customers from piracy. This implies that piracy leads to a decrease in revenues for more popular products, but less so for more popular products. As a baseline, we can expect that higher revenues are associated with higher incentives to undertake innovation, and therefore, we expect piracy to lead to lower levels of innovation overall. However, this effect may be muted or reversed for less popular products.

Hyp 1. *Piracy is associated with an overall decline in innovation.*

Hyp 2. *Piracy is associated with a greater decline for more popular products, in comparison to less popular products.*

3.2 Piracy and Conditions to "Escape Competition"

Existing studies of piracy have overwhelmingly argued that expected revenues map to incentives for innovation, as described in the hypothesis above. Yet, from the literature on industrial organization, there are theories that suggest that lower expected revenues may lead to higher incentives for innovation. In a setting without piracy, the exclusive distributor of a particular software title, the software owner may have little incentive to create an improved (updated) version or a new product, as doing so would cannibalize their existing revenues (Arrow, 1962). In a setting with piracy, the revenues of the software owner firm may be lower, potentially even down to zero, as consumers can choose between the legal and pirated versions of these products. This also implies that the firm would not cannibalize any existing revenues if it chooses to update its products or develop new products. This suggests that incentives for innovation may be higher based solely on the fact that there are less existing revenues to cannibalize in the case of piracy. However, there is an additional mechanism that may enhance incentives for innovation.

In a setting with piracy, firms may experience lower revenues. Yet, if they were to successfully develop a new product or revise their existing products such that consumers have a reason to purchase the new product rather than the existing pirated version, then firms could have an incentive to innovate. This is analogous to what Bloom et al. (2005) describe as the "escape competition" effect, whereby firms have an incentive to innovate in order to improve their offering with respect to competitors and thus face less competition. This concept can be traced back to Schumpeter (1939), who argued that firms will often undertake innovation to gain an advantage over its competitors and keep innovating in order to sustain that advantage.[4]

If this is a factor influencing innovation, then it directly shapes the types of innovations that are created. For instance, in order for a firm to create an updated product that consumers would prefer rather than the existing pirated version, they would have to create a product that is a substantial improvement on the existing pirated version. Ellison and Fudenberg (2000) documented this strategy in the early package software market as an avenue for companies to push consumers to upgrade to newer products. Minor or incremental innovations, would not be helpful in this regard, as they would be expensive to create, while only a small fraction of consumers would be willing to purchase an incrementally improved version rather than use a pirated version that may be available for free. This implies that piracy can create a disincentive to create minor or incremental innovations such as bug fixes, but an incentive to create new products or major revisions.[5]

[4]Schumpeter specifically discussed the case of escaping competition to become a monopolist in the marketplace. In this case, while the focal firm may not be a monopolist, the underlying mechanisms are similar to that of Schumpeter.

[5]In the appendix we present a simple theory model that arrives at the same set of predictions.

Hyp 3. *Piracy leads to a decline in the release of incremental innovations (minor bug fixes), but less so for more substantial innovations (feature updates or new products).*

The practice of releasing incremental innovations such as bug fixes is a critical aspect of software development, as firms often race to get their products out into the marketplace, often with bugs (minor problems in the code). These problems are resolved later with patches or updates (bug fixes) that help to improve the experience of consumers by eliminating these errors. These bug fixes are often released shortly after a product is updated, once users identify problems with the product. Therefore, firms often release bug fixes shortly following the release of their products. Alternatively, firms may not want to wait to release larger feature updates or new products, as releasing these could cannibalize the potential revenues that firms could capture from their existing products (Arrow, 1962). Therefore, the age (or maturity) of a product may influence the types of revisions that are released. Given that we expect piracy to lead to a decline in the release of incremental innovations such as bug fixes rather than more substantial innovations, we might expect that piracy leads to a decline in innovation shortly after the release of a product, but not in the longer term when new products or major product revisions might be expected.

Hyp 4. *Piracy leads to a decline in innovations released shortly after a product has been brought to market, but less so after products have been in the market for a longer period of time.*

4 Empirical Context and Data

4.1 Empirical Setting

We study the impact of piracy on innovation within a prominent mobile application marketplace (or App Store), known as the "cydia" marketplace and a piracy event that occurred within this community. Mobile application marketplaces have become an important area of economic activity and an important setting for digital innovation (Yin et al., 2014). Additionally, these marketplaces have been critical to the success of the smartphone platforms that they complement. The Jailbreak marketplace is a storefront for mobile applications, or apps, that were designed for iOS devices (iPhone, iPad, iPod Touch, etc.) but were not allowed by Apple to be sold through the App Store. Apple did not allow these applications on their devices because they modified a part of the functionality or operating system that Apple did not want to allow consumers to modify. For instance, in the early days of the iPhone, Apple did not allow consumers to modify their backgrounds, icons, or other cosmetic features of their products and would therefore disallow the App Store from listing these types of software.[6] The Jailbreak marketplace served as an outlet for these applications and existed in parallel to Apple's App store from 2009 to 2018. While this marketplace is only a fraction of the size of the official App Store marketplace (approximately fifty-thousand applications, and ten-thousand developers) it was an important context for innovation, as many of the software features now common in mobile telephony first appeared there.

The jailbreak marketplace was organized around a storefront, cydia through which customers could download applications directly through their phones as with other application marketplaces such as the App Store. The files that customers could download were stored on a number of repositories (online storage providers) that would host the files so that they could be downloaded. In July of 2014, an announcement was made that a hacking group named

[6]While in principle this could have facilitated the release of "illegal" software, this was, in fact, only a tiny fraction of the overall community, and not the set of observations studied in this paper. The products studies were applications utilizing functionality that Apple did not want to make customizable to consumers.

"Kim Jong Cracks" had hacked and copied all of the files held on the major repositories and released these pirated versions for free through their own app. This meant that the entire population of applications that had previously been sold had been pirated and made freely available on that day. This was not expected by the content developers.

This setting provides a unique context in which to evaluate how piracy influenced innovation outcomes in a way that was not possible in earlier studies and allows us to consider how piracy influences different innovation decisions–in particular, whether developers choose to continue updating their products or whether they stop updating and abandoned their products.

4.2 Data Collection

We assembled data from a number of different sources. We collected information on individual applications (products) directly from the cydia storefront, as well as a history of product updates online sources that track this information. The software titles that were pirated prior to the hacking attack were available on a number of file-sharing websites that allowed users to install these applications. From these websites, we extracted a list of files that allowed us to identify which titles had been pirated prior to the hacking shock. The hacking itself also exposed the internal data of the marketplace, including the number of times each file was downloaded prior to the hacking attack. Combining these different data sources, we were able to observe each title, the date of release, subsequent updates, the price of product, the file size and product description, the market niche (category), the identity of the developer, the number of downloads prior to the hacking attack, and whether the title had been pirated prior to the hacking attack. Data on the Android marketplace was collected from an online repository which contained corresponding information for all available applications.

4.3 Empirical Design

Our empirical approach is based on a difference-in-differences design around the piracy event described above. Below, we describe the main elements of our empirical strategy.

4.3.1 Measuring Product Revisions and Product Abandonment.

Our conceptualization of innovation is based on whether developers are likely to create a revision or improvement on their existing products. We operationalize this by looking at whether an update is released for a particular product. If an update is released, this indicates that a product had been revised, while if no updates were released, this indicates that a product has been abandoned. Empirically, we focus on whether a particular product has been updated in any given month. A decline in the number of updates (or probability of an updating event each month) indicates a decline in innovation and a shift towards product abandonment.[7]

4.3.2 Counterfactual (Control) Groups for Analysis.

We consider the affected (treated) products in our setting as those which were sold on the platform (non-zero price) prior to the hacking event but had not been previously pirated. We utilize several potential control groups for this analysis, as described below.

Previously Pirated Applications: There is a subset of paid (non-zero price) products that were pirated prior to the hacking event. When the hacking event occurred, these developers did not experience a drastic shift in their incentives through the mechanisms previously described, as they had already experienced this. Interviews with members of the community who had suffered from piracy confirmed that this hack did not affect them financially. At the same time, the products which were hacked had been comparable to those which were not previously hacked, as described in Table 1.

Free Applications: A considerable share of the Jailbreak marketplace is based on free titles. This piracy event did not harm the payoffs of these developers as they by definition were not charging for their products. Advertisements and other sources of revenues were limited in this marketplace. Additionally, the non-monetary benefits of participating in this community were not visibly impacted in a way that was different from other developers.

Android Applications: Android applications are, in many ways, comparable to those on

[7] There are very few instances where more than one product revision is released in the same period.

the Jailbreak marketplace. However, the Android marketplace was not at all affected by this piracy event and serves as a control group. One limitation of this control group is that the Android marketplace was in its ascent, while the Jailbreak community was in its decline, therefore, it is necessary to control for time trends when using this as a control group. Additionally, the Jailbreak marketplace is more volatile, while the Android marketplace follows a more stable cycle.

Time Window and Unit of Analysis. The unit of analysis for our econometric results is at the product-period level. We are interested in whether particular products are updated or abandoned, which makes this an appropriate unit on which to focus. We test the robustness of our analysis to various time windows.

One important factor to consider when selecting the time window is that developers on the Jailbreak marketplace often release their products to coincide with a new hardware (iPhone) release or an operating system (iOS) release. Therefore, it is important to consider whether we are including or excluding these events when selecting our time window. There are several hardware and operating system releases around the piracy event, creating fluctuations in the release of product updates (see Figure 1). We limited our sample for analysis to one year (12 months pre / post) around the piracy event in order to include only those events that occurred close to the piracy event and may have been relevant. This does include a hardware and software update release so that we can observe whether developers wait until that event to update their software. Our results are not sensitive to changing the time windows, as the strongest change occurs directly after the piracy event.

Accounting for Potential Bias. There are numerous sources of potential bias in the analysis that we attempt to address. First, there may be various unobserved factors in the environment or the overall success of the platform that influence the decisions of developers to update. The difference-in-differences design allows us to account for these given that they are equally likely to affect the control group. Additionally, there may be product or firm specific heterogeneity influences specific developers. We address this first by including fixed effects that account for much of the unobserved heterogeneity across different products, as well as testing for heterogeneity in the effects across different covariates, such as product age and popularity. Another potential issue may be that there exist differences between those products that were previously pirated and those that are affected by this piracy attack. This is something we attempt to account for by testing the robustness of our results across different control groups. Finally, we perform a number of econometric tests, such as placebo event robustness tests, and an aggregation to a single pre/post period in order to validate the results.

5 Analysis and Results

5.1 Descriptive Results

In Figure 1, we provide a breakdown of the average number of product updates over a period of 20 months, before and after the piracy attack, for each of the control groups. We also present a synthetic counterfactual of the different control groups. It is clear that there is considerable variability in terms of the intensity with which products are updated over time. This is largely due to the fact that new hardware (Apple iPhones) of software (iOS updates) at the platform level were released during this time window (in Periods -20, -5, and 5). This is shown with vertical lines that indicate when new hardware was released and when a hack (or jailbreak), enabling the use of jailbroken applications became available. It is clear that in anticipation of these jailbreaks, developers started creating applications to ensure that they would became available just after the jailbreak was released. It is reassuring that the volatility of the treatment and control groups moves in step for treatment and control groups prior to the piracy attack.

< Figure 1 About Here >

The overall decline can be attributed to the fact that many of the technological reasons that led innovators to join this platform (inability to change elements of iOS software) were now available on the official Apple App Store, prompting developers to shift to that platform. This further reinforces the rationale behind the need for a difference-in-differences framework that controls for the overall decline in the trend over this period. Looking at the change in the trend in Figure 1, it is clear that a decline occurs in the treatment group after the piracy shock (Period 0). Even though over time, the level of updates in the first two control groups also declines, the reduction in the treatment group is far more drastic. This can be clearly seen by looking at the observations around Period 5 when a new operating system was released, prompting developers to release new updates. There was a surge in updating by the control groups, but not by the treatment group. The Android control group remains flat during this period, which is in part why the other specifications represent the preferred control groups. While on the one hand, this provides descriptive evidence in support of H1, it also highlights the importance of our counterfactual for identifying the mechanisms purported in the remaining hypotheses.

< Table 1 About Here >

In Table 1, we present a breakdown of key characteristics of our treatment and control groups. The level of updating within the treatment group is between that of the two control groups. However, the treatment group does appear to release fewer major updates than the other two groups. Additionally, the firms in the treatment group appear to have slightly higher ratings than Free Titles (Control Group B), although less than previously pirated titles (Control Group A). Importantly, the treatment group appears to have received many more downloads (orders of magnitude greater) than the control group (particularly Control Group A) prior to the piracy shock. This difference suggests that piracy does lead to a decline in downloads, and therefore a decline in revenues. This also suggests that our treatment groups were unaffected by piracy prior to the piracy event, while our control groups (particularly Control Group A) were not likely to have been influenced by the piracy event since they have already lost many of their customers to piracy.

5.2 Regression Results

Our regression analysis is based on a difference-in-differences approach. The outcome variable is, as described above, an indicator for whether a given product (i), was updated in a given period (t), in this case, a month. The variable *Post* refers to the period after the piracy event. The variable $Treated$ distinguishes the treatment group from the different control groups: A) titles with a non-zero price which were previously pirated, B) titles which were sold for zero price (free), and C) Titles available on Android. The rationale for the control groups is that the treatment and control groups will be equally affected by unobserved factors, with the exception of the treatment (in this case, the piracy event) which only influenced the treated group. The difference between them can be thought of as a causal estimate of the impact of the piracy effect on a particular outcome, in this case, the number of updates released. Since this empirical design depends so heavily on the assumption that our control group is similar to our treatment group prior to the piracy event, we use a number of different control groups to validate our results. Our basic regression model is as follows.

$$Update_{it} = \alpha + \beta_1 Post_{it} + \beta_2 Post_{it} \times Treated_{it} + \mathbf{P}_{it}\delta + \mathbf{T}_{it}\gamma + \epsilon$$

We include product and time period fixed effects, denoted by $\mathbf{P}$ and $\mathbf{T}$, respectively. As described above, we restrict our analysis to one year (12-month periods) before and after the shock. The coefficient of *Post* captures the overall trend that is occurring for the control group across the two periods. The product fixed effects capture the differences in the propensity with which each application is updated, which is why the $Treated$ variable is not included on its own. The key variable of interest is the interaction term β_2, which captures the difference in the response of the treatment group relative to the control group. This initial analysis provides us with an estimate of whether piracy led to an overall increase or decrease in product updates. We extend this analysis in later sections to attempt to identify different mechanisms that may be driving these results.

