

Digital Piracy, Creative Productivity, and Customer Care Effort: Evidence from the Digital Publishing Industry

Xiaolin Li* Chenxi Liao† Ying Xie‡

Abstract

We empirically investigate how writers’ output is affected by copyright piracy using data from a Chinese digital publishing platform. We identify two measurements of writers’ output, creative productivity and customer care, which are also affected by readers’ feedback through purchasing, tipping, and commenting. We take advantage of an exogenous event—the termination of free personal storage service and search function by a leading Chinese cloud storage provider in June 2016—to causally identify the effects of the resultant reduced copyright piracy on writers’ efforts. Using a difference-in-difference modeling approach, we compare the changes in average writer behaviors before and after the event across two groups of writers: 1) writers who have profit-sharing contracts with the platform; and 2) those who do not. We find that after the termination, contracted writers increased their creative productivity efforts in terms of quantity without sacrificing quality, but reduced their customer care efforts. However, these effects are absent for non-contracted writers. Our study is among the first to provide empirical support for the positive effect of digital IPR infringement reduction on creative productivity.

Keywords: Intellectual property rights, digital piracy, emerging markets, creative production, customer care, DID model

1 Introduction

Advances in information technology have made it easier than ever to publish and consume creative works digitally. Thanks to the proliferation of digital publishing platforms, electronic reader devices,

*London School of Economics and Political Science, x.li166@lse.ac.uk

†The Chinese University of Hong Kong, chenxiliao@cuhk.edu.hk

‡The University of Texas at Dallas, ying.xie@utdallas.edu

and mobile reading apps, the digital publishing industry has enjoyed explosive growth in recent years. As one of the largest markets in digital publishing, the Chinese digital publishing industry had an estimated worth of 12.76 billion CHY (1.83 billion USD) in 2017 with more than 14 million published works and 380 million readers.¹ Moreover, many of these works (also referred to as digital intellectual property, or digital IP) have been adapted into popular TV dramas, films, audio books, comics, and video games, generating additional revenues for top writers and their publishing platforms.

Technological improvement has also lowered the cost of copyright piracy² of digital goods, which poses a daunting challenge to the intellectual property rights (IPR hereafter) protection of original creative works published in digital format. This is especially true in developing economies due to generally weak IPR legislature and enforcement. A prevalent means of copyright piracy in digital publishing is through online peer-to-peer file sharing.³ In China, the availability of free or low-cost cloud storage services in recent years has significantly reduced the cost of online file sharing and, as an unintended consequence, greatly facilitated digital piracy.

Although the creative process requires costly sunk investments, the marginal cost of reproducing and distributing the completed creative work digitally is virtually zero. Without strong digital IPR protection, pirated copies of these digital works would become rampant on the Internet and therefore negatively impact sales of the original works. Consequently, writers and publishers would have little financial incentive to undertake the investment to produce original creative works. In the long run, the insufficient supply of high-quality original creative works would severely hinder the growth of the digital publishing industry. For example, the Chinese online literature industry (which is a significant sub-sector of the Chinese digital publishing industry) suffered a loss of nearly 7.98 billion CHY (1.15 billion USD) in 2016 alone due to piracy (iResearch 2016). In particular, writers complain about feeling discouraged when readership of their original works is lost to pirated copies.⁴ Qunshi Liao, the deputy president of Guangdong Internet Writers Association, commented that when original works that writers took great pains to create can easily become someone else's

¹See *Revenue of China's Internet Literature Surged 32.1%* available at http://www.iresearchchina.com/content/details7_47047.html.

²Piracy refers to the duplication and sale of copyrighted goods without the permission of the copyright holder.

³See, e.g., iResearch (2016), and Chris Tribbey (2018) *Irdeto: P2P Piracy Still Dominant* available at <https://www.mesalliance.org/2018/08/08/irdeto-p2p-piracy-still-dominants-mesa/>.

⁴For example, <http://bbs.jjwxc.net/showmsg.php?board=17&boardpagemsg=822&id=543315>.

profit-making tool, it “hurts writers’ motivation and enthusiasm.”⁵

Existing empirical studies on piracy have largely focused on its impact on consumer demand; namely, how various forms of IPR infringement, regulation, and enforcement influence sales of authentic products versus pirated products (see Fink et al. 2016 for a comprehensive review). Although a number of studies predict that high levels of IPR infringement would undermine the incentives for innovation (Maskus 2000, Fink et al. 2016), very few empirical accounts have examined the impact of IPR infringement on the production of original creative or innovative works. Qian (2007) and Lerner (2009) are two exceptions. However, surprisingly, neither was able to find a statistically significant effect of strengthening IPR protection (through patent policy) in stimulating domestic innovation, which is against economic intuition. As Lerner (2009) summarizes, there are at least two reasons for finding a null or negative effect. First, the time frame may be too short. Both studies use patent filing events as a way to measure domestic innovation, which are one-time decisions made by the firm. To observe multiple patent policy shifts in the data, the researchers have to rely on data sets that span a very long time period. For example, Lerner (2009) collected 177 patent policy shifts events in 60 nations over the past 150 years to study the effect of IPR enforcement on innovation, which might still be too sparse. Second, the measures of innovative output are crude, based only on patent applications, and do not capture many other innovation-related activities that are also affected by IPR infringement. In this paper, we attempt to counter both challenges by using a novel dataset we collect from a leading Chinese online literature publishing platform that records individual writers’ frequent and repeated creative output when they update their works online.

Different from the digital publishing format prevalent in western countries facilitated by platforms such as Amazon Kindle Direct Publishing, most digital literature works published in China are serialized (i.e., writers publish their books chapter by chapter). This provides writers an opportunity to incorporate readers’ feedback into future writing, thus making it a dynamic process. Readers can access the first few chapters for free and then decide whether to continue reading by purchasing the following chapter(s). In addition to purchasing book chapters, readers can interact with writers by posting comments and sending tips. Because readers’ consumption enjoyment of online literature depends upon writers’ frequent updates of the books and these interactions with writers, two types of writer efforts are essential to the writers’ and platforms’ revenue: the creative

⁵See Chi Dehua (2018) *Online Literature Market Flourishing in China* available at <https://gbtimes.com/online-literature-market-flourishing-in-china>.

effort to update the books frequently and the customer care effort to interact regularly with readers.

In the current study, we randomly select 969 online fiction books published on a leading Chinese digital publishing platform in 2016. For each book update from these 969 books, we obtain purchase data from the publishing platform and combine it with scraped data on readers' tipping and commenting behaviors and writers' publishing and interactive activities. Using this dataset, we quantify how much and in what directions writers' creative production and customer care efforts are influenced by reduced digital piracy. We address our research questions by utilizing an exogenous event, the termination in June 2016 of the search function and free personal cloud storage service by V-Disk, an independent cloud storage provider in China. Like in many other countries, a major route of digital piracy in China is through unauthorized peer-to-peer file sharing. Among all major cloud storage platforms in China, V-Disk was the most visited by individual users for sharing pirated books. Its termination of the search function and free personal cloud storage service increased the search, storage, and redistribution cost of pirated books. We expect this to substantially reduce piracy for books published on our focal platform after the service termination.

We take advantage of this exogenous event to causally identify the effects of the resultant reduced copyright piracy on writers' efforts. More specifically, we employ a difference-in-difference (DID) modeling framework to compare the change in creative effort and customer care effort before and after the event across two groups of writers: contracted writers and non-contracted writers. Contracted writers are those who have signed profit-sharing contracts with the publishing platform and therefore are entitled to a proportion of sales revenue and readers' tips of their digital books published by the platform. Non-contracted writers do not have such contracts and, as a result, do not receive any pecuniary compensation for their digital books published by the platform. Because of the different monetary incentives faced by these two types of writers, we expect that the exogenous event of V-Disk service termination would increase the creative efforts exerted by contracted writers but not non-contracted writers. In other words, we use non-contracted writers as a control group to causally identify the effect of piracy reduction (through the exogenous shock of V-Disk service termination) on contracted writers, our treatment group. We measure the quantity of each writer's creative effort using the length of each book update measured by the number of Chinese characters and the quality of the creative effort by analyzing readers' comments on each update and extracting sentiment from these comments using Natural Language Processing. We use the number of replies to

readers’ comments from writers between two updates to measure each writer’s customer care effort. In addition, to obtain a clean identification of the treatment effects, we control for the influences of three different types of reader feedback (i.e., book chapter purchases, reader comments, and tips) in our DID model.

Thanks to the frequent and interactive nature of the writers’ publishing efforts, we are able to empirically detect changes in creative production output in both quantity and quality in reaction to the reduced digital piracy within a relatively short period of time. We find that after V-Disk’s termination of the search function and free personal cloud storage service (which we refer to hereafter as V-Disk’s termination event), contracted writers increased their creative effort in quantity without sacrificing quality. This result is consistent with the theoretical rationale in the existing literature that patent protection interventions encourage investment in research and innovation. Moreover, we find that contracted writers reduced their customer care efforts after V-Disk’s service termination event. We believe this finding is primarily driven by the reduced incentive for these writers to “augment” the reading experience (through frequent writer-reader interactions) on the authorized publishing platform once copyright piracy is no longer a serious threat. Not surprisingly, these effects are absent for non-contracted writers who do not have any pecuniary incentives to publish on the focal platform. In fact, our result reveals a decrease in creative production quality after the V-Disk termination event for non-contracted writers. In addition, we find that all three types of reader feedback (purchasing, tipping, and commenting) affect writer efforts. While pecuniary incentives such as book chapter sales and tips positively affect creative production efforts but not customer care efforts for contracted writers, non-pecuniary incentives such as comments tend to increase both types of efforts.

Our paper makes the following contributions to the marketing literature and practices. First, our study is one of the first to provide empirical support for the positive effect of digital IPR infringement reduction on creative productivity. Unlike the extensive research examining the effect of IPR infringement on consumer demand for authentic goods, empirical studies that directly examine the relationship between IPR infringements and innovation investment have been very sparse. More importantly, these few studies fail to find a positive effect of IPR protection on innovation and creative productivity, likely due to various data constraints. Utilizing a unique panel dataset of writers from a major online literature publishing platform and a DID modeling framework, we are

able to bridge this gap by showing increased writer creative effort in reaction to reduced copyright piracy (as the result of an exogenous event). Our findings provide empirical support for the positive effect of reduced digital piracy on creative production, but also caution its potential undesirable side-effect: that enhanced creative production due to IPR protection can crowd out writer customer care effort, which is essential for the long-term growth of customer equity for the platform and its writers.

Second, our study contributes to the growing literature on user-generated content (UGC) creation and consumption. In contrast to the vast majority of extant studies that focus on content creation in non-commercial settings (for example, reviews and social media posts), our study examines UGC creation in a commercial setting where writers act as product innovators as well as product managers, motivated by both pecuniary and non-pecuniary incentives (e.g., Sun and Zhu 2013). Our findings on the effectiveness of piracy reduction in encouraging creative productivity underscores the importance of IPR protection for online publishing platforms.

Third, our work is related to the literature on incentive and productivity (Holmstrom and Milgrom 1991). More specifically, it contributes to the recent research on multi-dimensional effort in multitasking settings where multi-dimensional incentive design is crucial in inducing agents' effort allocation between tasks (Kim et al. 2019, 2020). In our empirical setting, writers allocate their efforts between two tasks: maintaining creative productivity and providing customer care. We show that reduced piracy leads to an increased quantity in contracted writers' creative output but a decrease in their interactions with readers, suggesting that incentivizing one dimension of multiple tasks may lead to decreased effort towards the other performance dimension(s). This finding highlights the urgency for commercial UGC platforms to provide direct incentives for content creators' customer care efforts in order to better retain customers.

Lastly, our paper contributes to the marketing literature on emerging markets. As Sudhir et al. (2015) and Narasimhan et al. (2015) point out, research on these markets can augment the understanding of markets and marketing because the lack of regulation and weak enforcement of laws can lead to distortions in market outcomes that do not exist in developed markets. For example, compared with developed countries, digital piracy is extremely prevalent in emerging markets (Karaganis 2011). Nevertheless, policymakers in these countries wonder whether to allocate scarce resources to IPR enforcement (Fink et al. 2016). In this paper, we show that in the context

of emerging markets, interventions that increase the cost of distributing and consuming pirated content, such as the V-Disk event we describe in the current study, can be an effective alternative to curb digital piracy and thus help increase creative productivity. Policymakers in these countries may consider utilizing these alternative means in addition to patent and copyright laws to further enhance the IPR protection.

The rest of the paper is organized as follows. We summarize the existing literature on IPR infringement and innovation in Section 2. After introducing the background for our study in Section 3, including institutional details and the quasi-experiment, we describe the empirical setting and summary statistics of our data in Section 4. In Section 5 we present our model and estimation. We discuss our findings in Section 6, followed by robustness checks in Section 7. We conclude in Section 8 with a brief discussion of managerial implications.

2 Literature Review

In this section, we first discuss the basic economic implications of two types of IPR infringements, trademark counterfeiting and copyright piracy, and then summarize the existing empirical literature that documents and quantifies such effects.

The economic consequences of trademark counterfeiting depend upon whether consumers are deceived to believe that a counterfeit is produced by the owner of the trademark. In the presence of deceptive counterfeiting, a rational consumer does not want to pay the full price of a high-quality product due to her inability to distinguish the original from the fake, resulting in a decreased incentive for firms to invest in higher quality (Qian et al. 2014). In a similar vein, high levels of copyright piracy undermine incentives to produce original creative works such as books, music, films, and software (Fink et al. 2016). Because these products can be reproduced and distributed at virtually zero marginal cost, pirated copies of these products tend to flourish when the original creators charge a high price. Without active detection and meaningful punishment, such pirating behaviors can significantly deter the development and production of new creative works.

Beyond these basic economic implications of IPR infringement, several theoretical studies have made additional predictions regarding firms' pricing and product quality decisions in the face of piracy. Besen and Kirby (1989) develop a game theory model that suggests the price of original works can either rise or decline as a result of piracy. Qian (2008) suggests that prices of original works can increase in the presence of counterfeiters because unauthorized copying may induce copyright owners

to increase product quality to differentiate the originals from the counterfeits.⁶ Jain (2008) finds that digital piracy can reduce the price competition between firms producing authentic products and may lead to a higher incentive for innovation. Using a vertical differentiation setup, Qian (2014a) predicts that the producer of an original product will upgrade product quality and raise product prices in response to entry by counterfeiters if the quality of the fake products is sufficiently low.

