AN EXPLORATION OF THE ECONOMICS OF CYBER LAW AND POLICY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Pengfei Zhang

May 2022

AN EXPLORATION OF THE ECONOMICS OF CYBER LAW AND POLICY

Pengfei Zhang, Ph.D.
Cornell University 2022

This dissertation is my exploration on an economic analysis of some emerging cyber law and policy issues. I focus on three topics corresponding to some of the most contentious pieces of law: copyright infringement, intermediary immunity, and antitrust. Chapter I presents a welfare evaluation of the widely adopted content takedown policy that secures copyright in cyberspace. Chapter II studies the optimal mediation design where disputants are asymmetrically informed and hard evidence can be acquired or presented with a cost. Chapter III discusses how a profit cap, imposed via taxation on a group of firms, can improve efficiency for both vaccine sharing and platform competition.

BIOGRAPHICAL SKETCH

Pengfei Zhang grew up in Beijing, China and graduated from Beijing 101 Middle School in 2012. He earned a Bachelor of Science in Mathematics and Global China Studies from the Hong Kong University of Science and Technology in 2016. He joined the Ph.D. program in economics at Cornell University in the fall of 2016, and aspires to earn his Ph.D. in May 2022. His primary research interests are law and economics, economic theory, and digital economics.

To my grandparents, my parents, and Ruizhe

ACKNOWLEDGEMENTS

I am truly indebted to my advisor Kaushik Basu for his tremendous encouragement and inspiration. Kaushik is a teacher, a mentor, a leader, a friend, and a role model. I am also very grateful for my committee members, Steve Coate, Jed Stiglitz, and Yi Chen for their continuous guidance and support. Steve taught me the arts of building policy models. Jed led me to the legal academy. Yi showed me how to be a productive young Chinese scholar.

I would like to thank Marco Battaglini, Tommaso Denti, Ted O'Donoghue, Kris Iyer, Wooyoung Lim, Suraj Malladi, Seth Sanders, Kathy Spier, Chantal Thomas, and Mike Waldman for their valuable comments and suggestions that greatly improve my work at different stages. George Hay's intriguing lectures and the CRADLE reading group laid the foundation for my understanding of the field of law and economics. At the final stage of my doctoral training, James Grimmelmann and Chris Forman provided me invaluable help in navigating the scholarly community on cyber law and digital economics.

My peers at Cornell have also been an incredibly valuable resource: Luming Chen, Senan Hennessy, Zihan Hu, Chenyang Li, Fikri Pitsuwan, Yu She, Haokun Sun, Qinshu Xue, Nahim Zahur, Bin Zhao, Tony Zhou, and Si Zuo have been consistently providing helpful comments. Yizhe Zhang provided me pivotal help with python scripts for data collection. I also acknowledge the Department of Economics, the Graduate School, the Einaudi Center, and the Kahin Center for financial support.

None of these would have been possible without the support of my wife, Ruizhe He, who has always been the cheer leader for this unusual journey.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 0

INTRODUCTION

Digital technologies and global digital networks offer virtually anyone the ability to access, store, transmit, and manipulate vast amounts of information. Consequently, digital technology both brings new benefits and poses new challenges to the economy. On one hand, digital technology reduces the cost of many economic activities, including the cost of search, the cost of replication, the cost of verification, among others (Goldfarb and Tucker, 2019). On the other hand, digital technology magnifies economic harm at an unprecedented scale, including the harm of piracy, the harm of defamation, the harm of misinformation, etc.

This revolution in the use of information raises new and often complex legal disputes and governance issues. In the new millennium, an increasing number of laws and regulations are drafted, discussed, and ready to profoundly shape the future of the digital economy. The U.S. has been the leader in this, but many countries in both the developed and the developing world have moved in the same direction by either following the U.S. model or venturing into some novel recipes. The collection of statutes, regulations, disputes, and cases, which affects people and businesses interacting through computers and the internet, is defined as cyber law and policy. It includes both private actions and government actions in cyberspace, and covers a variety of areas, such as intellectual property, antitrust, online speech, privacy, e-commerce, e-governance, cybercrimes, and recent developments in blockchains and cryptography.

To an economist, this seems to be a field of inquiry where prescription so far has outstripped analysis by a wide margin. It is the aim of this dissertation to analyze some of the pressing cyber policy issues using an economist's toolbox of microeconomic theory, causal inference econometrics, and behavioral insights. I focus on three topics recurring in the policy debates: copyright and creativity, intermediary power, Big Tech. These three topics correspond to some of the most contentious pieces of law that are fundamental to the digital world: on copyright infringement, it is the Section 512 of the Digital Millennium Copyright Act; on intermediary immunity, it is the Section 230 of the Communication Decency Act; on platform antitrust, it is the Section 1 and 2 of the Sherman Act.

Despite the digital context, the policy questions regarding copyright, dispute resolution, and antitrust are by no means novel. An immediate economist's response is: what is different? do policies that work for the conventional offline markets continue to work in the online era? The following three chapters present a coherent view that appropriate changes to these policies should be made. The changes could be limitations imposed on the current policy, expansions of the policy, or alternatives to the policy. In the case of digital copyright, open-access content generated by users has been phenomenal, which has never been observed in the offline counterpart. Chapter I shows how accounting for the social motivation behind such creations changes the classical intellectual property trade-off and consequently the welfare conclusions. In the case of online dispute resolution, information technology eases the process of verification, lowers the burden for the intermediary, and opens up new opportunities for intelligent design. Chapter II, based on a co-authored work with Yi Chen, shows how incorporating costly evidence into the optimal design escalates the power of mediation to solve information asymmetry and overcome the bargaining impasse. In the case of platform competition, debates have been lasting on whether antitrust law is capable of regulating the Big Tech and the Big Pharma with their natural economies of scale. Chapter III, based on a joint work with Kaushik Basu and Fikri Pitsuwan, shows that an industry-level profit cap, done right, can be a virtual substitute for antitrust law, with the advantage that efficiency is achieved without breaking up firms and hurting consumers.

I see this dissertation as the first step to a long journey, and in doing so, the chapters restrain themselves from reaching conclusive answers. My economic analysis is constrained by the availability of data and the welfare criterion of efficiency; throughout the analysis, I rest on the rationality assumption on individual decision making. An unbiased understanding of cyber law and policy, I believe, has to be interdisciplinary, bringing social scientists, legal scholars, computer scientists to the same table. As hard as the task might sound, the distinctive feature of human interactions mediated by computers, the extraordinary level of innovation and creativity on the internet, and the role of the government in the online space will keep fascinating generations of academic interests.

0.1 The Human Side of Cyber Takedowns: Theory and Evidence from GitHub

This paper presents a welfare evaluation of the widely adopted content takedown policy that secures copyright in cyberspace. In our analytical framework, unauthorized reproduction occurs in an open-access platform, either as an unproductive investment to steal the content or as a cost-saving method to expand the content. The first kind of reproduction reduces original contributions, whereas the second kind of reproduction invites them. The takedown policy is only socially beneficial in removing the former, but it allows a profit-maximizing copyright owner to block competition for the latter. An event study design using data from GitHub provides evidence that reproduction on this platform is most likely driven by the motive for content expansion, and enforcing the takedown policy is thus inefficient. This welfare loss can be addressed by an appropriate fair-use condition.

0.2 Mediation and Costly Evidence

This paper studies the optimal mediation design where disputants are asymmetrically informed and hard evidence can be acquired or presented with a cost. The model encompasses both facilitative mediation where the mediator only transmits information, and evaluative mediation where the mediator bases recommendation on verifiable information. We compare two modes of evaluation: the adversarial procedure where the disputants present evidence, and the inquisitorial procedure where the mediator acquires evidence. Any optimal mediation plan can be characterized by a simple threshold partitioning the ordered state space into a truth set and a pooling set. A welfare-maximizing mediator prefers the adversarial procedure, whereas a profit-maximizing mediator prefers the inquisitorial procedure. Full settlement is always the outcome irrespective of the procedures or the objectives. The derived mediation design matches the practice of professional mediators on mediation style, settlement rate, settlement terms, and satisfaction ratings.

0.3 Profit-Cap Policy: Theory and Applications

A known policy dilemma occurs between the need to curb extra-large profits by some industries, like pharmaceuticals, and the need to ensure the incentive to produce is not damaged. This paper shows that a profit cap, imposed via taxation on a group of firms, can simultaneously eliminate inefficiency and excess profit by intensifying competition among oligopolistic firms. The result has a direct bearing on current policy debates on COVID-19 vaccine sharing and antitrust problem of e-commerce platforms.

CHAPTER 1

THE HUMAN SIDE OF CYBER TAKEDOWNS: THEORY AND EVIDENCE FROM GITHUB

1.1 Introduction

A digital era has fallen upon us. By means of information technology, the massive interaction of people, goods, and services on the internet has given rise to a social environment called cyberspace. A key defining feature of cyberspace is the phenomenal sharing of knowledge through rich content provisions. Much of the digital content that users experience and enjoy is characterized by the notion of open access – the practice of making creative works freely available online to everyone (Suber, 2010). [1] The common theme of open-access platforms is that they are public goods provided to a large group of internet users and rely heavily on voluntary group contributions. [2] However, as internet users explore the abundance of knowledge shared online, the practice of unauthorized reproduction in open access is highly controversial. [3] Some reproduction practices in open-access content are plainly piracy, whereas many others are creative content expansions. [4] Distinguishing expanded content from stolen content is a critical issue that will profoundly shape the use and accumulation of knowledge for future generations.

[1]Such successful projects include open-source software (e.g., Linux and Mozilla), open educational resources (e.g., MIT OpenCourseWare and Coursera), open knowledge provision (e.g., Project Gutenberg and Wikipedia), open-access research and scientific collections (e.g., the Public Library of Science), and open media and news (e.g., citizen journalism).

[2]Benkler (2008) describes this production process as commons-based peer production – "a decentralized, collaborative, and nonproprietary process, based on sharing outputs among widely distributed, loosely connected individuals".

[3]Reproduction of content refers to the use of original content as input without the owner's permission. See Rob and Waldfogel (2006) on music file sharing. See Marron and Steel (2000) on software piracy.

[4]A famous example is Unix and Linux. Unix is one of the first operating systems developed and copyrighted by AT&T. In the 1990s, Linus Torvalds took the basics of Unix and made a better operating system called Linux. He posted his codes publicly so that people around the globe could add new features to it. As a result, Unix experienced revenue loss, and AT&T devoted substantial legal effort to stop Linux and its variants on the grounds of property rights. See Section 1.2 for more notable cases.

Before fully appreciating the delicacy of the task, policy makers began to draw legal boundaries on the use of knowledge and content in cyberspace. In the new millennium, there has been a concerted global effort toward enforcing authorized use and preventing copyright infringement in the digital environment. One of the most notable milestones was the content takedown policy that was first initiated by the United States under the 1998 Digital Millennium Copyright Act (DMCA hereafter) and has since served as a model for other countries. [5] The content takedown policy enables copyright owners to send takedown notices to request the removal of content that they believe to be infringing, and it requires online service providers to expeditiously honor such requests and remove infringing content in order to avoid legal liability. The notice and takedown system has been used extensively by copyright-based industries. According to the Lumen database, more than $40,000$ takedown notices per week are processed by major digital platforms (e.g., Google, Twitter, YouTube, Wikipedia), resulting in millions of removal requests addressed to internet users. The individual experiences of affected users occasionally make headlines, and estimates are emerging to give us a sense of the magnitude of its impact. A report from the French Ministry of Culture (2020) shows that 13% of French internet users (or, in absolute terms, more than 7 million French internet users) who have contributed content to some platform have had at least one upload blocked.

The welfare consequences of enforcing the takedown policy are a subject of contentious debate. Advocates, mostly corporations and lobby groups, argue that conventional copyright protection is inadequate for the digital environment as illegal copying is increasingly easy; thus, the celebrated takedown policy would restore the incentive to create. Critics contend that the takedown policy gives copyright owners excessive protection against future creators, which could severely constrain content reuse and future creativity (Lessig, 2004).

[5] Similar policies have been implemented in EU Article 13, China's 2010 Copyright Law Amendment, and India's 2012 Copyright Amendment Bill.

This paper proposes an analytical framework to evaluate the welfare consequences of the takedown policy and applies it to an empirical assessment of reproduction behavior on a major open-access platform: GitHub. We consider a canonical model of mixed-motive platform competition in which one platform, developed by the owner, is a profit-maximizing entity and the other platform, relying upon user-generated contributions, is freely available to any user. The novelty of the model is that a user has two ways to contribute to the open-access platform: she can contribute by original authorship at the same cost as the copyright owner, or she can choose to copy any portion of the proprietary content and contribute a reproduction. Reproduction is at a lower cost, but duplicate efforts are not useful for the platform. We model the takedown policy as a *zero* quantity constraint on the total amount of reproduction. We ask whether the resulting new equilibrium leads to an efficiency-improving allocation of content provision and consumption.

We start with a benchmark assumption on user preference in which users care only about their own consumption of platform content; we call this the standard neoclassical preference. Under the standard preference, users have a self-interest in strategically investing in reproduction to avoid a higher price. The laissez-faire equilibrium is always inefficient because reproduction crowds out original contributions and undermines the owner's incentive to provide content. Enforcing the takedown policy thus increases welfare by cutting unproductive and wasteful investments in reproduction. In fact, because of the free-riding problem, a ban on the open-access platform can achieve an efficient outcome.

However, the standard preference cannot explain the rich content users voluntarily provide on the internet. Based on empirical evidence regarding content creation and public good provision, we argue that the community preference, in which users care (to varying degrees) about the community consumption of open-access content conditional on whether they are part of the community, is a better model grounded on empirical evidence. Under the community preference, a user gains more satisfaction if either there is more content in the open-access platform or the size of the open-access community is larger. The reason for reproduction changes dramatically once we incorporate the assumption of community preference. Reproduction is a cost-saving way for content expansion. As more users are attracted to the open-access community, the marginal utility of original contributions increases as the platform content accumulates, which leads the users to contribute beyond the proprietary output, an amount no user would otherwise find incentive compatible without such reproduction. Enforcing the takedown policy decreases welfare by leading to a *contraction* in content provision.

We then attempt a welfare evaluation of enforcing the takedown policy on software developers' reproduction behavior using data from GitHub, the largest online host of open source code and repositories in the world. Based on our analytical framework, we devise a sufficient statistic test to overcome the difficulties caused by the unobservable nature of the distribution of the community preference and by the need to calibrate the entire model. The test relates the underlying preference to the nature of the game and leads to competing hypotheses regarding the effect of a takedown on undisputed original contributions. The "crowding out" hypothesis predicts that user i's contribution would *increase* with an exogenous removal of user j's reproduction if the reproduction were self-serving and thus were prescribed to be taken down. The "crowding in" hypothesis instead predicts that user i's contribution would *decrease* with an exogenous removal of user j's reproduction if reproduction were community driven and thus were not prescribed to be taken down.

In identifying an exogenous change only in the amount of reproduction, we exploit an institutional feature of GitHub's takedown protocol regarding fork repositories. A fork repository is a copy of a parent repository, and it usually contains both the content of its parent repository and its own newly added content. Upon receiving a takedown notice, GitHub automatically disables the content of a parent repository, but when disabling the parent repository, it does *not* disable the content of forks unless it receives a separate notice. [6] This allows us to estimate the effect of removing infringing content (caused by the takedown of parent repositories) on original contributions to fork repositories. The econometric specification is an event study with all treated units based on the exogenous random timing of takedown notices. We find a persistent decline in contributions to fork repositories following a takedown of the parent repository. We interpret this finding as evidence consistent with the "crowding in" hypothesis that reproduction on this platform is most likely driven by the community preference, and enforcing the takedown policy is therefore inefficient.

The rest of this paper proceeds as follows. The remainder of this section reviews the related literature. Section 1.2 reviews the copyright law relevant for understanding the content takedown policy and outlines notable legal cases of content takedown. Section 1.3 formally presents the model. Section 1.4 characterizes the laissez-faire equilibrium under both the standard and community preferences and contrasts them with the equilibrium resulting from the takedown policy, which forms the basis of the empirical hypothesis. Section 1.5 describes the GitHub data and empirical strategies. Section 1.6 presents the empirical results. Section 1.7 concludes.