< Table 2 About Here >

In Table 2, we present regression results, reporting bootstrapped standard errors in all columns. We begin with OLS regressions of only the baseline variable in Column 1 and the interaction in Column 2. Here, we utilize only those products that were previously pirated (Control Group A) as the control group in the sample along with the treatment group. While we find a decline in the baseline term ($\beta = -0.218$, $S.E. = 0.04$, $p = 0.000$), which we might expect given the decline in updates that we observe over time, the interaction term is also negative ($\beta = -0.062$, $S.E. = 0.020$, $p = 0.001$) suggesting a further decline following the piracy event. In Columns 3, we present the results of a placebo test by including a "placebo shock" at a period of six months (halfway through the pre-shock period) in the regression. The coefficient for the interaction term for the placebo test ($\beta = 0.052$, $S.E. = 0.03$, $p = 0.152$), suggesting that the negative effect is not driven by spurious differences in the trend between the treatment and control groups. In Columns 4 and 5, we repeat the models from Columns 1 and 2, but with a logistic model. The results are comparable in Columns 1 and 2 (See Figure 3 for marginal Effects). As a further robustness test, we repeat the analysis with the alternative control groups in Columns 6 through 9. The results remain consistent across these specifications.

5.2.1 Heterogeneous Response by Product Popularity

H2 predicts that piracy may increase innovation for less popular products, in comparison to more popular products. It is unclear what is the most appropriate way of quantifying the popularity of a product including that variable in a regression. To ensure that we have selected the correct variable construction would require experimenting different variables. However, this is not straightforward as testing different interactions will provide unreliable estimates.[8] To do this in a statistically valid way, we utilize the approach developed by Wager

[8] This is because the test statistics for linear regression are not designed for "data mining" where a number of different variables are tested and explored.

and Athey (2018), which provides an unbiased way of estimating the average treatment effect for a large number of potential covariates. This approach uses a random forest model to estimate the impact of the policy event on individual outcomes. Since the random forest approach is non-parametric, it allows us to predict the outcome of the model across different subgroups. We can then plot these predicted values with respect to different variables to understand how these variables influence the estimates. This enables us to estimate the average treatment effect (ATE) and confidence intervals at different levels of any chosen variable, without performing data mining in a way that would invalidate our inference.

< Figure 2 About Here >

In terms of choosing the covariates to analyze, we constructed several variables that may conceptually be related to the different mechanisms, such as measures of popularity (Average Daily Downloads) and quality (Product Quality Rating). We include measures of developer and product age, developer downloads, number of products released, number of different market niches entered, as well as various measures of the dispersion (maxima, minima, etc.) of product ratings and downloads.

In Figure 2, we present the magnitude of the treatment effect at different levels of Product Age, along with the magnitude of the effect at different levels of Average Daily Download and product ratings. Importantly, from Figure 2, it is clear that the magnitude of the treatment effect is not changing at different levels of either of these variables; however, for Product Age, we find that for younger products (lower age) the treatment effect is stronger (more negative) than for older products. After five months, the magnitude of the treatment effect becomes indistinguishable from zero (Supporting H4). However, we do not find any differences across the popularity or quality variables, suggesting that more niche products are differently affected by piracy (Not Supporting H2).

< Table 3 About Here >

To validate whether our non-parametric evidence is consistent with a difference-in-differences regression, in Table 3, we repeat the earlier regressions from Table 2, but include an interaction for *product downloads*, *product ratings,* and *product age.* Here, the value is collapsed to one observation pre/post the piracy event to avoid having multiple observations for time-invariant variables such as downloads or ratings.[9] We do not find that any of the effects for popularity are statistically distinguishable from zero, consistent with the results from Figure 2. However, we do find evidence treatment effect is stronger for younger (more recently released) products ($\beta = -0.662$, $S.E. = 0.210$, $p = 0.002$) most likely due to the fact that most product updates are released within four months of product release.

5.2.2 Feature Updates versus Bug Fixes

H3 predicts that piracy can create an incentive for firms to revise their products in order to create a new product that will lead consumers to purchase the updated version, instead of using the freely available pirated version, and that this will manifest itself through fewer incremental innovations and more substantial innovations such as product revisions. We define incremental innovations as minor tweaks that are released in order to correct minor issues or account for incompatibility issues. These are often referred to as bug fixes – for instance, shifting from Version 1.0.0 to Version 1.0.1 of a product. We define substantial innovations as feature updates or other major product revisions (such as a shift from Version 1.0 to Version 2.0) where the software application experiences a considerable improvement in quality or user experience. We distinguish between incremental and more substantial innovations based on the version number, whether it was a substantial shift in the first digit of the version, or a shift in later numbers, indicating only a minor change according to version numbering conventions.

We repeat the analysis from the earlier regression for both Feature Updates and Bug Fixes, and estimate the results together using a seemingly-unrelated-regression estimator. This approach allows us to account for the correlation between the two models over time.

[9] Variables constructed as ratings and downloads at the time of piracy rather than varying over time.

For instance, if the release of a major update was unlikely to coincide with the release of a minor update, this approach would account for this correlation.

< Table 4, Figure 3 About Here >

In Table 4, we present the regression results comparable to those in Table 2, but stratified for feature updates and bug fixes. We present these results with each of the different control groups in our analysis. Across the different specifications, we find that piracy is associated with decline in the propensity that minor updates are released. However, we do not find evidence that there is a decline in major updates across any of the specifications. Marginal Effects Reported in Figure 3.

There is an overall decline in updating for both the treatment and control groups, as indicated by the earlier descriptives and regression results. However, the decline in absolute terms of the number of bug fix updates in relation to the number of feature updates appears to be considerably greater. For instance, while prior to the piracy event the probability of a major and minor update release is comparable, following the piracy event there is no decline the release of feature updates, while there is a 25% greater decrease in the release of minor bug fixes. This suggests that the piracy event had a much stronger negative effect on minor updates such as bug fixes but an ambiguous effect on larger feature updates.

5.3 Robustness Tests and Additional Specifications

One concern with the earlier set of results is that none of the control groups follow the treatment group perfectly prior to the piracy event. As a robustness check, we constructed a single "synthetic" control. The synthetic control is plotted in comparison to the treatment group in Figure 1. We use the weights from the synthetic control variables to reweigh the observations in our regression analysis, and use all three control groups within the same regression. These results are shown in Table 5 and are consistent with the earlier results. As an additional check, we use matching on the trend prior to the having event

< Table 5 About Here >

As a further check, we reconstruct our dataset, so that for each title, there is only one period before and after the piracy shock. This is done to ensure that the results are not being biased by the large number of observations for each product (Bertrand et al., 2004). We find that the results are consistent with the earlier analysis, across each of the control groups and for both feature and bug fix updates. (See Appendix for all robustness check results)

An additional concern may be that developers are responding to an overall decline in the platform. While this is a concern, the baseline (or control) group in the difference differences analysis would capture this overall effect. The interaction term, would capture if our treatment group was subject to an additional reduction controlling for the overall decline, due to the loss of revenues that resulted as a function of piracy. Furthermore, the particular hacking attack may be a unique case of piracy where all products are pirated simultaneously. This might lead to a competitive effect between multiple pirated products, rather than simply an effect of piracy. While this is an obvious concern, it is again something that is also true of the control groups as well as the treatment group. Therefore, our control group would capture the resulting impact of increased competition, while our main interaction terms capture the additional decline in updating that occurs as a result of piracy.

5.4 Impact of Piracy on New Product Development

One potential factor that is important to consider is whether the piracy event led to new product development rather than product updates or revisions. On the one hand, a decline in new product development would be consistent with the idea that piracy is leading to a decline in innovation. If, on the other hand, piracy is leading to an increase in new product development, this provides evidence consistent with H3, whereby piracy leads to an erosion of monopoly power of existing products and creates an incentive for new product development. To test whether this is the case, we replicated our difference-in-differences regression, with the number of products launched, by developer, as the outcome variable. In Table 6, we present the results of these regressions for each of the control groups. We include time trends to account for a potential decrease in new product releases, especially with the third control group where there is potentially a considerable divergence in the pre-shock trends.

< Table 6 About Here >

Columns 1 and 4 present the results for the difference-in-differences regressions for both the unaffected control group (A), and the free application control group (B). In both cases, we find that there was an increase in product releases relative to the control groups ($\beta = 0.733, p = 0.000; \beta = 1.137, p = 0.000$). This suggests that the piracy event led to an increase in new product development. Columns 5 and 6, present the results for the third control group of Android applications. Here, the coefficient is positive but the standard errors are quite large ($\beta = 0.062, p = 0.694$). This suggests that the impact of the piracy event on innovation is noisy and we cannot conclude that it leads to new product development. This does provide support for H3, suggesting that piracy does lead to a decline of minor bug fixes, but an increase (or less of a decrease) of major updates such as feature updates.

6 Discussion and Conclusion

In this paper, we study how piracy influence incentives for innovation, and particularly whether the violation of digital copyright through piracy leads to product abandonment and a decline in innovation, or whether it leads firms to undertake new innovation efforts. Our theoretical predictions follow the mechanisms purported by earlier studies (Peukert et al., 2017; Givon et al., 1995) and look at whether piracy simply leads to a reduction in revenues, or whether it creates diffusion and, in turn, an incentive for innovation. We also consider a mechanism that had not been considered in the case of piracy: while piracy may reduce the profits of companies, it may also remove the prospect of cannibalizing existing profit streams.

This might, in turn, lead firms to develop new products or seek more radical reinventions of their current products. Our results indicate that piracy is associated with an overall decline in product updates, but mostly in terms of minor bug fixes and tweaks. Rather than lending support to one view (increasing or decreasing innovation), these results provide nuance into how piracy is actually influencing software innovation and new product development. The fact that we find across multiple specifications that piracy does lead to a decline in minor bug fixes, suggests that once they are pirated firms stop providing support and maintenance for their pirated products. Instead, they delay their efforts until another major update is released.

We do not find evidence consistent with the idea that piracy increases product diffusion, as earlier studies have found. One reason may be that in our empirical setting, all products are simultaneously being pirated rather than any single product. Therefore, it is not any particular product that is receiving an increase in diffusion, but competing products are similarly diffused to consumers (Peukert et al., 2017; Givon et al., 1995). While this is an important feature of our empirical design that allows us to test many of our hypotheses, it does suggest something about the limits of when piracy can create product diffusion, unlike conditions where piracy creates an incentive for firms to innovate in order to attract consumers to purchase new products.

The results of this paper provide evidence constistent with a small number of earlier studies of how piracy may influence innovation rather than, simply, expected payoffs of firms (Telang and Waldfogel, 2018; Danaher and Smith, 2017). These results contribute to our understanding of how property rights shape innovation outcomes in digital settings (Luo and Mortimer, 2017; Nagaraj, 2018). This in turn informs a broader set of questions about whether property rights simulate innovation and under which conditions (Teece, 1986). The present study contributes to a number of recent papers that have considered the implications of property rights on incentives for innovation (MacGarvie and Moser, 2015; Moser, 2005; Cockburn and MacGarvie, 2011). However, unlike these earlier papers which focus on property rights from the perspective of policy, the present paper considers the implications of digital copyright, and in turn, the violation of those property rights through piracy, a question that affects the strategy of platform firms which try to optimize innovation, "on top of their platforms" (Jacobides et al., 2018; Boudreau and Jeppesen, 2015). The insights from this paper have less to do with optimal strategies for pirated firms themselves to regulate piracy (Conner and Rumelt, 1991; Dey et al., 2019), but more for the platform itself that can benefit or be harmed by the presence of piracy. For instance, the results of the present study suggest that piracy may shape innovative effort, but that this may be observed mostly through a decline in incremental innovations such as bug fixes. This suggests that the decline of incremental innovations could lead to more unrefined products with potential bugs and errors. While this might not influence larger innovations, it could influence consumer experiences because of bugs or other performance issues, which may indirectly lead to fewer users.

6.1 Generalizability and Limitations.

Our results provide insights into how piracy influences innovation, and a test of an exogenous piracy event to understand how it influences innovation. While our analysis does provide a more stringent test of this effect than earlier studies, it does look at a unique setting of a platform based marketplace, on a relatively niche platform. One of the features of these platforms is that they provide democratized access to a "long tail" of third parties, including many highly specialized or niche complementers. While these complementers are much smaller than the types of firms whose products are typically pirated, the products studied in this context are representative of "long tail" products.

The magnitude of the effects from studying small developers on top of a platform may differ from those of larger firms in the broader economy. Yet, the mechanisms identified in this study are a more general phenomenon that are likely to apply in broader settings. For instance, the results identified here are consistent with the patterns found by Branstetter et al. (2016), where the entry of Paragraph IV generic drugs shifted the innovative output of branded pharmaceutical companies, towards expanding their product portfolio. Similarly, Buccafuso and Heald (2013) take up the question of whether anyone would invest in developing audiobooks on content once the copyright term has expired, as this may be difficult to protect and easier to imitate, but instead find increased investment once the copyright term has expired.

Relating to digital piracy more broadly the mechanisms studied in this paper may provide insights for digital property rights in different settings. Consider, for instance, the case of open-source software, where property rights such as copyright are not readily enforced. Insights from this paper could apply in this context, and could help explain why open source software is often buggy (containing errors which are not fixed by bug fixes), while being highly innovative (similar to software that has been pirated). Yet, commercial software, on the other hand, may not receive major revisions as frequently in order to avoid affecting their revenues, while their products are often less buggy (have more frequent bug fix updates) than many open source applications.

6.2 Implications for Platform Strategy

Software piracy is often discussed as a policy issue. However, as we have argued thus far, the growth of platforms and software innovation, has made this an important strategy issue for platforms as they devise policies to foster innovation. The present paper suggests that the actual implications of piracy are more nuanced.