The majority of empirical work on IPR infringement focuses on documenting the effects of various forms of trademark counterfeits and copyright piracy on consumer demand for the authentic product versus the counterfeit or pirated product. However, the findings are rather mixed. Some studies have found a negative effect of unauthorized peer-to-peer file sharing on authentic music sales (see Liebowitz 2006 for a summary of findings from various studies). For example, using survey data of 15,000 European consumers, Zentner (2006) estimates that peer-to-peer file sharing reduces the probability of purchasing music by 30 percent. Rob and Waldfogel (2006) survey 412 college students and find that album downloads negatively affect music record purchases. Hui and Png (2003) use aggregate-level data for music CDs across different countries and find that demand for music CDs decreased with piracy.⁷ Meanwhile, other studies report a null effect. Oberholzer-Gee and Strumpf (2007) find that music downloading does not have a statistically significant effect on music record sales using panel data of weekly album sales and downloads for the United States. Bhattacharjee et al. (2007) analyze the relationship between peer-to-peer file sharing and the popularity of music albums. They find that file sharing does not seem to hurt the top albums' survival on the charts. Despite the negative "direct" substitution effect, counterfeiting or piracy may also have a positive effect on sales of authentic products. For example, in the presence of positive consumption or network externality, piracy may increase consumer demand for certain software by increasing consumer valuation (Katz and Shapiro 1986). In addition, a few studies propose an advertising effect of counterfeit or piracy. For example, using the monthly shipment data on DOS-based microcomputers and two software products (spreadsheets and word processors) in the U.K. between 1987 and 1992, Givon et al. (1995) show that pirated versions significantly influence the legal diffusion

⁶This strategy is also suggested by Varian (2005).

⁷Despite the possible negative effect of piracy on consumer demand, some studies caution against anti-piracy or anti-counterfeit actions. For example, digital rights management (DRM) in the music industry may be harmful when too many restrictions lead legitimate consumers to value the product less (Sinha et al. 2010, Vernik et al. 2011). Also, the overt anti-counterfeiting technologies (OACTs) in the pharmaceutical industry may lead to higher counterfeit sales when the easy-to-replicate OACT makes consumers believe the drugs are genuine even though they are coming from a dubious source (Gao 2018).

of the software. More recently, Qian (2014b) combines field data collected from Chinese shoe companies with lab experiments to investigate the sales impacts of counterfeits on authentic products. Her analysis shows that counterfeits have both a positive advertising effect by increasing awareness for the brand and a negative substitution effect by replacing some consumers' demand for the authentic products. She further finds that the advertising effect dominates the substitution effect for high-end authentic products while the opposite holds for low-end authentic products. Similarly, Lu et al. (2020) examine the impact of piracy on word-of-mouth (WOM) and revenue in the U.S. movie industry. They find a positive correlation between post-release movie piracy and WOM volume and further quantify a 3% increase in box office revenue due to the presence of post-release piracy.⁸

In comparison to the extensive research examining the effect of counterfeit or piracy on consumer demand for authentic goods, empirical studies that directly examine the relationship between IPR infringements and innovation investment are scant. Two exceptions are Qian (2007) and Lerner (2009), in which the authors utilize patent policy changes to gauge the extent of IPR infringements and study their influences on innovation activities. Qian (2007) evaluates the effects of patent law protection on pharmaceutical innovations using panel data from 26 countries. Her results show that the implementation of patent laws does not have a statistically significant effect on stimulating domestic innovation. However, domestic innovation is accelerated in countries with a higher level of economic development, education, and economic freedom after implementing patent law. Lerner (2009) examines the impact of 177 major patent policy changes on the number of patent applications in 60 countries over the past 150 years. He also does not find a positive effect of strengthening patent protection on domestic innovation. In a very recent study, Giorcelli and Moser (2020) collect historical information on new operas across eight Italian states between 1770 and 1900, and show that the adoption of copyright laws at the state level led to an increase in new opera production within the state. Our study aims to add to this limited literature by using individual-level panel data of writers' creative output and sales records made available to us by a major online publishing platform to explore the effects of digital IPR infringement on the level of creative effort exerted by independent writers. Unlike the previous three papers that relied on patent law changes or policy shifts to establish the effects, our main identification relies on an exogenous event in the private

⁸Apart from studying the direct effect of piracy on consumer demand, some studies (e.g., Papiés and van Heerde 2017) also examine its effect on sales of different forms of authentic products. For example, recorded music and live concerts are different forms of copyrighted music.

sector that factually increased the cost of peer-to-peer file sharing and thus led to reduced piracy of the original creative works.

Lastly, we note that the extant empirical literature on the effect of piracy mostly focuses on developed markets. However, the findings and implications from the studies that use data from developed markets may not intuitively apply to the case of developing markets due to the very different market conditions. As one of the few studies examining the effect of piracy in an emerging market setting, our paper aims to help bridge this gap.

3 Background

In this section, we provide background information for our study. First, we provide institutional details about the Chinese online literature industry. Second, we discuss the issue of digital piracy in this industry. Third, we describe an exogenous event in June 2016 that led to reduced digital piracy of online novels in the Chinese market, which is key to our identification strategy.

3.1 Chinese Online Literature Industry

Publishing original fiction books online became popular in China about twenty years ago in 1998 when a romantic story, *The First Intimate Contact*, was posted on a bulletin board system by an unknown amateur writer, became widely popular, and was later acquired and published by a publishing house. Since then, the number of online writers has grown rapidly for two main reasons. First, unlike in the traditional publishing industry where only books that pass the publishing house’s screening can be published, the entry barriers are considerably lower for online publishing due to the ease of digital distribution. Second, with more people switching to digital devices as their main reading means, it is much easier to attract fans and form “buzz” online for ones’ original digital publication. As a result, the low entry cost and the ability to draw a wide audience have attracted many writers to publish their books online. According to Jo Lusby, managing director of Penguin China (a publishing house), “there are no authors under the age of 35 who were not discovered on the Internet.”⁹

After two decades of growth, along with the emergence of several big online literature publishing platforms, a mature industry chain has formed. Writers usually publish their original books, mostly fiction, chapter by chapter after obtaining a writer’s account (with free registration) at one of the

⁹See C.S.-M (2013) *Voices in the wilderness - Chinese online literature* available at <https://www.economist.com/prospero/2013/03/24/voices-in-the-wilderness>.

online publishing platforms. To attract and retain good writers, platforms may offer profit-sharing contracts, which allows them to charge readers a subscription fee to read book chapters written by these writers and also entitles these writers to obtain a fixed proportion of revenues generated from their book sales and other income sources such as monetary tips from readers. The IPR of the most popular books may be acquired by outside publishers and media companies to adapt into printed books, audio books, comics, video games, TV dramas, and even films, thus generating additional revenue for the publishing platforms and the writers (iResearch 2016). Also, many of the works and derivative products are translated into other languages and attract millions of fans across the world.¹⁰ In 2017, the Chinese online literature industry had an estimated worth of 12.76 billion CHY (1.83 billion USD), a 32.1% increase from the previous year (iResearch 2018). When China Literature, one of the largest players in the industry, had its IPO debut in Hong Kong in late 2017, its market capitalization hit 93 billion HKD (around 12 billion USD).¹¹

3.2 Threat from Digital Piracy

IPR infringement, including copyright piracy, is a serious concern in most emerging markets.¹² In China, despite improved legislation protecting digital copyrights, the online literature industry has experienced severe copyright piracy. In a survey of more than 1000 online fiction writers, the majority think their works suffer from *extremely* (45.8%) or *rather* (26.6%) serious piracy issues (iResearch 2016). In 2016, when the estimated market worth of the entire Chinese online literature industry was 9.66 billion CHY (iResearch 2018), piracy led to an estimated loss of nearly 7.94 billion CHY (equivalent to 1.15 billion USD) (iResearch 2016).

There are two major types of digital piracy in the Chinese online literature industry: 1) end-user piracy; and 2) commercial piracy. End-user piracy refers to readers saving popular digital books from authorized sources using their personal accounts and sharing them with others by sending links to the books stored in their accounts or simply setting the content as “public” if saved on a cloud storage platform. As a result, many readers can enjoy reading these books for free either by

¹⁰See Prime Sarmiento (2018) *China’s Online Literature Leaping off the Page to Go Global* available at <https://www.telegraph.co.uk/china-watch/culture/chinas-online-literature-goes-global/>.

¹¹See He and Yeung (2017) *Tencent’s China Literature most-profitable IPO debut in a decade after value soars as much as 100pc* available at <https://www.scmp.com/business/companies/article/2118854/china-literature-shares-soar-63-cent-hong-kong-trading-debut>.

¹²In *2018 Special 301 Report* (available at <https://ustr.gov/issue-areas/intellectual-property/special-301/2018-special-301-review>), which is conducted by the Office of the United States Trade Representative to review the global state of IPR protection and enforcement, China, Indonesia, India and Russia are listed in the priority watch list, whereas Vietnam, Pakistan, Turkey and Brazil are listed in the watch list due to serious IPR concerns.

clicking the links shared by friends or searching the books on the cloud storage platform (if the platform supports a search function). Note that the search function for public content on cloud storage platforms is critical for the wide spread of digital piracy in China. Without this function, even though users could still share unauthorized copyrighted content with their friends (as they can do on Dropbox), it would be much more costly and difficult to disseminate the content beyond their friends, therefore preventing the rapid and wide spread of digital piracy. Commercial piracy involves website operators “stealing” books originally published on authorized platforms and posting the content on their own webpages without permission. To obtain the content, the pirating websites first try to scrape the content directly from the authorized publishing platforms. However, this approach has become more and more difficult due to the wide adoption of “anti-scrape” technology by major publishing platforms. In this case, many pirating websites have turned to other sources to scrape the books, including files uploaded to cloud storage accounts as described for end-user piracy.

Two issues are important to note. First, although pirated books from both sources are accessible, many readers are reluctant to read fiction on pirating websites. A study reports more than 98% of readers who have read pirated books on unauthorized pirating websites have encountered unpleasant experiences, including excessive spam ads, pop-ups, typos, missing words, and bad layout (iResearch 2016). In other words, reading fiction on these pirating websites can be quite “costly” due to the unpleasant reading experience (although readers are not charged by the websites). Second, the books uploaded to readers’ personal cloud storage accounts have been crucial sources for both types of piracy.

3.3 The Termination of Search Function and Free Personal Cloud Storage Service at V-Disk

Among all available cloud storage platforms in China, V-Disk was the one most visited by individual users for sharing pirated books for the following reasons. First, it is a reliable cloud storage platform with millions of active users.¹³ Second, it provided free personal cloud storage service and supported a convenient and efficient search function for public and/or shared contents, even for unregistered users, before June 2016 when most other major cloud storage platforms in China did not or were very

¹³As of January 2016, V-Disk had 2.4 million active app users, according to <http://www.bigdata-research.cn/content/201603/115.html>. This figure did not include active web users of V-Disk. In addition, all registered Weibo users are given accounts on V-Disk by default. In December 2016, there were about 313 million active Weibo users (<https://www.jianshu.com/p/287c9a621f68>).

restrictive in providing such services. Among the major cloud storage platforms in China, Baidu Cloud has had a strict ban on sharing copyrighted content since 2014¹⁴ while others including Huawei Cloud, Tianyi Cloud, and Tencent Cloud never provided a search function for public content on their platforms. Only one major cloud storage service provider, 360 Cloud, had supported a function that allowed its users to search pirated content, but in a much less convenient way than V-Disk: Users could search for “groups” that shared pirated content, but they had to send join requests to such “groups” and could only download the pirated content shared within the group after the join request was granted. Due to the additional hassle to access and disseminate the shared content, 360 Cloud was much less a popular destination for finding pirated books than V-Disk. 360 Cloud disabled the group search function in November 2016.¹⁵ Third, V-Disk was founded by Sina, the parent company that operates Weibo, the most popular microblogging platform in China, where lots of WOM about popular online novels is originated and disseminated. All registered Weibo users are given accounts on V-Disk by default and their V-Disk accounts are automatically linked to their Weibo accounts. As a result, it is easy for Weibo users to add a V-Disk link for downloading pirated content to their Weibo posts when making recommendations. Consequently, readers became used to searching for pirated source on Weibo, then following the link to V-Disk to download the pirated content.

In 2016, V-Disk decided to discontinue the search function for public content and terminate the free personal service. The policy change was announced on April 25, 2016 and officially launched on June 30, 2016. (See V-Disk’s announcement on social media in Figure 1.) Although V-Disk attributed the decision as complying with a government regulation to “purge sexual content, copyrighted content, and improper content on the cloud,”¹⁶ it is commonly believed that “the high operating cost and the low ratio of paying users are the main reason” for its free cloud storage service termination.¹⁷ Whatever the main reason for the termination, we point out that this is an exogenous event; in other words, it is not a decision made or influenced by our focal online literature platform nor by writers or readers. Nevertheless, storing digital copies of online novels on V-Disk

¹⁴Please find Baidu Cloud’s announcement at <http://tieba.baidu.com/p/3004751370>.

¹⁵To fully remove the potential effect of 360 Cloud as an alternative source for pirated books, we conducted a robustness check in Section 7 using only the data within the period when 360 Cloud’s operating strategy remained unchanged.

¹⁶See <https://vdisk.weibo.com/useHelpNotice> for details.

¹⁷See an article on Chinese government’s official website for the campaign against IPR infringements and counterfeits available at <http://www.ipraction.gov.cn/article/xxgk/mtbd/mtgc/201611/20161100112461.shtml> for details.

(and sharing them with their friends and/or the public) became more costly after the termination of the free personal cloud storage service.¹⁸ Moreover, without the sorting and searching functions, neither readers nor the pirating websites could easily find the pirated material stored on V-Disk anymore. As a result, the cost of both types of piracy mentioned in Section 3.2 greatly increased after the event.

Insert Figure 1 about here.

V-Disk’s policy change was a big surprise to readers and writers. Some readers even openly complained on social media about the suddenly increased difficulty to obtain pirated novels (see Appendix A for examples). Although not necessarily an action taken for IPR protection concerns, we expect that V-Disk’s service termination event led to a reduction of digital piracy of online novels because it incidentally increased the cost of searching for and consuming pirated novels. As a result, many readers may have returned to authorized platforms and paid for novels, although others may have decided not to read online novels anymore. In Appendix B, we describe a preliminary analysis that shows readership of authorized novels increased after the V-Disk service termination event.