[6]As mentioned in the literature review, platforms usually do not take an active role in checking the validity of requests. The current design of DMCA does not give them an incentive to do so.

1.1.1 Literature Review

Despite its manifest importance and prevalence, there has been very little work on content takedown policy. Most of the anecdotes about its effect seem to suggest that the DMCA takedown procedure removes more content than it should. Urban and Quilter (2005) presents the first set of descriptive statistics on the notice and takedown process based on the Lumen database. The authors found that corporations and business entities were the primary senders of the notices, a majority of the notices were sent for competition purposes, one third of the notices were questionable regarding the validity of the copyright infringement claim, and very few individual users responded with a counter-notice. [7] Penney (2019) surveys 1,296 panelists with hypothetical scenarios on receiving a takedown notice, indicating some chilling effects of the policy. Respondents broadly reported being less likely in future not only to share the same content again, but also to share content they themselves had created (72%). Only 34% said they would counter-notice or challenge a takedown they believed was wrong or mistaken. None of the cited empirical legal studies takes up the question what kinds of reproduction would have prevailed in absence of the policy, and this is critical in answering whether such removal of content is efficient or not from a social welfare perspective. This paper advances the literature by providing a framework for welfare evaluation of the takedown policy that serves as a basis for empirical applications to different platforms. To the best of our knowledge, this paper is the first to initiate a theoretical analysis of the takedown policy, and use the theory as a guide for a welfare evaluation of the policy in an empirical setting.

This paper contributes to the nascent and growing literature on digital economy within the large literature on copyright law. This literature consists of both legal case stud-

[7] In a follow-up study by Urban et al. (2017), the authors emphasize the role of automation in sending complaints, and compare how the automated notices differ from the manual notices by small rightholders. Seng (2020) compiles a larger dataset and shows more detailed statistics, questioning the validity of many takedown notices, especially those generated by automated systems.

ies and empirical economic research, with the key focus being the optimal enforcement of copyright in digital platforms. The pioneering work of Lessig (2004, 2009) argues that the expanded scope and overly broad interpretation of copyright law (e.g., DMCA) gives the corporate interests an unbalanced power against individual online content creators. The author advocates for balancing the incentive that comes from granting an exclusive right against the burdens such an exclusive right brings on the users and the would-be innovators. A similar position is shared by Benkler (2008) and Zittrain (2008), both of which highlight the importance of an open and creative cyberspace and warn how DMCA might threaten it. The empirical literature in economics can be classified into three categories: (i) the revenue consequences of online copying, (ii) how copyright affects the creation of new works, (iii) how copyright affects reuse and incentives to build on prior works (see Waldfogel (2017) and Goldfarb and Tucker (2019) for an extensive review, and the references therein). Most studies on (i) show that online copying displaces sales and reduces revenue in many copyright-protected industries (Liebowitz, 2006; Rob and Waldfogel, 2006; Oberholzer-Gee and Strumpf, 2007), but research on (ii) shows that quality of the content has not declined although copying is made easier by digital technology (Waldfogel, 2012a; Waldfogel, 2012b; Waldfogel and Reimers, 2015). Stricter enforcement of copyright, however, does appear to reduce incentives by others to build on copyrighted work (Nagaraj, 2018).

Novos and Waldman (1984) and Johnson (1985) are two leading models of copying and copyright, though their models exclusively focus on reproduction. Novos and Waldman (1984) consider a model where reproduction is costly and private, and they show that copyright protection can increase social welfare by providing higher profits for production, and perhaps surprisingly it will not decrease social welfare by restricting access because users are expending more resources in copying than what would be incurred if the good were purchased from the monopoly. [8] Johnson (1985) consider reproduction with differentiated products and heterogeneous tastes, and his finding is in line with Novos and Waldman (1984) that unlimited copying reduces social welfare and copyright could enhance social surplus. Neither papers consider original contribution and its relation with reproduction as we did in this paper, which turns out to be important for the welfare conclusion we can draw for the takedown policy.

This paper also closely relates to the literature of content provision on digital platforms. The literature so far highlights the social motivation in content provision and open source software (Lerner and Tirole, 2002, Lerner and Tirole, 2005a, Zhang and Zhu, 2011), illustrates the public good nature of open access platform (Johnson, 2002), and investigates the mixed oligopoly nature of competition between open access platform and proprietary platform (Mustonen, 2003, Casadesus-Masanell and Ghemawat, 2006, Athey and Ellison, 2014). Casadesus-Masanell and Ghemawat (2006) look at competition between the two forms of software in the presence of demand-side network externalities. They show that, while competition from open access software may cause the proprietary firm to lower its price, users may be worse-off from this competition as developers make fewer applications compatible with the proprietary software. Mustonen (2003) examines a model where the two platforms simultaneously compete for developers in the labor market and the consumption market market. Most related is Athey and Ellison (2014) who use a model of reciprocal altruism among open access users to examine competition between commercial software and open access software. They show that when developers care about the altruism displayed by other developers through their contributions, a "critical mass" hurdle needs to be passed for open access software development to occur. This paper features both a competing public goods model and community prefer-

[8] Even though Novos and Waldman (1984) did not analyze takedown policy explicitly, we could infer from their proposition 2 and 3 that if reproduction cost is uniformly distributed, the maximum amount of copyright protection is optimal.

ence similar in spirit to the one outlined by Athey and Ellison (2014). However, the focus of Athey and Ellison (2014) is on understanding the competition dynamics, whereas the key question we are interested in is the welfare effect of the takedown policy and copyright enforcement, thus the co-existence of reproduction and original contribution in user-generated content plays an indispensable role in our analysis in addition to the competition between the two platforms.

1.2 Institutional Background

1.2.1 Open Access Movement

The notion of open access appeared before the digital age. With the spread of the internet and the ability to copy and distribute electronic data at no cost, the idea of open access gained new importance. Open-access projects thrived at a time when major problems arose with the by-then dominant proprietary copyright model. Most commercial software vendors provide users only with object or binary code, which is difficult for programmers to interpret or modify.[9] At the same time, university librarians around the world found themselves in the middle of a major problem known as the "serials crisis" in which subscription costs for publications rose much faster than inflation for years. Libraries no longer had money to access all of the publications they wanted and were forced to make difficult choices between journals.

Both the scale and formalization of open access and peer production expanded dra-

[9]Software can be transmitted in either source code or object code. Source code uses languages such as Basic, C and Java. Object/binary code is a sequence of 0s and 1s that directly communicates with the computer.

matically with the widespread diffusion of the internet in the early 1990s. With the rapid development of open-access projects, there was also the proliferation of alternative approaches to licensing that supported cooperative peer production. The Free Software Foundation introduced a formal licensing procedure, called a General Public License, for a computer operating system called GNU. In keeping with the philosophy of the organization that the software should be free to use, free to modify and free to redistribute, the license aimed to preclude the assertion of copyright concerning cooperatively developed software.[10] Alternative approaches in the same spirit include the Berkeley Software Distribution (BSD) license and the Creative Commons license. These kind of licenses are under the "copyleft" umbrella because if copyright seeks to keep intellectual property private, copyleft seeks to keep intellectual property free and available. In the last few decades, as the international rules on cyberspace went global, so did the development of the open-access movement. Open source software has been developed in many parts of the world: Linux originated in Helsinki, Ruby was developed in Japan, and Ubuntu hails from South Africa.

1.2.2 Legal Framework

Copyright law provides protection for literary, scientific, and artistic works, giving creators the ability to control certain uses of their works. Copyright, in its current form, is a bundle of two major rights: (i) the exclusive right to make and distribute copies and (ii) the exclusive right for further derivative works.[11] The first and foremost international

[10] In exchange for being able to modify and distribute GNU software, software developers had to agree to a) make their source code freely available (or available at a nominal cost) to whomever the program was distributed and b) insist others who used the source code agreed to do likewise. Furthermore, all enhancements to the code—and even in many cases code that intermingled cooperatively developed software with software developed separately— had to be licensed on the same terms.

[11] A "derivative work" is defined as a work based upon one or more preexisting works, such as translation, musical arrangement, dramatization, fictionalization, motion picture conversion, sound recording, art

agreement governing copyright is the 1886 Berne Convention, which required automatic protection for all creative works in a fixed medium and enforced a minimum duration of protection (at least 50 years after the author's death for any work). The Berne Convention set up a bureau to handle administrative tasks, which later evolved to be the World Intellectual Property Organization (WIPO).

To address the new challenges posed by digital technologies, particularly the dissemination of protected material over the internet, WIPO introduced what they referred to as the "Internet treaties". The "Internet treaties" aimed at setting international norms of preventing unauthorized access to and use of creative works on the internet and any other digital networks.

One of the most prominent "Internet treaties" is the WIPO Copyright Treaty (WCT), effective in 2002. The WCT has broad coverage, dealing with protection for authors of literary and artistic works, such as writings and computer programs, original databases, musical works, audiovisual works, and works of fine art and photographs. The WCT clarifies that existing rights continue to apply in the digital environment and also creates new online rights. The WCT requires countries to provide frameworks of basic rights allowing creators to control and/or be compensated for the various ways in which their creations are used and enjoyed by others. Most importantly, the WCT ensures that the owners of those rights continue to be adequately and effectively protected when their works are disseminated through new technologies and communications systems such as the internet.

Calling upon a concerted effort from all member nations to curb digital copyright infringement, the WCT has attracted wide adherence, with 109 signatory countries. The

reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted.

Treaty does not by itself grant rights; rather, signatory countries implement the provisions of the WCT through national copyright and intellectual property rights legislation. In the U.S., such legislation includes the 1997 No Internet Theft (NET) Act, the 1998 Digital Millennium Copyright Act, and the 1999 Digital Theft Deterrence and Copyright Damages Improvement Act. The European Union enacted the 2001 Information Society Directive. China and India passed the 2010 Copyright Law Amendment and the 2012 Copyright Amendment Bill, respectively. By design, the WCT allows reasonable flexibility in establishing exceptions or limitations to rights in the digital environment. Countries may, in appropriate circumstances, grant exceptions for uses deemed to be in the public interest, such as for non-profit educational and research purposes. In reality, countries vary in their discretion in their commitment to the Treaty and their ability to enforce the legal measures.

Digital Millenium Copyright Act

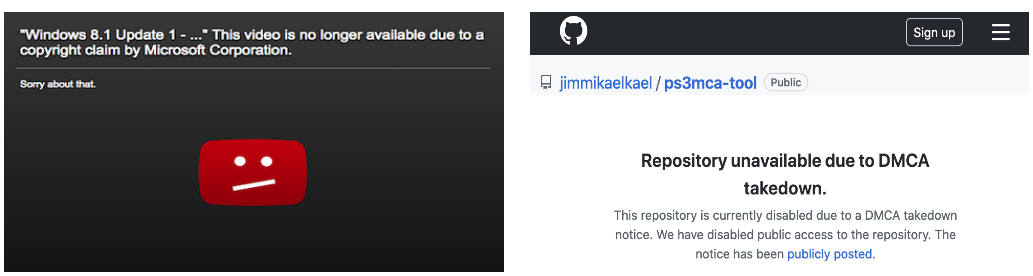
The DMCA provides a notice and takedown procedure for complaints about copyright infringement, including (i) a takedown-notice procedure for copyright holders to request that content be removed and (ii) a counter-notice procedure for users to have content reenabled when content is taken down by mistake or due to misidentification.

DMCA takedown notices are used by copyright owners to ask GitHub to take down content they believe to be infringing. Upon an initial investigation, the plaintiff prepares and sends a takedown notice to GitHub. Assuming the takedown notice meets the minimum requirements of the DMCA, GitHub posts the notice to its public repository and passes the complaint along to the affected user. [12] GitHub disables access to the defen-

[12]The minimum requirements include (a) identifying copyrighted works that are allegedly being infringed, (b) claiming under penalty of perjury that the plaintiff owns the copyright to the original work,

dant user's content if (i) the copyright owner has alleged copyright over the user's entire repository or package; (ii) the user has not made any changes after being given an opportunity to do so; or (iii) the copyright owner has renewed the takedown notice after the user had a chance to make changes.

Counter-notices can be used to dispute a takedown notice. If a user believes that her content was disabled as a result of a mistake or misidentification, she may respond with a counter-notice. GitHub posts it to the public repository and passes the counter-notice back to the copyright owner. If a copyright owner wishes to keep the content disabled after receiving a counter-notice, she needs to initiate legal action seeking a court order to restrain the user from engaging in infringing activity relating to the content on GitHub. If the copyright owner does not give GitHub notice within 10-14 days by sending a copy of a valid legal complaint filed in a court of competent jurisdiction, GitHub will reenable the disabled content.



(a) an example of takedown in YouTube (b) an example of takedown in GitHub

Figure 1.1: Examples of Content Takedown in Major Digital Platforms

Notable Cases of Content Takedown

The tension between open access and commercial interests was first exemplified by the UNIX litigation in the 1980s, when ATT began enforcing its (purported) intellectual prop-

and (c) demonstrating that the content on GitHub is unauthorized and infringing. GitHub exercises little discretion in the process, and it is up to the parties (and their lawyers) to evaluate the merit of their claims.

erty rights related to the operating system software UNIX, to which many academics and corporate researchers at other firms had made contributions. Berkeley Software Distribution (BSD) was an open-source operating system developed by Bill Joy and other researchers based on the Unix operating system jointly invented by American Telephone & Telegraph Co. (AT&T) and UC Berkeley. In 1991, AT&T sued the University of California at Berkeley for infringing copyrights on its Unix software, which the two parties ultimately settled. The settlement required that anybody who uses the university's Unix software also must pay royalties to the company. The key creator of BSD Bill left Berkeley and started his own operating system, Sun, which was later purchased by AT&T.

In 1997, MP3.com was launched by Mike Robertson to facilitate the sharing of new music, and it was soon the internet home to many independent musicians. To recommend new artists, MP3.com enabled users to insert their favorite CDs into their computer and discover similar content. Five major labels headed by the Recording Industry of American Association (RIAA) filed a lawsuit against MP3.com. Controversially, MP3.com was found guilty of willful infringement. A year later, one of the plaintiffs, Vivendi Universal, purchased MP3.com. The RIAA employed the same strategy in suits against individual users. The RIAA has sued more than 20,000 people suspected of distributing copyrighted works and has settled approximately 2,500 of these cases (Electronic Frontier Foundation, 2007).

1.3 Theory: Model

Primitives. We consider a canonical model of mixed-motive platform competition where one platform is a profit-maximizing entity, and the other platform is open-access to all users. [13] The model features two kinds of agents: an institutional copyright owner and a set of n users $\mathcal{N}$. The owner develops a proprietary platform with amount of content $\hat{x} \geqslant 0$ at cost $c\hat{x}$, and its contents are distributed only for profits. [14] Content in the open-access platform comprises creations competing with those of the proprietary platform; they are substitute goods. The platform is indivisible; that is, every user can consume at most one platform. Every user is endowed with $\overline{y}$ units of numeraire consumption goods, which can be used to purchase access to the proprietary platform, and $\overline{l}$ units of leisure, which can be used to contribute to the open-access platform at the opportunity cost of earning numeraire consumption goods.

An open-access platform relies upon user-generated content, and its output $x^* \geqslant 0$ is determined by all users' contributions. There are two ways that a user can contribute content to the open-access platform. The first method is to make any amount of original contribution z_i^o at unit cost c. The second method is to copy any portion of the proprietary output from the owner and make a reproduction z_i^r at unit cost cq, where $q \in (0, 1)$ is a scalar reflecting the difficulty of unauthorized use. We think of q as the extent of the technological barrier to copying; the higher the extent of the barrier, the higher the cost of reproduction.[15] The cost of reproduction is always lower than the cost of original production (i.e., $q < 1$), and the amount of individual reproduction is capped by the output of the proprietary source, i.e., $z_i^r \leqslant \hat{x}$ for any i. Furthermore, duplicates of reproduction do not count as additional content, and only the highest amount of individual reproduction

[13] Readers can find examples of mixed-motive platform competition documented in 1.2, where platforms can be source code repositories (e.g., UNIX vs. Berkeley Software Distribution) or online music archives (e.g., Vivendi vs. MP3.com), Britannica vs. Wikipedia, etc.