From the perspective of platform strategy, this implies that allowing software piracy may not stifle innovation and creativity. Yet, it also implies that there will be a decline in incremental innovations such as bug fixes. While these may not be critical to innovation, they improve the experience for customers. This suggests that markets that are subject to piracy will inherently have more buggy and less user-friendly products, even though they may be innovative.

Acknowledgements. The authors are grateful to Joseph Raffiee, Florenta Teodoritis, Daniel Fehder, Nan Jia, Paul Adler, the participants of the O&S Workshop and Yingying

Fan, and the participants of the iOrb Statistics Mini-Conference for their helpful comments. Jinhong Lu provided excellent research assistance for this project.

References

Adermon, A. and Liang, C. Y. (2014). Piracy and music sales: The effects of an anti-piracy law. *Journal of Economic Behavior and Organization*, 105:90–106.

Afuah, A. (2013). Are Network Effects Really All About Size? The Role of Structure and Conduct. *Strategic Management Journal*, 34:257–273.

Aguiar, L., Claussen, J., and Peukert, C. (2018). Catch Me If You Can: Effectiveness and Consequences of Online Copyright Enforcement. *Information Systems Research*, 29(3):656–678.

Appel, G., Libai, B., and Muller, E. (2018). On the monetary impact of fashion design piracy. *International Journal of Research in Marketing*, 35(4):591–610.

Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pages 609–626. Princeton University Press, New Jersey, NY.

Athey, S. and Stern, S. (2015). The Nature and Incidence of Software Piracy: Evidence from Windows. In Goldfarb, A., Greenstein, S., and Tucker, C., editors, *Economic Analysis of the Digital Economy*, pages 443–477. University of Chicago Press.

Bai, J. and Waldfogel, J. (2012). Movie piracy and sales displacement in two samples of Chinese consumers. *Information Economics and Policy*, 24(3-4):187–196.

Batikas, M., Claussen, J., and Peukert, C. (2019). Follow the money : Online piracy and self-regulation in the advertising industry. *International Journal of Industrial Organization*, 65:121–151.

Belleflamme, P. and Peitz, M. (2012). Digital Piracy : Theory. In *The Oxford Handbook of the Digital Economy*. Oxford University Press.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should we Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, (February):249–275.

Bloom, N., Blundell, R., Howitt, P., Boone, J., Carlin, W., David, P., Eberly, J., Ranque, D., Shankerman, M., Solow, R., and Trajtenberg, M. (2005). Competition and Innovation : An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2):701–728.

Boudreau, K. J. (2012). Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation. *Organization Science*, 23(5):1409–1427.

Boudreau, K. J. and Hagiu, A. (2008). Platform Rules: Multi-Sided Platforms as Regulators.

Boudreau, K. J. and Jeppesen, L. B. (2015). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal*, 36(12):1761–1777.

Bradley, W. A. and Kolev, J. (2019). Can Piracy Increase Innovation? The Software Industry's Response to Online File Sharing. *Working Paper*.

Branstetter, L., Chatterjee, C., and Higgins, M. J. (2016). Regulation and welfare: evidence from paragraph IV generic entry in the pharmaceutical industry. *RAND Journal of Economics*, 47(4):857–890.

BSA (2016). BSA Global Software Survey.

Buccafuso, C. and Heald, P. J. (2013). Do Bad Things Happen When Works Enter the Public Domain?:Empirical Tests of Copyright Term Extension. *Berkeley Technology Law Journal*, 28(1):1–43.

Cennamo, C. and Santalo, J. (2013). Platform Competition: Strategic Trade-offs in Platform Markets. *Strategic Management Journal*, 34:1331–1350.

Cockburn, I. and MacGarvie, M. (2009). Patents, Thickets and the Financing of Early Stage Firms : Evidence from the Software Industry. *Journal of Economics & Management Strategy*, 18(3):729–773.

Cockburn, I. M. and MacGarvie, M. J. (2011). Entry and Patenting in the Software Industry.

Management Science, 57(5):915–933.

Cohen, J. E. and Lemley, M. A. (2001). Patent scope and innovation in the software industry. *California Law Review*, 89(1):3–4.

Cohen, W. M., Nelson, R. R., and Walsh, J. P. (2000). Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not).

Conner, K. R. and Rumelt, R. P. (1991). Software Piracy: An Analysis of Protection Strategies. *Management Science*, 37(2):125–139.

Danaher, B. and Smith, M. D. (2014). Gone in 60s: The impact of the Megaupload shutdown on movie sales. *International Journal of Industrial Organization*, 33(1):1–8.

Danaher, B. and Smith, M. D. (2017). Digital Piracy, Film Quality and Social Welfare. *George Mason Law Review*, 23:923–938.

Dey, D., Kim, A., and Lahiri, A. (2019). Online Piracy and the "Longer Arm" of Enforcement. *Management Science*, 65(3):1173–1190.

Eisenmann, T., Parker, G. G., and Van Alstyne, M. (2011). Platform Envelopment. *Strategic Management Journal*, 32:1270–1285.

Ellison, G. and Fudenberg, D. (2000). The Neo-Luddite's Lament: Excessive Upgrades in the Software Industry. *The RAND Journal of Economics*, 31(2):253–272.

Ethiraj, S. and Zhou, Y. M. (2019). Fight or flight? Market positions, submarket interdependencies, and strategic responses to entry threats. *Strategic Management Journal*, pages 1545–1569.

Ethiraj, S. K., Levinthal, D., and Roy, R. R. (2008). The dual role of modularity: Innovation and imitation. *Management Science*, 54(5):939–955.

Galasso, A. and Schankerman, M. (2015). Patents and Cumulative Innovation: Causal Evidence From The Courts. *The Quarterly Journal of Economics*, 130(1):317–369.

Giorcelli, M. and Moser, P. (2017). Copyrights and Creativity - Evidence from italian Operas.

Givon, M., Mahajan, V., and Muller, E. (1995). Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion. *Journal of Marketing*, 59(1):29.

Godinho de Matos, M., Ferreira, P., and Smith, M. D. (2017). The Effect of Subscription Video-on-Demand on Piracy: Evidence from a Household-Level Randomized Experiment. *Management Science*, 64(12):5461–5959.

Goldfarb, A., Greenstein, S., and Tucker, C. (2014). Introduction to "Economics of Digitization". In *Economics of Digitization*, pages 13–14. University of Chicago Press.

Goldfarb, A., Greenstein, S. M., and Tucker, C. E. (2015). Introduction. In Goldfarb, A., Greenstein, S., and Tucker, C., editors, *Economic Analysis of the Digital Economy*, pages 1–35. University of Chicago Press.

Graham, S. J. H., Merges, R. P., Samuelson, P., and Sichelman, T. (2009). High Technology Entrepreneurs and the Patent System: Results of the 2008 Berkeley Patent Survey. *Berkeley Technology Law Journal*, 24(4):1255–1327.

Greenstein, S., Lerner, J., and Stern, S. (2013). Digitization, innovation, and copyright: What is the agenda? *Strategic Organization*, 11(1):110–121.

Hall, B. H. and Ziedonis, R. H. (2001). The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995. *The RAND Journal of Economics*, 32(1):101.

Heald, P. (2014). How Copyright Keeps Works Disappeared. *Journal of Empirical Legal Studies*, 11(4):829–866.

Huang, P., Ceccagnoli, M., Forman, C., and Wu, D. (2012). Appropriability Mechanisms and the Platform Partnership Decision: Evidence from Enterprise Software. *Management Science*, 59(1):102–121.

Ishihara, M. and Muller, E. (2020). Software Piracy and Outsourcing in Two-Sided Markets. *Quantitative Marketing and Economics*, 18:61–124.

Jacobides, M. G., Cennamo, C., and Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, 39(8):2255–2276.

Kapoor, R. and Lee, J. M. (2013). Coordinating and Competing in Ecosystems: How Organizational Forms Shape New Technology Investments. *Strategic Management Journal*, 34:274–296.

Koetsier, J. (2018). The Mobile Economy Has A $17.5B Leak: App Piracy.

Kretschmer, T. and Peukert, C. (2020). Video Killed the Radio Star? Online Music Videos and Recorded Music Sales. *Information Systems Research*.

Lee, J. F. (2018). Purchase, Pirate, Publicize: Private-Network Music Sharing and Market Album Sales. *Information Economics & Policy*, 42:35–55.

Lerner, J. (2009). American Economic Association The Empirical Impact of Intellectual Property Rights on Innovation : Puzzles and Clues. *American Economic Review*, 99(2):343–348.

Leung, T. C. (2013). What is the True Loss Due to Piracy? Evidence from Microsoft Office in Hong Kong. *Review of Economics and Statistics*, 95(3):1018–1029.

Lieberman, M. B. and Asaba, S. (2006). Why Do Firms Imitate Each Other ? *Academy of Management Review*, 31(2):366–385.

Luo, H. and Mortimer, J. H. (2017). Copyright Enforcement: Evidence from Two Field Experiments. *Journal of Economics and Management Strategy*, 26(2):499–528.

MacGarvie, M. and Moser, P. (2015). Copyright and the Profitability of Authorship: Evidence from Payments to Writers in the Romantic Period. In Goldfarb, A., Greenstein, S., and Tucker, C., editors, *Economic Analysis of the Digital Economy*, pages 357–379. University of Chicago Press.

McIntyre, D. P. and Srinivasan, A. (2013). Networks, Platforms, And Strategy: Emerging Views and Next Steps. *Strategic Management Journal*, (38):141–160.

McIntyre, D. P. and Subramaniam, M. (2009). Strategy in Network Industries: A Review and Research Agenda. *Journal of Management*, 35(6):1494–1517.

Miric, M., Boudreau, K. J., and Jeppesen, L. B. (2019). Protecting Their Digital Assets. *Research Policy*, 48(8).

Moser, P. (2005). American Economic Association How Do Patent Laws Influence Innovation ? Evidence from Nineteenth-Century World's Fairs. *The American Economic Review*, 95(4):1214–1236.

Moser, P. (2013). Patents and Innovation: Evidence from Economic History. *Journal of Economic Perspectives*, 27(1):23–44.

Nagaraj, A. (2018). Does Copyright Affect Reuse? Evidence from Google Books and Wikipedia. *Management Science*, 64(7):2973–3468.

Nan, G., Wu, D., Li, M., and Tan, Y. (2018). Optimal Pricing of Information Goods under the Freemium Strategy in the Presence of Piracy and Network Effects. *Journal of the Association for Information Systems*, 19(4):266–305.

Novos, I. and Waldman, M. (1984). The Effects of Increased Copyright Protection: An Analytic Approach. *Journal of Political Economy*, (92):236–246.

Oberholzer-Gee, F. and Strumpf, K. (2007). The Effect of File Sharing on Record Sales : An Empirical Analysis. *Journal of Political Economy*, 115(1):1–42.

Parker, G. and Van Alstyne, M. (2017). Innovation, Openness, and Platform Control. *Management Science*, 64(7):3015–3032.

Parker, G. G., Van Alstyne, M. W., and Jiang, X. (2017). Platform Ecosystems: How Developers Invert the Firm. *MIS Quarterly*, 41(1).

Peukert, C., Claussen, J., and Kretschmer, T. (2017). Piracy and box office movie revenues: Evidence from Megaupload. *International Journal of Industrial Organization*, 52:188–215.

Posen, H. E. and Martignoni, D. (2018). Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity. *Strategic Management Journal*, 39:1350–1369.

Reimers, I. (2016). Can Private Copyright Protection be Effective? Evidence from Book Publishing. *The Journal of Law and Economics*, 59(2):411–440.

Rob, R. and Waldfogel, J. (2006). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *Journal of Law & Economics*, (49):29–62.

Rob, R. and Waldfogel, J. (2007). Piracy on the silver screen. *The Journal of Industrial Economics*, 55(3):379–395.

Sakakibara, M. and Branstetter, L. (2001). Do Stronger Patents Induce More Innovation? Evidence from the 1998 Japanese Patent Law Reforms. *Quarterly Journal of Economics*, 32(1):77–100.

Schumpeter, J. A. (1939). *Capitalism, Socialism and Democracy.* McGraw-Hill, New York.

Shapiro, C. and Varian, H. R. (1998). *Information rules: a strategic guide to the network economy.* Harvard Business Press, Cambridge, MA.

SIIA (2018). Internet Anti-Piracy Program.

Sinha, R. K., Machado, F. S., and Sellman, C. (2010). Don't Think Twice , It's All Right : Music Piracy and Pricing in a DRM-Free Environment. *Journal of Marketing*, 74(2):40–54.

Smith, M. D. and Telang, R. (2019). Competing with Free: The Impact of Movie Broadcasts on DVD Sales and Internet Piracy. *MIS Quarterly*, 33(2):321–338.

Sundararajan, A. (2004). Managing digital piracy: Pricing and protection. *Information Systems Research*, 15(3):287–308.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6):285–305.

Telang, R. and Waldfogel, J. (2018). Piracy and new product creation : A Bollywood story. *Information Economics and Policy*, 43:1–11.

Tunca, T. I. and Wu, Q. (2013). Fighting Fire with Fire : Commercial Piracy and the Role of File Sharing on Copyright Protection Policy for Digital Goods. *Information Systems Research*, 24(2):436–453.

Wager, S. and Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Waldfogel, J. (2012a). Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster. *The Journal of Law and Economics*, 55(4):715–740.

Waldfogel, J. (2012b). Copyright Research in the Digital Age : Moving from Piracy to the Supply of New Products. *American Economic Review: Papers & Proceedings*, 102(3):337–342.

Waldfogel, J. and Aguiar, L. (2018). Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music. *Journal of Political Economy*, 126(2):492–524.