4 Data

We obtain our data from a leading Chinese online literature publishing platform.¹⁹ Similar to other major literature platforms, the focal platform was created to allow literature lovers to publish or consume books in a much more convenient way than through the traditional publishing industry. All users are allowed to publish their books for free after registration at the platform. Writers usually publish their books chapter by chapter in a sequential format. As a common practice in the industry, writers earn experience scores for their writing and other activities on the platform.²⁰ After accumulating experience scores to reach a certain threshold, writers may receive an invitation from the platform to sign an exclusive contract with the platform.²¹ The crucial difference between

¹⁸Before the termination event, V-Disk provided each user with free storage of up to 100 GB. (The size of a book is usually less than 1 MB.) Since the termination event, a user needs to pay 12 CHY per month.

¹⁹We cannot disclose the identity of the platform due to a non-disclosure agreement.

²⁰The experience score is calculated by the focal platform in the form of $score = \vec{\beta} \cdot \vec{X}$, where \vec{X} is a vector consisting of writer performance variables such as length of the book, number of clicks, number of times the book or the writer is added as a favorite by readers, and number of comments, and $\vec{\beta}$ is a coefficient vector influenced by writer’s contract status and whether the book has violated platform rules.

²¹Contracted writers cannot post their works on other platforms. Moreover, contracted writers grant the focal platform exclusive rights to negotiate with other parties (e.g., publishing houses and studios) for the use of copyrighted materials and share the copyright royalty with the platform.

contracted and non-contracted writers lies in the revenue they receive from books they publish through the platform. First, books published by contracted writers are called “contracted books,” of which only the first few chapters (usually less than one-third of the total book) are accessible by all readers for free. After “sampling” these free chapters, readers who want to continue reading the book have to log in to their accounts and purchase each one of the remaining chapters (which are called “pay-chapters”) from the platform using in-site tokens. The price of each pay-chapter is determined by its length (i.e., measured by the number of Chinese characters in that chapter) and contracted writers receive a portion of each purchase of the pay-chapters of their books.²² In contrast, non-contracted writers can only publish “non-contracted books,” of which all chapters are available to the public for free. Second, although readers can “tip” a contracted or non-contracted writer, only contracted writers receive a portion of such tips from the platform. Third, while non-contracted writers can publish their works on any platform, contracted writers can only publish their books on the focal platform. Due to these differences, we expect that although digital piracy is a big threat for contracted writers that undermines their economic incentive for writing and publishing, it is much less a concern for non-contracted writers who do not receive any revenue for books they publish on the focal online literature platform.

Table 1 provides a summary of the writers’ incentive structure.

Insert Table 1 about here.

To study writers’ creative productivity and customer care effort under different levels of IPR infringement, we take advantage of the exogenous event of V-Disk’s free service termination in June 2016. Because we expect the reduced digital piracy as a result of the V-Disk free service termination event affects contracted writers but not non-contracted writers due to the economic incentive differences, we use non-contracted writers as a control group to compare with contracted writers, our treatment group. By comparing the changes in average writer behaviors before and after the event across these two groups of writers using a *Difference-in-Difference* model, we are able to gauge how writers alter their efforts in reaction to the reduced copyright piracy resulting from the exogenous event.

²²Roughly, the reader pays four tokens for each 1000 Chinese characters in the pay-chapter and the writer receives three tokens. Each token is worth 0.01 CHY.

4.1 Data Description

Our sample consists of 969 books published on the focal platform in 2016. To make sure the sample collected is randomly distributed across time and types of writers, we tag each book based on the month of the year in which its first chapter was published. Among the books with no fewer than 30 chapters published in 2016, we randomly select 50 contracted books and 50 non-contracted books from each month, which constitutes our initial sample of 1,200 books. After removing 231 books that have been locked and/or deleted from the platform,²³ we obtain our final sample of 969 books, of which 564 are contracted books and 405 are non-contracted books. For each book, we scrape the public data on writers’ publishing and interactive activities as well as readers’ feedback including comments and tips for each book chapter. Next, we match the public data with the proprietary sales data from January to November of 2016, which we obtained from the platform, for each book. Table 2 presents summary statistics of the 969 books in our final sample. For each book, we observe the writer’s ID, the genre, and whether it is contracted or not. In our sample, the 969 books are written by 942 individual writers, of which 543 (57.64%) have exclusive contracts with the platform. Each book is listed as one of 16 genres, such as romantic, fantasy, legends, martial heroes, immortal heroes, etc. In our sample, a book consists of 68 chapters, on average, and there are more chapters in contracted books than non-contracted books (on average, 80 vs. 53 chapters, respectively). On average, it takes 175 days for a writer to complete posting all book chapters online. By the end of November 2016, 34.18% of the sampled books were complete, while the rest were still in progress (being updated by writers). On average, a chapter consists of 3,269 Chinese characters and chapters in contracted books are typically longer than those in non-contracted books (3,534 Chinese characters per contracted chapter vs. 2,713 Chinese characters per non-contracted chapter).

Insert Table 2 about here.

Next, we describe how we construct our main variables of interest—writers’ efforts. We focus on two types of efforts: the “creative effort” to create and publish new book chapters (which is divided into quantity and quality dimensions) and the “customer care effort” to interact with readers by

²³A book can be locked or deleted from the platform because of pornographic content or plagiarism, for example. These are excluded from the analysis because we could not scrape all their information for model analysis.

replying to their comments.

Creative effort—quantity: On our focal platform, readers follow books they are interested in and determine whether to purchase pay-chapters each time writers update a book by uploading new chapters. To best reflect such practices, we measure the quantity of writers’ creative effort for each update by the length (i.e., the number of Chinese characters in the newly uploaded chapter(s) within the same update occasion).²⁴ Because readers can “drop out” and switch to other books at any time, to maintain their customers’ interest, most writers update their works on a daily basis.²⁵ On average, there are 3,893 Chinese characters in each update (4,041 for contracted books and 3,539 for non-contracted books). Because the number of Chinese characters in an update is right-skewed, we take a natural logarithm transformation of the variable in our analysis.

Creative effort—quality: To retain readers, it is also important to maintain the quality of book chapters on each update. Similar to other creative works, the quality of a book chapter is subjectively evaluated by readers and reflected in readers’ feedback, especially comments. Therefore, we measure the quality of writers’ creative effort by analyzing all readers’ comments each chapter received by August 2018. Together, we collect the texts of 1,648,508 reader comments received by 914 books.²⁶ We construct the measure for the quality of creative effort following four steps. First, we randomly select 5,000 reader comments and hire annotators to tag the sentiment of each comment manually. Second, we randomly choose 2,000 tagged comments as the training set to train the customized version of Baidu’s Sentiment Classification API.²⁷ Third, we use the trained Sentiment API to tag the sentiment of each of the remaining reader comments into three categories: positive (with a score of 2), neutral (with a score of 1), or negative (with a score of 0). Out of all reader comments in our

²⁴For example, if a writer uploads two chapters the same day (i.e., during the same update), we count all Chinese characters in the two chapters as a measure of her creative effort.

²⁵In our sample, in 75.36% cases, the gap between two adjacent updates is no more than one day, while in another 13.65% cases, the gap is two days.

²⁶Records for the remaining 55 books were missing. We excluded these books in the analysis of creative effort quality.

²⁷Based on a deep learning method, Baidu’s Sentiment Classification API tackles the analysis tasks at multiple granularities: sentence level, entity level, and document level. The sentence-level analysis is based on Bidirectional Long Short-Term Memory Recurrent Neural Networks (BiLSTM), which significantly outperforms traditional methods, such as heuristic methods and classification models based on hand-crafted features. In the entity-level analysis, the API attempts to identify the sentiment towards one of the entities in the text using a semantic hierarchy. Aided by the advanced algorithm and large training set, Baidu’s Sentiment Classification API is believed to be the best sentiment classification tool for Chinese texts. Compared with its generic version that is pre-trained with texts in various contexts, Baidu’s customized version of sentiment classification API further allows users to improve the accuracy rate in specific setting by customizing the API with a user supplied training set. Refer to the following articles for more details: 1) <https://developer.baidu.com/announcement/628> and 2) <https://ai.baidu.com/forum/topic/show/942825>.

sample, 62.09% are classified as positive, 5.10% neutral, and the remaining 32.81% negative. Lastly, we calculate the average sentiment score of all readers' comments given to chapters within the same update occasion for all update occasions. We use this average sentiment score as the measure of the quality of creative effort for each book update. Please refer to Appendix C for a detailed description of our human tagging procedure and the performance of our sentiment classification method.

Customer care effort: Given the interactive nature of the online literature industry, customer care management is as important as, if not more than, creative production because writers essentially act as product managers cultivating long-term relationships with readers. After reading each chapter, readers may post comments about their thoughts related to the story line, main characters, writing style, or to compliment the writer's quality, which are all publicly viewable by the writers and other readers. To address readers' concerns, answer their questions, acknowledge their support, and maintain a good relationship with them, writers often leave replies under readers' comments. The writers' replies are also viewable to all, not just those readers to whom the writers reply. Therefore, these replies may affect a writer's relationship with all readers. Given that replying to readers' comments is the major and most important way writers interact with readers, we use the total number of a writer's replies to readers since the last update to measure the writer's customer care effort for each update. Because the number of writer replies to reader comments is right-skewed and its minimum is 0, we use $\log(\text{Number of Writer Replies} + 1)$ transformation for the variable in our analysis.

Lastly, we note that there are three major ways through which readers can incentivize writers on the focal platform (and also in the industry): purchasing pay-chapters using tokens they purchase from the platform (sales), tipping the writers by awarding them extra tokens (tips), and leaving comments to interact with the writers (comments). While sales and tips from readers provide writers with pecuniary incentives, reader comments provide non-pecuniary incentives. While a contracted writer can obtain both pecuniary and non-pecuniary incentives, a non-contracted writer does not have any pecuniary incentives. In this interactive environment, all three types of readers' activities affect contracted writers' efforts in both creative production and customer care, which we account for in our analysis. In our sample, a chapter of a contracted book receives 13,378 tokens in sales (if it is a pay chapter), 366 tokens in tips, and 21 comments, on average. In contrast, a chapter of a non-contracted book receives 2 comments, on average.

4.2 Model-free Evidence

In this section, we show model-free evidence that V-Disk’s service termination event, which we believe effectively increased the cost of piracy, led to a change in readers’ behaviors as shown in their book chapter purchases, tips, and comments.

Figures 2 to 4 illustrate how sales, tips, and comments changed over time. There are two main patterns. First, there is an obvious increase in average sales and tips a writer received in each newly uploaded chapter after V-Disk’s service termination event. Second, the average number of comments for contracted books increased considerably after the termination event, while there was no obvious change for non-contracted books (see Figure 4). This pattern is consistent with our expectation stated in Section 3.3 that many readers switched back to the authorized platform after the event and started purchasing contracted books. (For additional evidence, refer to Appendix B.)

Insert Figure 2 about here.

Insert Figure 3 about here.

Insert Figure 4 about here.

5 Model

In this section, we describe how we model and estimate the effects of digital piracy on two types of writer efforts: creative effort (quantity and quality) and customer care effort. We take advantage of an exogenous event, the termination of V-Disk’s search function and free personal cloud storage service that resulted in a reduced level of IPR infringement, to identify the effects of digital piracy at the writer-update level through a *Differences-in-Differences* (DID) model. We study the creative production and customer care activities of two groups of writers: contracted writers and non-contracted writers. Because the two types of writers have different monetary incentives, we expect that the exogenous event of terminating the V-Disk search function and free storage would increase the efforts exerted by the contracted writers but not the non-contracted writers. Therefore, we can use non-contracted writers as a control group to compare with contracted-writers, our treatment group. In other words, we utilize a *DID* modeling framework and compare the difference in quantity

and quality of creative productivity and customer care effort across our treatment and control groups before and after V-Disk’s termination. This approach is similar to Hermosilla et al. (2018), where the authors use the movies less likely to enter the Chinese market and thus less likely to be affected by the relaxation of China’s foreign film importation policy as the control group to identify the effect of Chinese society’s aesthetic preference on movies in the treatment group, although movies in both the treatment and control groups are exposed to the same policy intervention.

Although we believe it is valid to use non-contracted writers as the control group since they do not receive any monetary incentives for their published works throughout the study period, all writers and all books in our study sample were exposed to the V-Disk intervention. Therefore, one could still argue that there might be other differences between contracted and non-contracted writers that could affect their responses to the exogenous event and therefore confound the treatment effect. For example, compared to non-contracted writers, contracted writers tend to be more experienced and tend to write books with more chapters and that attract more readers’ comments.

To reduce the self-selection issue, we apply the Propensity Score Matching method (Rosenbaum and Rubin 1983) to better match our control group with our treatment group based on observed characteristics before the V-Disk service termination event and use the matched sub-sample for our model estimation. The propensity score is the predicted probability that a writer has a contract with our focal platform (and therefore is subject to the treatment effect). In our empirical setup, whether a writer is offered a contract by the platform is largely determined by his/her status and experience on the focal platform. Therefore, we collect data on three status and experience related writer characteristics to use as explanatory variables in our propensity score model: 1) how many books each writer has written before 2016; 2) the total number of words he/she has published; and 3) experience score calculated by the focal platform. (Refer to footnote 20 for how the experience score is calculated.) When the propensity scores for a contracted book and a non-contracted book are close enough to each other, signing the contract is considered random (i.e., driven by some unobservable writer characteristics that are unrelated to writer experience and status on the platform). Following this method, we formed a *matched control-group* from non-contracted books for our treatment group of contracted books.

Specifically, we use a logistic regression to estimate the probability of signing a contract as a function of the covariates and then calculate the propensity score for each book based on the

covariates and parameter estimates. We adopt a 1:1 nearest neighbor matching algorithm without replacement and a caliper of 0.01 to match a contracted book with a non-contracted book with the closest propensity score. This algorithm is one of the most common and easiest methods to implement matching with minimal sacrifice in estimation power (Stuart 2010, Li et al. 2019). The resulting matched sample contains 322 pairs of contracted books and non-contracted books. We tried different sets of covariates, including the maximum number of words across all published books for each writer, and get very similar matched samples. (Fewer than five pairs of books will add or drop from the final matched sample using different methods.)