[14] We assume that the variable cost of distributing a copy is zero. This assumption is reasonable in the context of digital technology; see Goldfarb and Tucker (2019).

[15] The assumption that q is exogenous to individual users may appear reasonable enough to many readers. In rare cases, the action of easing and removing technological barriers to copying does happen and is called circumvention. However, unlike unauthorized use, circumvention is a criminal offense under U.S. copyright law, and banning circumvention does not seem to be a controversial issue. In that case, we can also think of q as a combination of the existing barriers and the extent of anti-circumvention enforcement against copying.

enters the production function. The production function of the open-access platform is

$$x^* = \sum_i z_i^o + \max_i z_i^r \tag{1.1}$$

such that the total amount of content is the sum of all users' original contributions plus the maximum of all users' reproduction efforts.

Timing. Figure 3.3 describes the sequence of events. The owner chooses $\hat{x}$, and users then generate open-access content x^*. Upon observing the output of the two platforms, the owner charges price $t \geqslant 0$, after which every user $i \in \mathcal{N}$ independently and simultaneously chooses whether to use the proprietary platform, the open-access platform, or neither of the platforms.

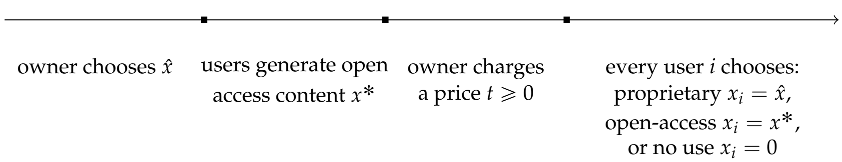


Figure 1.2: The Timing of the Model

Payoffs. The institutional copyright owner cares only about profits Π_0. We start with the benchmark assumption on user preference, which we call the standard preference. The standard preference is one where a user derives satisfaction from her consumption of the numeraire good y_i, her leisure l_i, and her consumption of the chosen platform content $x_i \in \{\hat{x}, x^*, 0\}$,

$$U_i(y_i, l_i, x_i) = y_i + l_i + \varphi(x_i), \tag{1.2}$$

where $\varphi(\cdot)$ is continuously differentiable, increasing, and concave in x with $\varphi(0) = 0$. However, this standard assumption on user preference cannot explain the rich content users voluntarily provide on the internet. Here, we borrow from the behavioral and experimental literature in which the other-regarding preference is the leading explanation for public good provision (Charness and Rabin, 2002; Andreoni, 2006). In our context, any such formulation must accommodate two crucial features grounded in empirical evidence: (i) an open-access contributor directly takes the benefits to other open-access users into account. [16] (ii) The utility of an open-access contributor is increasing in the number of other open-access users. [17]

We propose one such formulation of the user utility function where conditional on being members of the online open community, users care to varying degrees about the total consumption of the open-access content; we call this preference profile the *community preference*. The community preference of a user $i \in \mathcal{N}$ is represented by the following utility function $U_i(y_i, l_i, \mathbf{x})$:

$$U_i(y_i, l_i, \mathbf{x}) = y_i + l_i + \varphi(x_i) + \theta_i \sum_{j \in \mathcal{N}} \varphi(x_j) \mathbb{1}\{x_j = x_i = x^*\}, \tag{1.3}$$

such that every individual derives satisfaction from her consumption of the numeraire good y_i, her leisure l_i, her consumption of the chosen platform content x_i, and similar

[16] There is plenty of experimental evidence on the effect of the social preference in public good provision (Andreoni, 2006; 2007). The empirical finding in Fershtman and Gandal (2007) and the simple fact that open-source platforms establish hierarchies and are careful to credit contributors suggest that such motivations are important. Open-access communities explicitly recognize and honor altruistic behavior. Open-source platforms, for example, are scrupulous about keeping track of contributors, which reflects the fact that giving authors credit is essential in the open-source movement. This principle is included as part of the nine key requirements in the "Open Source Definition". This point is also emphasized by Raymond (1999), who points out "surreptitiously filing someone's name off of a platform is, in cultural context, one of the ultimate crimes." This point is also emphasized in Lerner et al. (2006), where they conduct interviews with open-source platform managers and SourceForge officials.

[17] Surveys of open-source participants indicate that programmers want to have an impact with their contributions, much as academics do. They appear to enjoy being part of important platforms, especially platforms that have a large user base. This suggests a relationship between altruism and the extent to which code is helpful to casual users. Shah (2006) quotes one programmer on this: "Why work on something that no one will use? There's no satisfaction there." Other supporting evidence includes that some programmers report that they monitor discussions of features they have developed even though they rarely take part in them. Lakhani and von Hippel (2003) similarly find, in a survey of forum participation in the Apache project, that a desire for enhanced job prospects is only a weak driver, with the key drivers instead being a desire to reciprocate for past help, a desire to promote OSS, or simply fun. In the context of Wikipedia, Zhang and Zhu (2011) show that the number of other users on the Wikipedia platform, or the audience size, positively influences the amount of editing. Gorbatai (2011) shows that experienced editors become more active when observing prior edits by casual users, which signal an interest in the topic. Kummer (2013) shows how exogenous shocks in readership lead to increased editing behavior on articles.

users' consumption of the open-access content x_j if $x_j = x_i = x^*$. Under the community preference, a user gains more satisfaction if either there is more content in the open-access platform or the size of the open-access community is larger. Users are heterogeneous in their weights on community consumption, indexed by their social type $\theta_1 \leqslant \theta_2 \leqslant \ldots \leqslant \theta_n$, and this is common knowledge.

The budget constraint of user i is $y_i = \overline{y} - t\mathbb{1}\{x_i = \hat{x}\} + \bar{l} - l_i - cz_i^o - cqz_i^r$ such that it balances the spending of the numeraire consumption good with disposable income, consisting of after-purchase income and labor income. Let $\mu = \sum_{j \in \mathcal{N}} \mathbb{1}\{x_i = x^*\}$. Given other users' choice, the end-node payoff of user i's is thus

$$U_i = \begin{cases} \overline{y} + \bar{l} + \varphi(\hat{x}) - t - cz_i^o - cqz_i^r & \text{if } i \text{ chooses the proprietary platform,} \\ \overline{y} + \bar{l} + \varphi(x^*) + \theta_i \mu \varphi(x^*) - cz_i^o - cqz_i^r & \text{if } i \text{ chooses the open-access platform,} \\ \overline{y} + \bar{l} - cz_i^o - cqz_i^r & \text{otherwise.} \end{cases}$$

Note that the option of no use is dominated by the use of open access.

Strategy and Solution Concept. In this game, the owner's pure strategy is a tuple $(\hat{x}, t(\hat{x}, x^*))$, where the pricing strategy $t(\hat{x}, x^*)$ is a complete contingent plan that picks the competition price for every history of the total contributions x^*, and user i's pure strategy is a triple $(z_i^o(\hat{x}), z_i^r(\hat{x}), x_i(\hat{x}, x^*, t)$ where a contribution plan is a function that picks the pair of original contribution and reproduction for every history of proprietary output $\hat{x}$, and a consumption plan is a function that picks whether and which platform to use for every history of the outputs of the two platforms and the price. We use the (pure strategy) subgame perfect Nash equilibrium as the solution concept.

Definition 1.1. *A strategy profile* $\{(\hat{x}, t(\hat{x}, x^*)), (z_i^o(\hat{x}), z_i^r(\hat{x}), x_i(\hat{x}, x^*, t))_{i \in \mathcal{N}}\}$ *is a subgame perfect Nash equilibrium such that* *i) For any $i \in \mathcal{N}$, user i chooses $x_i(\hat{x}, x^*, t)$ to maximize her payoff.*

ii) For any $i \in \mathcal{N}$, $(z_i^o(\hat{x}), z_i^r(\hat{x}))$ is a best response to $(z_j^o(\hat{x}), z_j^r(\hat{x}))_{j \neq i}$.

iii) The owner chooses $(\hat{x}, t(\hat{x}, x^*))$ to maximize his payoff.

Content Takedown Policy. The takedown policy is modeled as a *zero* quantity constraint on the total amount of reproduction. Under the takedown policy, the definition of an equilibrium remain the same except that the constraint $\sum_i z_i^r = 0$ is added. The takedown policy increases welfare if it increases the total surplus, i.e., $\Pi_0 + \sum_i U_i(y_i, l_i, \mathbf{x})$, subject to the constraints that the allocation $\{y_i, l_i, x_i\}_{i \in \mathcal{N}}$ can be implemented in the equilibrium.

1.4 Theory: Policy and Equilibrium Analysis

In this section, we characterize the social planner's solution, and the subgame perfect Nash equilibrium with and without the takedown policy for standard preference and community preference. We compare the welfare effects of the takedown policy under the two preference profiles, which hinges critically on which kind of reproduction would have prevailed in absence of the takedown. All proofs are delegated to the appendix.

1.4.1 Standard Preference: Inefficiency of Equilibrium

Denote μ as the number of open-access users, and consequently $n - \mu$ is the number of proprietary users. Under the quasi-linear preference, any efficient allocation must maximize aggregate surplus. The planner's problem is thus

$$\begin{aligned} \max_{\hat{x},(z_i^o,z_i^r,x_i)_{i\in\mathcal{N}}} \quad & (n-\mu)\varphi(\hat{x}) + \mu\varphi(x^*) - c\hat{x} - c\textstyle\sum_i z_i^o - (cq)\sum_i z_i^r \\ \text{s.t.} \quad & x^* = \max_i z_i^r + \textstyle\sum_i z_i^o, \\ & \hat{x} \geqslant 0, \\ & z_i^r \geqslant 0, z_i^o \geqslant 0 \quad \forall i, \\ & z_i^r \leqslant \hat{x} \quad \forall i. \end{aligned}$$

The three constraints are the feasibility constraints described in Section 1.3.

Define the inverse function of $\varphi'(\cdot)$ as $\phi(\cdot)$. A feasible allocation $\{\hat{x}, (z_i^o, z_i^r, x_i)_{i\in\mathcal{N}}\}$ is efficient if and only if [18]

$$\textit{either } \hat{x} = 0, \sum_i z_i^o = \phi(\frac{c}{n}), x_i = x^* \textit{ or } \hat{x} = \phi(\frac{c}{n}), \sum_i z_i^o = 0, x_i = \hat{x}, \tag{1.4}$$

$$z_i^r = 0 \quad \forall i \in \mathcal{N}. \tag{1.5}$$

To interpret, condition 1.4 states that the planner is indifferent between having all the users to use the open-access platform or to use the proprietary platform instead. In either case, the planner provides the efficient level of content by Samuelson rule (sum of the marginal benefits equal to the marginal cost) on that platform, and sets the content on the other platform to zero. Condition 1.5 says that in both cases, the efficient allocation involves no reproduction since reproduction is duplicating the cost of production without increasing the utility of any user, thus a waste of resources.

The market equilibrium under standard preference, however, is such that there is

[18] We assume $\varphi'(0) > \frac{c}{n}$ such that the planner always find it optimal to produce a positive amount. Otherwise, the problem becomes uninteresting.

always reproduction. The equilibrium can be solved by the standard technique of backward induction. At the final consumption stage, given the outputs of the two platforms x^* and $\hat{x}$, and the price t, a user will choose the proprietary platform over the open access if [19]

$$\varphi(x^*) \leqslant \varphi(\hat{x}) - t \Leftrightarrow t \leqslant \varphi(\hat{x}) - \varphi(x^*). \tag{1.6}$$

The availability of a substitute platform not only gives users more options but also disciplines the pricing of the copyright owner. The maximum price the owner can charge is the gap of user value between the two platforms. Given $\hat{x}$ and x^*, the owner's optimal pricing problem is as follows:

$$\begin{aligned} \max_{t} \quad & (n-\mu)t \\ \text{s.t.} \quad & t \leqslant \varphi(\hat{x}) - \varphi(x^*) \end{aligned}$$

Since the users are homogeneous, the owner either gets all users to accept the offer or none of the users. If $\varphi(\hat{x}) \leqslant \varphi(x^*)$, any positive price is not incentive compatible and $t = 0$. The profit-maximizing price in competition is $t = \max\{\varphi(\hat{x}) - \varphi(x^*), 0\}$. [20]

Contributions of users are contingent on which platform they will use. If user i is a proprietary user, given $t > 0$, her objective function is

$$\max_{z_i^r, z_i^o} [\varphi(\hat{x}) - t] - cz_i^o - (cq)z_i^r \Leftrightarrow \max_{z_i^r, z_i^o} \varphi(x^*) - cz_i^o - (cq)z_i^r$$

[19] We make the tie-breaking assumption that whenever a user is indifferent, she will choose the proprietary platform. Otherwise, the price can always be ϵ lower to make the user choose the proprietary platform.

[20] If t is such that the inequality is strict, then the owner can always raise t to $t + \varepsilon$ and earn a higher profit. Therefore, the constraint is binding if and only if $t > 0$.

The motivation behind the contribution made by the proprietary user is selfishness. A proprietary user contributes to limit the price that the owner can charge, such that she gets a better bargain with the owner by engaging in unauthorized copying. An open-access user, on the other hand, contributes for her subsequent enjoyment of the open-access content. If user i is an open-access user, she will solve

$$\max_{z_i^r, z_i^o} \quad \varphi(x^*) - cz_i^o - (cq)z_i^r$$

It so happens that under standard preference, the two objective functions are the same because the owner is able to extract all the rents from the users. Therefore, given $\hat{x}$, user i will solve the following optimization problem

$$\begin{aligned}
\max_{z_i^r, z_i^o} \quad & \varphi(x^*) - cz_i^o - (cq)z_i^r \\
\text{s.t.} \quad & x^* = \max\{z_i^r, \max_{j \neq i} z_j^r\} + z_i^o + \textstyle\sum_{j \neq i} z_j^o \\
& z_i^r \geqslant 0, z_i^o \geqslant 0 \\
& z_i^r \leqslant \hat{x}
\end{aligned}$$

Comparing with a typical public good game, there are two noteworthy differences: (i) a strategy of a user involves a choice of the platform in addition to the amount of contributions, (ii) the amount of contribution is a two-dimensional decision because a user has access to the cheaper technology of contributing a reproduction.

In Appendix B, we prove lemma 1.2 that in equilibrium there will be a unique user i^* that reproduces a positive amount, and which user is i^* is indeterminate. It is a result of the fact that the production function of open access only takes the maximum amount of all reproduction efforts. Intuitively speaking, if another user reproduces a lower amount, she would be better off not reproducing. If another user reproduces a higher amount, then

i^* would not reproduce. This lemma is a reminiscent of the pure strategy equilibrium in games of volunteer dilemma where one volunteer is sufficient for the provision of a public good.

Let λ^r be the Lagrange multiplier on the feasibility constraint of reproduction $z_i^r \leqslant \hat{x}$. An optimal (z_i^r, z_i^o) has to satisfy the following conditions:

$$\varphi'(z_i^o + \sum_{j \neq i} z_j^o + \bar{z}^r) = c \text{ or } z_i^o = 0 \quad (1.7)$$

$$\varphi'(z_i^r + \sum_{j} z_j^o) - \lambda^r = cq \text{ or } z_i^r = 0 \quad (1.8)$$

The two conditions are first order conditions of original contribution and reproduction respectively. In Appendix B, we prove lemma 1.3 that the optimal amount of user's reproduction either copies the entire work or balances the marginal utility of providing open-access content with the replication cost, i.e., $z_i^r = \min\{\hat{x}, z^r(q)\}$ where $\varphi'(z^r(q)) = cq$ for $i = i^*$. However, the pirated content crowds out any incentive to contribute original content as the self-interested users already reach their sanitation point by the cheaper technology of reproduction.

The owner chooses the output of the proprietary platform to maximize profits

$$\max_{\hat{x}} (n - \mu)t - c\hat{x}$$

$$\text{s.t. } t = \max\{\varphi(\hat{x}) - \varphi(x^*), 0\}$$

$$\Leftrightarrow \quad \max_{\hat{x}} \mathbb{I}\{\hat{x} > x^*(\hat{x})\} n[\varphi(\hat{x}) - \varphi(x^*(\hat{x}))] - c\hat{x}$$

The profit function depends critically on whether the proprietary platform has higher amount of content than the open-access platform.