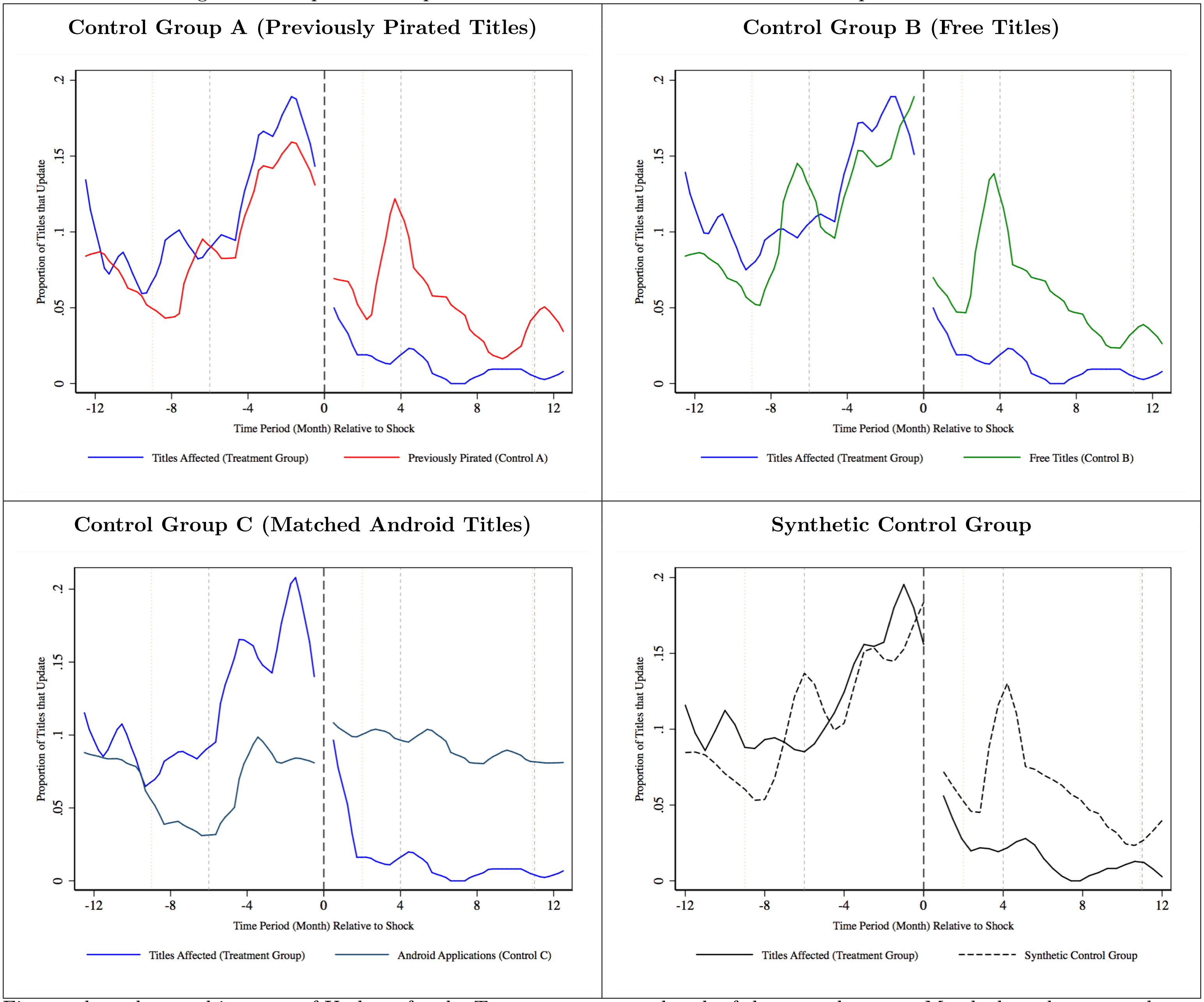
Wang, Q., Li, B., and Singh, P. V. (2015). Copycats versus Original Mobile Apps: A Machine Learning Copycat Detection Method and Empirical Analysis. *Information Systems Research*, 29(2):273–291.

Wang, R. D. and Shaver, J. M. (2014). Competition Driven Repositioning. *Strategic Management Journal*, 35(11):1585–1604.

Wen, W. and Zhu, F. (2019). Threat of platform-owner entry and complementor responses: Evidence from the mobile app market. *Strategic Management Journal*, 40(9):1336–1367.

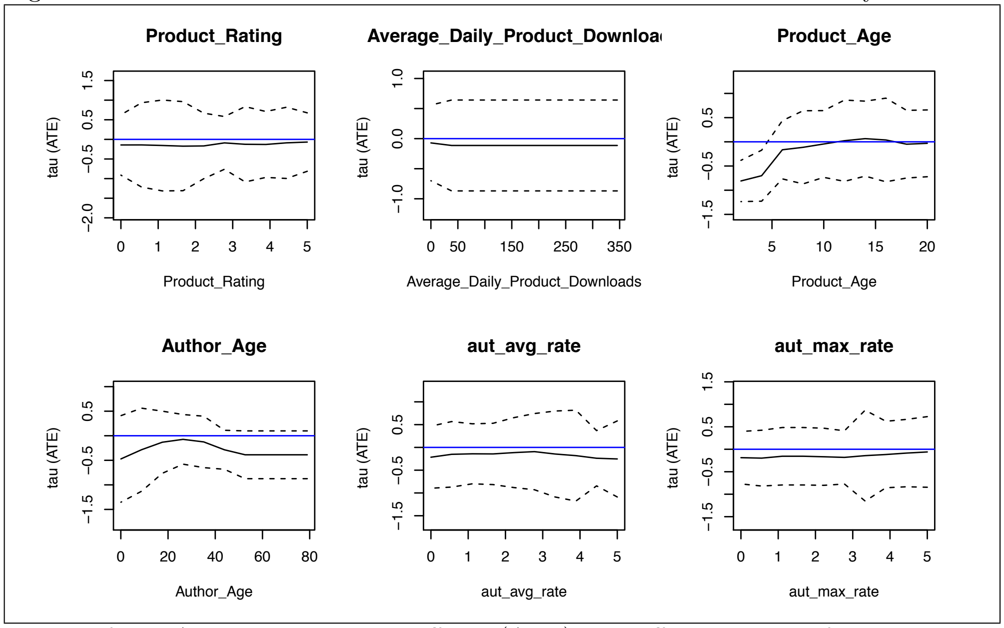
Yin, B. P.-l., Davis, J. P., and Muzyrya, Y. (2014). Entrepreneurial Innovation : Killer Apps in the iPhone Ecosystem. *American Economic Review: Papers and Procedings*, 104(5):255–259.

Figure 1: Comparison of Update Releases for Treatment and Control Groups Over Time



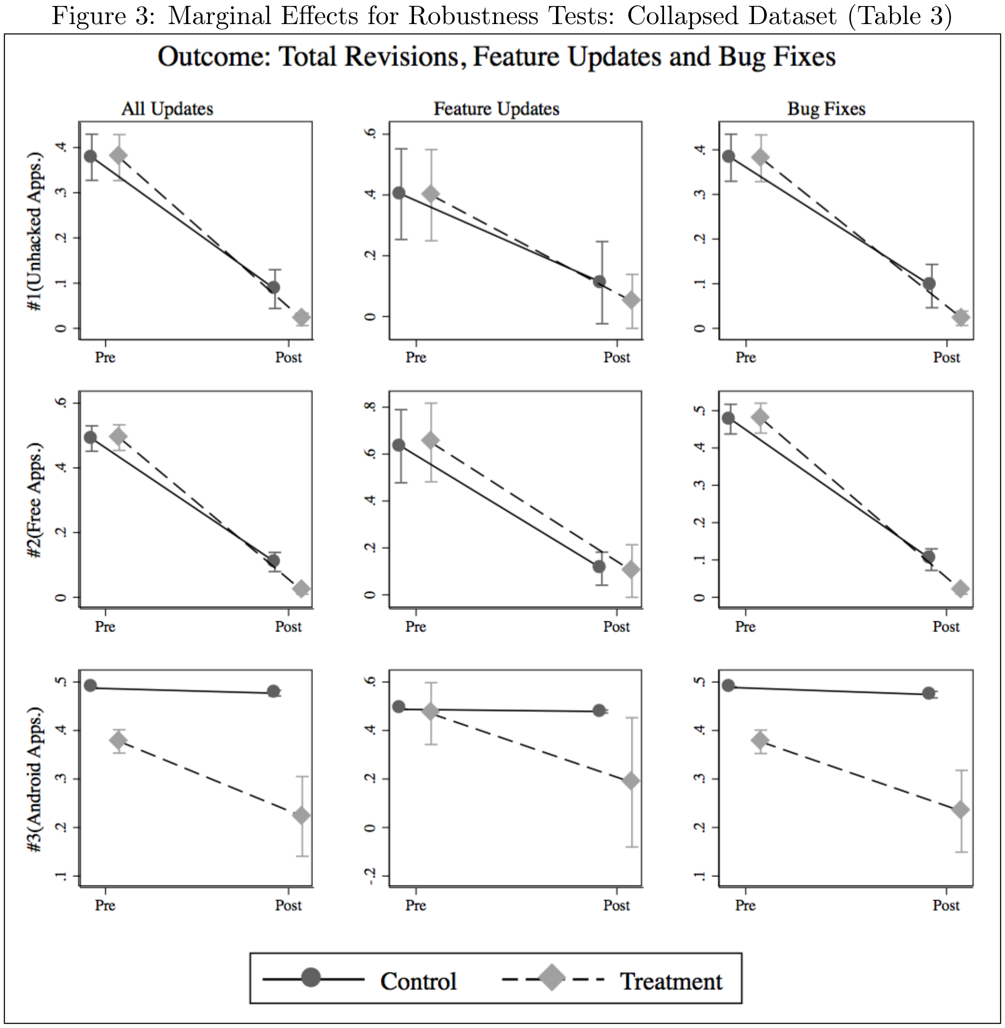
Figures show the trend in terms of Updates for the Treatment group and each of the control groups. Matched results reported. Matching performed separately for each control group. Figures with longer trend and unmatched groups shown in Appendix.

Figure 2: Results of Random Forest Estimate of Treatment Effects for Key Variables



Representation of the Average Treatment Effects (ATE), at different levels of a particular outcome variable based on the approach of Wager and Athey (2018). Previously unhacked products as the control group. Plotted values can be interpreted as marginal effects (magnitudes and Standard Errors) of the treatment variable β. Horizontal blue line indicates the zero level. Magnitudes and standard errors can be interpreted as statistically different from zero if the confidence intervals (dashed lines) lie outside of the zero level (blue line).

Figure 3: Marginal Effects for Robustness Tests: Collapsed Dataset (Table 3)



Marginal effects for logit regressions in Table 3.

Table 1: Comparison of Sample Means for Treatment and Control Groups

		Previously Non-Pirated Titles *(Treatment)*	**Previously Pirated Titles** *(Control A)*	**Free Titles** *(Control B)*
A)	*# of Total Revisions*	1.07 (1.03)	1.29 (1.30)	0.85 (1.13)
B)	*# of Feature Updates (Substantial Revisions)*	0.09 (0.371)	0.12 (0.47)	0.14 (0.47)
C)	*# of Bug Fix Updates (Minor Innovations)*	1.65 (2.46)	2.13 (2.91)	1.18 (2.45)
D)	*Rating (1 - 5)*	2.26 (2.09)	2.44 (1.80)	1.49 (1.96)
E)	*Daily Downloads*	44.48 (104.28)	4.30 (24.07)	5.52 (27.53)
F)	*Product Age (Months Since Release)*	7.26 (3.87)	7.68 (3.77)	7.65 (3.63)

Mean values prior to the piracy shock reported. Standard deviation in brackets.

Table 2: Results of Difference in Difference Regressions
Unit of Observation: *Title (Application) - Period (Month)*
Outcome Variable: *Indicator Product Update Released*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Group	**(A) Previously Un-hacked Applications**					**(B) Free Applications**		**(C) Android Applications**	
Econometric Model	OLS			LOGIT		OLS	LOGIT	OLS	LOGIT
Post	-0.236	-0.218	-0.096	-2.002	-0.853	-0.108	-2.359	-0.108	-0.042
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Post X Treated		-0.062	-0.077	-1.594	-1.599	-0.082	-1.888	-0.123	-0.928
		(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)
Placebo (Six Months Prior)			-0.126		-1.150				
			(0.004)		(0.000)				
Placebo X Treated			0.052		0.030				
			(0.152)		(0.915)				
Constant	0.312	0.314	0.312			0.138		0.251	
	(0.000)	(0.000)	(0.000)			(0.000)		(0.000)	
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,494	11,494	11,494	11,494	11,494	33,051	33,051	494,155	494,155
F/χ^2	839.58	2180.57	1764.55	917.82	1334.12	2572.71	1609.05	724.28	3023.73
$p-score$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.07	0.07	0.07			0.03		0.01	
$log-likelihood$	-2616.53	-2609.49	-2607.83	-2715.77	-2715.76	321.28	-6589.11	-2.4×10^5	-2.0×10^5

p values reported in parentheses, based on boostrapped standard errors. Odds ratio reported for LOGIT results

Application (product) level fixed effects, and time period (month) fixed effects are included in all models. Sample limited to only those titles that were released prior to the hacking event.

Table 3: Results of Difference-in-Differences Regressions for Feature versus Bug-Fix Updates
Unit of Observation: *Title (Application) - Period (Month)*
Outcome Variable: *Indicator Product Update Released (Feature Update and Bug Fixes)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome Variable	**Feature Updates**				**Bug Fix Updates**			
Control Group	(A)	(A)	(B)	(C)	(A)	(A)	(B)	(C)
Econometric Model	OLS	LOGIT			OLS	LOGIT		
Post	-0.023	-2.063	-3.538	-0.036	-0.196	-1.912	-2.409	-0.600
	(0.096)	(0.377)	(0.604)	(0.008)	(0.000)	(0.000)	(0.00)	(0.000)
Post X Treated	-0.002	-0.945	-0.230	-1.337	-0.059	-1.559	-1.894	-2.140
	(0.799)	(0.696)	(0.753)	(0.199)	(0.000)	(0.000)	(0.000)	(0.008)
Constant	0.028				0.288			
	(0.050)				(0.000)			
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,138	2,138	6,226	429,217	11,494	11,494	30,233	481,199
F/χ^2	338.17	100.11	233.86	1407.62	29.43	935.80	1417.99	1488.16
$p-score$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.01				0.06			
$log-likelihood$	-316.42	9919.48	-867.36	-1.5×10^5	-2394.02	-2611.33	-6076.46	-1.9×10^5

p values reported in parentheses, based on boostrapped standard errors.
Marginal effects reported in Figure 3.