We validate that after matching, the pseudo-treatment group (contracted books) and pseudo-control group (non-contracted books) are not significantly different in any of the covariates. Table 3 summarizes t-test results for these two types of books on the variables mentioned above before and after matching. Results indicate that contracted books have significantly larger values on all explanatory variables before matching, while the differences are not significant after matching. We also provide a graph for the pre-treatment trends in writer efforts and a formal test in Appendix D to show that our treatment and control groups in the matched sample satisfy the parallel trend assumption for DID analysis.

Insert Table 3 about here.

Based on the matched sample, we model writer behaviors at the writer-book update level using a DID modeling framework. Let $i = 1, \dots, N$ denote writers and $t = 1, \dots, T$ denote their book chapter update events. We jointly model the two types of writer efforts, creative effort (quantity and quality) and customer care effort, in each writer update (t) as follows:

$$\begin{aligned}
Y_{\kappa,i,t} = & \alpha_{\kappa,i} + \beta_{1\kappa} I_t + \beta_{2\kappa} Contract_i + \beta_{3\kappa} I_t * Contract_i + \gamma_{1\kappa} Sales_{i,t-1} + \gamma_{2\kappa} Tips_{i,t-1} + \gamma_{3\kappa} Comments_{i,t-1} \\
& + \phi_{1\kappa} I_t * S_{i,t-1} + \phi_{2\kappa} I_t * Tips_{i,t-1} + \phi_{3\kappa} I_t * Comments_{i,t-1} + \phi_{4\kappa} Contract_i * Comments_{i,t-1} \\
& + \phi_{5\kappa} I_t * Contract_i * Comments_{i,t-1} + \delta_{\kappa} * C_{i,t} + \epsilon_{\kappa,i,t} \quad (1)
\end{aligned}$$

$Y_{\kappa,i,t}$ represents the quantity and quality of writer i 's creative effort and customer care effort in chapter update t , with $\kappa = \{\text{Quantity of Creative Effort, Quality of Creative Effort, Customer Care Effort}\}$. $\alpha_{\kappa,i}$ is a fixed effect that accounts for writer i 's average effort level in κ , which absorbs the

systematic time-unvarying differences across writers and their works (Datta et al. 2018). I_t is an indicator for the exogenous event, which equals 1 if the V-Disk’s search function and free storage service is terminated at the time the writer posts her book update t , and 0 otherwise. $Contract_i$ is an indicator for contracted writers, our treatment group.²⁸ Therefore, under the DID modeling framework, $\beta_{3\kappa}$, which captures the interaction effect of the termination event with the contracted writer dummy, is the primary parameter of interest.

Further, as discussed in Section 4.1, a writer’s effort in update t can be affected by readers’ feedback on her last update at $t - 1$, including pecuniary feedback such as book chapter purchases ($Sales_{i,t-1}$) and tips ($Tips_{i,t-1}$), both measured by number of in-site tokens, and also non-pecuniary feedback such as number of comments ($Comments_{i,t-1}$). Therefore, we include all three types of feedback in our model to capture their corresponding effects on writer efforts. We use the lagged versions of reader feedback to avoid the reverse causality problem. Because all three feedback variables are heavily skewed, we use the natural logarithm transformation of these variables in our analysis. In addition, we control for the interaction effects between the event of V-Disk service termination and reader feedback. Note that while only contracted writers can receive sales and tips from readers, all writers can receive comments. Therefore, we also include the interaction effect between contracted writers and comments and a three-way interaction effect between the V-Disk service termination event, contracted writer indicator, and reader comments.

We also include several other covariates $C_{i,t}$ in our model. First, to control for seasonality and other possible market trends affecting the popularity of online novels, we collect the number of monthly searches using key words “online novel” on the major search engine in China (Baidu.com) in the same year as a proxy ($Trend_t$). We use the natural logarithm transformation of the variable in our analysis to adjust for its skewness. (We also include several other measures to capture the seasonality/market trend in Section 7 as robustness checks and produce very similar results.) Second, because most writers in our platform are amateur writers, they only write during their spare time. Therefore, we also include dummy variables $v_{\kappa,t}$ to indicate school summer vacations and Chinese New Year holiday as a control for the variation in writers’ leisure time. Lastly, we allow for the correlation between equations by using the seemingly unrelated regression (SUR); that is, the error terms $\epsilon_{\kappa,i,t}$ are multivariate normally distributed with means of 0 and a covari-

²⁸The coefficient of $Contract_i$ cannot be estimated due to the inclusion of the book and writer specific fixed effects. In other words, the main effect of $Contract_i$ is absorbed by the estimates of $\alpha_{\kappa,i}$.

ance matrix V to capture unobserved features that may influence more than one type of writer output.²⁹ When estimating SUR, we need to avoid an identical structure of covariates between equations in order to estimate the covariance matrix V . For this purpose, we add one additional explanatory variable, number of tips received for each update ($NumTips_{i,t}$), in the equation for writer’s customer care effort because it directly affects the number of writers’ interactions with readers, but not writers’ creative productivity because we already account for the value of tips. Lastly, $\theta = \{\alpha_{\kappa,i}, \beta_{1\kappa}, \beta_{2\kappa}, \beta_{3\kappa}, \gamma_{1\kappa}, \gamma_{2\kappa}, \gamma_{3\kappa}, \phi_{1\kappa}, \phi_{2\kappa}, \phi_{3\kappa}, \phi_{4\kappa}, \phi_{5\kappa}, \delta_{\kappa}, V\}$ is the set of parameters to be estimated.

6 Results and Discussion

We present the estimation results in Table 4. In addition to our main model (i.e., Model III in Table 4), we also estimate two reduced models. Model I does not account for the effects of the three types of reader feedback (direct effects of sales, tips, and comments) nor their interactions with the termination event or contracted writers. Model II accounts for the effects of reader feedback and their interactions with the termination event, but not the three-way interactions between reader comments, the termination event, and contracted writers. In all three models, the estimates of the treatment effects are qualitatively consistent.

Insert Table 4 about here.

Next, we discuss the estimation results of our main model. We first discuss the effect of the termination event on non-contracted writers. Our results show that the termination event did not have a significant effect on the quantity of creative effort and the customer care effort of non-contracted writers. These results are consistent with our expectation that because non-contracted books are free to readers on the focal platform, they are unlikely to be targets of digital piracy. Furthermore, even if there is piracy of their books, non-contracted writers most likely do not care from which source readers access their books because they do not have exclusive contracts with the focal platform nor do they receive any revenue from the focal platform. Interestingly, our results also reveal that the quality of work by non-contracted writers decreased after the termination event. This could be a result of more intense competition from contracted writers’ work.

²⁹For example, a writer might be traveling or busy with his/her full-time job. As a result, the writer’s creative production and/or customer care efforts may decrease.

In contrast, we find that after the V-Disk free service termination event, contracted writers increased their creative efforts by writing more without sacrificing quality,³⁰ but decreased their customer care efforts by replying less to readers’ comments. This result can be explained by the change in market competition before and after the V-Disk free service termination event. Before the termination event, contracted writers faced immediate threats from copyright piracy because their books available on the authorized platform had to compete for readership with pirated copies available through peer-to-peer file sharing and/or pirating websites. Therefore, their creative efforts were not completely compensated by sales because of lost purchases to pirated copies. When the content and updating frequency were almost identical between the official versions and pirated copies, writers’ customer care effort to interact and engage readers was not only a way to increase the likelihood of receiving reader tips, but also an effective means to differentiate their authentic work from pirated versions. This finding is consistent with previous literature showing that companies of authentic products strive to upgrade product quality after counterfeiters enter (Qian, 2008, 2014b). After the termination event, due to the reduced supply of pirated books, readers could not so easily read pay-chapters for free anymore. Therefore, some readers returned to the authorized publishing platform and the marginal pecuniary incentive from sales per update became higher for contracted writers. As a result, competition mainly comes from other writers and books within the same focal platform. Contracted writers then directed their effort towards creative production, which is more effective than customer care effort at differentiating their works from other writers. Furthermore, this finding suggests that an increased incentive in one dimension of multiple tasks (i.e., creative productivity) may lead agents to decrease their efforts in the other performance dimension (i.e., customer care), echoing recent research in incentive design and effort allocation in multi-tasking settings (for example, Kim et al. 2019, 2020). In particular, Kim et al. (2019) investigate the sales performance-moral hazard trade-off in response to multidimensional performance (acquisition and maintenance) incentives in a CRM setting. They find that without the maintenance incentive, the adverse selection effect of acquisition incentives overwhelms the sales-enhancing effect, leading to salespeople acquiring low-quality customers, which hurts future payoffs. Their findings highlight the importance of multidimensional incentives for CRM, which also applies to our setting if the

³⁰Since both $\beta_{1,quality}$ and $\beta_{3,quality}$ are significant, the effect of the termination event on contracted writers’ quality of creative production is the sum of these two parameters (i.e., $\beta_{1,quality} + \beta_{3,quality} = -0.1444 + 0.1845 = 0.0401$), which is not statistically significant. Therefore, contracted writers’ quality of creative production remain unchanged after the termination event.

focal platform wants to maintain a balance between writers’ creative productivity and customer care quality.

In addition, we find that all three types of reader feedback (sales, tips, and comments) affect writer effort for contracted writers. This finding validates our concern about controlling for all three types of feedback and interaction effects in our model to identify the effect of our focal event. In general, we find that lagged sales and tips positively affects creative efforts but not customer care efforts for contracted writers while lagged reader comments tend to increase both types of efforts. We also find that lagged sales became less effective at encouraging writers’ creative effort after the termination event, suggesting content creators react less to short-term sales incentives after the event leading to a permanent reduction of digital piracy.

Lastly, we evaluate the joint effect of the termination event on contracted writers by summing all effects related to the termination event.³¹ We find that the joint effect is positive on creative production quantity and negative on customer care, both qualitatively consistent with the direct treatment effects.

7 Robustness Checks

In this section, we conduct three sets of robustness checks to further strengthen the validity of our findings.

Alternative Controls for Demand-Side Seasonality and Market Trends. In our main model presented in Section 5, we control for demand-side seasonality and other possible market trends affecting the sales on our focal platform by including search records of the keyword “online novel” in the same year to represent the popularity of the online literature industry. However, it is possible that our focal platform follows a different market development trend from the online literature industry in general. To address this concern, we conduct three robustness checks using alternative controls for demand-side seasonality and market trends. In particular, in the first robustness check, we include the number of searches for our focal platform in each month of 2015 (i.e., the year before our study period). In the second robustness check, we include the monthly average

³¹In our main model, the V-Disk termination event directly affects writers’ effort levels and also indirectly affects the effort levels through the interactions with reader feedback. We calculate the joint effect of the termination event on contracted writers by summing all direct effects (the main effect of the termination event and the interaction between the termination event and contract writers) and indirect effects from the changed influence of sales, tips, and comments after the termination event (2-way interactions with sales, tips, and comments and 3-way interaction with comments and contracted writers) for an average contracted writer (by taking the average for all reader feedback variables).

number of searches for our focal platform from 2013 to 2015 (i.e., the three years before our study period). In the third robustness check, we include month fixed effects to control for unobserved market trends and seasonality. Table 5 summarizes the results.³² In all three models, we find very similar results as in our main analysis—contracted writers increased their creative efforts in quantity without sacrificing quality, and decreased their customer care effort by replying less to readers’ comments.

Insert Table 5 about here.

Control for Supply-Side Seasonality. It is possible that writers have seasonal fluctuations in their writing productivity absent of the demand side influence. To address this concern, we collected writers’ publishing activities on our focal platform in the year of 2015. Among the 942 individual writers in our sample (before matching), 306 writers were active in at least one month in 2015 and had the same contract status as in our study period, of which 243 are contracted writers and 63 are non-contracted writers.³³ Using this information, we calculate the average productivity (i.e., average number of Chinese characters in each update occasion) in each month of 2015 for contracted writers and non-contracted writers, respectively. We add these writer-type specific productivity seasonality in 2015 to our main model as the supply-side seasonality control. Table 6 summarizes the results. We find consistent results as our main model.

Insert Table 6 about here.

Chapter Fixed Effects. The ease of creative productivity may depend on the stage of the creative project. For example, writers might find it easier or more difficult to write as the storyline progresses. To rule out the effect of creative flow, we include chapter fixed effects in our main model as an additional robustness check. The number of chapters in a book in our matched sample ranges from 30 to 333, with 89.13% of books consisting of fewer than 100 chapters. To adjust for this long-tail distribution and to avoid adding too many fixed effects, we construct a book chapter dummy variable for each 10 chapters up to chapter 100 and group all chapters beyond chapter 100

³²The main effect of the termination event is not identified and hence not reported in Robustness III in Table 5, due to the inclusion of month fixed effects.

³³Ideally, we can control for a specific writer’s seasonality by including his/her productivity in 2015 as controls. However, this writer-specific seasonality control is not available for most of the writers in our matched sample due to their inactivity or different contract status in 2015.

as one dummy. The estimation of this alternative model yields results consistent with our main model (see Table 7 for more details).

Insert Table 7 about here.

Platform Front-Page Recommendation. If a fiction book is recommended on the platform front page, the writer may adjust her efforts either because she feels encouraged and/or she wants to take advantage of this rare opportunity to develop more readership. Unfortunately, we did not track the front-page recommendation information during our study period nor were we able to obtain this information from the platform. Instead, we scrape the platform’s front-page content archived by Wayback Machine (web.archive.org) in 2016. The front page of the focal platform was archived four to eight times each month by Wayback Machine. From the scraped front-page content, we know which fiction books were recommended on the front page on these four to eight days of each month. Note that the front-page recommendation on the focal platform has a continuous nature: a book usually remains in the recommendation section for several days once it appears in the section. Therefore, if a book was on the platform front page on two adjacent archived days, we can infer that the book was also very likely recommended between these two days. (For example, if the platform frontpage is archived by Wayback Machine on February 2 and 5, but not February 3 and 4, and a book is recommended on the frontpage on both February 2 and 5, then the book is also likely to be recommended on February 3 and 4.) Given this information, we construct two variables—whether the book was recorded to have been recommended on the front page since the last update and whether the book was recorded or expected to have been recommended since the last update—and include them separately in the analysis as robustness checks. Table 8 summarizes the results. Again, we find very similar results as in our main analysis.