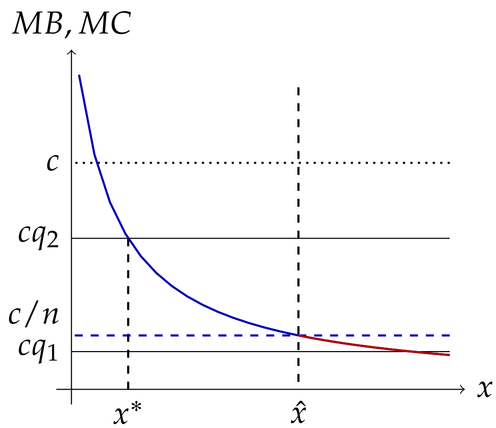


Figure 1.3: Equilibrium under Standard Preference

Figure 1.3 depicts a graphical illustration of the equilibrium under standard preference. It plots the marginal benefit and marginal cost of contribution as a function of the amount of content x in the open-access platform. Because of the concavity of the $\varphi(\cdot)$ function, the marginal benefit curve decreases as contents accumulate. The curve has a blue portion for $x \leqslant \hat{x}$ representing the marginal utility when the user is a proprietary user, and a red portion $x > \hat{x}$ representing the marginal utility when the user is an open-access user. The horizontal lines cq_1 and cq_2 represent two replication costs that lead to two different scenarios. Suppose the replication cost is at cq_1, the horizontal line crosses the marginal benefit curve at a point beyond $\hat{x}$, meaning that the user will copy the entire proprietary output, and consequently the owner has no incentive produce anything on the proprietary platform in the first place, i.e., $\hat{x} = 0$. This case generalizes to all $q \leqslant \frac{1}{n}$ such that if copying is extremely easy, reproduction crowds out proprietary output. The owner produces no content, and there is severe under-provision in the open access. Suppose the replication cost is at cq_2, the equilibrium amount of content x^* is identified by the intersection of the marginal benefit curve and the replication cost line. To see why x^* is an equilibrium, notice that at cost cq_2 the user can always increase the amount of reproduction incrementally without hitting the $\hat{x}$ cap, and consequently the user has no incentive to contribute any original content as she can always substitute original contribution with cheaper reproduction. This case generalizes to all $q > \frac{1}{n}$ such that if the cost of copying is bounded below, reproduction crowds out original contributions. The proprietary platform has superior amount of content, and the open access platform only contains an inferior copy of the proprietary output, i.e., $x^* = z^r = \phi(cq) \leqslant \phi(\frac{c}{n}) = \hat{x}$. As $n \to \infty$, the first case becomes unimportant, and the only relevant case is the second one.

Proposition 1.1. *Under standard preference, the market equilibrium is inefficient for $q \in (0, 1)$.*

The inefficiency of the market equilibrium comes from the fact that reproduction is an unproductive and wasteful investment into evading barriers. In absence of social preference, the free-riding problem is severe for public good provision.

1.4.2 Standard Preference: Takedown Policy

Let us now consider how a regulatory authority guided by copyright law can intervene to improve efficiency of the content platforms. Since reproduction is the source of inefficiency, it appears sensible that the takedown policy should target at reproduction by removing all the pirated content of the open-access platform. Under the content takedown policy, the total amount of reproduction will be pushed to zero, i.e., $z_i^r = 0$ for all i. Accordingly, user i's optimization problem changes to

$$\begin{aligned} \max_{z_i^o} \quad & \varphi(x^*) - cz_i^o \\ \text{s.t.} \quad & x^* = z_i^o + \textstyle\sum_{j \neq i} z_j^o \\ & z_i^o \geqslant 0. \end{aligned}$$

User-generated content becomes a typical public good game, and the free-riding problem of under provision rightfully applies. Following the steps in the last subsection, we can show that the takedown policy results in a new equilibrium where the owner provides the Samuelson level of content (i.e., $\hat{x} = \phi(\frac{c}{n})$), users strategically invest in original open-access content to avoid higher price (i.e., $x^* = \phi(c)$), and all users use the proprietary platform (i.e., $x_i = \hat{x}$). Comparing the social welfare of the policy-induced equilibrium and that of the market equilibrium, we conclude

Proposition 1.2. *Under standard preference, the takedown policy increases welfare.*

The removal of reproduction has two effects on the equilibrium: (i) the profits of the copyright owner increase, (ii) original contributions increase, $\sum_i z_i^o = \phi(c) \geqslant \max\{\phi(c), 0\}$. Since the profits motivate the owner to provide higher amount of content, both effects are welfare-improving.

In fact, under standard preference, one of the social planner's solutions is achieved if we ban the open-access platform completely, even though that makes the proprietary platform a monopoly. Suppose the proprietary platform is a monopoly, let the price be t_m, a user will pay for the platform if $t_m \leqslant \varphi(\hat{x})$. The owner's profit maximization problem is

$$\begin{aligned} \max_{\hat{x}, t_m} \quad & (n - \mu) t_m - c\hat{x} \\ \text{s.t.} \quad & t_m \leqslant \varphi(\hat{x}) \end{aligned}$$

It is easy to see that the profit-maximizing monopoly price is $t_m = \varphi(\hat{x})$, and the amount of content is $\hat{x} = \phi(\frac{c}{n})$, which coincides with the efficient level by Samuelson rule. There is no monopoly distortion because users are homogeneous. Although the monopoly solution is efficient, it has undesirable distributional consequence as the copyright owner extracts all the rents from the users. The takedown policy, in this extreme form, can be viewed as assigning full property rights to the copyright owner, and this efficiency result can be understood as an application of the Coase theorem (Coase, 1960). Coase theorem states that if property rights are clearly specified and there are no transactions costs, bargaining will lead to an efficient outcome. Likewise in our model, the owner internalizes the positive spillover effects of the proprietary content to all users through monetary transfers. The difference is that to which party the property rights are assigned to is important in our model. Suppose we assign property rights to a group of users, they will not contribute up to the Samuelson level as open access does not allow them to internalize the marginal social benefit via transfers. Property rights thus solve the "missing market" problem of public good provision under standard preference. From the perspective of transaction cost economics, the transaction cost within a group of users is high, whereas the transaction cost with a single copyright owner is low. Assigning full property rights to the owner is an institution that minimizes the transaction cost.

1.4.3 Community Preference: Equilibrium

This sub-section examines the implications of community preference for reproduction behavior and the market equilibrium. We will delegate the discussion on social planner problem to the next sub-section. The current model is flexible enough to allow for predictions on all three possible outcomes of the mixed-motive platform competition: (i) the proprietary platform dominates and all the users choose the proprietary platform, (ii) the two platforms coexist and the users self-select into them respectively based on their types, (iii) the open-access platform eliminates the proprietary platform and all the users choose the open-access platform. Nevertheless, (iii) distinguishes itself from the other two by having positive and substantial original contributions that outperforms the proprietary output.

Recall μ is the number of users that choose open access. User i will pay for the proprietary platform instead of using the open access if

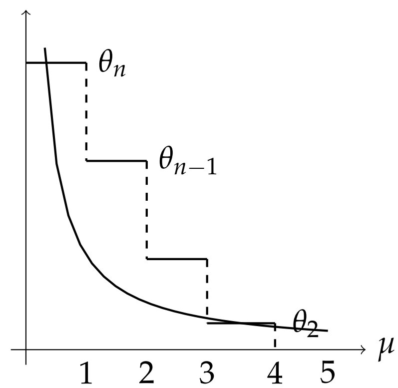
$$\varphi(\hat{x}) - t \geqslant \varphi(x^*) + \theta_i \mu \varphi(x^*). \tag{1.9}$$

Contrasting with equation 1.6, the community preference gives rise to network effects where a user's content consumption decision depends on how many users are using the open-access platform. Given a price t, users will self-select into the two platforms depending on θ_i.

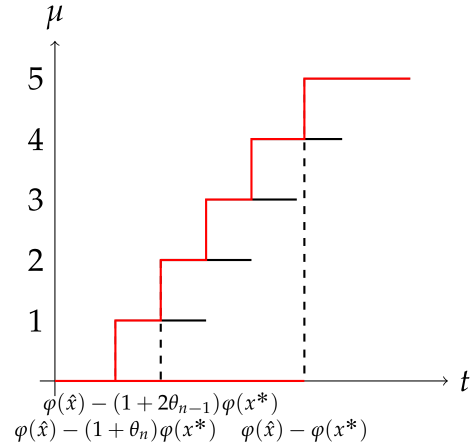
Equation 1.9 leads us to three cases of the content consumption stage. If $x^* > \hat{x}$, then $\mu = n$. If $x^* \leqslant \hat{x}$ and $\mu = 0$, then the analysis is similar to that of standard preference, and $t \leqslant \varphi(\hat{x}) - \varphi(x^*)$. Suppose $x^* \leqslant \hat{x}$ and $\mu > 0$, the above equation can be written as

$$\theta_{n-\mu} \leqslant \frac{\varphi(\hat{x}) - \varphi(x^*) - t}{\mu \varphi(x^*)} \leqslant \theta_{n-\mu+1}$$

This inequality states that users whose types are below the cutoff will choose the proprietary platform, and users whose types are above the cutoff will choose the open access platform. Figure 1.4(a) plots this inequality in relation to μ taking the price t as given where the hyperbola of μ is the term in the middle, and the step function is the type θ_i ordered from the highest to the lowest. The equilibrium μ is a solution of the fixed point equation, which can be identified as the intersection of the hyperbola and the step function. In the graph, there are three fixed points: $\mu = 0$, $\mu = 3$, and $\mu = 4$. Each of them is a possible outcome of self-confirming rational expectation. For example, suppose the user with θ_2 joins the open access, the value of the hyperbola will be lower with a larger size of community $\mu = 4$, which makes it indeed optimal for θ_2 to join. Suppose instead the user with θ_2 does not join the open access, the value of the hyperbola will be higher with a smaller size of community $\mu = 3$, which makes it indeed optimal for θ_2 not to join.



(a) fixed point of number of open-access users



(b) equilibrium number of open-access users as a function of price

Figure 1.4: Optimal Price and Equilibrium Content Consumption

Figure 1.4(b) varies the price t and plots the possible equilibrium number of open-access users μ as a correspondence of the price t. Consistent with figure 1.4(a), a price t corresponds to three candidate μ, including $\mu = 0$, a smaller positive μ, and a larger positive μ. A higher price has two offsetting effects on the owner's profits: as t increases, it increases the per-user revenue but pushes more users to choose open access. Thus, there is an interior optimal price given the outputs $\hat{x}$ and x^*.

The owner's profit can be written as a function of μ, and the owner's pricing problem can be recast as equivalently choosing μ as follows:

$$\max_{\mu \in [0,n]} \quad (n - \mu)t(\mu)$$
$$\text{s.t.} \quad t(\mu) \leqslant \varphi(\hat{x}) - \varphi(x^*) - \mu\theta_{n-\mu}\varphi(x^*).$$

Denote the solution to the problem as $\mu(x^*)$, and $\mu(x^*)$ must satisfy the first order condition

$$[1 + r(\mu(x^*)) - r(\mu(x^*)+1)]\varphi(x^*) < \varphi(\hat{x}) \leqslant [1 + r(\mu(x^*)-1) - r(\mu(x^*))]\varphi(x^*),$$

where $r(\mu) = \mu(n-\mu)\theta_{n-\mu}$. The above optimality condition states that $\mu(x^*)$ is optimal if moving from $\mu(x^*) - 1$ to $\mu(x^*)$ strictly increase profits (by charging a higher price) and moving further from $\mu(x^*)$ to $\mu(x^*) + 1$ does not. We have now derived the equilibrium number of open access users and equilibrium price given x^*:

$$(\mu, t)(x^*) = \begin{cases} (1, \varphi(\hat{x}) - (1 + 2\theta_{n-1})\varphi(x^*)) & \text{if } \varphi(x^*) \in (\frac{\varphi(\hat{x})}{1+r(0)-r(1)}, \frac{\varphi(\hat{x})}{1+r(1)-r(2)}], \\ (2, \varphi(\hat{x}) - (1 + 3\theta_{n-2})\varphi(x^*)) & \text{if } \varphi(x^*) \in (\frac{\varphi(\hat{x})}{1+r(1)-r(2)}, \frac{\varphi(\hat{x})}{1+r(2)-r(3)}], \\ \ldots & \\ (n, 0) & \text{if } \varphi(x^*) > \frac{\varphi(\hat{x})}{1+r(n-1)-r(n)} = \varphi(\hat{x}). \end{cases}$$

As the open-access content x^* increases, higher-type users will be attracted to use the open-access platform, which forces the owner to focus on the lower-type segment of the users. As the open-access option becomes more attractive, the optimal price charged for the remaining proprietary users has to be lower.

The consumption stage implies two different objective functions for user-generated content. If user i is a proprietary user, given $t > 0$ and $\{z_j^r, z_j^o\}_{j \neq i}$, she will solve

$$\begin{aligned} & \max_{z_i^r, z_i^o} \varphi(\hat{x}) - t - cz_i^o - (cq)z_i^r \\ \Leftrightarrow \quad & \max_{z_i^r, z_i^o} [\varphi(x^*) + \theta_{n-\mu(x^*)}\mu(x^*)\varphi(x^*)] - cz_i^o - (cq)z_i^r \end{aligned}$$

Observe that the objective function of a proprietary user is a concave function. As x^* goes higher, the price the owner can charge is increasingly constrained, and the return to marginal reproduction for a proprietary user gets smaller. Because of that, contributions of a proprietary user is decreasing in others' contributions. If user i is an open access user, given $\{z_j^r, z_j^o\}_{j\neq i}$, she will solve

$$\max_{z_i^r, z_i^o} \quad [\varphi(x^*) + \theta_i \mu(x^*)\varphi(x^*)] - cz_i^o - (cq)z_i^r$$

We know that $\mu(x^*)$ is increasing and convex in x^*. We prove in Appendix B that the objective function of an open access-user, given $\theta_i > \theta_{n-\pi}$, is instead a convex function. Intuitively, a marginal contribution to open access not only increase its user value but also attracts marginally more users into the platform, which raises the marginal utility of contribution for an open-access user. As individual contributions become higher, even more users find the open-access platform attractive, which further increases the marginal utility of contribution. When all the users join the open-access platform, the marginal utility of contribution for an open-access user reaches the peak. Because of that, contributions of an open access user is increasing in others' contributions.

Given $\hat{x}$, user i's optimization problem is as follows:

$$\begin{aligned} \max_{z_i^r, z_i^o} \quad & \max_{x_i}\{\theta_i \mu(x^*)\varphi(x^*), \theta_{n-\mu(x^*)}\mu(x^*)\varphi(x^*)\} + \varphi(x^*) - cz_i^o - (cq)z_i^r \\ \text{s.t.} \quad & x^* = \max\{z_i^r, \max_{j\neq i} z_j^r\} + z_i^o + \textstyle\sum_{j\neq i} z_j^o \\ & z_i^r \geqslant 0, z_i^o \geqslant 0 \\ & z_i^r \leqslant \hat{x} \end{aligned}$$

Comparing with the users' optimization problem under standard preference, the critical difference under community preference is that the sub-objective of an open-access user diverges from that of a proprietary user. Accounting for the self-selection into platforms, the user's objective function is thus the maximum of the two sub-objectives, and an optimal plan has to compare the utility of choosing open access and contributing *as* an open-access user with the utility of choosing proprietary and contributing *as* a proprietary user.