Table 4: Results of Difference in Difference Regressions - Testing for Heterogeneity
Unit of Observation: *Title (Application) - Period (Pre/Post)*
Outcome Variable: *Number of Product Updates Released*

	(1)	(2)	(3)
Post	-1.201	-0.711	-0.654
	(0.013)	(0.000)	(0.000)
Post X Treated	-2.230	-1.340	-0.312
	(0.117)	(0.002)	(0.501)
Post X Downloads	0.100		
	(0.048)		
Post X Treatment X Downloads	0.129		
	(0.316)		
Post X Ratings		0.126	
		(0.001)	
Post X Treatment X Ratings		0.052	
		(0.681)	
Post X Product Age			0.290
			(0.000)
Post X Treatment X Product Age			-0.662
			(0.002)
N	400	1,168	1,186
χ^2	23.44	83.40	186.35
$p-score$	0.00	0.00	0.00
$log-likelihood$	-219.96	-582.85	-557.27

p values reported in parentheses, based on robust standard errors.

Table 5: Results of Difference in Difference Regressions with Synthetic Control Group
Unit of Observation: *Title (Application) - Period (Month)*
Outcome Variable: *Indicator Product Update Released*

	(1)	(2)	(3)	(4)
Control Group	**Synthetic Control Group**			
Outcome Variable	**All Updates**		**Feature**	**Bug Fix**
Sample	**UnWeighted**	**Weighted**	**Weighted**	**Weighted**
Post	-0.098	-0.125	-0.061	-0.105
	(0.000)	(0.000)	(0.00)	(0.000)
Post X Treated	-0.138	-0.134	0.020	-0.132
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.280	0.254	0.159	0.204
	(0.000)	(0.000)	(0.000)	(0.000)
Time Dummies	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes
N	558,185	558,185	558,185	558,185
F	125.07	150.42	83.12	112.58
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.01	0.02	0.00	0.01
$log-likelihood$	-2.5×10^5	-1.9×10^5	-8.9×10^4	-1.3×10^5

p values reported in parentheses, based on boostrapped standard errors. OLS model used.

Following the procedure developed by Abadie, Diamond & Hainmueller (2010),we arrive the the following sampling weights: Group A (Previously Pirated), 0.242; Group B (Free Applications), 0.692; Group C (Android Applications), 0.066. Weights introduced in the OLS Fixed Effects regression to reflect a "synthetic" control group. Weights for the treatment group set to 1.

Table 6: Results of Count Model (Poisson) Regressions for New Product Releases
Unit of Observation: *Author (Developer Firm) - Period (Month)*
Outcome Variable: *Count of Product Releases (New Products)*

	(1)	(2)	(3)	(4)	(5)	(6)
Control Group	**A) Un-hacked Apps.**		**B) Free Apps.**		**C) Android Apps.**	
Post	-0.260	-0.891	-0.350	-0.574	-0.005	-0.011
	(0.006)	(0.000)	(0.000)	(0.000)	(0.830)	(0.487)
$Affected\,Developer\ \times Post$		0.733		1.137		0.062
		(0.000)		(0.000)		(0.694)
$Time\,Trends$	Yes	Yes	Yes	Yes	Yes	Yes
$Firm\,FE$	Yes	Yes	Yes	Yes	Yes	Yes
N	11,731	11,731	39,542	39,542	47,358	47,358
χ^2	1003.87	1696.65	6507.82	3437.44	43.43	44.37
	0.00	0.00	0.00	0.00	0.00	0.00
$log-likelihood$	-4.6×10^4	-6.6×10^4	-1.6×10^4	-1.1×10^4	-3.2×10^4	-3.2×10^4

p values reported in parentheses, based on boostrapped standard errors.

Data aggregated to firm (developer) level, including an indicator for whether the developer was affected by the piracy event ($Affected\,Developer$). Results are consistent accounting for zero-inflation. Controlling for time trends helps to equalize the trends prior to piracy event.

Appendix A: Summary of Robustness Tests

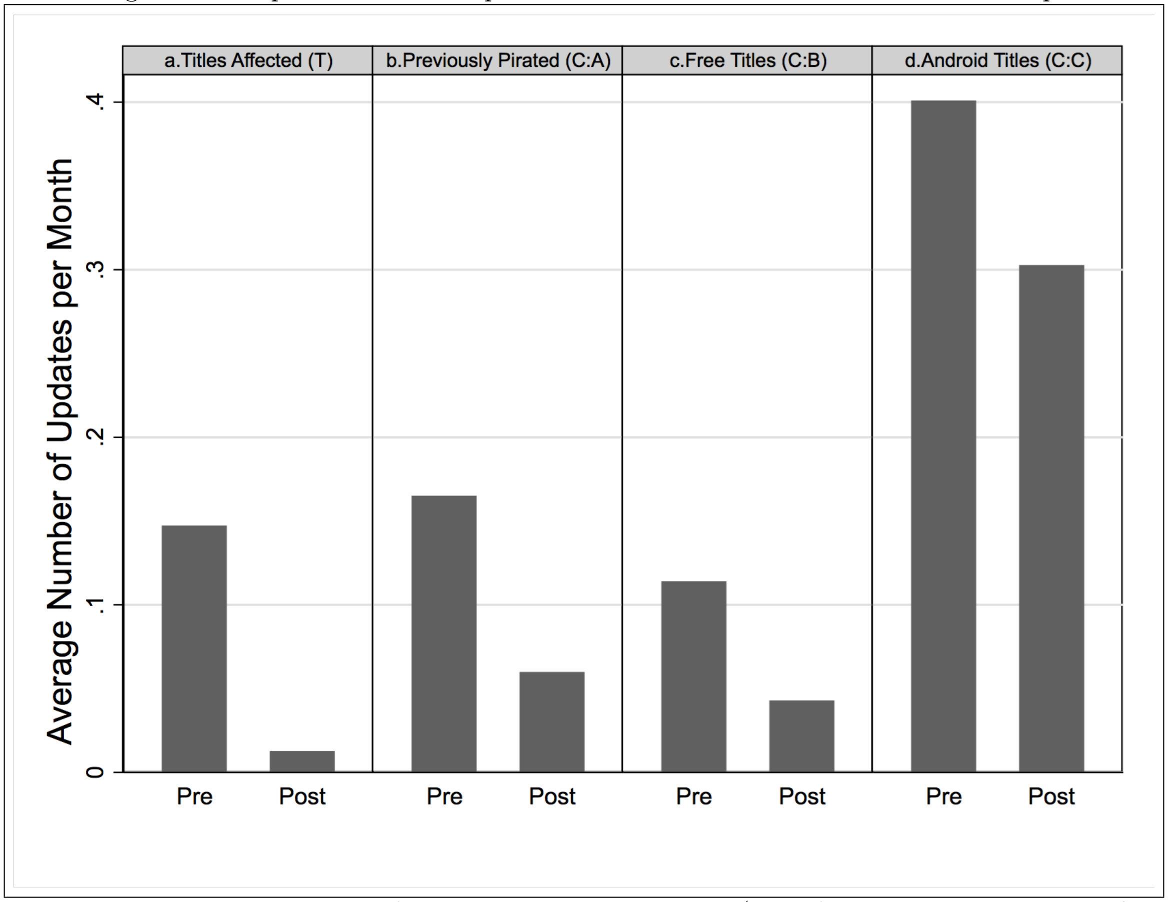
In the Figure below we provide an overview of the various robustness tests performed in this analysis.

Table 1: Summary of Robustness Tests

Concern	**Solution**	**Supported**
Choice of time window influencing outcome.	Test results with different time windows: 6 months, 12 months and 18 months pre / post hacking event. Results remain consistent.	Hyp 1, 3, and 4
Multiple observations for each variable could inflate statistical significance, as described by Bertrand et al. (2006).	Following suggested approach of Bertrand et al. (2006), we use bootstrapped standard errors and collapse the data to a single observation pre-post shock.	Hyp 1, 3, and 4
Similarity of treatment and control groups, particularly in pre-trend.	Tested robustness of results with multiple control groups. Comparison of treatment and control groups on observables (Table 1) and trend (Figure 1 & 2). Placebo test six months prior to hacking event, as test for violation of parallel trend assumption.	Hyp 1, 3, and 4
Modeling of outcome variables (Variables can be constructed as counts, continuous variables, discrete outcomes) may influence statistical outcomes.	Results discussed are robust to count, linear (OLS) and discrete (LOGIT) specifications. Results are reported using the most appropriate model, but robustness was confirmed with each of the alternative specifications described.	Hyp 1, 3, and 4
Heterogeneous impact of piracy is sensitive to variable construction (measuring experience differently may alter significance and direction of results)	We use the Wager and Athey (2018) approach to test a battery of different potential variable constructions as described in later sections. For those variables where there appears to be a relationship, we test this in a parametric regression (Table 4).	Hyp 1, 3, and 4

Appendix B: Supplementary Tables and Figures

Figure 1: Comparison of Total Update Releases for Treatment and Control Groups



Here we present the Proportion of Products that are updated (both feature updates and bug fixes) in a given month (12 Months Pre / Post) for the treatment group and the three control groups. The treatment group declines by 93%, while the previously pirated control group (A) declines by 60%, the free titles control group (B) declines by 42%, the android control group (C) declines 24%. This provides very stark evidence that following the piracy event there was a greater decline in the release of product revisions by the affected developers. It is also important to note, that these declines in the two control groups are also quite substantial. This decline may have likely been caused by the uncertainty and potential threat of future piracy, that the hacking attack symbolized (as well as potentially a decline in the platform). This highlights the importance of using these two control groups for our analysis, because these conditions were common across the treatment and control groups. Therefore, the effect that explains why piracy may have had a 33% higher effect on our treatment group, may be attributed to the variables that we hypothesize about.

Figure 2: Comparison of Trend In Total Updates Released for Treatment and Control Groups
Extended Time Window (18 Month Window). Raw (Unmatched) Sample.

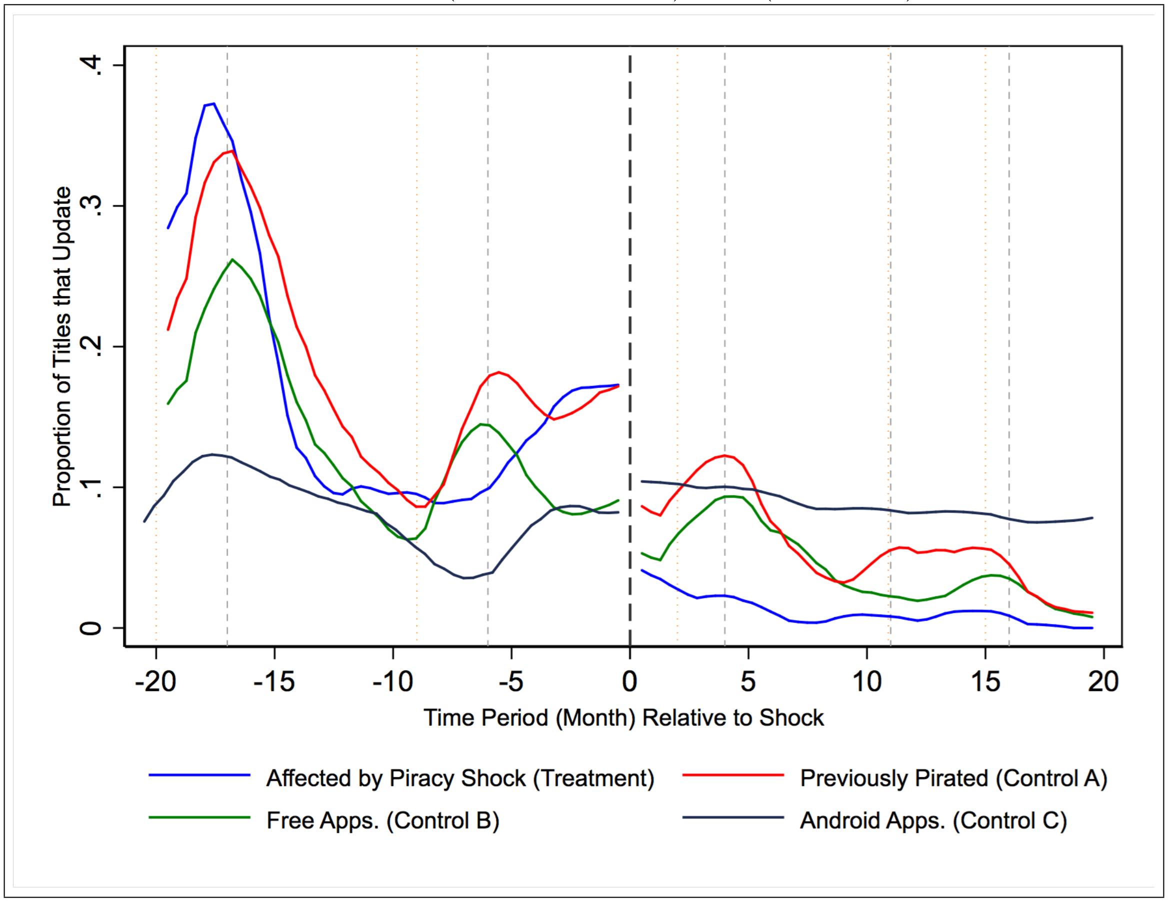


Figure 3: Comparison of Update Releases for Treatment and Control Groups Over Time
Extended Time Window (18 Month Window) with Matched Sample