Insert Table 8 about here.

Removal of Potential Effect from Another Event. As explained in Section 3, among the big Chinese cloud storage service providers, other than V-Disk, 360 Cloud also allowed users to search for pirated content once they join a “group.” On November 1, 2016, 360 Cloud disabled this function. To eliminate the potential influence of the 360 Cloud shut-down event on our treatment effect, we conduct another robustness check using data between January 1, 2016 and October 31,

2016. Table 9 summarizes the results. Again, the estimation results are consistent with our main model.

Insert Table 9 about here.

8 Conclusion

Facilitated by advances in information technology, the past decade has witnessed an explosive growth of creative user-generated content (UGC) and platforms that specialize in publishing and distributing various forms of creative UGC (e.g., YouTube and TikTok specialize in video, SoundCloud and Ximalaya FM in audio, DeviantArt in images, and blogs and online literature appear on platforms such as WordPress, Wattpad, and Qidian). Although traditional UGC is free, many content creators have been working with publishing platforms to monetize the intellectual property of their creative work. However, due to the low cost of reproducing and distributing digital content, the digital publishing platforms and especially the writers of creative works have been seriously impaired by widespread copyright piracy. The situation is exacerbated in developing economies due to inadequate IPR regulation and poor law enforcement.

In this paper, we take advantage of an exogenous event—the termination of the free personal storage service and search function by a leading Chinese cloud storage provider in June 2016—to causally identify the effects of the resultant reduced copyright piracy on writers’ efforts. Using a difference-in-difference modeling approach, we compare the changes in average writer behaviors before and after the event across two groups of writers: 1) writers who have profit-sharing contracts with the platform; and 2) writers who do not. Our empirical analysis reveals that once the level of IPR infringement decreases, contracted writers exert more effort in creative production. To the best of our knowledge, this is one of the first studies to provide empirical support for the positive effect of lessened IPR infringement on the production of creative works. Given that the industry’s profits greatly depend on the richness and quality of the creative content, this result highlights the importance of strengthening IPR protection. Furthermore, we find that the increased creative effort may come at the cost of decreased customer care effort. That is, with reduced digital piracy, the contracted writers exert less effort interacting with readers. Though this act is reasonable from the writers’ perspective (because of the decreased need to compete with pirated versions of their own work by augmenting readers’ experience on the authorized platform with interpersonal interaction),

it can reduce reader engagement and therefore negatively affect the platform's customer equity in the long run. If this is the case, the platform may want to consider devising a multi-dimensional compensation plan that provides direct incentive for writers to maintain interactions with fans in addition to compensation from book sales and reader tips.

This paper also provides implications for policymakers in emerging markets where the lack of IPR regulations and weak law enforcement leaves plenty of room for improvement in IPR protection. In this paper, we look at the digital publishing industry in China, one of the largest emerging markets in the world. We show that interventions, such as the V-Disk event we describe in the current study, can be an effective means to reduce digital piracy through which to encourage innovative efforts. As many emerging markets transition from imitation to innovation, our study is timely in providing direct empirical evidence for the benefits of having an innovation-conducive environment with strong IPR protection. Moreover, our findings suggest that in addition to laws and regulations, policymakers in these markets may utilize economic incentives or interventions that increase the cost of piracy distribution and consumption as alternative means to enhance IPR protection.

Our study is not without limitations. First, we do not have measures of traffic changes in other platforms and pirating websites. Therefore, we cannot model readers' choices between multiple authorized publishing platforms and pirating websites. Second, we have data from our focal platform, but not from similar platforms that are not subject to the influence of the V-Disk free service termination event (e.g., a similar platform whose readers do not use V-Disk). We hope future research can further investigate this question through a controlled field experiment. Third, in our analysis, we use the length of each book update to measure the quantity of writer creative effort and the reader feedback sentiment to measure the quality of writer creative effort in the simplest and most straightforward way. Future research might construct multiple dimensions of writer creative activities using more metrics or more advanced text-mining techniques.

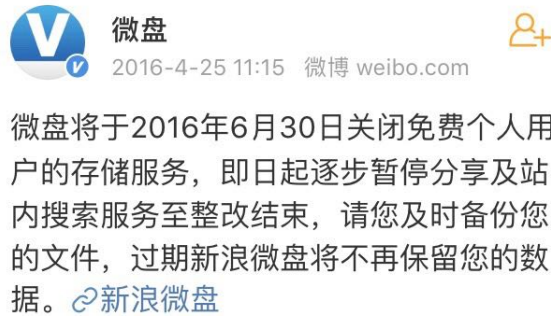
References

- Besen, S. M., S. N. Kirby. 1989. Private copying, appropriability, and optimal copying royalties. *The Journal of Law and Economics* **32**(2, Part 1) 255–280.
- Bhattacharjee, S., R. D. Gopal, K. Lertwachara, J. R. Marsden, R. Telang. 2007. The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Management Science* **53**(9) 1359–1374.
- Datta, H., G. Knox, B. J. Bronnenberg. 2018. Changing their tune: How consumers’ adoption of online streaming affects music consumption and discovery. *Marketing Science* **37**(1) 5–21.
- Fink, C., K. E. Maskus, Y. Qian. 2016. *The economic effects of counterfeiting and piracy: a review and implications for developing countries*. The World Bank.
- Gao, Y. 2018. On the use of overt anti-counterfeiting technologies. *Marketing Science* **37**(3) 403–424.
- Giorcelli, M., P. Moser. 2020. Copyrights and creativity: Evidence from Italian operas. *Journal of Political Economy*, forthcoming.
- Givon, M., V. Mahajan, E. Muller. 1995. Software piracy: Estimation of lost sales and the impact on software diffusion. *Journal of Marketing* **59**(1) 29–37.
- Hermosilla, M., F. Gutierrez-Navratil, J. Prieto-Rodriguez. 2018. Can emerging markets tilt global product design? Impacts of Chinese colorism on Hollywood castings. *Marketing Science* **37**(3) 356–381.
- Holmstrom, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *JL Econ. & Org.* **7** 24.
- Hui, K.-L., I. Png. 2003. Piracy and the legitimate demand for recorded music. *Contributions in Economic Analysis & Policy* **2**(1).
- iResearch. 2016. The white paper of internet literature copyright protection.
- iResearch. 2018. 2018 China’s digital reading sector report.
- Jain, S. 2008. Digital piracy: A competitive analysis. *Marketing science* **27**(4) 610–626.
- Karaganis, J. 2011. *Media piracy in emerging economies*. Social Science Research Council.
- Katz, M. L., C. Shapiro. 1986. Technology adoption in the presence of network externalities. *Journal of political economy* **94**(4) 822–841.
- Kim, M., K. Sudhir, K. Uetake. 2020. A structural model of a multitasking salesforce: Job task allocation and incentive plan design. *Working Paper*.
- Kim, M., K. Sudhir, K. Uetake, R. Canales. 2019. When salespeople manage customer relationships: Multidimensional incentives and private information. *Journal of Marketing Research* **56**(5) 749–766.
- Lerner, J. 2009. The empirical impact of intellectual property rights on innovation: Puzzles and clues. *American Economic Review* **99**(2) 343–48.
- Li, Y., Y. Xie, Z. E. Zheng. 2019. Modeling multichannel advertising attribution across competitors. *Management Information Systems Quarterly* **43**(1) 263–286.
- Liebowitz, S. J. 2006. File sharing: creative destruction or just plain destruction? *The Journal of Law and Economics* **49**(1) 1–28.
- Liu, X., D. Lee, K. Srinivasan. 2019. Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning. *Journal of Marketing Research* **56**(6) 918–943.

- Lu, S., X. Wang, N. Bendle. 2020. Does piracy create online word of mouth? An empirical analysis in the movie industry. *Management Science* **66**(5) 2140–2162.
- Maskus, K. E. 2000. Intellectual property rights and economic development. *Case W. Res. J. Int'l L.* **32** 471.
- Narasimhan, L., K. Srinivasan, K. Sudhir. 2015. Marketing science in emerging markets. *Marketing Science* **34**(4) 473–479.
- Oberholzer-Gee, F., K. Strumpf. 2007. The effect of file sharing on record sales: An empirical analysis. *Journal of political economy* **115**(1) 1–42.
- Papies, D., H. J. van Heerde. 2017. The dynamic interplay between recorded music and live concerts: The role of piracy, unbundling, and artist characteristics. *Journal of Marketing* **81**(4) 67–87.
- Qian, Y. 2007. Do national patent laws stimulate domestic innovation in a global patenting environment? a cross-country analysis of pharmaceutical patent protection, 1978–2002. *The Review of Economics and Statistics* **89**(3) 436–453.
- Qian, Y. 2008. Impacts of entry by counterfeiters. *The Quarterly Journal of Economics* **123**(4) 1577–1609.
- Qian, Y. 2014a. Brand management and strategies against counterfeits. *Journal of Economics & Management Strategy* **23**(2) 317–343.
- Qian, Y. 2014b. Counterfeiters: Foes or friends? How counterfeits affect sales by product quality tier. *Management Science* **60**(10) 2381–2400.
- Qian, Y., Q. Gong, Y. Chen. 2014. Untangling searchable and experiential quality responses to counterfeits. *Marketing Science* **34**(4) 522–538.
- Rob, R., J. Waldfogel. 2006. Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students. *The Journal of Law and Economics* **49**(1) 29–62.
- Rosenbaum, P. R., D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1) 41–55.
- Sinha, R. K., F. S. Machado, C. Sellman. 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.
- Stuart, E. A. 2010. Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics* **25**(1) 1.
- Sudhir, K., J. Priester, M. Shum, D. Atkin, A. Foster, G. Iyer, G. Jin, D. Keniston, S. Kitayama, M. Mobarak, et al. 2015. Research opportunities in emerging markets: An inter-disciplinary perspective from marketing, economics, and psychology. *Customer Needs and Solutions* **2**(4) 264–276.
- Sun, M., F. Zhu. 2013. Ad revenue and content commercialization: Evidence from blogs. *Management Science* **59**(10) 2314–2331.
- Varian, H. R. 2005. Copying and copyright. *Journal of Economic Perspectives* **19**(2) 121–138.
- Vernik, D. A., D. Purohit, P. S. Desai. 2011. Music downloads and the flip side of digital rights management. *Marketing Science* **30**(6) 1011–1027.
- Zentner, A. 2006. Measuring the effect of file sharing on music purchases. *The Journal of Law and Economics* **49**(1) 63–90.

Tables and Figures

Figure 1: The Quasi-Experiment



V-Disk's Announcement on Weibo on April 25th, 2016: V-Disk will terminate the free personal storage service on June 30, 2016, and gradually suspend the sharing and sorting function within the platform. Please back up your files in a timely manner. V-Disk will not keep your data afterwards.

Figure 2: Monthly Sales Volume of Each Chapter: Before and After V-Disk Service Termination

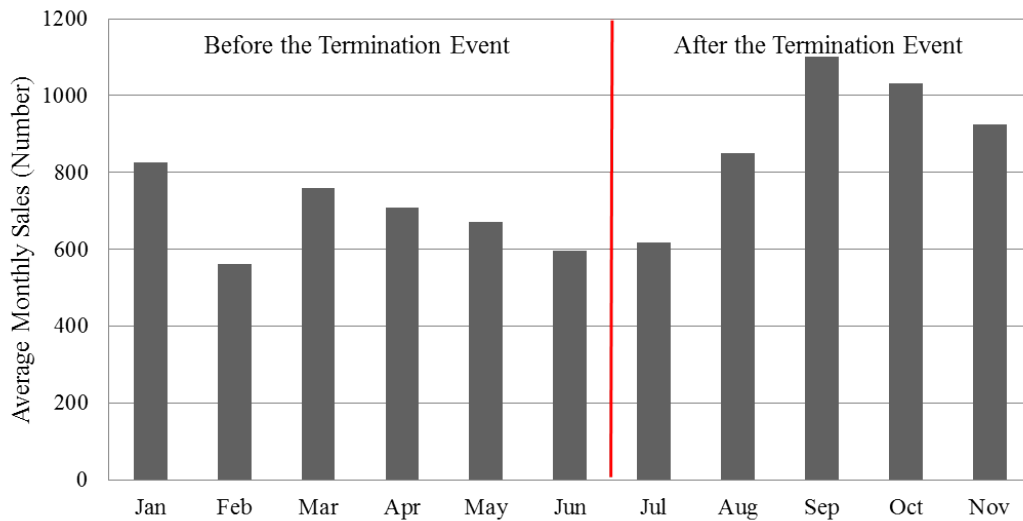


Figure 3: Monthly Tips for Each Chapter: Before and After V-Disk Service Termination

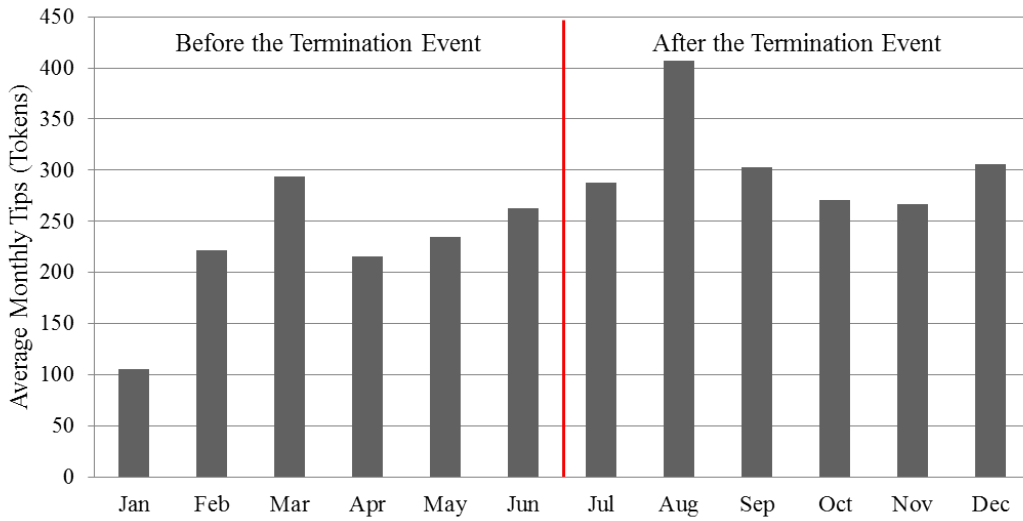


Figure 4: Monthly Comments Received for Each Chapter for Two Types of Books: Before and After V-Disk Service Termination