Denote $\hat{MB}(x)$ as the marginal utility of contribution for a proprietary user $x_i = \hat{x}$, and $MB^*(x)$ as the marginal utility of contribution for an open access user $x_i = x^*$. By lemma 1.2, we know there exists a user i^* such that $z_i^r > 0$ for $i = i^*$ and $z_i^r = 0$ for any $i \neq i^*$. An optimal (z_i^r, z_i^o, x_i) has to satisfy the following conditions:

$$\mathbb{1}\{x_i = x^*\}MB^*(z_i^o + \sum_{j \neq i} z_j^o + \bar{z}^r) + \mathbb{1}\{x_i = \hat{x}\}\hat{MB}(z_i^o + \sum_{j \neq i} z_j^o + \bar{z}^r) = c \text{ or } z_i^o = 0, \tag{1.10}$$

$$\mathbb{1}\{x_i = x^*\}MB^*(z_i^r + \sum_i z_i^o) + \mathbb{1}\{x_i = \hat{x}\}\hat{MB}(z_i^r + \sum_i z_i^o) - \lambda^r = cq \text{ or } z_i^r = 0, \tag{1.11}$$

$$U_i(z_i^r, z_i^o, x^*) \geqslant U_i(z_i^r, z_i^o, \hat{x}) \text{ if and only if } x_i = x^*, \tag{1.12}$$

$$MB^*(x) = \varphi'(x) + \theta_i[\mu(x^*)\varphi'(x) + \mu'(x^*)\varphi(x))], \tag{1.13}$$

$$\hat{MB}(x) = \varphi'(x) + [r(x)\varphi'(x) + r'(x)\varphi(x)]. \tag{1.14}$$

Condition 1.10 and 1.11 are first order conditions of original contribution and reproduction respectively, taking into account the fact that the marginal benefit of contribution varies with the platform choice of user i. Condition 1.12 disciplines the sequential rationality of the users, which states that if user i chooses the open-access platform, the complete plan of choosing and contributing to the open-access platform gives her a higher utility than the complete plan of choosing the proprietary platform and contributing to the open-access platform.

In Appendix B, we show that given $\hat{x}$, a reproduction and consumption scheme $\{z_i^r, x_i\}_{i \in \mathcal{N}}$ is characterized by a triple (i^*, z^r, μ) where i^* is the unique user who reproduce a positive amount, z^r is the amount of her reproduction, and μ is the number of open-access users. Define $\bar{x}(\theta_i, \mu, c)$ as the individual satiation level of the amount of content such that equation 1.10 holds with equality for an open-access user i, i.e., $(1 + \theta_n \mu)\varphi'(\bar{x}(\theta_n, n, c)) = c$. It is easy to show by implicit function theorem, $\bar{x}(\theta_i, \mu, c)$ is increasing in θ and μ, and decreasing in c. A key quantity is $\bar{x}(\theta_n, n, c)$, the ideal level of the highest-type user given all the users choose the open-access platform.

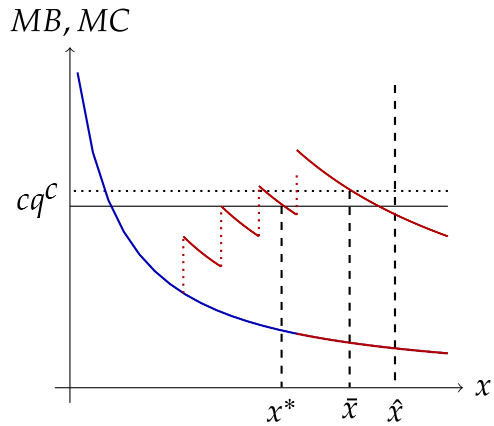


Figure 1.5: An Equilibrium of Content Duplication under Community Preference

Figure 1.5 and figure 1.6 extend 1.3 to graphically illustrate the equilibrium under the community preference. Similar as before, the solid blue curve represents $\hat{MB}(x)$, and the solid red curve represents $MB^*(x)$. The red curve consists of segments of a family of curves $(1 + \theta_i \mu(x^*))\varphi'(x)$ with curve of $\mu(x^*) + 1$ on top of curve of $\mu(x^*)$, and it jumps at each x where $\mu(x^*)$ is increased to $\mu(x^*) + 1$. Figure 1.5 demonstrates an equilibrium of content duplication (i.e., $x^* \leqslant \hat{x}$) if community preference is weak (i.e., $\bar{x}(\theta_n, n, c) \leqslant \hat{x}$). Suppose the strength of community preference is such that given all the users join the open-access platform, the ideal amount of the highest-type user does not go beyond the proprietary output i.e., $\bar{x}(\theta_n, n, c) \leqslant \hat{x}$, then it must be that no user would contribute any original content because reproduction is always feasible and at a lower cost. Thus, the equilibrium amount of content x^* can be identified by the intersection of the $MB^*(x)$ curve and the replication cost line. In this case, reproduction crowds out original contributions, low-type users (i.e., $\theta_i \leqslant \theta_{n-\mu}$) choose the proprietary platform, and high-type users choose the open-access platform (i.e., $\theta_i \geqslant \theta_{n-\mu+1}$).

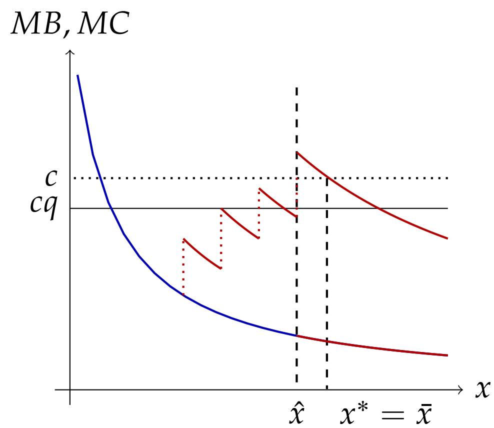


Figure 1.6: An Equilibrium of Content Expansion under Community Preference

However, if community preference is strong enough (i.e., $\bar{x}(\theta_n, n, c) > \hat{x}$), the equilibrium changes dramatically to an equilibrium of content expansion. Figure 1.6 demonstrates this equilibrium. In this equilibrium, the reason for reproduction changes: reproduction is a cost-saving way to attract users toward open access, which motivates a higher amount of original contributions beyond $\hat{x}$. Reproduction induces the marginal type to choose open access, and expands the size of the community at low cost. As more people join the community, marginal utility of original contributions increases, which leads the highest-type user to contribute beyond the proprietary output and closer to the (modified) Samuelson rule, an amount she otherwise would not find incentive compatible without such reproduction. Thus, the equilibrium amount of content x^* can be found at the intersection of the $MB^*(x)$ curve and the original-contribution cost line. In this case, reproduction crowds in original contributions, and all the users choose the open-access platform (i.e., $\mu = n$).

The above analysis leads to the following simple characterization of the open-access content provision.

Proposition 1.3. *Given $\hat{x}$, user-generated content in the open-access platform is*

$$x^*(\hat{x}) = \begin{cases} \min\{\hat{x}, \bar{x}(\theta_n, \mu, cq)\} & \text{if } \bar{x}(\theta_n, n, c) \leqslant \hat{x}\ , \\ \bar{x}(\theta_n, n, c) & \text{if } \bar{x}(\theta_n, n, c) > \hat{x}. \end{cases} \tag{1.15}$$

The profit-maximization problem of the copyright owner now faces a strategic production constraint: he can earn a positive profit only if $\hat{x} \geqslant \bar{x}(\theta_n, n, c)$. If the constraint is not binding, he will choose $\hat{x}$ to balance the revenue and the cost. If the constraint is binding, he has to make strategic investment in additional proprietary content as a response to the entry of the open-access platform. The owner has no option but to exit the market if $\bar{x}(\theta_n, n, c)$ is high enough such that the content expansion leads to zero profit of the platform. Formally, a critical threshold $\hat{\theta}$ is defined by the zero-profit condition of the owner $\Pi_0(\bar{x}(\hat{\theta}, n, c)) = 0$. The equilibrium proprietary content is then

$$\hat{x} = \begin{cases} \max\{\bar{x}(\theta_n, n, c), \phi(\frac{c}{n-\mu})\} & \text{if } \theta_n \leqslant \hat{\theta}\ , \\ 0 & \text{if } \theta_n > \hat{\theta}\ . \end{cases}$$

The existence of a content expansion equilibrium hinges on (i) the community preference of the highest-type user, (ii) the self-confirming belief of the users, as shown by the following proposition.

Proposition 1.4. *Under community preference, $x^* > \hat{x}$ if and only if $\theta_n > \hat{\theta}$ and $x_i = x^*$ for some i.*

Proposition 1.4 states that the market equilibrium is one of content expansion if and only if there is at least one user of the open-access platform, and the the community preference of the highest-type user lies above the threshold $\hat{\theta}$. For any distribution of

$(\theta_1, \ldots, \theta_n)$, $x_i = \hat{x}$ for all i is a self-fulfilling equilibrium. If no users expect others to use the open access, they will be poorly motivated and only contribute reproduction to the open-access platform. With little content, the open-access platform cannot attract any user and end up with zero user participation. If instead $x_i = x^*$ for some i, as we have seen above, the market equilibrium will be one of content duplication if $\theta_n \leqslant \hat{\theta}$ and one of content expansion if $\theta_n > \hat{\theta}$. As we shall see next, enforcing the takedown policy has different welfare implications depending on whether the market is in a content duplication equilibrium or a content expansion equilibrium.

1.4.4 Community Preference: Takedown Policy

There is a debate on whether we should include social preference into the social welfare function. The clear argument for inclusion is that this is the preference that determines behavior and therefore they should be respected by a social welfare evaluation. Argument for exclusion makes a distinction between utility terms that individuals experience and utility terms that motivate their decisions [21]. Without the community preference component, social planner's solution is characterized by condition 1.4 and 1.5. Including the community preference component, social planner's solution still requires condition 1.5 such that the efficient level of reproduction remains zero. However, providing content through the proprietary platform is no longer efficient because it fails to generate the social utility of a community. The open-access platform has a salient role because of community preference, and the efficient level of content follows the modified Samuelson rule internalizing the other-regarding parameter across users (i.e., $\sum_i z_i^o = \phi(\frac{c}{n+(n-1)\sum_i \theta_i})$). For our sake, we do not need to make a value judgement here, as our welfare conclusion on takedown policy is robust under both assumptions on social welfare functions.

[21] See Diamond (2006) for a discussion

Once we impose the takedown policy constraint (i.e., $z_i^r = 0$ for all i), user i's optimization problem becomes

$$\begin{aligned} \max_{z_i^r, z_i^o} \quad & \max_{x_i}\{\theta_i \mu(x^*)\varphi(x^*), \theta_{n-\mu(x^*)}\mu(x^*)\varphi(x^*)\} + \varphi(x^*) - cz_i^o \\ \text{s.t.} \quad & x^* = z_i^o + \textstyle\sum_{j \neq i} z_j^o \\ & z_i^o \geqslant 0 \end{aligned}$$

The takedown policy can lead to the same equilibrium as that under standard preference where $\hat{x} = \phi(\frac{c}{n}), x^* = \phi(c), x_i = \hat{x}$. This equilibrium remains self-fulfilling: the little content provision on the open-access platform, resulting from zero user participation, confirms the belief that there is no open-access users in equilibrium. However, it can also lead to equilibrium with positive number of open-access users. If $x_i = x^*$ for some i, we can derive a similar threshold $\overline{\theta}$ above which the open-access platform outperforms the proprietary platform under the takedown policy, and as we show in Appendix B, $\overline{\theta} > \hat{\theta}$. Figure 1.7 depicts the effect of content takedown in relation to the strength of the community preference θ_n by comparing the market equilibrium before the takedown and the policy-induced equilibrium after the takedown. If $\theta_n \in [0, \hat{\theta}]$, the proprietary platform outperforms the open-access platform for both equilibria (i.e., $x^* \leqslant \hat{x}$), and original contributions would *increase* in response to the removal of reproduction because the market is in a content duplication equilibrium. If $\theta_n \in (\hat{\theta}, \overline{\theta}]$, the open-access platform outperforms the proprietary platform only before the takedown, and original contributions would *decrease* in response to the removal of reproduction because the market is in a content expansion equilibrium. If $\theta_n \in (\overline{\theta}, 1]$, the open-access platform outperforms the proprietary platform for both equilibria (i.e., $x^* > \hat{x}$), and original contributions would not change in response to the removal of reproduction because θ_n is so high that the optimal plan of an open-access user is unaffected by the cost-saving technology of reproduction.

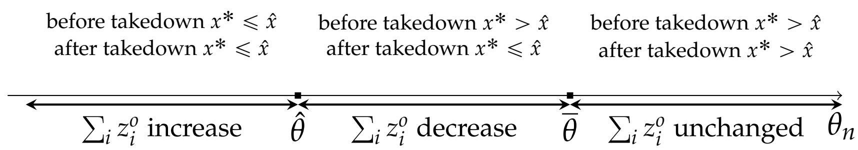


Figure 1.7: Community Preference and the Effect of Takedown

The welfare consequence of the takedown policy is determined by the prevailing equilibrium in absence of the policy. If the market is in a content expansion equilibrium, content takedown reduces welfare, as shown by the following proposition.

Proposition 1.5. *Under community preference, the takedown policy decreases social welfare if and only if $x^* > \hat{x}$ in the market equilibrium.*

Since reproduction crowds in original contributions in a content expansion equilibrium, the takedown policy discourages users' participation in open access platform, and thereby takes away the incentive for the high-type users to make original contributions. The removal of reproduction still secures higher profits for the copyright owner to provide content, but in the meantime it suppresses original contributions by forcing users toward proprietary platform. The second effect outweighs the first effect, and enforcing the takedown policy strictly decreases welfare by leading to a *contraction* of content provision.

We conclude this section by summarizing our main theoretical results. Under standard preference, the takedown policy increases welfare by cutting wasteful investment and restoring the owner's incentive to create. Importantly, reproduction crowds out original contribution in equilibrium, and original contribution would increase once reproduction is removed. Under community preference, there remains the self-fulfilling equilibrium of zero open-access users. But when community preference is strong, there is another subgame perfect Nash equilibrium with substantial original contributions and all users' participation in open access. In this equilibrium, reproduction crowds in original contribution, and original contribution would decrease once reproduction is removed. If this equilibrium is the one that would have prevailed, the takedown policy decreases welfare by forcing users to use the proprietary platform and in effect deterring original contributions. The theory gives us competing hypotheses on the welfare effect of the content takedown. It guides us to examine how the content takedown affects original contributions to the platform, which forms the basis of the following empirical section.

1.5 Data Description

In this section, we apply our analytical framework to study program developers' reproduction and contribution behavior in GitHub, the largest online host of open-source code and repositories in the world. We focus on whether the takedown policy on unauthorized content should be strictly enforced on this platform. The difficulty of this welfare evaluation is that we cannot observe the distribution of the community preference or which equilibrium prevails in absence of the policy. In fact, even if such survey data were available, we would worry about bias in self-reports. Our theory provides guidance for overcoming this difficulty by presenting competing hypotheses for an empirical test. The "crowding out" hypothesis states that if the takedown policy increases welfare, it must be that reproduction is selfish and crowds out original contributions. Therefore, given an exogenous removal of user j's reproduction, we should observe an *increase* in user i's original contribution. The "crowding in" hypothesis states that if the takedown policy decreases welfare, it must be that reproduction is community driven and crowds in original contributions. Therefore, given an exogenous removal of user j's reproduction, we should observe a *decrease* in user i's original contribution.

Founded in 2008, GitHub is an online platform hosting open-source code and repositories. By 2015, GitHub reported over 9 million users and over 20 million repositories, making it the largest host of source code in the world. GitHub is a collaborative code-hosting site built on top of the Git version control system. Users can have three possible roles: administrators/owners, contributors, and testers. An owner initiates and tracks the development of a repository, and contributors can make commits to the repository that improve its content. A commit is a collection of several lines of codes that adds to the functionality of a program. The owner can then review the codes for correctness and complicity with standards and approve the commits by merging them into the codebase of the repository. In addition to code hosting, collaborative code review, and integrated issue tracking, GitHub integrates social features. Users are able to subscribe to information by "watching" projects and "following" users, resulting in a feed of information on those projects and users of interest. Users also have profiles that can be populated with identifying information and contain their recent activity within the site.

The institutional feature that relates two users' contribution is the forking function of GitHub. A fork repository is a copy of its parent repository, and it contains both the content of the parent repository and its own newly added content. Any repository can be forked. Most commonly, forks are used to build upon another user's project for one's own project. Users can make changes to their forks without affecting the parent repository, but changes to the parent repository usually update their part in the forks.