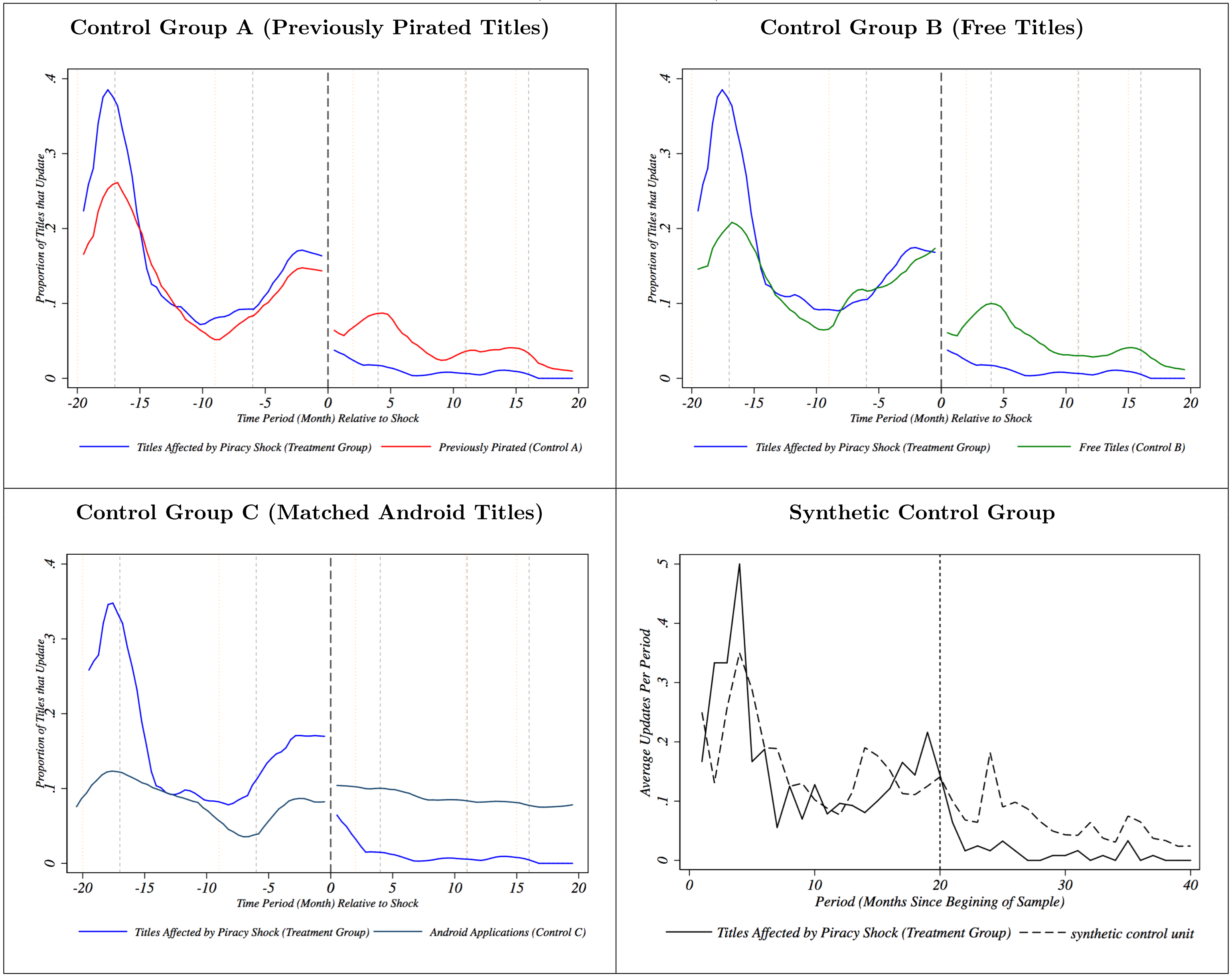
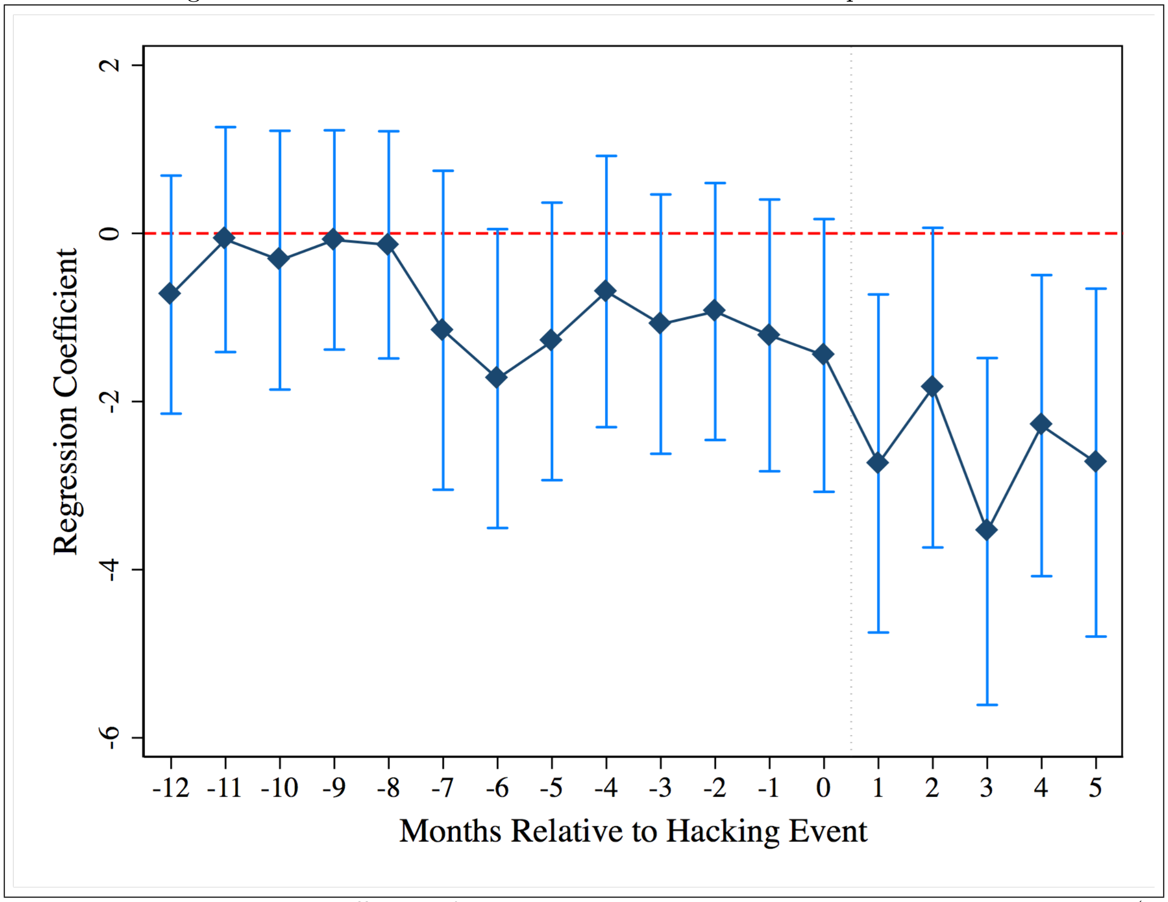


Figure 4: Coefficient Plot for Treatment and Control Groups Over Time



Here we present regression coefficients for the interaction between Period and Treatment groups (i.e. The regression we estimate can be thought of as $Pr(Update) = \textbf{Month}\gamma + \textbf{Treatment} \times \textbf{Month}\beta + \epsilon$.) Plotted values are the vector of β values for the logit model that capture the changes in the outcome for the treatment variable holding constant monthly variation (i.e. Period Fixed effects). There is a decline in the coefficients following the hacking event suggesting that the piracy event lead to a decrease in revisions in the treatment group. This appears to be particularly strong in periods 1 and 3, but less so in Period 2. Looking at the raw breakdown in Figure 4 of the appendix, helps provide us with an explanation. In Period 2 there was a general decline in updating across both the treatment and control groups prior o a new hardware release. In the periods that followed, there was a surge in Updating by the control groups to create products for the new hardware release, but there was no such increase for the treatment group. From the above figure, there does appear to be a downward trend in the coefficients from Period 6, onwards. We use a placebo test with a "Placebo Shock" in Period 6 in Table 2, to test whether this is in fact the case and whether this is driving the effects. We find that this is in fact not the case (placebo coefficients positive and significant).

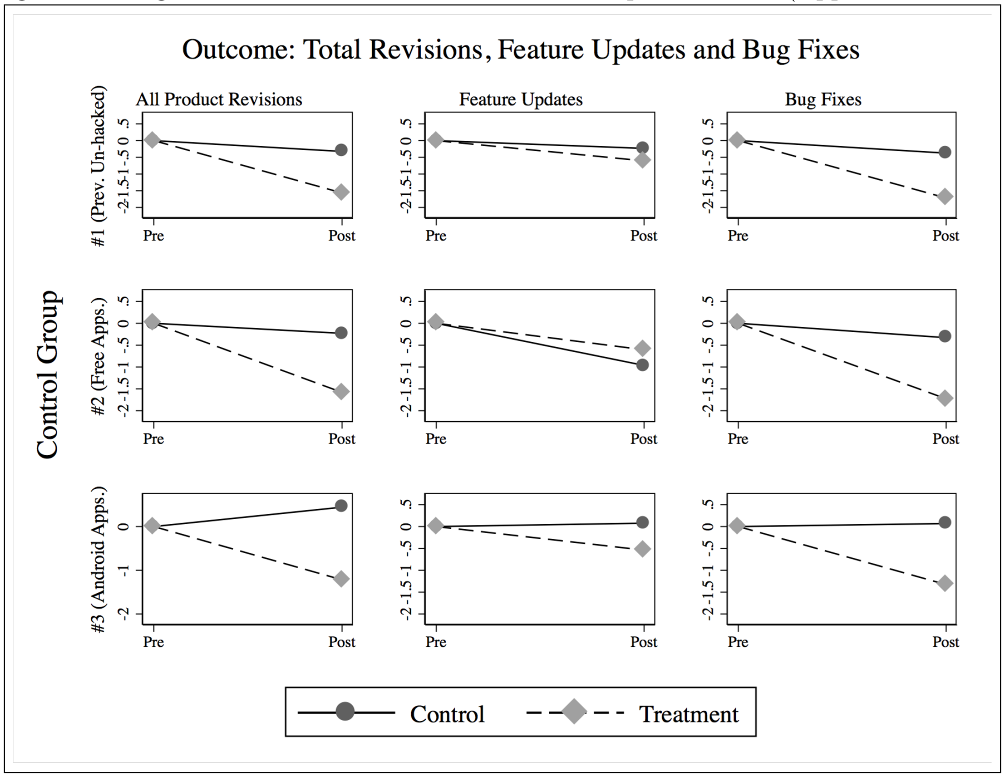
Table 2: Results of Collapsed Difference in Difference Regressions
Unit of Observation: *Title (Application) - Period (Pre/Post)*
Outcome Variable: *Number of Product Updates Released (Feature Update and Bug Fixes)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Control Group	**Un-hacked Applications**				**Free Applications**				**Android Applications**			
Outcome Variable	**All Updates**		**Feature**	**Bug Fix**	**All Updates**		**Feature**	**Bug Fix**	**All Updates**		**Feature**	**Bug Fix**
Post	-0.43	-0.32	-0.23	-0.38	-0.29	-0.23	-0.96	-0.33	0.44	0.44	0.08	0.07
	(0.07)	(0.07)	(0.20)	(0.08)	(0.04)	(0.04)	(0.12)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)
Post X Treated		-1.24	-0.37	-1.34		-1.35	0.36	-1.40		-1.66	-0.62	-1.40
		(0.26)	(0.73)	(0.31)		(0.25)	(0.71)	(0.30)		(0.22)	(0.64)	(0.28)
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1186	1186	226	1138	3350	3350	588	3090	49,998	49,998	54,492	49,432
χ^2	43.30	62.71	2.06	55.81	46.69	69.26	62.87	73.04	3566.48	3618.33	32.70	47.37
	0.00	0.00	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R^2												
$log - likelihood$	-617.69	-597.23	-92.85	-852.11	-1547.11	-1521.02	-206.57	-1823.29	-4.6×10^4	-4.6×10^4	-3.1×10^4	-1.3×10^4

Robust standard errors in parentheses

These results provide a robustness tests for the results in Table 2. One criticism of difference in difference specifications is that having multiple observations pre / post the hacking event can inflate significance statistics. The results here follow the suggested robustness tests of Bertrand et al. (2006), which involve collapsing the data to a single observation before and after the hacking event. Collapsing the data in this way greatly reduces the number of observations, which is why the number of observations differs from those in Tables 2 - 3. Results are consistent with those of Tables 2 - 3 in sign and statistical significance. Results consistent with alternative control groups (Shown in Appendix)

Figure 5: Marginal Effects for Robustness Tests: Collapsed Dataset (Appendix Table 2)



Appendix C: Matching and Synthetic Controls

Table 3: Results of Matched Difference-in-Differences Regressions
Unit of Observation: *Title (Application) - Period*
Outcome Variable: *Number of Product Updates Released (Feature Update and Bug Fixes)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Control Group	(A) Previously Un-hacked Apps.			(B) Free Apps.			(C) Android Apps.		
Outcome Variable	All	Feature	Bug Fix	All Updates	Feature	Bug Fix	All	Feature	Bug Fix
Post	-0.29	-0.02	-0.27	-0.11	-0.01	-0.10	0.04	-0.10	-0.05
	(0.04)	(0.01)	(0.04)	(0.02)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)
Post X Treated	-0.06	-0.00	-0.05	-0.05	0.00	-0.05	-0.16	-0.01	-0.20
	(0.02)	(0.01)	(0.02)	(0.02)	(0.00)	(0.01)	(0.03)	(0.02)	(0.02)
Constant	0.36	0.03	0.33	0.15	0.01	0.14	0.15	0.20	0.12
	(0.03)	(0.01)	(0.03)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8650	8650	8650	36725	36725	36725	228695	228695	228695
F	24.61	2.76	23.25	40.07	6.43	34.93	74.71	60.75	32.67
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.07	0.01	0.07	0.04	0.01	0.04	0.00	0.01	0.01
$log-likelihood$	-1494.02	8340.32	-1319.28	-2848.28	35968.89	-1685.71	-7.7×10^4	-6.0×10^4	1495.48

Robust standard errors in parentheses.
Matching performed using CEM for each control group, based on trend in outcome variable prior to hacking event.