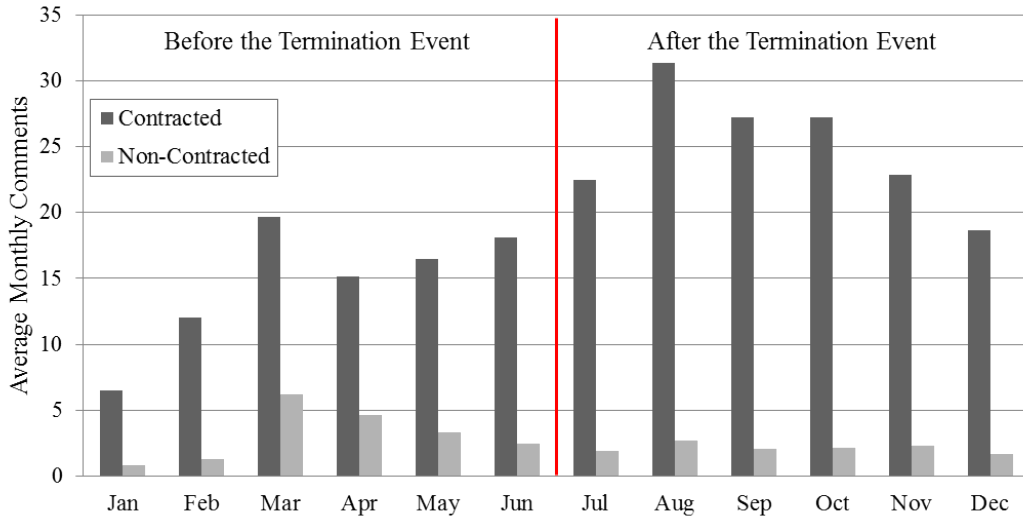


Table 1: Incentive Structure for Contracted vs. Non-Contracted Writers

	Contracted Writers	Non-contracted Writers
Contract with Platform	Yes	No
Payment from Sales	Yes	No
Tips from Readers	Yes	No
Comments from Readers	Yes	Yes

Table 2: Summary Statistics

Statistic	Book Type	N	Mean	Std Dev	Min	Max
<i>Book-specific</i>						
Number of Chapters	All	969	68.37	39.35	30	333
	Contracted	564	79.55	43.43	30	333
	Non-Contracted	405	52.80	25.78	30	219
Number of Updates	All	969	57.41	33.47	1	273
	Contracted	564	69.56	35.01	17	273
	Non-Contracted	405	40.48	21.96	1	128
<i>Update-specific</i>						
Number of Chapters	All	55,627	1.19	1.14	1	119
	Contracted	39,234	1.14	0.48	1	19
	Non-Contracted	16,393	1.30	1.97	1	119
Number of Chinese Characters	All	55,627	3,893.19	3,334.72	109	310,072
	Contracted	39,234	4,041.12	2,124.87	130	59,215
	Non-Contracted	16,393	3,539.14	5,172.29	109	310,072
<i>Chapter-specific</i>						
Number of Chinese Characters	All	66,252	3,268.83	1,315.87	107	48,156
	Contracted	44,868	3,533.68	1,262.72	130	20,563
	Non-Contracted	21,384	2,713.11	1,250.25	107	48,156
Sales (in number)	Contracted	23,324	1,194.14	2,664.86	0	40,564
Tips (in tokens)	Contracted	44,868	365.70	3,560.14	0	404,500
Number of Comments in 2016	All	66,252	15.22	49.45	0	2,365
	Contracted	44,868	21.32	58.11	0	2,365
	Non-Contracted	21,384	2.42	15.77	0	1,366
<i>Key Variables</i>						
Quantity of Creative Effort	All	55,627	8.15	0.45	4.69	12.64
	Contracted	39,234	8.21	0.39	4.87	10.99
	Non-Contracted	16,393	7.99	0.54	4.69	12.64
Quality of Creative Effort	All	44,822	1.27	0.53	0	2
	Contracted	36,808	1.26	0.49	0	2
	Non-Contracted	8,014	1.33	0.69	0	2
Customer Care Effort	All	54,658	0.90	1.14	0	6.52
	Contracted	38,670	1.13	1.22	0	6.52
	Non-Contracted	15,988	0.34	0.67	0	5.30

Table 3: Propensity Score Matching: T-test for Covariates of Matched Books

	Before Matching			After Matching		
	Contracted	Non-contracted	T-test (p-value)	Contracted	Non-contracted	T-test (p-value)
Writers Experience Score	12.0984	5.3632	0.0000	6.6622	6.3383	0.6278
Number of Books Each Writer Has Written Before 2016 (<i>ln</i> value)	1.2505	0.4871	0.0000	0.6104	0.6055	0.9468
Total Number of Words Published Before 2016 (<i>ln</i> value)	8.3641	4.0009	0.0000	4.6784	4.7605	0.8647

Table 4: Estimation Results

	Model I (no indirect effect)		Model II (no 3-way interaction)		Model III (full model)	
	Creative Effort	Customer Care Effort	Creative Effort	Customer Care Effort	Creative Effort	Customer Care Effort
	Quantity	Quality	Quantity	Quality	Quantity	Quality
Termination Event	-0.0318 (0.0266)	-0.1166 (0.0389)	-0.0257 (0.0276)	-0.2134 (0.0671)	-0.0113 (0.0344)	-0.1444 (0.0520)
Termination * Contracted	0.0765 (0.0278)	0.1238 (0.0406)	0.0990 (0.0287)	-0.1279 (0.0695)	0.0904 (0.0382)	0.1845 (0.0577)
<i>Sales_{i,t-1}</i>			0.0188 (0.0015)	-0.0035 (0.0036)	0.0183 (0.0015)	0.0077 (0.0023)
<i>Tip_{s,i,t-1}</i>			0.0300 (0.0068)	-0.0025 (0.0103)	0.0283 (0.0069)	-0.0048 (0.0104)
<i>NumTips_{i,t-1}</i>				0.0585 (0.0461)		0.0552 (0.0462)
<i>Comments_{i,t-1}</i>			0.0338 (0.0065)	0.0235 (0.0098)	0.0051 (0.0136)	-0.0082 (0.0206)
Contracted * <i>Comments_{i,t-1}</i>				0.3221 (0.0162)		0.2526 (0.0329)
Termination * <i>Sales_{i,t-1}</i>			-0.0089 (0.0018)	-0.0072 (0.0027)	-0.0086 (0.0151)	-0.0065 (0.0366)
Termination * <i>Tips_{i,t-1}</i>			-0.0195 (0.0083)	0.0222 (0.0125)	-0.0187 (0.0083)	0.0248 (0.0126)
Termination * <i>NumTips_{i,t-1}</i>				0.0398 (0.0563)		0.0340 (0.0564)
Termination * <i>Comments_{i,t-1}</i>			-0.0165 (0.0075)	-0.0087 (0.0113)	-0.0201 (0.0171)	0.0213 (0.0258)
Termination * Contracted * <i>Comments_{i,t-1}</i>					0.0007 (0.0185)	-0.0380 (0.0450)
Online Novel Search	0.0001 (0.0001)	-0.0002 (0.0001)	0.0000 (0.0001)	-0.0002 (0.0001)	0.0000 (0.0001)	-0.0002 (0.0001)
Summer Vacation	-0.0510 (0.0111)	-0.0433 (0.0162)	0.1356 (0.0276)	-0.0423 (0.0170)	-0.0317 (0.0112)	0.1631 (0.0170)
Chinese New Year	-0.0282 (0.0238)	0.0548 (0.0348)	0.1718 (0.0591)	0.0665 (0.0355)	-0.0071 (0.0235)	0.0659 (0.0355)
Book Dummies	Included	Included	Included	Included	Included	Included
Correlation Matrix of Residuals:						
Quantity of Creative Effort	1.0000		1.0000		1.0000	
Quality of Creative Effort	0.0111	1.0000	0.0028	1.0000	0.00285	1.0000
Customer Care Effort	0.1223	0.0278	1.0000	0.0173	0.1000	0.0172
Breusch-Pagan test of independence	chi2(3)=275.533, p=0.0000		chi2(3)=181.152, p=0.0000		chi2(3)=177.190, p=0.0000	

Notes. Standard errors in parentheses.
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Table 5: Robustness Checks I-III: Alternative Controls for Market Trend

	Robustness Check I				Robustness Check II				Robustness Check III			
	Control for Market Trend Using Searches in 2015		Control for Market Trend Using Average Searches from 2013 to 2015		Control for Market Trend Using Searches in 2015		Control for Market Trend Using Average Searches from 2013 to 2015		Month Fixed Effects			
	Creative Effort	Customer Care Effort	Creative Effort	Customer Care Effort	Creative Effort	Customer Care Effort	Creative Effort	Customer Care Effort	Quantity	Quality	Quantity	Quality
Termination Event	-0.0084 (0.0338)	-0.1249* (0.0511)	0.0033 (0.0817)	-0.0112 (0.0337)	-0.1244* (0.0510)	0.0063 (0.0816)	0.0000 (0.0000)	0.0000 (0.0000)	0.1020*** (0.0337)	0.1825**** (0.0513)	-0.6801**** (0.0819)	
Termination * Contracted	0.0887** (0.0382)	0.1858*** (0.0578)	-0.3233**** (0.0925)	0.0901** (0.0382)	0.1855**** (0.0577)	-0.3254**** (0.0924)			0.0201*** (0.0016)	0.0071**** (0.0024)	0.0008 (0.0039)	
<i>Sales_{t,t-1}</i>	0.0184**** (0.0015)	0.0078**** (0.0023)	-0.0041 (0.0036)	0.0183**** (0.0015)	0.0078**** (0.0023)	-0.0046 (0.0036)			0.0256**** (0.0016)	-0.0039 (0.0024)	0.0008 (0.0039)	
<i>Tips_{t,t-1}</i>	0.0285**** (0.0069)	-0.0050 (0.0104)	0.0194 (0.0341)	0.0284**** (0.0069)	-0.0049 (0.0104)	0.0202 (0.0341)			0.0137 (0.0069)	-0.0039 (0.0105)	0.0137 (0.0340)	
<i>NumTips_{t,t-1}</i>			0.0550 (0.0462)			0.0548 (0.0462)			0.0511 (0.0069)		0.0511 (0.0460)	
<i>Comments_{t,t-1}</i>	0.0053 (0.0136)	-0.0081 (0.0206)	0.2529**** (0.0329)	0.0051 (0.0136)	-0.0082 (0.0206)	0.2521**** (0.0329)			-0.0001 (0.0133)	-0.0084 (0.0202)	0.3233**** (0.0323)	
Contracted * <i>Comments_{t,t-1}</i>	0.0358* (0.0151)	0.0396* (0.0228)	0.0850** (0.0366)	0.0359** (0.0148)	0.0398** (0.0225)	0.0863** (0.0366)			0.0415*** (0.0148)	0.0409** (0.0225)	0.0159 (0.0361)	
Termination * <i>Sales_{t,t-1}</i>	-0.0085**** (0.0018)	-0.0073*** (0.0027)	0.0055 (0.0044)	-0.0085**** (0.0018)	-0.0072*** (0.0027)	0.0063 (0.0044)			-0.0059*** (0.0020)	-0.0062** (0.0030)	-0.0023 (0.0048)	
Termination * <i>Tips_{t,t-1}</i>	-0.0192** (0.0083)	0.0251** (0.0126)	-0.0742* (0.0416)	-0.0188** (0.0084)	0.0250* (0.0126)	-0.0748* (0.0416)			-0.0170** (0.0083)	0.0235* (0.0127)	-0.0593 (0.0415)	
Termination * <i>NumTips_{t,t-1}</i>			0.0343 (0.0564)			0.0332 (0.0564)			0.0284 (0.0083)		0.0284 (0.0562)	
Termination * <i>Comments_{t,t-1}</i>	-0.0200 (0.0170)	0.0199 (0.0258)	-0.0895*** (0.0413)	-0.0200 (0.0170)	0.0200 (0.0258)	-0.0878** (0.0413)			-0.0117 (0.0161)	0.0230 (0.0244)	-0.2161**** (0.0390)	
Termination * Contracted * <i>Comments_{t,t-1}</i>	0.0008 (0.0185)	-0.0359 (0.0280)	0.1223*** (0.0449)	0.0006 (0.0185)	-0.0361 (0.0280)	0.1205*** (0.0449)			-0.0141 (0.0174)	-0.0369 (0.0265)	0.2517**** (0.0423)	
Online Search of Focal Platform in 2015	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)						
Online Search of Focal Platform from 2013 to 2015												
Summer Vacation	-0.0384*** (0.0118)	-0.0511*** (0.0178)	0.1435**** (0.0285)	-0.0327*** (0.0116)	-0.0521**** (0.0175)	0.1351**** (0.0280)						
Chinese New Year	-0.0098 (0.0236)	0.0655* (0.0357)	0.0930 (0.0570)	-0.0075 (0.0236)	0.0651* (0.0356)	0.0898 (0.0570)						
Month Fixed Effects												
Book Dummies	Included	Included	Included	Included	Included	Included			Included	Included	Included	
Correlation Matrix of Residuals:												
Quantity of Creative Effort	1.0000			1.0000					1.0000			
Quality of Creative Effort	0.0025	1.0000		0.0025	1.0000				0.0016	1.0000		
Customer Care Effort	0.1000	0.0173	1.0000	0.1000	0.0173	1.0000			0.1053	0.0192	1.0000	
Breusch-Pagan test of independence			chi2(3)=177.016, p=0.0000			chi2(3)=177.174, p=0.0000					chi2(3)=196.875, p=0.0000	

Notes: Standard errors in parentheses.
 * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