The major treatment variable is the DMCA takedown notice. Such notices arrive in a staggered and unanticipated fashion, allowing us to employ an event study design. The key institutional feature of GitHub's protocol for processing takedown requests is that GitHub automatically disables the content of a parent repository. Two sources of variations in copyright law are explored to explain variations in open-source behavior including the number of repositories shared and coding contributions: the timing of DMCA takedown notice and the longitudinal software piracy rate of countries participating in the WIPO Copyright Treaty. The goal of the empirical section is twofold: (i) test the implications of the theory and support the relevance of the model and (ii) provide evidence suggesting that the current copyright law discourages open-access activities.

Data on GitHub users' behavior is obtained from the GHTorent project (Gousios (2013)). GHTorrent is an offline persistent mirror of GitHub's event streams, including all user events, such as commits, comments, and pull requests. It uses the GitHub API to collect raw data and extract, archive and share queryable metadata. The GHTorent toolset allows retrieving the full history of a single project and the full list of actions for a single developer, which makes it popular in the software data mining community. In particular, we have raw data on which user commits to which project at what time.

We begin with the 2016-01-18 dump. In processing the data, several key decisions are made. Git commits contain information about both the author (the person who originally changed the code) and the committer (the person who last applied the change), each with their own timestamp. The two are not necessarily one and the same person (e.g., they can differ when someone rebases or cherry picks a commit). In this paper, we consider only commits that record the same person as both the author and committer (97.8% of the commits in our dataset) and record the date at which a commit was authored (rather than committed). In addition, Git allows commit metadata, including the authorship date, to be overwritten. For instance, commits with 1969-12-31 or 2050-07-18 timestamps underwent such a history-rewriting process. We excluded these observations.

Using the methodology described above, we assembled a panel dataset of GitHub users in which each observation contains the user's time-invariant characteristics and the

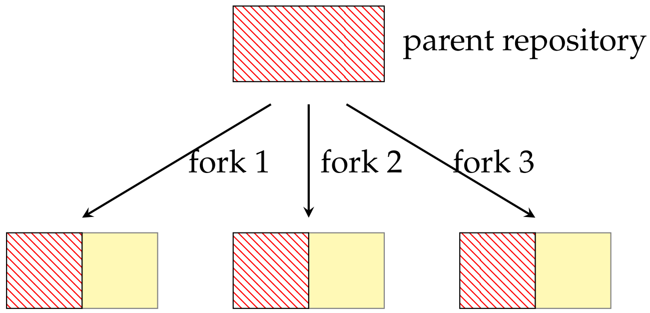


Figure 1.8: Identification Graph

user's GitHub activities in several dimensions for each month upon registration. The main individual characteristics are the location of the user's IP address (from which nationality is inferred), the registration time, the user's preferred programming languages, the user's company affiliation, and whether the user is an individual account or a group account. The activities observed over time are the number of unique commits made by the user, the number of repositories shared by the user, the user's cumulative number of followers, and the total number of watchers for each repository the user shared.

Variable/Year	2008	2009	2010	2011	2012	2013	2014	2015
Number of repository per user	0.10 (0.72)	0.23 (1.03)	0.77 (2.02)	1.65 (3.34)	1.88 (3.62)	2.05 (3.99)	2.23 (4.38)	2.55 (5.72)
Number of forks per repository		11.67 (5.87)	13.52 (10.74)	14.68 (24.73)	13.04 (12.91)	11.86 (7.30)	11.47 (5.58)	11.01 (4.31)
Number of commits per repository	94.82 (433.71)	66.61 (462.16)	44.54 (350.67)	28.54 (195.62)	23.72 (150.28)	16.43 (77.51)	15.84 (103.27)	15.28 (159.43)
Followers per user	0.94 (9.11)	1.36 (11.75)	2.24 (28.12)	2.84 (66.20)	1.93 (50.38)	0.47 (34.10)	0.52 (40.21)	0.16 (20.47)
Watchers per repository	0.63 (30.28)	1.21 (53.81)	3.74 (74.02)	3.26 (45.14)	5.62 (77.17)	8.47 (97.34)	10.91 (120.66)	14.78 (180.62)
Proprietary software piracy rate	60.79 (21.00)	60.83 (20.99)	60.33 (20.96)	59.82 (21.33)		58.89 (21.56)		57.66 (21.74)
Number of countries	104	106	106	106		106		106

Table 1.1: Summary Statistics of the GHTorrent Sample and Country Aggregates

The DMCA takedown notice is a request to remove content of individual users whose shared repository is found to be infringing by claimed copyright holders. We collect data from the GitHub DMCA repository. Every unit of observation is a notice event, and for each notice, we collected the following information: the date of the DMCA notice addressed to the user, the name of the claimed copyright-holder who sent the notice, the URL of the repository in question, the login name of the GitHub user who received the notice, whether the user is an individual or a group, whether the repository has been taken down by GitHub at the time of data collection, whether the GitHub user replies a counter-notice and the date of a counter-notice if there is one, whether the plaintiff retracts the notice and the date when plaintiff retracts the notice. There are $2,213$ complaints filed between 2011 and 2014. $1,091$ of all cases can be matched with a login name in GHTorrent, which results in a total of $1,712$ repositories. The $1,712$ parent repositories are then matched to a total of $16,404$ forks.

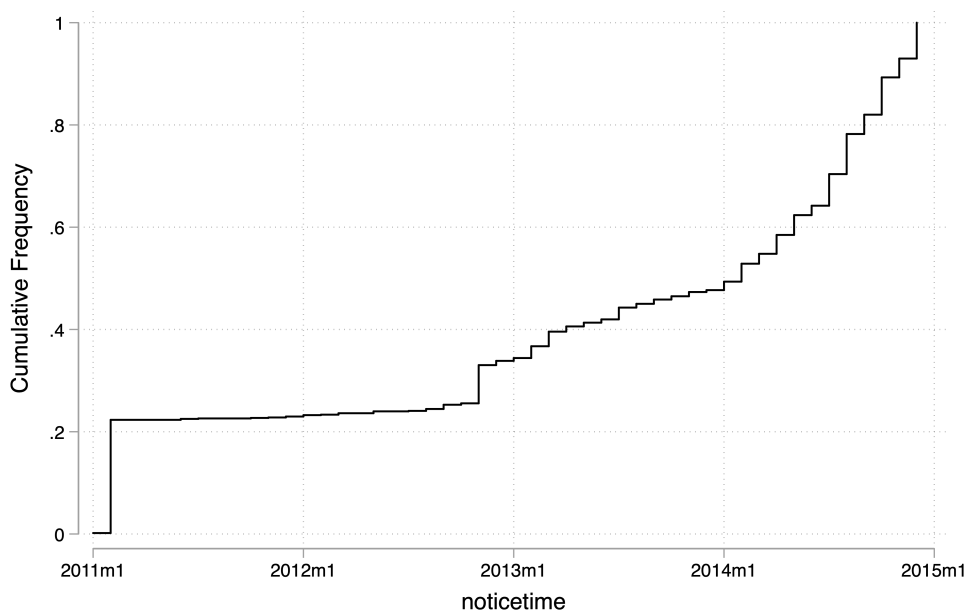


Figure 1.9: Distribution of notice time

This figure shows the distribution of event dates. The earliest notice dates back to Jan. 2011, and the latest notice comes in Dec. 2014, by which month all users have received their notice. After an initial spike of complaints in Feb. 2011 by Sony Computer Entertainment America LLC, the notices are filed at a smooth yet slightly increasing rate.

In addition, we use cross-country software piracy rates as a proxy for the difficulty of unauthorized use. Data on the rates of software piracy were provided by the Business Software Alliance (BSA). BSA is a network of software manufacturers and includes companies such as Microsoft and Novell. Software piracy rates are calculated by estimating the demand for software based on the worldwide number of PC shipments and the sale of U.S. business applications. By assuming that there is a set of software sales accompanying each new PC sold, the difference between expected demand and supply (in the form of sales) is attributed to software piracy. Piracy rates are reported as percentages, with 0% indicating no software is pirated and 100% indicating all software is pirated. Starting from 2011, BSA reports the piracy rates every two years. Despite the limitations of this indicator, BSA piracy rates are one of the most commonly accepted indicators in the industry.

As shown in Table 1.1, average piracy rates across individual countries steadily decline from 61% in 2008 to 57% in 2015. There are sizable regional variations. The highest regional software piracy rates are seen in Eastern Europe and Latin America, with rates of approximately 60%, suggesting six of every 10 software packages in use are pirated copies. North America and Western Europe have the lowest regional software piracy rates, at approximately 30%. Several developing countries reduced their software piracy rates by more than 8% during the sample period, with China, India, Sri Lanka, Vietnam, and Mexico at the top of the list.

1.5.1 Hypotheses and Empirical Strategy

This section describes the main econometric specification to test the competing hypothesis outlined in Section 1.4 and the identification assumption required for consistent estimates.

The main specification is an event study that allows the effect of a takedown notice to vary over subsequent periods. The identification relies on the fact that the timing of a takedown notice is staggered and unanticipated for users, so it is as good as random assignment. Thus, early-treated users and late-treated users can serve as counterfactuals for one another.

The "crowding out" hypothesis of reproduction states the following:

Hypothesis 1.1. *The takedown of a parent repository has a positive effect on open-source contributions to fork repositories.*

The competing "crowding in" hypothesis of reproduction states that

Hypothesis 1.2. *The takedown of a parent repository has a negative effect on open-source contributions to fork repositories.*

We match the data file of DMCA notice events with the relevant GitHub user information. The two files are merged using users' login names, and each GitHub user is then identified by his or her ID assigned by GHTorrent.

Let r be a parent repository, f be a fork repository, t be a calendar quarter where we observe the outcomes, and τ be the time-since-event; e.g., $\tau = 3$ means three months after the notice.

$$commit_{frt} = \sum_{\tau=-10,\tau\neq-1}^{\tau=10} \beta_\tau takedown_{r,t-\tau} + \alpha_f + \eta_t + \varepsilon_{frt}$$

where $commit_{frt}$ is the logarithm of the number of commits contributed to fork f of parent repository r in quarter t. $takedown_{r,t-\tau}$ is an indicator variable for the takedown taking place in period $t - \tau$, so $takedown_{r,t-\tau} = 1$ if as of time t, the parent repository r received a takedown notice τ quarters ago. α_f is the fork fixed effect, and η_t is the quarter fixed effect. We drop any observation beyond the $[-10, 10]$ time window. We follow the convention to normalize the effect of a notice in the reference quarter $\tau = -1$ to be zero. The standard errors are clustered at the parent repository level.

The coefficients β_τ for $\tau \geqslant 0$ capture the persistent effects of a takedown notice. Hypothesis 1.1 expects them to be zero or positive, whereas hypothesis 1.2 expects them to be significantly negative. The terms β_τ for $\tau < 0$ provide a placebo or falsification test. In the absence of anticipating effects, model misspecification, or omitted confounding variables, we expect them not to be significantly different from zero.

We then extend the main specification to other repository-level outcomes, user-level outcomes, and possible heterogeneous effects. In our user-level analysis, i denotes a user, and the regression equation we use is as follows:

$$numrepo_{it} = \sum_{\tau=-10,\tau\neq-1}^{10} \beta_\tau takedown_{i,t-\tau} + \alpha_i + \eta_t + \varepsilon_{it}$$

where $numrepo_{it}$ is the number of repositories shared by user i in quarter t, α_i is the user fixed effect, and η_t is the quarter fixed effect. The standard errors are clustered at the user level.

We also match the data file of country-level longitudinal software piracy rates with the relevant GitHub user whose nationality information is available. The two files are merged using variable ccTLD, a two-digit country code technically known as country code top-level domain.

To test any possible heterogeneous effect of takedown with respect to difficulty of unauthorized use, we use the following regression equation:

$$commit_{frit} = \sum_{\tau=-10,\tau\neq-1}^{10} \beta_\tau takedown_{r,t-\tau} + \sum_{\tau=-10,\tau\neq-1}^{\tau=10} \gamma_\tau q_i takedown_{r,t-\tau} + \alpha_f + \alpha_r + \eta_t + \varepsilon_{frit}$$

where $commit_{frit}$ is the logarithm of the number of commits contributed to fork f of infringing repository r reproduced by user i in quarter t, q_i is user i's difficulty of unauthorized use approximated by the software piracy rate of his or her home country, and α_r is the repository fixed effect. The standard errors are clustered at the user level.

The coefficient γ captures the heterogeneous effects of takedown on open-source contributions with respect to different technological constraints on unauthorized use.

1.6 Empirical Results

1.6.1 The Effect of Takedown on Original Contributions

Figure 1.10 shows that the number of commits contributed to fork repositories drops immediately in the post-notice period, and the decrease lasts for 10 quarters after the notice within the event window studied. Compared to the average pre-notice level, the number of commits on average decreases by 9.8% during the first five quarters after the notice and by more than 10% during the second five quarters after the notice. The figure also shows that the pre-trend is relatively stable and shows no sign of trending downward during the 10 quarters prior to the notice. A similar pattern is observed for the number of contributors to fork repositories (see Figure 1.19).

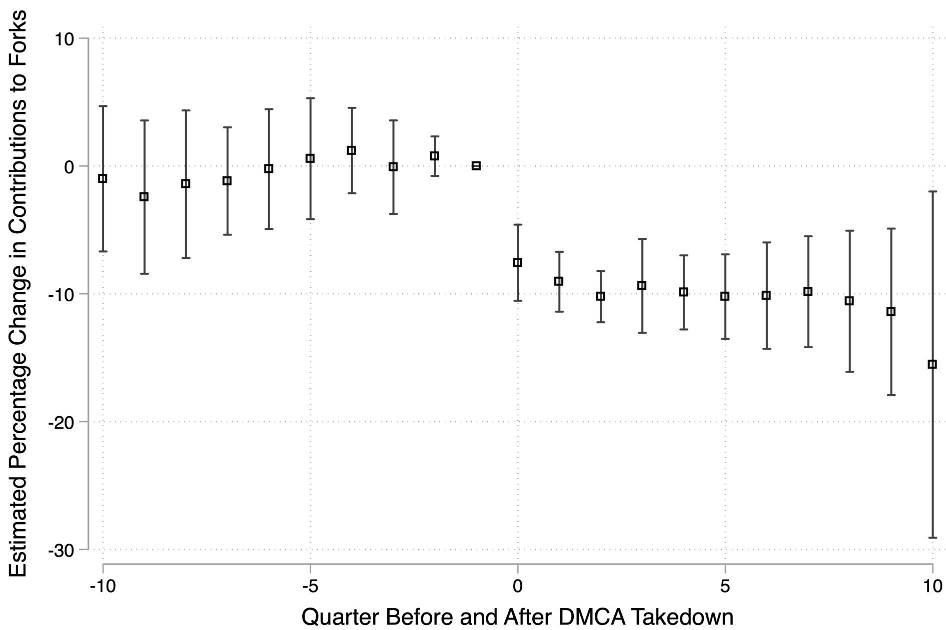


Figure 1.10: The Effect of Takedown Notices on the Number of Commits to Fork Repositories

As a robustness check, we find that the takedown notice has no significant effect on contributions to fork repositories if the notice did not result in a takedown of the parent repository. Figure 1.21(a) shows that if the takedown notice is rebutted with a counter-notice from the user, we do not observe significant changes in the number of commits in the post-notice period within the event window studied. Compared to the average pre-notice level, the change in the number of commits on average is not statistically different from zero during the ten quarters after the notice. Likewise for the countered notice, we do not observe significant changes in the number of authors in the post-notice period within the event window studied (see Figure 1.21(b)).

Figure 1.22(a) shows that if the takedown notice is retracted by the copyright owner, we do not observe significant changes in the number of commits in the post-notice period within the event window studied. Compared to the average pre-notice level, the change in the number of commits on average is not statistically different from zero during the ten quarters after the notice. Likewise for the retracted notice, we do not observe significant changes in the number of authors in the post-notice period within the event window studied (see Figure 1.21(b)).

Further specifications (figure 1.16) reveal that the number of repositories shared by the received GitHub users also drops sharply following a takedown notice, and the decrease lasts for 10 quarters after the notice within the event window studied. Compared to the average pre-notice level, the number of repositories on average decreases by 1.9 repositories during the first five quarters after the notice and by 2.2 repositories during the second five quarters after the notice. The figure also shows that the pre-trend is relatively stable and shows no sign of trending downward during the 10 quarters prior to the notice.