Table 4: Results of Matched Difference in Difference Regressions

Unit of Observation: *Title (Application) - Period*

Outcome Variable: *Number of Product Updates Released (Feature Update and Bug Fixes)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Control Group	**(A) Un-hacked Applications**				**(B) Free Applications**				**(C) Android Applications**			
Outcome Variable	**All Updates**		**Feature**	**Bug Fix**	**All Updates**		**Feature**	**Bug Fix**	**All Updates**		**Feature**	**Bug Fix**
Sample	Full	Matched	Matched	Matched	Full	Matched	Matched	Matched	Full	Matched	Matched	Matched
Post	-0.22	-0.29	-0.02	-0.27	-0.05	-0.11	-0.01	-0.10	-0.01	0.04	-0.10	-0.05
	(0.03)	(0.04)	(0.01)	(0.04)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)
Post X Treated	-0.06	-0.06	-0.00	-0.05	-0.08	-0.05	0.00	-0.05	-0.12	-0.16	-0.01	-0.20
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)
Constant	0.31	0.36	0.03	0.33	0.11	0.15	0.01	0.14	0.25	0.15	0.20	0.12
	(0.03)	(0.03)	(0.01)	(0.03)	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)	(0.01)
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11494	8650	8650	8650	42357	36725	36725	36725	513986	228695	228695	228695
F	31.00	24.61	2.76	23.25	220.56	40.07	6.43	34.93	495.78	74.71	60.75	32.67
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.07	0.07	0.01	0.07	0.01	0.04	0.01	0.04	0.01	0.00	0.01	0.01
$log-likelihood$	-2.6×10^4	-1.4×10^4	8.3×10^4	-1.3×10^4	-89.95	-2.8×10^4	-3.5×10^5	-1.6×10^41	-2.4×10^5	-7.7×10^4	-6.0×10^4	1495.48

Robust standard errors in parentheses

These results provide a robustness tests for the results in Table 2, by using coarsened exact matching (King et al., 2009) to construct a matched sample. One concern with the earlier results may be the lack of a consistent pre-trend between the treatment and control groups. Therefore, we match individually each sample to each control group based on the number of updates launched in three month windows, prior to the piracy event. This matching on the outcome prior to the piracy event is meant to find observations which follow a similar pre-trend in order to satisfy the parallel trend assumption. Groups / Cuts for exact matching were automatically generated by the algorithm. Weights generated by matching algorithm used as weights for regression. Matching performed at the level of the individual application. Results are consistent with those of Tables 2 - 3 in sign and statistical significance. OLS Regressions reported as weights cannot be implemented with Fixed-Effects Logit regression.

Appendix D: Formal Model of Software Piracy and Innovation

Setup of Theoretical Model

We develop a model of a (i) a profit-maximizing monopolist, and (ii) utility-maximizing consumers, in a setting where a freely available, "pirated" version of the product offered by the monopolist may exist. Our model is meant to provide a stylized representation of how piracy may influence incentives for innovation.

The timeline in the model is as follows. First, the software developer offers a product of quality $s > 0$ at a price $p > 0$. The existence of piracy or not, is an exogenous feature of the environment when these decisions are made. The pirated good is offered at zero price and is of equal quality, s, as the product offered by the software developer. Then, each consumer decides whether to buy, use the pirated version, or to not use any product. The decision of the consumer depends on whether the consumers is an ethical type, not willing to use pirated goods, or an un-ethical type willing to use pirated goods, and the price of the good. We solve this game through backward induction, starting with the adoption decisions of consumers.

Consumer Choice

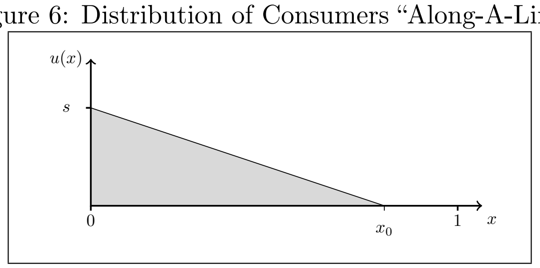
The consumers in the model vary on two dimensions. First, they vary in the degree to which they are "ethical" or willing to use pirated software. We assume that of a mass of 1 consumers, there is fraction β that are ethical and not willing to use the pirated software if it exists, and a fraction $1 - \beta$ that are "not ethical" and willing to use pirated software if it is available.[1] Second, our consumers vary in their "taste" for a product of quality s. Therefore, we model the product being sold as being located on a line at position 0, with consumers uniformly distributed along the line from 0 to 1. As consumers move away from 0 by a distance x, they encounter an adjustment cost of λx.[2] Since software is an experience good, consumers often trust or understand the value of the product, as it is used by a larger number of other consumers. Therefore, we define the "informational benefit" that consumers gain from utilizing a particular software application as δN, where N is the number of both ethical and unethical consumers that utilize the product.[3]

The utility of consumers under each of the three scenarios:

$$u = \begin{cases} s + \delta N - \lambda x - p & if \;\; Consumer\ Buys \\ s + \delta N - \lambda x & if \;\; Consumer\ Pirates \\ 0 & if \;\; Consumer\ Does\ Not\ Use \end{cases}$$

The utility of consumers at different positions, can be observed in the following figure. The key insight of this figure is that consumer utility from using a particular version of the software declines as we move from left to right (increase adjustment cost), until x_0 where the consumer is indifferent between using a particular product and forgoing its use.

Figure 6: Distribution of Consumers "Along-A-Line"



Case of Market With Piracy. In the case of a market without piracy, both ethical and non-ethical consumers will chose between buying a product, and not using any product at all. For both types of consumers, there quantity of consumers that adopt the software will be the point between 0 and the point x_0, where $u_b = u_0$. Since both ethical and non-ethical consumers are symmetrical in their utility functions, the share of consumers N that will adopt the software is $2x_0$. Equating $u_b = 0$, substituting this in, and solving for x_0 yields, $x_0 = \frac{s-p}{\lambda - 2\delta}$. If we substitute $Q = x_0$ into the payoff equation $\pi = \beta p Q_e + (1 - \beta) p Q_{ne}$, and solve for p and Q in equilibrium we arrive at $p* = s/2$, $Q = \frac{s}{2(\lambda - 2\delta)}$.

[1] This is a common characterization in many models of piracy (Belleflamme and Peitz, 2012; Conner and Rumelt, 1991).

[2] This is a common practice for city-on-a-line or hoteling models of competition.

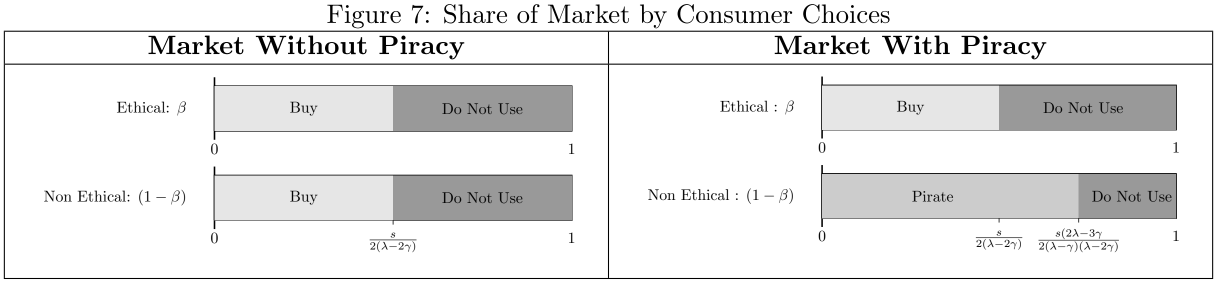
[3] This is an important aspect of many models of piracy (Belleflamme and Peitz, 2012). Empirical evidence also suggests that this is an important variable in shaping consumer decisions under piracy (Peukert et al., 2017).

Case of Market Without Piracy. In the case of a market without piracy, ethical consumers will chose between purchasing a particular product and not using any product at all, while unethical consumers will choose between purchasing a particular product, using a pirated version, and not using any product at all. It follows, that since the quality of the product is s for both types of consumers, and transportation costs are equivalent, that unethical consumers will all choose the pirated version instead of buying the product.[4] If the transportation costs are too high, consumers may still choose to forgo using any product instead of using the pirated version. Since a different share of ethical and non-ethical consumers may use the products, $N = \delta(x_e + x_{ne})$. Solving for the point where $u_p = u_0$,we derive $x_{ne} = \frac{\delta x_e + s}{\lambda + \delta}$. We similarly derive the region where ethical consumers will be indifferent between purchasing a product and forgoing the use of the product ($u_b = u_0$), $x_e = \frac{s - p + \delta x_2}{\lambda - \delta}$. Using the basic payoff function $\pi = \beta p Q_e$, solving for equilibrium prices and quantities yields $p^* = \frac{\lambda s}{2(\lambda - \delta)}$, $Q_e = \frac{s}{2(\lambda - 2\delta)}$. Equilibrium prices, quantities and profits are summarized in the Table below. Consumer choices are illustrated in Figure 2.

Table 5: Summary of Equilibrium Prices, Quantities and Profits

	Market Without Piracy		**Market With Piracy**
p	$\frac{s}{2}$		$\frac{\lambda s}{2(\lambda-\delta)}$
Q_e,Q_n	$\frac{s}{2(\lambda-2\delta)}, \frac{s}{2(\lambda-2\delta)}$		$\frac{s}{2(\lambda-2\delta)}, 0$
x_e, x_{ne}	$\frac{s}{2(\lambda-2\delta)}, \frac{s}{2(\lambda-2\delta)}$		$\frac{s}{2(\lambda-2\delta)}, \frac{s(2\lambda-3\delta)}{2(\lambda-\delta)(\lambda-2\delta)}$
π	$\frac{s^2}{4(\lambda-2\delta)}$		$\frac{\lambda s^2}{4(\lambda-\delta)(\lambda-2\delta)}$

Figure 7: Share of Market by Consumer Choices



The timeline we consider here is that a firm lives for two periods. In the first period, a market exists with or without piracy. Software developers have the possibility of updating their products for the following period. The benefit of updating their products is that they may enhance the value of their products or to limit competition. This innovation decision can be considered an update or revision, whereby the firms releases a product that is different from the pirated version that is available.

Firm Choice and Piracy

Information Diffusion Mechanism. Software is in large part an experience good, whereby consumers need to use the software to fully understand it's value (Shapiro and Varian, 1998). Therefore, firms often haven an incentive to provide free trials, or beta tests in order for consumers to understand the value of the product (Cheng and Liu, 2012; Cheng and Tang, 2010; Niculescu et al., 2018). One potential benefit of piracy is that it provides a free version of the software that customers can use to understand the value of the product, before they purchase the legal version of the software. This is often touted as one of the main mechanisms behind how piracy influences the revenues of firms and in turn their incentives to innovate (Belleflamme and Peitz, 2012; Nan et al., 2018; Smith and Telang, 2010; Tunca and Wu, 2013).

While this is a driving mechanism in many theoretical models, several studies have also found empirical evidence in support of this phenomenon. For instance, Peukert et al. (2017) found that piracy lead to an increase in viewership of more niche products, that would not have gotten exposure without the existence of piracy. Similarly, Waldfogel and Aguiar (2018) found that despite the growth of music piracy in recent years, there was an increase in the supply of music. This can be attributed to the diffusion and discovery of music, that has occurred as a result of music piracy. This diffusion effect creates a boost to revenues from piracy that can offset the "Threat of Future Piracy" and "Loss of Current Revenues", mentioned earlier. Therefore, if this mechanic is at play we would expect that piracy has a positive (or less negative) effect of piracy, for more niche, less popular titles that would be difficult to otherwise observe (Peukert et al., 2017).

[4]This will be the case so long as $p > 0$, such that consumers will always be better of pirating rather than purchasing the paid version.

Formal Representation. The focal firm may choose to innovate on the basis of the magnitude of future payoffs π_{np}, π_p. If the market is not covered (i.e. $Q_e, Q_{ne} < 1$) and $\delta > 0$, then piracy can increase the payoffs. The firm generates higher revenues in the presence of piracy if δ is sufficiently large, namely $\delta > \lambda(1-\beta)$. However, if the market is fully covered then, there is no increase in the size of the consumer base as a result of δ and therefore, firms are worse of under piracy. This directly maps to incentives to undertake innovation in the presence of piracy.

Lemma 1. *Developer revenues are higher under piracy if knowledge diffusion is sufficiently strong ($\delta > \lambda(1-\beta)$) and the baseline value of the product is sufficiently low ($s < 2(\lambda - \delta)$), otherwise developers are better off without piracy.*

Proof. Software developer revenues are higher under piracy than without piracy if $\pi_p > \pi_{np}$. This is true when $\delta > \lambda(1-\beta)$, and when $Q_e \leq 1$. This condition is met if $x_e \leq 1$, which is true if $s < 2(\lambda - \delta)$. □

A naive characterization of incentives for innovation, would suggest then that higher expected revenues map to higher incentives for innovation (We expand on this below). Therefore, the influence of piracy on incentives to innovate through the "information diffusion" channel would suggest that: 1) Piracy lowers revenues for highly valuable products, but 2) piracy can lead to higher revenues for less valuable products if diffusion channel is sufficiently strong.

The intuition here is that if a product is not sufficiently valuable that it is consumed by all consumers, than a firm may generate more revenue in the presence of piracy. This occurs because of this "information diffusion" externality that leads consumers to extract more utility from consumers if they are used by others.

Escaping Competition. A firm may have little incentives to innovate as a monopolist, since after they innovate they will again be a monopolist in a similar situation (Arrow, 1962). However, in the presence of piracy, the software developer may have a greater incentive to innovate in order to "escape competition" and become a monopolist in the future period (Bloom et al., 2005). This also creates an incentive for the types of innovations that are created. In order to "escape competition", firms have an incentive to develop, larger or more drastic innovations that can allow them to become a monopolist in future periods. Therefore, if this mechanism is at play, we would expect that piracy would shift innovation towards larger innovations, and away from smaller more incremental innovations.

Formal Representation. As we have described the case of information diffusion in the paper, we focus on the instance where $\delta = 0$, such that revenues under piracy are always lower than without piracy. This creates a lower bound for our results.