Table 6: Robustness Check IV: Control for Supply-Side Seasonality

	Creative Effort		Customer Care Effort
	Quantity	Quality	
Termination Event	-0.0384 (0.0354)	-0.1485 ^{***} (0.0535)	0.0018 (0.0856)
Termination * Contracted	0.1090 ^{***} (0.0386)	0.1873 ^{***} (0.0584)	-0.3410 ^{****} (0.0934)
<i>Sales_{i,t-1}</i>	0.0185 ^{****} (0.0015)	0.0077 ^{****} (0.0023)	-0.0046 (0.0036)
<i>Tips_{i,t-1}</i>	0.0284 ^{****} (0.0069)	-0.0048 (0.0104)	0.0192 (0.0341)
<i>NumTips_{i,t-1}</i>			0.0550 (0.0462)
<i>Comments_{i,t-1}</i>	0.0023 (0.0136)	-0.0087 (0.0206)	0.2553 ^{****} (0.0330)
Contracted * <i>Comments_{i,t-1}</i>	0.0384 ^{**} (0.0151)	0.0405 [*] (0.0228)	0.0830 ^{**} (0.0366)
Termination * <i>Sales_{i,t-1}</i>	-0.0090 ^{****} (0.0018)	-0.0065 ^{**} (0.0028)	0.0068 (0.0044)
Termination * <i>Tips_{i,t-1}</i>	-0.0185 ^{**} (0.0083)	0.0248 ^{**} (0.0126)	-0.0738 [*] (0.0416)
Termination * <i>NumTips_{i,t-1}</i>			0.0335 (0.0564)
Termination * <i>Comments_{i,t-1}</i>	-0.0171 (0.0171)	0.0217 (0.0258)	-0.0908 ^{**} (0.0414)
Termination * Contracted * <i>Comments_{i,t-1}</i>	-0.0025 (0.0185)	-0.0385 (0.0281)	0.1226 ^{***} (0.0450)
Online Novel Search	0.0000 (0.0001)	-0.0002 [*] (0.0001)	-0.0002 (0.0001)
Writer Productivity in 2015	0.0001 ^{***} (0.0000)	0.0000 (0.0000)	-0.0001 (0.0001)
Summer Vacation	-0.0162 ^{***} (0.0122)	-0.0397 ^{**} (0.0184)	0.1479 ^{****} (0.0295)
Chinese New Year	-0.0019 (0.0235)	0.0667 [*] (0.0356)	0.0919 (0.0570)
Book Dummies	Included	Included	Included
Correlation Matrix of Residuals:			
Quantity of Creative Effort	1.0000		
Quality of Creative Effort	0.0025	1.0000	
Customer Care Effort	0.1003	0.0172	1.0000
Breusch-Pagan test of independence		chi2(3)=178.199, p=0.0000	

Notes. Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

Table 7: Robustness Check V: Control for Chapter Fixed Effects

	Creative Effort		Customer Care Effort
	Quantity	Quality	
Termination Event	-0.0542 (0.0351)	-0.1625*** (0.0532)	0.1061 (0.0842)
Termination * Contracted	0.0973** (0.0383)	0.1905*** (0.0580)	-0.1867** (0.0919)
<i>Sales_{i,t-1}</i>	0.0132**** (0.0017)	0.0066*** (0.0025)	0.0104** (0.0040)
<i>Tips_{i,t-1}</i>	0.0300**** (0.0069)	-0.0038 (0.0104)	0.0224 (0.0337)
<i>NumTips_{i,t-1}</i>			0.0369 (0.0457)
<i>Comments_{i,t-1}</i>	0.0029 (0.0136)	-0.0118 (0.0206)	0.2442**** (0.0327)
Contracted * <i>Comments_{i,t-1}</i>	0.0368** (0.0151)	0.0412* (0.0228)	0.0495 (0.0363)
Termination * <i>Sales_{i,t-1}</i>	-0.0067**** (0.0019)	-0.0060** (0.0028)	0.0027 (0.0044)
Termination * <i>Tips_{i,t-1}</i>	-0.0202** (0.0083)	0.0245* (0.0126)	-0.0636 (0.0412)
Termination * <i>NumTips_{i,t-1}</i>			0.0363 (0.0558)
Termination * <i>Comments_{i,t-1}</i>	-0.0213 (0.0170)	0.0217 (0.0258)	-0.0990** (0.0409)
Termination * Contracted * <i>Comments_{i,t-1}</i>	0.0020 (0.0185)	-0.0380 (0.021)	0.1590**** (0.0446)
Online Novel Search	0.0000 (0.0001)	-0.0002* (0.0001)	0.0000 (0.0001)
Summer Vacation	0.0144 (0.0119)	0.0372** (0.0181)	0.0300 (0.0286)
Chinese New Year	-0.0033 (0.0235)	0.0688* (0.0357)	0.0573 (0.0565)
Chapter Fixed Effects	Included	Included	Included
Book Dummies	Included	Included	Included
Correlation Matrix of Residuals:			
Quantity of Creative Effort	1.0000		
Quality of Creative Effort	0.0018	1.0000	
Customer Care Effort	0.1060	0.0173	1.0000
Breusch-Pagan test of independence		chi2(3)=198.262, p=0.0000	

Notes. Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

Table 8: Robustness Check VI & VII: Control for Front-Page Recommendation Information

	Robustness Check VI		Robustness Check VII	
	Control for Front-Page Recommendation Information (Recorded)		Control for Front-Page Recommendation Information (Recorded or Inferred)	
	Creative Effort Quantity	Customer Care Effort Quality	Creative Effort Quantity	Customer Care Effort Quality
Termination Event	-0.0109 (0.0344)	-0.1446** (0.0520)	-0.0247 (0.0832)	-0.0248 (0.0832)
Termination * Contracted	0.0904** (0.0382)	0.1845*** (0.0577)	-0.3228*** (0.0924)	-0.3225*** (0.0924)
<i>Sales_{i,t-1}</i>	0.0184*** (0.0015)	0.0077*** (0.0023)	-0.0044 (0.0036)	-0.0045 (0.0037)
<i>Tips_{i,t-1}</i>	0.0283*** (0.0069)	-0.0048 (0.0104)	0.0193 (0.0341)	0.0194 (0.0341)
<i>NumTips_{i,t-1}</i>			0.0551 (0.0462)	0.0549 (0.0462)
<i>Comments_{i,t-1}</i>	0.0057 (0.0136)	-0.0084 (0.0206)	0.2526*** (0.0329)	0.2522*** (0.0329)
Contracted * <i>Comments_{i,t-1}</i>	0.0356** (0.0151)	0.0402* (0.0228)	0.0855** (0.0366)	0.0855** (0.0366)
Termination * <i>Sales_{i,t-1}</i>	-0.0086*** (0.0018)	-0.0065** (0.0028)	0.0064 (0.0044)	0.0064 (0.0044)
Termination * <i>Tips_{i,t-1}</i>	-0.0185** (0.0083)	0.0247** (0.0126)	-0.0739* (0.0416)	-0.0742* (0.0416)
Termination * <i>NumTips_{i,t-1}</i>			0.0341 (0.0564)	0.0342 (0.0564)
Termination * <i>Comments_{i,t-1}</i>	-0.0204 (0.0171)	0.0214 (0.0258)	-0.0879** (0.0413)	-0.0876** (0.0413)
Termination * Contracted * <i>Comments_{i,t-1}</i>	0.0009 (0.0185)	-0.0381 (0.0280)	0.1195*** (0.0450)	0.1193*** (0.0450)
Online Novel Search	0.0000 (0.0001)	-0.0002* (0.0001)	-0.0002 (0.0001)	-0.0002 (0.0001)
Front-page Recommendation (Recorded)	-0.0272** (0.0137)	0.0094 (0.0208)	-0.0025 (0.0332)	(0.0001)
Front-page Recommendation (Recorded or Inferred)				0.0131 (0.0261)
Summer Vacation	-0.0322*** (0.0112)	-0.0419** (0.0170)	0.1630*** (0.0272)	0.1628*** (0.0272)
Chinese New Year	-0.0067 (0.0235)	0.0658* (0.0355)	0.0970* (0.0569)	0.0963* (0.0569)
Book Dummies	Included	Included	Included	Included
Correlation Matrix of Residuals:				
Quantity of Creative Effort	1.0000		1.0000	
Quality of Creative Effort	0.0026	1.0000	0.0026	1.0000
Customer Care Effort	0.1000	0.0172	0.1001	0.0171
Breusch-Pagan test of independence		chi2(3)=177.202, p=0.0000		chi2(3)=177.328, p=0.0000

Notes. Standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; **** $p < 0.001$.

Table 9: Robustness Check VIII: Analysis with Data before November 1, 2016

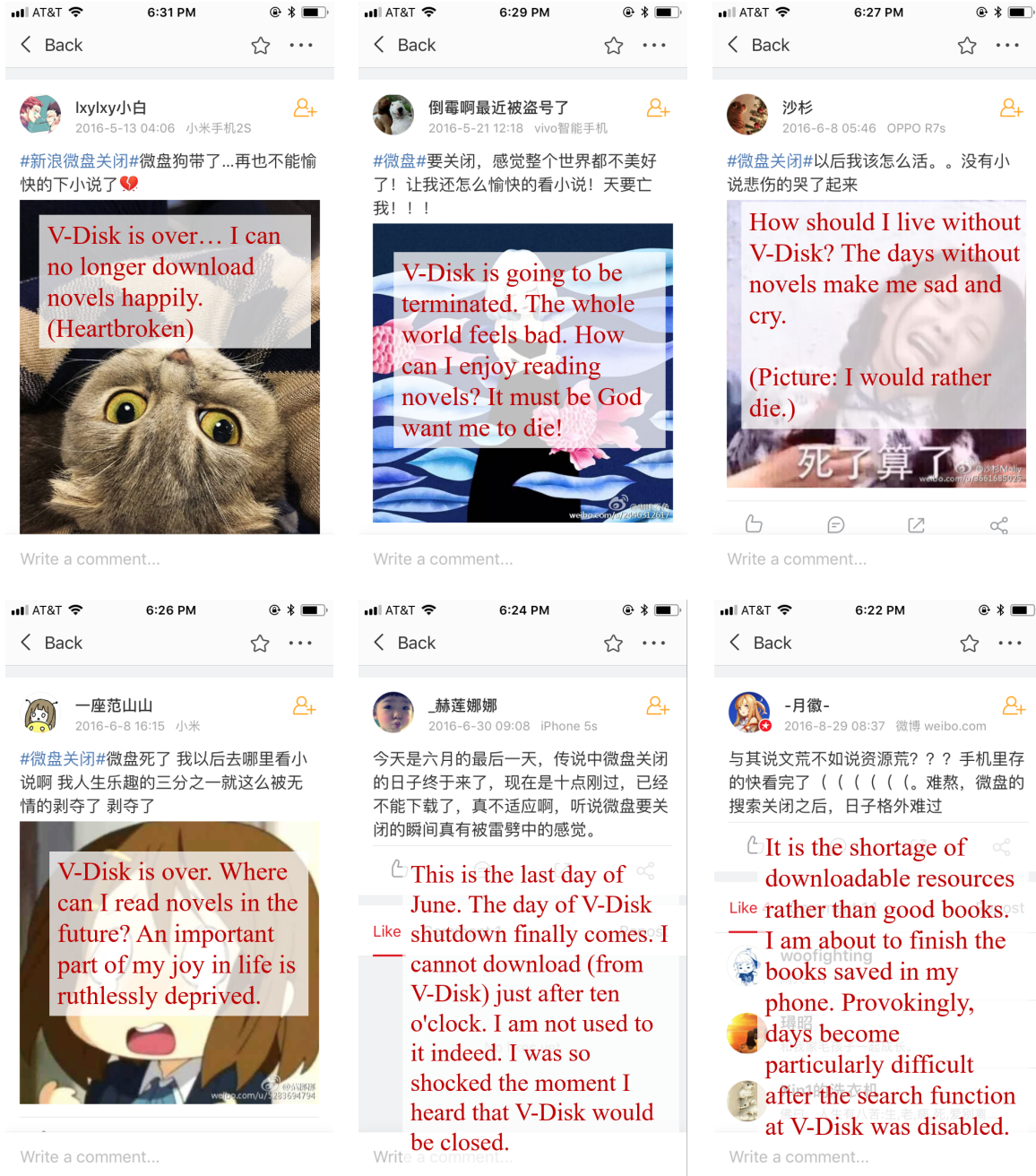
	Creative Effort		Customer Care Effort
	Quantity	Quality	
Termination Event	0.0169 (0.0387)	-0.1458** (0.0596)	0.0071 (0.0924)
Termination * Contracted	0.0832** (0.0405)	0.1990*** (0.0623)	-0.2115** (0.0967)
$Sales_{i,t-1}$	0.0188**** (0.0015)	0.0079**** (0.0023)	-0.0055 (0.0036)
$Tips_{i,t-1}$	0.0296**** (0.0068)	-0.0050 (0.0105)	0.0204 (0.3322)
$NumTips_{i,t-1}$			0.0556 (0.0450)
$Comments_{i,t-1}$	0.0040 (0.0135)	-0.0080 (0.0208)	0.2559**** (0.0323)
Contracted * $Comments_{i,t-1}$	0.0359** (0.0151)	0.0444* (0.0231)	0.0747** (0.0360)
Termination * $Sales_{i,t-1}$	-0.0069**** (0.0020)	-0.0059* (0.0031)	-0.0089* (0.0048)
Termination * $Tips_{i,t-1}$	-0.0085 (0.0091)	0.0249* (0.0140)	-0.0538 (0.0451)
Termination * $NumTips_{i,t-1}$			0.0422 (0.0609)
Termination * $Comments_{i,t-1}$	-0.0203 (0.0185)	0.0172 (0.0284)	-0.1062** (0.0441)
Termination * Contracted * $Comments_{i,t-1}$	-0.0049 (0.0201)	-0.0304 (0.0310)	0.1309*** (0.0481)
Online Novel Search	0.0003** (0.0001)	0.0001 (0.0002)	-0.0001 (0.0003)
Summer Vacation	-0.0397*** (0.0133)	-0.0467** (0.0204)	0.1172**** (0.0316)
Chinese New Year	-0.0074 (0.0233)	0.0648* (0.0358)	0.0957* (0.0555)
Book Dummies	Included	Included	Included
Correlation Matrix of Residuals:			
Quantity of Creative Effort	1.0000		
Quality of Creative Effort	0.0068	1.0000	
Customer Care Effort	0.1168	0.0235	1.0000
Breusch-Pagan test of independence		chi2(3)=180.602, p=0.0000	

Notes. Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

Appendix A Posts on Social Media Complaining about V-Disk Service Termination and Increased Difficulties to Obtain Pirated Novels

Figure A1: Examples of Social Media Posts



Appendix B Evidence: Readership of Authorized Copies Increased after V-Disk Service Termination Event

Here, we show preliminary evidence that the readership of books on authorized platforms increased after the termination of the search function and free personal cloud storage service at V-Disk. We focus on the monthly sales volume of the first pay chapter of each contracted book in our sample. This is a good measure of how many readers choose to pay for authorized copies (although some of them may drop out after buying a few chapters due to lack of interest) because we expect most readers will follow the narrative order of the story while reading a fiction book and are unlikely to skip the first pay chapter if they ever purchase any pay chapter. To show that the V-Disk service termination event increased the readership of authorized copies, we regress the monthly sales volumes of the first pay chapters on the event indicator variable.