Figure 1.17 shows that if the takedown notice is rebutted with a counter-notice from the user, we do not observe changes in the number of repositories following a takedown notice within the event window studied. Compared to the average pre-notice level, the change in the number of repositories on average is not statistically different from zero during the ten quarters after the notice. Figure 1.18 shows that if the takedown notice is retracted by the copyright owner, we do not observe changes in the number of repositories following a takedown notice within the event window studied.

1.6.2 The Effect of Takedown on the Size of the Community

Figure 1.11 shows that the number of followers of the user also drops following a takedown notice, and the decrease lasts for 10 quarters after the notice within the event window studied. Compared to the average pre-notice level, the number of followers on aver-

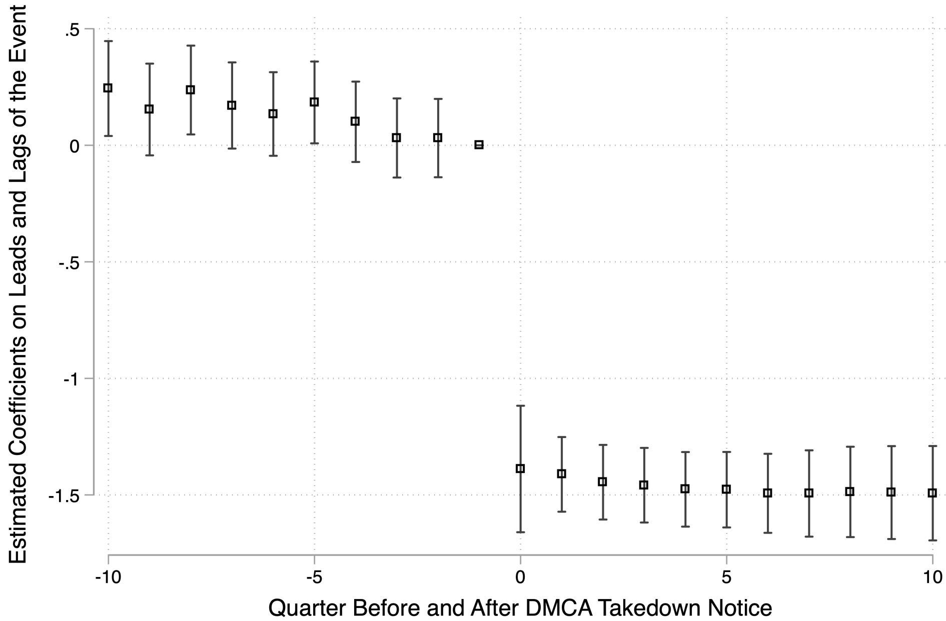


Figure 1.11: The Effect of Takedown Notice on the Number of Followers

age decreases by 1.5 users during the consecutive ten quarters after the notice. The figure also shows that the pre-trend is relatively stable and shows no sign of trending downward during the 10 quarters prior to the notice. A similar pattern is observed for the number of watchers to fork repositories (see Figure 1.20). For the indirectly affected repositories (either forked or shared by the receiving user), the size of the community shifts from $\mu = n$ to $\mu < n$. As a result, the motivation of contribution is greatly undermined under community preference. Since the cost of contribution is higher without the reproduction, contributions are greatly reduced. In addition, as no one is expected to contribute to the repository being shared, the receiving user has no incentive to share further open-access repository in the first place.

Taken together, these findings suggest an explanation consistent with the "crowding in" hypothesis of reproduction: users' contributions decrease in an exogenous removal of reproduction because the community-driven users are poorly motivated by the shrinking size of the open-access community. Therefore, it is most likely that prior to content takedown, the reproduced content is used for content expansion. Going back to our theory, the event study estimates provide evidence that reproduction on this platform is most likely driven by the motive to expand content, and enforcing the takedown policy is thus inefficient. [22]

1.6.3 Checking for Heterogeneous Effect of Takedown

Table 1.2 reports the cross-country findings of copyright enforcement on peer contributions. For each dependent variable we report two specifications, where the sets of covariates are sequentially expanded. Columns (1)-(2) consider average commits, and columns (3)-(4) consider total commits. The regression results suggest a significant positive relationship between country piracy rates and contributions to repositories. An improvement in software piracy by the home country reduces the commits to repositories shared by its citizens. The estimated effect is robust to the inclusion of users' experience in addition to the two-way fixed effects. The estimates in column (4) tell us that holding all other things constant, a one-percent increase in piracy rates is associated with a 0.5% rise in total commits. To put it in context, countries aggressive in copyright enforcement experience a 2% annual decline in piracy rates (China, India, etc.), and users from these countries attract 1 fewer commit every three years purely because of the litigation risk.

[22] Some commentaries call this a "chilling effect" in that the takedowns appear to shock and frighten internet users, and limit all their future open-access sharing. The GitHub team also explicitly points this out in their readme file of the DMCA repository.

1.7 Conclusion

This paper presents the first welfare evaluation of the widely adopted content takedown policy that secures copyright in cyberspace. In general, content takedown appears to be a sensible policy. When users steal content without providing owners fair compensation, the policy can be justified on efficiency grounds even though it has undesirable distributional consequences. This kind of self-interested reproduction is predicted to reduce original contributions. However, in cases where the community preference is strong such that reproduction crowds in original contributions, content takedown strictly decreases social welfare. Our data analysis indicates that reproduction behavior on GitHub is one such case. This suggests a fair-use condition that would impose limitations on copyright in cyberspace. Such legal rule would require a more active role of digital platforms.

Our findings caution against the strict enforcement of content takedown, but it is too early to conclude how we should extrapolate these results to the global efforts on the authorized use of knowledge in cyberspace. After all, many of open-access content in emerging markets appears to be pirated content without expansions or further derivative works (Marron and Steel, 2000). For such efforts to be counterproductive to the creativity of cyberspace, the social motivation to reproduce and create is essential. However, when we apply this framework to content takedown in emerging markets, we should be mindful of the following two key differences. First, the development of open access in emerging markets is in its infancy. Although highly skilled developers may be willing to contribute, they may not be motivated to do so because of the limited number of internet users. In fact, if users are given the opportunity, user-generated content from underdeveloped regions can be quite substantial, as seen by the phenomenal success of Ubuntu, an open-source software primarily developed by African users. Second, social norms on the use of online knowledge have yet to be formulated in emerging markets.

In the US, several non-profit organizations, including Creative Commons, the Free Software Foundation, and the Electronic Frontier Foundation, largely recognize and promote the community preference of internet users. By contrast, in developing countries, the emphasis on community production is lagging.

An important aspect of the policy debate on content takedown is reminiscent of the broader discussion on the relationship between intellectual property rights and knowledge commons. It is generally believed that intellectual property rights are a useful tool to motivate knowledge creation. By enforcing a *temporal* exclusive rights to creations, copyright law provides expected profits to creators as incentives for productive economic activities; as the exclusive rights expires, the creations can *perpetually* augment the knowledge commons, which allows for general public use and feeds into the next generation of creations (Posner, 2005). As a result, economic development benefits. The real relationship between intellectual property and economic development is far more complicated and well beyond the scope of this paper. Nonetheless, we observe that the inefficiency of content takedown this paper presents, where the enforcement of such exclusive rights does not match its intended outcome, is not unique to the GitHub case. In other contexts, researchers have found that the creation of property rights may crowd out otherwise socially desirable behaviors (Bohnet et al. (2001), Gneezy and Rustichini (2000)). As Samuel Bowles puts it succinctly (2008), “the critical assumption in the conventional approach is not that other-regarding motives are absent but that policies that appeal to economic self-interest do not affect the salience of ethical, altruistic, and other social preferences.”

We conclude our paper by discussing limitations and promising extensions. Several intriguing questions regarding content takedown remain unexplored. In cases where takedowns are inefficient, why do they persist? What is the rationale for sending invalid copyright claims? Are other types of takedown requests (for instance, trademark and defamation) efficient? Because of data limitations, we cannot observe the pricing response to content or determine whether the two parties proceed to court if there is a counter-notice. It would be interesting to supplement our theoretical predictions on owners' behavior with such data. In addition, the rich dataset of GitHub allows us to study the social network of open-source users and understand what that local structure implies for copying and copyright policy. Importantly, alternative policies to content takedown that may achieve Pareto efficiency are worthy of further investigation. This paper suggests the fair-use condition in one possibility, but the way platforms implement this rule or whether alternatives such as compulsory licensing may perform better remain open questions. The analysis would be more involved if we further augmented our welfare criterion to include equity concerns.

1.A Appendix A. Extensions and Robustness

Appendix A. Extensions and Robustness

Notice and Takedown Procedure

The main model of the paper focuses on the takedown aspect of the policy. In this section, we augment the baseline model to incorporate the notice and counter-notice procedure of the policy.

The owner files a lawsuit against the creator, claiming that the creator's platform is used for copyright infringement. Settlement negotiations are then conducted against the background of a possible trial. The owner proposes a settlement agreement specifying that the owner will drop the lawsuit if the creator agrees to terminate the open access platform. The creator decides whether to accept the settlement. If the creator accepts, there will be no open access platform, and the owner's proprietary platform will be a monopoly. If the creator rejects, the owner will have to choose whether to proceed to trial or drop the case. If the owner chooses not to litigate, his proprietary platform has to compete with the open access platform. The creator will respond to this proposal only if the owner's threat to litigate is credible: the owner's expected gain in profits must cover his litigation costs.

The outside option of the settlement is a trial. The trial involves costs (legal fees, discovery costs) for both parties, summarized by the monetary amount $L > 0$. If a trial does take place, there is a probability $p \in [0, 1]$ that the owner will prevail and the open access platform will be taken down. In that case, the owner's proprietary platform is a monopoly. If the creator prevails, the court dismisses the claim, and the owner's proprietary platform has to compete with the open access platform. Figure 1.12 describes the sequence of events where a black node is a decision node of the owner, a blue node is a decision node of the creator or users, a red node is a nature's move.

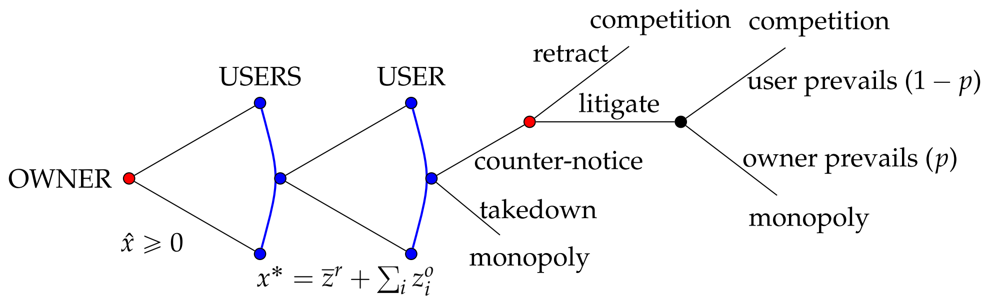


Figure 1.12: The Timing of the Game Under Three-Step Takedown Procedure

Copyright law is announced ex-ante; it is summarized as a pair $(p, q) \in [0,1] \times [0,1]$ where p is the probability that the owner wins an injunction against the open access platform, and q is the extent of anti-circumvention enforcement against unauthorized copying. [23] (p, q) is common knowledge for all players. A copyright law (p, q) is optimal if it maximizes the total payoffs of all users, i.e., $\int_0^{\bar{\theta}} U_\theta(y(\theta), x(\theta))dF(\theta)$, subject to the constraints that the allocation $\{y(\theta), x(\theta)\}_{\theta \in \Theta}$ can be implemented in a perfect Bayesian equilibrium.

When the owner's platform is a monopoly, the owner charges a price $t_m \geqslant 0$, after which every user $i \in \mathcal{N}$ independently and simultaneously chooses to accept or reject, and rejection leads to no use $x_\theta = 0$. Given other users' choice, the payoff of user i's is

[23] As mentioned in Section 1.2, copyright is a bundle of two major rights: (i) the exclusive right to make copies and distribute them, (ii) the exclusive right for further derivative works. We thus think of p as measuring the extent of a country's enforcement over the exclusive right to make copies and distribute them, and q as measuring the extent of a country's enforcement over the exclusive right for further derivative works. The latter usually increases in how broad the provisions on the fair use doctrine are, how stringent the requirements for derivative works are, how friendly the court is to "copyleft", etc.

thus

$$U_i = \begin{cases} y - t_m + \varphi(\hat{x}) & \text{if } i \text{ chooses the proprietary platform,} \\ y & \text{otherwise.} \end{cases} \tag{1.16}$$

Let the choice set of a user be $\mathcal{X}_\theta := \{\hat{x}, x^*, 0\}$, the choice set of the creator be $\mathcal{B}_A := \{settle, not\ settle\}$, the choice set of the owner be $\mathcal{B}_O := \{litigate, not\ litigate\}$. In this game, denote the owner's pure strategy as the tuple $(t_m, t_c, \hat{x}, b_O(x^*))$ where $t_m \in \mathbb{R}_{\geqslant 0}$, $t_c \in \mathbb{R}_{\geqslant 0}$, $\hat{x} \in \mathbb{R}_{\geqslant 0}$, and $b_O \colon \mathbb{R}_{\geqslant 0} \to \mathcal{B}_O$, the creator's pure strategy as $b_A : \Theta \times \mathbb{R}_{\geqslant 0} \times \mathbb{R}_{\geqslant 0} \to \mathcal{B}_A$, the user i's pure strategy as the triple $(x_i(\hat{x}, x^*), z_i^o(\theta_i, \hat{x}), z_i^r(\theta_i, \hat{x}))$ where $x_i \colon \mathcal{H} \times \Theta \to \mathcal{B}$, $z_i^o \colon \mathcal{H} \times \Theta \to \mathbb{R}_{\geqslant 0}$, and $z_i^r \colon \mathcal{H} \times \Theta \to \mathbb{R}_{\geqslant 0}$. To interpret, a pure strategy of the owner is a complete contingent plan that picks the production, the monopoly price, the competition price, and the litigation choice for every realization of the total contributions x^*. A pure strategy of the creator is a contingent plan that picks the settlement choice for every realization of her social type θ_A and the output of both platforms $\hat{x}$ and x^*. A pure strategy of any user i is a contingent plan that picks whether and which platform to use and her pair of original contribution and reproduction for every realization of her social type θ_i and the output of the proprietary platforms $\hat{x}$.

We use (pure strategy) perfect Bayesian equilibrium as the solution concept.

Definition 1.2. *A strategy profile* $\{(t_m, t_c, \hat{x}, b_O(x^*)), b_A(\theta_A, \hat{x}, x^*), (x(\theta, \hat{x}, x^*), z^o(\theta, \hat{x}), z^r(\theta, \hat{x}))_{\theta \in \Theta}\}$ *is a (pure strategy) perfect Bayesian equilibrium such that*

i) For any $\theta \in \Theta$*, given* $\hat{x} \in \mathbb{R}_{\geqslant 0}$*,* $(x(\theta, \hat{x}, x^*), z^o(\theta, \hat{x}), z^r(\theta, \hat{x}))$ *is a best response to* $(x(\theta', \hat{x}, x^*), z^o(\theta', \hat{x}), z^r(\theta', \hat{x}))_{\theta' \neq \theta}$*.*

ii) Given $b_A(\theta_A, \hat{x}, x^*)$*, and* $(x(\theta, \hat{x}, x^*), z^o(\theta, \hat{x}), z^r(\theta, \hat{x}))_{\theta \in \Theta}$*, the owner chooses* $(t_m, t_c, \hat{x}, b_O(x^*))$ *to maximize his payoff.*

iii) Given $(t_m, t_c, \hat{x}, b_O(x^*))$, *and* $(x(\theta, \hat{x}, x^*), z^o(\theta, \hat{x}), z^r(\theta, \hat{x}))_{\theta \in \Theta}$, *the creator chooses* $b_A(\theta_A, \hat{x}, x^*)$ *to maximize her payoff.*

The actual process of enforcing the takedown policy is more elaborate and one where against the background of a possible trial. the aim is to foster an efficient bargaining between the copyright owner and a representative user.