Let, v_d be the improvement in quality that results from innovating if an innovation is drastic, where $v_d > s$. Let, v_s be the improvement in quality that results from innovating if an innovation is not drastic, where $v_s > s$. The returns or benefits from innovating, defined as the difference between revenues without any change, and revenues if an innovation occurs.

		With Piracy	**Without Piracy**
Drastic Innovation		$\frac{v_d s^2 - \beta s^2}{4\lambda}$	$\frac{v_d s^2 - s^2}{4\lambda}$
Non-Drastic Innovation		$\frac{\beta\left[v_{nd} s^2 - s^2\right]}{4\lambda}$	$\frac{v_{nd} s^2 - s^2}{4\lambda}$

Lemma 2. *If $\beta < 1$, then $I_{p,d} > I_{m,d} > I_{m,nd} > I_{p,nd}$*

Proof. Since $\beta s^2 < \beta$, then$I_{p,d} > I_{m,d}$. Since $v_h > v_l$, then $I_{m,d} > I_{m,nd}$. Since $\beta < 1$, then $I_{m,nd} > I_{p,nd}$. □

The intuition here is that with piracy, firms have a greater incentive to innovate and escape competition, than without piracy, simply because it cannibalizes less of their existing revenue.(This is the same argument of Arrow, 1962). However, if the innovation is small such that it does not "escape competition", then there is very little incentive to innovate under piracy the gains are greatly reduced.

References

Belleflamme, P. and Peitz, M. (2012). Digital Piracy : Theory. In *The Oxford Handbook of the Digital Economy*. Oxford University Press.

Bloom, N., Blundell, R., Howitt, P., Boone, J., Carlin, W., David, P., Eberly, J., Ranque, D., Shankerman, M., Solow, R., and Trajtenberg, M. (2005). Competition and Innovation : An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2):701–728.

Cheng, H. K. and Liu, Y. (2012). Optimal software free trial strategy: The impact of network externalities and consumer uncertainty. *Information Systems Research*, 23(2):488–504.

Cheng, H. K. and Tang, Q. C. (2010). Free trial or no free trial: Optimal software product design with network effects. *European Journal of Operational Research*, 205(2):437–447.

Conner, K. R. and Rumelt, R. P. (1991). Software Piracy: An Analysis of Protection Strategies. *Management Science*, 37(2):125–139.

Nan, G., Wu, D., Li, M., and Tan, Y. (2018). Optimal Pricing of Information Goods under the Freemium Strategy in the Presence of Piracy and Network Effects. *Journal of the Association for Information Systems*, 19(4):266–305.

Niculescu, M. F., Wu, D. J., and Xu, L. (2018). Strategic Intellectual Property Sharing : Competition on an Open Technology Platform Under Network Effects. *Information Systems Research*, 29(2):498–519.

Peukert, C., Claussen, J., and Kretschmer, T. (2017). Piracy and box office movie revenues: Evidence from Megaupload. *International Journal of Industrial Organization*, 52:188–215.

Shapiro, C. and Varian, H. R. (1998). *Information rules: a strategic guide to the network economy*. Harvard Business Press, Cambridge, MA.

Smith, M. D. and Telang, R. (2010). Piracy or promotion? The impact of broadband Internet penetration on DVD sales. *Information Economics and Policy*, 22(4):289–298.

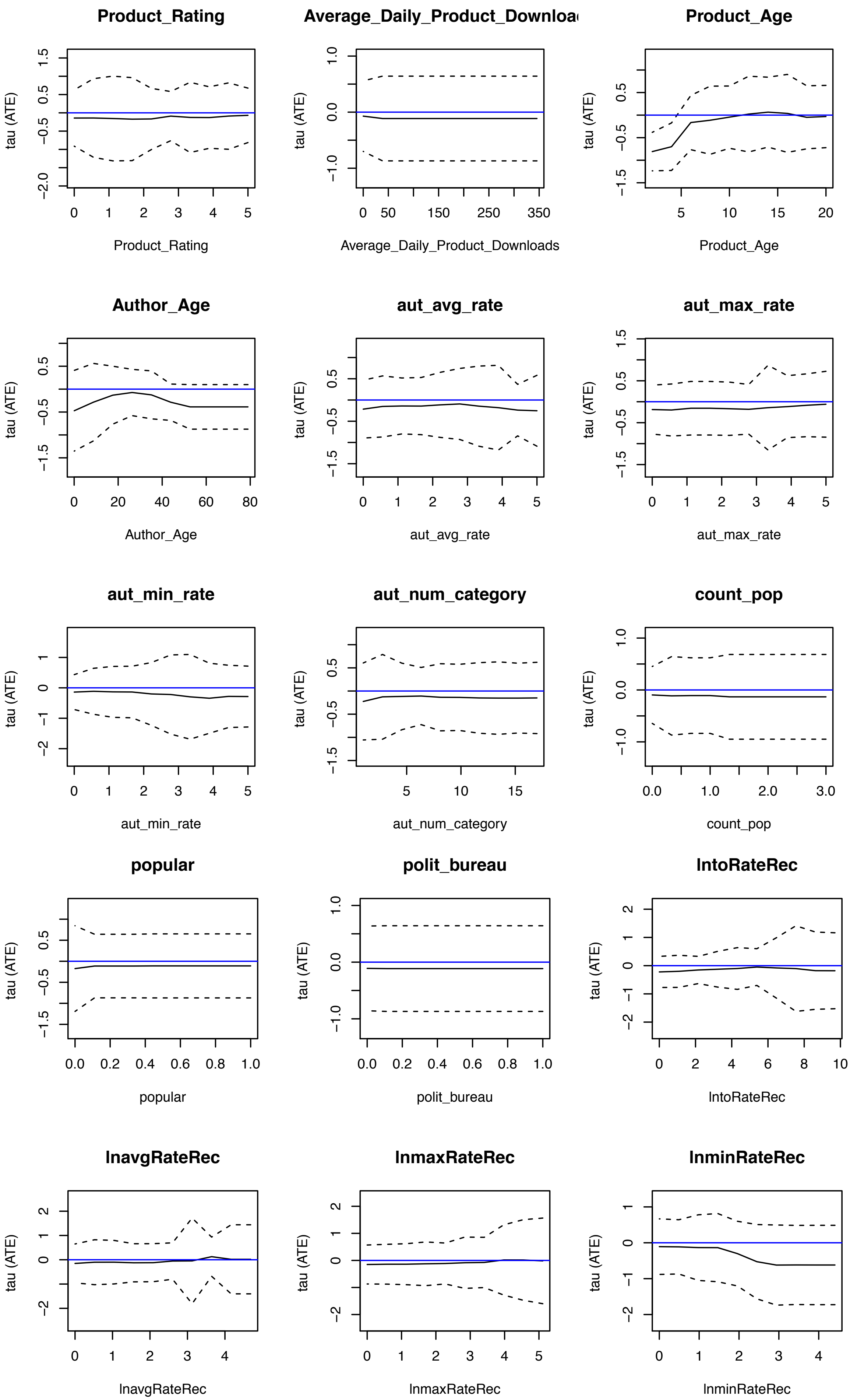
Tunca, T. I. and Wu, Q. (2013). Fighting Fire with Fire : Commercial Piracy and the Role of File Sharing on Copyright Protection Policy for Digital Goods. *Information Systems Research*, 24(2):436–453.

Waldfogel, J. and Aguiar, L. (2018). Quality Predictability and the Welfare Benefits from New Products: Evidence from the Digitization of Recorded Music. *Journal of Political Economy*, 126(2):492–524.

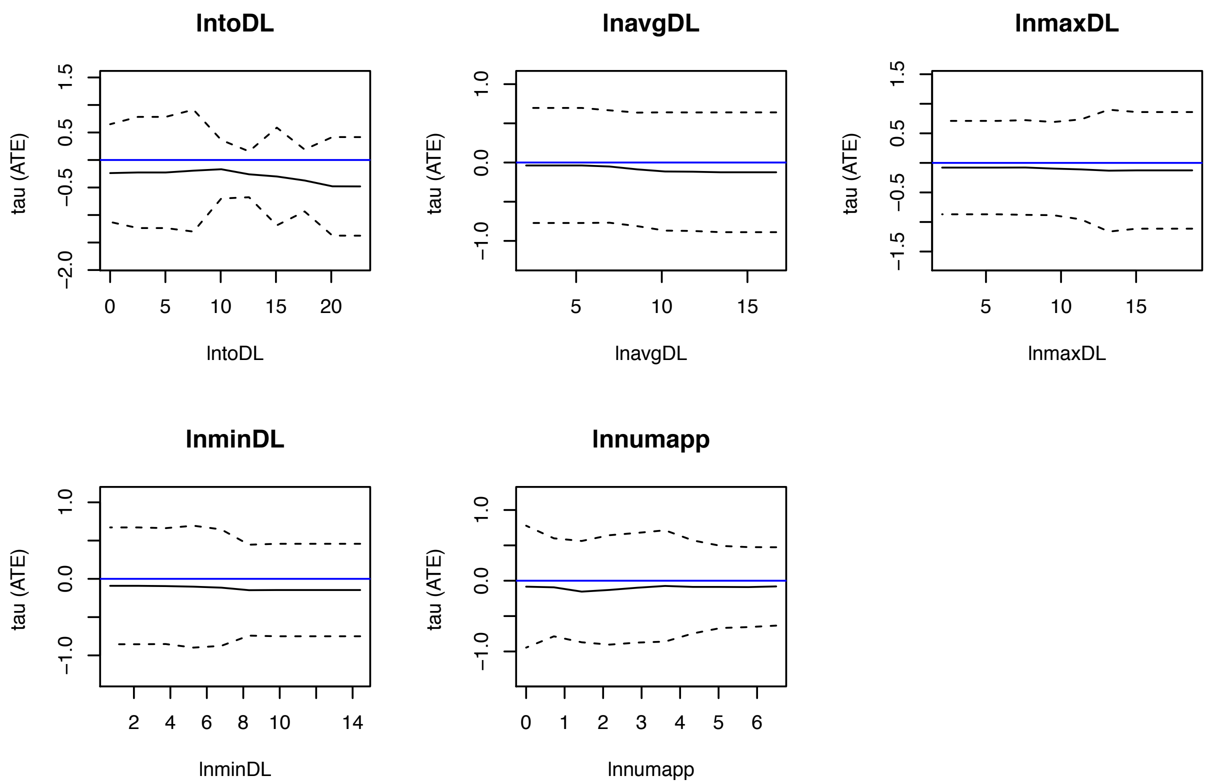
Appendix: Heterogeneous Treatment Effects

The empirical analysis in Section 5.2.1 seeks to explore how the magnitude of the effects estimate vary by product popularity. There are numerous measures of popularity that could be used to quantify these effects. Therefore, we use the approach developed by Wager and Athey (2018), which allows us to explore how the magnitude of the treatment effect varies across different covariates, without “data mining” by testing different variables in a regression. We test the following set of potential variables.

- **Product Age:** Days the title was on the market prior to the hacking attack.
- **Product Rating:** Average product rating (out of 5).
- **lnDL:** Total Number of Product Downloads, log transformed.
- **aut_avg_rate:** Average ratings of (all) products for developer.
- **aut_max_rate:** Highest rating of (all) products for developer.
- **aut_num_category:** Total number of different market niches in which developer has released products.
- **count_pop:** Total number of “popular” titles released.
- **popular:** Indicator if a title is in the top 10% of downloads.
- **polit_bureau:** Indicator if a developer is among the influential or core members of the community.
- **MinMoCount:** Earliest month that developer released a product in the marketplace.
- **lntoRateRec:** Total ratings of (all) products for developer
- **lnavgRateRec:** Average ratings of (all) products for developer (log transformed).
- **lnmaxRateRec:** Highest ratings of (all) products for developer (log transformed).
- **lnminRateRec:** Lowest ratings of (all) products for a developer (log transformed).
- **lntoDL:** Total number of downloads a developers products have received (log transformed).
- **lnavgDL:** Average number of downloads a developer’s products have received (log transformed).
- **lnmaxDL:** Maximum number of downloads any developer’s products have received (log transformed).
- **lnminDL:** Minimum number of downloads any developer’s products have received (log transformed).
- **lnnumapp:** Total Number of applications released by developers (log transformed).



Product_Rating
tau (ATE)
1.5 0.5 −0.5 −2.0
0 1 2 3 4 5
Product_Rating
Average_Daily_Product_Downloa
tau (ATE)
1.0 0.0 −1.0
0 50 150 250 350
Average_Daily_Product_Downloads
Product_Age
tau (ATE)
0.5 −0.5 −1.5
5 10 15 20
Product_Age
Author_Age
tau (ATE)
0.5 −0.5 −1.5
0 20 40 60 80
Author_Age
aut_avg_rate
tau (ATE)
0.5 −0.5 −1.5
0 1 2 3 4 5
aut_avg_rate
aut_max_rate
tau (ATE)
1.5 0.5 −0.5 −1.5
0 1 2 3 4 5
aut_max_rate
aut_min_rate
tau (ATE)
1 0 −1 −2
0 1 2 3 4 5
aut_min_rate
aut_num_category
tau (ATE)
0.5 −0.5 −1.5
5 10 15
aut_num_category
count_pop
tau (ATE)
1.0 0.0 −1.0
0.0 1.0 2.0 3.0
count_pop
popular
tau (ATE)
0.5 −0.5 −1.5
0.0 0.2 0.4 0.6 0.8 1.0
popular
polit_bureau
tau (ATE)
1.0 0.0 −1.0
0.0 0.2 0.4 0.6 0.8 1.0
polit_bureau
lntoRateRec
tau (ATE)
2 1 0 −1 −2
0 2 4 6 8 10
lntoRateRec
lnavgRateRec
tau (ATE)
2 1 0 −2
0 1 2 3 4
lnavgRateRec
lnmaxRateRec
tau (ATE)
2 1 0 −2
0 1 2 3 4 5
lnmaxRateRec
lnminRateRec
tau (ATE)
1 0 −1 −2
0 1 2 3 4
lnminRateRec



lntoDL
lnavgDL
lnmaxDL
lnminDL
lnnumapp
tau (ATE)
1.5
0.5
−0.5
−2.0
0
5
10
15
20
1.0
0.0
−1.0
−1.5
2
4
6
8
14
1
3