Before running the regression, we note that the monthly sales volume of a pay chapter depends on the time elapsed since it was published. As Figure A2 illustrates, the longer since a chapter was published, the lower the monthly sales volume. Moreover, monthly sales volume decreases over time at a decreasing rate. To understand why the monthly sales volume decreases over time, consider the following scenario. Once a book is published, readers who are enthusiastic about the writer or interested in the genre are likely to actively look for the book, discover it early, and quickly decide whether to read it. Other readers will discover the book later, possibly through word-of-mouth. But the number of readers who newly discover this book and consume it will gradually decrease because readers are attracted to a huge number of new books and new chapters published on the platform every day.

For the first pay chapter of a contracted book i at time t , we model its monthly sales volume as follows:

$$Sales_{i,t} = \alpha_i + \beta_1 I_t + \beta_2 NMonthElapsed_{i,t} + \beta_3 NMonthElapsedSquare_{i,t} + \epsilon_{i,t} \quad (A1)$$

Here, I_t is the indicator for the event. $NMonthElapsed_{i,t}$ is the number of months elapsed since the chapter was published. $NMonthElapsedSquare_{i,t}$ is the square term to capture the nonlinear pattern of monthly sales volume over time. We include the book dummies α_i to capture

Figure A2: Monthly Sales Volume of the First Pay Chapters



the heterogeneous popularity across the books. The key coefficient of interest is β_1 . We present the estimation results in Table A1.

Table A1: Effect of V-Disk Service Termination on First Pay Chapter Sales

	Sales of the First Pay Chapter
Termination Event	110.1482 ^{**} (50.7042)
Number of Months Elapsed	-275.4074 ^{****} (19.0997)
Number of Months Elapsed, Square	23.7112 ^{****} (2.1991)
Book Dummies	Included

Notes. Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.

Consistent with the pattern shown in Figure A2, as time goes by, monthly sales volume of the first pay chapter of a contracted book decreases ($\beta_2 = -275.4074$, $p < 0.001$) at a decreasing rate ($\beta_3 = 23.7112$, $p < 0.001$). The coefficient of the event indicator variable is significantly positive ($\beta_1 = 110.1482$, $p < 0.05$), meaning that sales increased after V-Disk terminated the search function and free personal cloud storage service. Given that the contracted books in our sample were randomly selected from all contracted books on the platform, we expect the increased readership to come from the pirated versions of the same books, instead of other books on the platform. In other words, many readers switch to the authorized platform and start paying for the novel after the V-Disk service termination event.

Appendix C Details of Sentiment Classification: Preparation and Performance

C.1 Human Tagging Procedure

We used the following strategies to improve the quality of the training set we supply to Baidu’s customized version of sentiment classification API.

1. We randomly selected 5,000 reader comments for human tagging. For each fiction writing style (e.g., drama, tragedy, humor, comedy, horror), we selected at least 250 reader comments to ensure the performance of our sentiment classification tool on any type of work.
2. We randomly split the 5,000 comments into two sets of 2,500 comments. We repeated this random split process four times and obtained a total of eight sets of 2,500 comments. Then we hired eight annotators and assigned each to tag one set of comments. Following Liu et al. (2019), we asked the annotators to tag the sentiment of each comment on a 7-Likert Scale, from strongly negative (with a score of 1) to strongly positive (with a score of 7). There were altogether 19,983 sentiment scores rated by the eight annotators, with 17 scores missing due to errors in assigning the scores or negligence. For those 17 comments with missing scores, we hired a ninth annotator to tag the sentiments following the same rule.
3. It took nine to twelve hours for an annotator to tag 2,500 reader comments.
4. As a result, each one of the 5,000 comments received sentiment scores from four different annotators. Moreover, we minimized the influence of heterogeneity in subjective annotating criteria across the annotators with the random split process explained in point 2. We obtained the final classification by taking the average of scores and compared it with 4 (i.e., the score corresponds to “neutral” in a 7-Likert Scale). Among the 5,000 comments, 2,748 were classified as positive, 786 as neutral, and 1,466 as negative.
5. As suggested by the user instructions of Baidu’s customized sentiment classification API, we supplied 1,000 randomly-chosen positive comments and 1,000 randomly-chosen negative comments to train the customized sentiment classification model.

C.2 Performance of the Sentiment Classification Method

Baidu’s Sentiment Classification API, because of its advanced algorithm and large training set, is considered the best sentiment classification tool for Chinese texts.

In Table A2, we summarize the performance of Baidu’s Sentiment Classification API. As shown there, our customized sentiment classification model with the help of a training sample of 2,000 tagged comments greatly increased the accuracy rate, compared to Baidu’s generic sentiment classification model that is trained on texts not necessarily in the context of novel reader comments. Moreover, based on the performance on the hold-out sample, the accuracy rate of our customized sentiment classification API is rather consistent across books of different writing styles (Table A3) and non-contracted or contracted writers (Table A4).

Table A2: Performance of Baidu’s Sentiment Classification API

Keep the distribution of true positive/neutral/negative comments the same as in the original sample with 5,000 comments	Number of Comments	Accuracy Rate	
		Generic Sentiment Classification API	Customized Sentiment Classification API
No	3,000	53.77%	62.37%
Yes	1,590	58.24%	68.24%

Table A3: Accuracy Rate Across Books of Different Writing Styles

Writing Style	Number of Comments	Accuracy Rate
A	145	0.65
B	1,081	0.65
C	1,305	0.62
D	178	0.58
E	291	0.56

Notes. We do not disclose the specific writing styles assigned by the focal platform in order not to conceal its identity.

Table A4: Accuracy Rate Across Books of Contracted or Non-Contracted Writers

Writer Type	Number of Comments	Accuracy Rate
Contracted	2,632	0.62
Non-Contracted	368	0.62

To further show that the customized version of Baidu’s Sentiment Classification API yields a satisfactory prediction accuracy, we consider two sets of criteria: (1) the consensus between any pair

of annotators on the classification of positive, neutral, and negative reader comments; and (2) the accuracy rate of a random guess in the three-level classification, (i.e., 33.3%). While the consensus between human annotators, a measure of the intrinsic ambiguity of novel comment sentiment and the difficulty of the sentiments classification task, serves as a rough upper bound of the machine prediction accuracy, the accuracy rate of a random guess provides a lower bound of the accuracy rate. As listed in Table A5, the consensus between any pair of annotators on these 5,000 comments is between 54.1% and 73.3% (we omit the consensus rate associated with Annotator No. 9 due to the small sample size), with an average of 67.0%. On average, the performance of our sentiment classification model is very close to the upper bound defined by the human annotator consensus, which is well above the lower bound of a random guess.

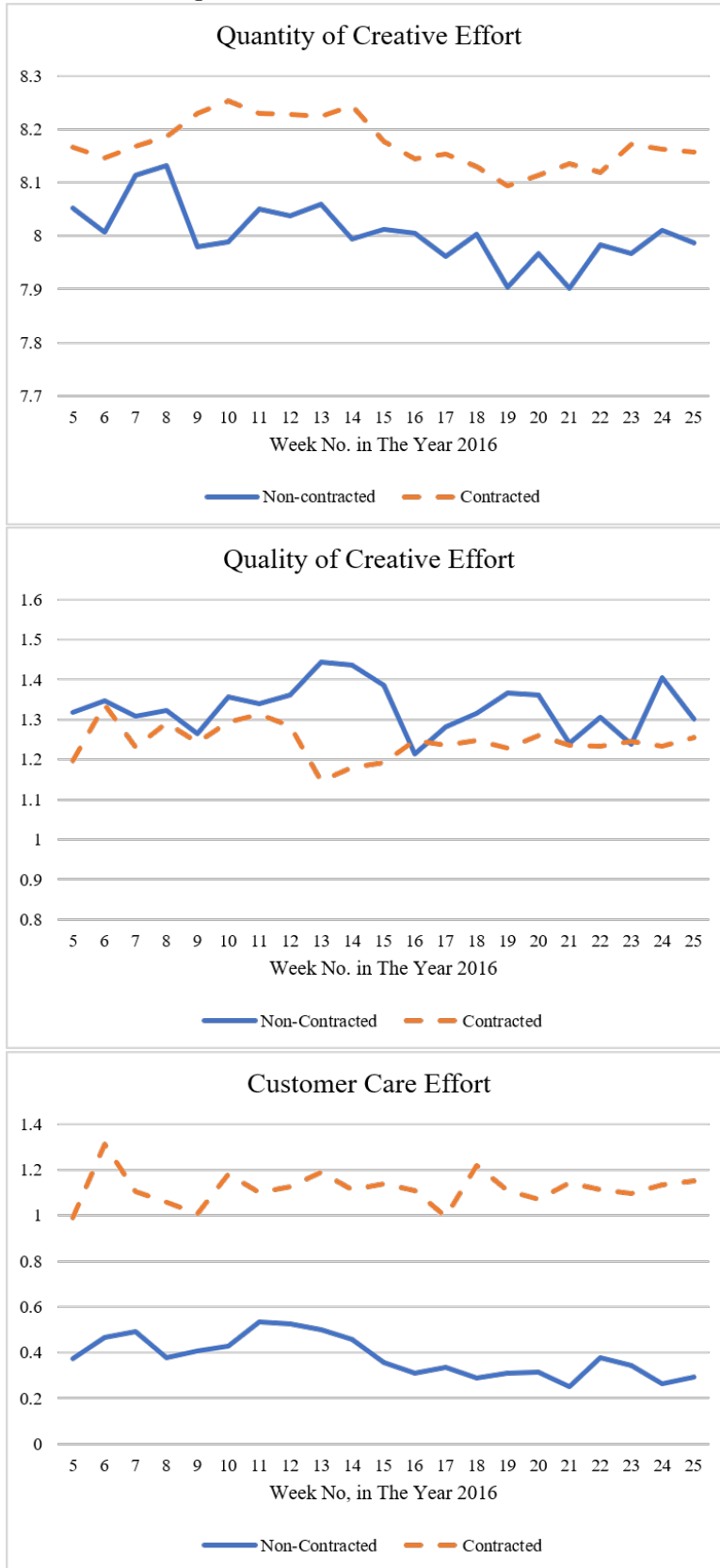
Table A5: Consensus between Annotators

	No. 1	No. 2	No. 3	No. 4	No. 5	No. 6	No. 7	No. 8	No. 9
No. 1	–	–	$\frac{846}{1207} \approx 70.1\%$	$\frac{905}{1282} \approx 70.6\%$	$\frac{731}{1271} \approx 57.5\%$	$\frac{863}{1218} \approx 70.9\%$	$\frac{844}{1240} \approx 68.1\%$	$\frac{855}{1246} \approx 68.6\%$	$\frac{2}{3} \approx 66.7\%$
No. 2	–	–	$\frac{938}{1286} \approx 72.9\%$	$\frac{892}{1214} \approx 73.5\%$	$\frac{720}{1224} \approx 58.8\%$	$\frac{912}{1276} \approx 71.5\%$	$\frac{904}{1253} \approx 72.1\%$	$\frac{846}{1244} \approx 68.0\%$	$\frac{2}{3} \approx 66.7\%$
No. 3	–	–	–	–	$\frac{799}{1281} \approx 62.4\%$	$\frac{882}{1219} \approx 72.4\%$	$\frac{856}{1260} \approx 67.9\%$	$\frac{834}{1235} \approx 67.5\%$	$\frac{7}{12} \approx 58.3\%$
No. 4	–	–	–	–	$\frac{710}{1219} \approx 58.2\%$	$\frac{939}{1281} \approx 73.3\%$	$\frac{870}{1240} \approx 70.2\%$	$\frac{875}{1259} \approx 69.5\%$	$\frac{3}{5} = 60\%$
No. 5	–	–	–	–	–	–	$\frac{699}{1264} \approx 55.3\%$	$\frac{668}{1234} \approx 54.1\%$	$\frac{3}{7} \approx 42.9\%$
No. 6	–	–	–	–	–	–	$\frac{856}{1236} \approx 69.3\%$	$\frac{830}{1260} \approx 65.9\%$	$\frac{6}{10} = 60\%$
No. 7	–	–	–	–	–	–	–	–	$\frac{4}{7} \approx 57.1\%$
No. 8	–	–	–	–	–	–	–	–	$\frac{1}{4} = 25\%$
No. 9	–	–	–	–	–	–	–	–	–

Appendix D Parallel Trend Test Before Termination Event

Before doing the DID analysis with the matched sample, we show that our treatment and control group satisfy the parallel trend assumption in the following ways. First, we plot the pre-treatment trends over time for the non-contracted and contracted books in the matched sample in Figure A3, which exhibits a pattern that is consistent with the parallel trend assumption for DID analysis. Second, we run a regression using data from the pre-treatment period and include the following explanatory variables: fixed effect for writer/book, a linear trend variable for updates, and the interactions between contracted or not and the update trend. According to the results summary in Table A6, we do not find significant interaction effects for any of the three dependent variables.

Figure A3: Pre-Treatment Trends



Notes. Due to the small sample size in the first four weeks (i.e., fewer than 25 books in control/treatment group), we do not display the information during this period.

Table A6: Parallel Trend Test on Efforts Before Termination Event

	Creative Effort		Customer Care Effort
	Quantity	Quality	
Contracted Books*t	-0.0013 (0.0009)	0.0015 (0.0020)	0.0015 (0.0040)
Books Fixed Effect	Included	Included	Included
Week Fixed Effect	Included	Included	Included

Notes. Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$.