There are three possible outcomes of the notice and takedown process: (i) retraction: the owner retracts his notice, and there is neither takedown nor trial, (ii) takedown: the owner does not retract his notice, and the user complies with takedown, (iii) trial: the owner does not retract his notice but the user chooses to send a counter-notice, so the two parties proceed to the trial.

The owner will retract the notice if his threat of litigation is not credible. The rent from pursuing litigation is the gain in profits ($pn(t_m - t)$) plus the expected court award ($(2p-1)L$). The litigation is credible if and only if

$$pn(t_m - t) \geqslant (1 - 2p)L.$$

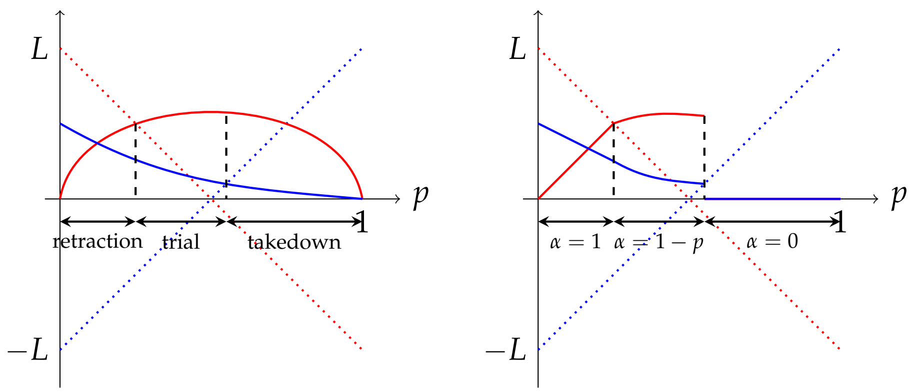
Substituting optimal (t_m, t), we have

$$pn \min\{\varphi(x^*), \varphi(\hat{x})\} \geqslant (1 - 2p)L. \tag{1.17}$$

The user will comply with the takedown if the expected benefit of keeping open access is less than the expected cost of a trial, i.e.,

$$(1-p)\varphi(x^*) \leqslant (2p-1)L.$$

In the graph below, we plot the four functions. the solid red curve is the owner's expected gain in profits, the dotted red curve is the owner's expected cost of trial. the solid blue curve is the user's expected benefit of keeping open access, the dotted blue line is the user's expected cost of trial.



(a) Assume $\alpha = p$ (b) Equilibrium α

Figure 1.13: Copyright Law and the Equilibrium Outcome of the Takedown Policy

The takedown policy makes contributions more risky.

The trade-off behind the takedown policy: if p is large, then we have a monopoly. if p is small, then we have a free-rider problem.

Proposition 1.6. *The takedown policy achieves efficiency if the court imposes no limitation on copyright (i.e., $p = 1$).*

Proof. If $p > p_H$, we have $pn\varphi'(x^*)$ thus the user will comply with the takedown even if

$x^* = z^r(p)$. Note that z^r is decreasing with p. z^r is such that $(1-p)\varphi'(z^r) \leqslant cq$. we need $0 > (1-2p)L$ and $0 \leqslant (2p-1)L$, we need $p > \frac{1}{2}$. we need to prove $p_H \geqslant \frac{1}{2}$

the threat of litigation is credible and , thus $\alpha = 0$. verify the threat and comply monopoly: $\hat{x} = \phi(\frac{c}{n})$, $x^* = 0$. □

The next lemma characterizes the necessary and sufficient condition for a settlement agreement.

Lemma 1.1. *Given x^* and $\hat{x}$, a settlement agreement is reached if and only if (i) the threat of litigation is credible, i.e., $pn\min\{\varphi(x^*), \varphi(\hat{x})\} > L$, (ii) the social type of the creator is bounded above by the user value ratio, i.e., $\theta_A \in [0, \frac{p}{1-p}\min\{\frac{\varphi(\hat{x})}{\varphi(x^*)}, 1\}]$.* [24]

Let α be the equilibrium probability that the open access platform survives (neither taken down by the court nor settled by the creator). Now we can determine the probability α in equilibrium. In the case where litigation is not credible, the open access platform always survives, i.e., $\alpha = 1$. If the expected rent from pursuing a trial exceeds the cost of litigation such that the threat of litigation is credible, then it follows from lemma 1.1 that the open access platform can only survive if the creator has a social type above the threshold and she prevails in the court; therefore, the equilibrium probability is $\alpha(\hat{x}, x^*, p) = (1-p)(1-F(\frac{p}{1-p}\min\{\frac{\varphi(\hat{x})}{\varphi(x^*)}, 1\}))$. We collects these observations as follows:

$$\alpha(\hat{x}, x^*, p) = \begin{cases} 1 & \text{if } pn\min\{\varphi(x^*), \varphi(\hat{x})\} \leqslant L, \\ (1-p)(1-F(\frac{p}{1-p})) & \text{if } pn\min\{\varphi(x^*), \varphi(\hat{x})\} > L \text{ and } \hat{x} \geqslant x^*, \\ (1-p)(1-F(\frac{p\varphi(\hat{x})}{(1-p)\varphi(x^*)})) & \text{if } pn\min\{\varphi(x^*), \varphi(\hat{x})\} > L \text{ and } \hat{x} < x^*. \end{cases} \tag{1.18}$$

[24] We make the tie-breaking assumption that if the owner is indifferent between a trial and no trial, then he will not litigate.

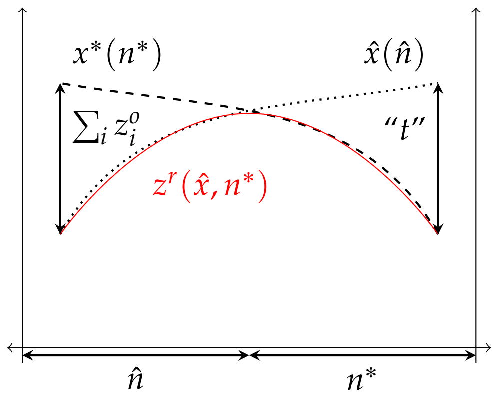


Figure 1.14: Welfare Effect of Moving Users Across Platforms

The output of the proprietary platform satisfies

$$\max_{\hat{x}} nt + I\{pn \min\{\varphi(x^*), \varphi(\hat{x})\} > L\}[(1 - \alpha(\hat{x}, x^*, p))n(t_m - t) - L] - \hat{x},$$

which states that in absence of the litigation strategy, the owner earns profits at the competition price; if litigation is credible and the open access platform does not survive, he can earn rents by charging every user an extra fee $t_m - t$.

In reproduction equilibrium, litigation complements anti-circumvention enforcement against unauthorized use. In original contribution equilibrium, litigation is purely for rent-seeking though under the name of justice.

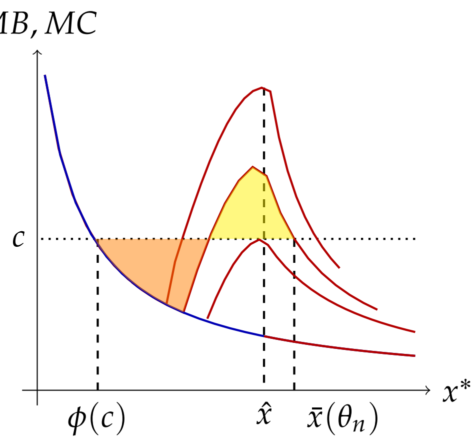


Figure 1.15: Community Preference and Original Contributions

1.B Appendix B. Proofs

Appendix B. Proofs

Lemma 1.2. *If there exists i^* such that $z_i^r > 0$ for $i = i^*$, then $z_i^r = 0$ for any $i \neq i^*$.*

Proof. Suppose there is a $i' \neq i^r$ such that $z_{i'}^r(\theta_{i'}) > 0$. If $z^r(\theta') \leqslant z^r(\theta^r)$, θ' has a profitable deviation towards $z^r = 0$, a contradiction. If $z^r(\theta^r) < z^r(\theta')$, θ^r has a profitable deviation towards $z^r = 0$, a contradiction. □

Lemma 1.3. *For $i = i^*$, $z_i^r = \min\{\hat{x}, z^r(q)\}$ where $\varphi'(z^r(q)) = cq$.*

given $\hat{x}$, a reproduction and consumption scheme $\{z_i^r, x_i\}_{i \in \mathcal{N}}$ is characterized by a triple (i^*, z^r, μ) where

(i) $z_i^r > 0$ for $i = i^*$, and $z_i^r = 0$ for any $i \neq i^*$.

(ii) $x_i = \begin{cases} \hat{x} & \text{if } \theta_i \leqslant \theta_{n-\mu}\ , \\ x^* & \text{if } \theta_i \geqslant \theta_{n-\mu+1}. \end{cases}$

If $\bar{x}(\theta_n, n, c) > \hat{x}$, if and only if the following three condition holds

$$(1+\theta_n n)\varphi(\bar{x}(\theta_n, n, c)) - cqz^r - c[\bar{x}(\theta_n, n, c) - \hat{x}] \geqslant \varphi(\phi(cq)) - cq\phi(cq) \tag{1.19}$$

$$z^r = \hat{x} \tag{1.20}$$

$$\mu = n \tag{1.21}$$

Original contributions satisfiy $z_n^o = \bar{x}(\theta_n, n, c) - \hat{x}$ and $z_i^o = 0$ for any $i < n$.

The copyright owner's profit-maximization problem is

$$\begin{aligned}
\max_{\hat{x}\geqslant 0} \quad & (n-\mu)t - c\hat{x} \\
\text{s.t.} \quad & \hat{x} \geqslant \bar{x}(\theta_n, n, c) \\
\Leftrightarrow \max_{\hat{x}\geqslant 0} \quad & [n-\mu][\varphi(\hat{x}) - (1+(\mu+1)\theta_{n-\mu})\varphi(x^*(\hat{x}))] - c\hat{x} \\
\text{s.t.} \quad & \hat{x} \geqslant \bar{x}(\theta_n, n, c)
\end{aligned}$$

The planner's problem is

$$\begin{aligned}
\max_{\hat{x},(z_i^o,z_i^r,x_i)_{i\in\mathcal{N}}} \quad & \textstyle\sum_{i\in\mathcal{N}/\mathcal{N}*} \varphi(\hat{x}) + \sum_{i\in\mathcal{N}*}(1+\sum_{j\in\mathcal{N}/\{i\}}\theta_j)\varphi(x^*) - c\hat{x} - c\sum_i z_i^o - (cq)\sum_i z_i^r \\
\text{s.t.} \quad & x^* = \max_i z_i^r + \textstyle\sum_i z_i^o \\
& \hat{x} \geqslant 0 \\
& z_i^r \geqslant 0, z_i^o \geqslant 0 \qquad \forall i \\
& z_i^r \leqslant \hat{x} \qquad \forall i
\end{aligned}$$

Proof. Complementary slackness requires

$$\eta^r z^r(\theta) = 0$$
$$\eta^o z^o(\theta) = 0$$
$$\lambda[1 - cz^o(\theta) - cqz^r(\theta)] = 0$$
$$\lambda^r[\hat{x} - z^r(\theta)] = 0$$

□

Proof. For any $\theta \neq \theta^r$, $z^r(\theta) = 0$. The first order condition of original contribution is

$$(1 + \theta n^*)\mathbb{E}\varphi'(z_i^o + \sum_{j \neq i} z_j^o + \bar{z}^r) + \eta^o - c\lambda = c$$

Complementary slackness requires

$$\eta^o z^o(\theta) = 0$$
$$\lambda[1 - cz^o(\theta) - cqz^r(\theta)] = 0$$

There exists a $\theta^o \in [0, \bar{\theta}]$ such that $(1 + \theta n^*)\mathbb{E}\varphi'(x^*) < c$ for $\theta < \theta^o$, $(1 + \theta n^*)\mathbb{E}\varphi'(x^*) > c$ for $\theta > \theta^o$, and $(1 + \theta^o n^*)\mathbb{E}\varphi'(x^*) = c$. It follows that for $\theta < \theta^o$, $\eta^o > 0$, $\lambda = 0$, and $z^o(\theta) = 0$; for $\theta > \theta^o$, $\eta^o = 0$, $\lambda > 0$, and $z^o(\theta) = \frac{1}{c}$.

The first order condition of reproduction is

$$(1 + \theta^r n^*)\mathbb{E}\varphi'(z^r(\theta^r) + \int_{\theta'} z^o(\theta')dF(\theta')) + \eta^r - cq\lambda - \lambda^r = cq$$

The first order condition of original contribution is

$$(1+\theta^r n^*)\mathbb{E}\varphi'(z^o(\theta^r)+\sum_{j\neq i} z^o(\theta')dF(\theta')+\bar{z}^r)+\eta^o-c\lambda=c$$

Complementary slackness requires

$$\begin{aligned}
\eta^r z^r(\theta) &= 0\\
\eta^o z^o(\theta) &= 0\\
\lambda[1-cz^o(\theta)-cqz^r(\theta)] &= 0\\
\lambda^r[\hat{x}-z^r(\theta)] &= 0
\end{aligned}$$

Taking the ratio of the two first-order conditions, we have

$$c(1-q)(1+\lambda)=\eta^o-\eta^r+\lambda^r.$$

By definition, $z^r(\theta^r)>0$, thus $\eta^r=0$. It follows that $\eta^o+\lambda^r>0$.

The feasibility constraints require that $z^r(\theta)\leqslant\min\{\hat{x},1/(cq)\}$. Suppose $z^r(\theta^r)<\min\{\hat{x},1/(cq)\}$. It follows that $\lambda=\lambda^r=0$. If $\theta^r\geqslant\theta^o$. We have $(1+\theta^r n^*)\mathbb{E}\varphi'(x^*)\geqslant c$. It follows that $c\lambda\geqslant\eta^o$ and $\lambda^r+cq(1+\lambda)\geqslant c$. That is, $cq\geqslant c$, a contradiction.

If $\theta^r<\theta^o$. It follows from the assumption that $(1+\theta^r n^*)\mathbb{E}\varphi'(x^*)=cq$. There exists $\theta'\geqslant\theta^o$ such that $z^r(\theta')=\min\{\hat{x},1/(cq)\}>0$. A contradiction with 1.2. Therefore, $z^r(\theta^r)=\min\{\hat{x},1/(cq)\}$.

□

Proof. Since $\bar{x}(\theta, n^*; p, q)$ is increasing in θ, below $\overline{\theta}^r$ the reproduction constraint is not binding, and above $\overline{\theta}^r$ the reproduction constraint is binding. For $\theta < \overline{\theta}^r$, $\bar{x}(\theta, n^*; p, q) < \hat{x}$ so reproduction alone reaches their satiation level, and therefore $z_i^o(\theta_i) = 0$ for $\theta_i < \overline{\theta}^r$. It follows that if $\overline{\theta}^r(\hat{x}) \geqslant \overline{\theta}$, $z_i^o(\theta_i) = 0$ for any $\theta_i \in [0, \overline{\theta}]$. Since $q \leqslant 1$, we have $z_i^o > 0$ for some i if and only if $\int_i z_i^r = \hat{x}$. For $\theta \geqslant \overline{\theta}^r$, it must be that the original contribution constraint is only binding for $\overline{\theta}$, and therefore $z_i^o(\theta_i) = 0$ for $\theta_i < \overline{\theta}$ and $z_i^o(\overline{\theta}) > 0$ for at least some i. □

Proof of Lemma 1.1

Proof. Suppose the threat of litigation is credible, the creator of type θ_A will settle if

$$\begin{aligned} & y + (1 + \theta_A n)[\varphi(\hat{x}) - t_m] \\ \geqslant & y - L + p(1 + \theta_A n)[\varphi(\hat{x}) - t_m] + (1 - p)(1 + \theta_A n) \max\{\varphi(x^*), \varphi(\hat{x}) - t_c\}, \end{aligned}$$

substituting optimal (t_m, t_c), we get

$$L \geqslant (1 - p)(1 + \theta_A n)\varphi(x^*).$$

The owner's gain from blocking open access is his additional rent exceeding the profits in competition, i.e., $n(t_m - t_c)$. His threat of litigation is credible if

$$pn(t_m - t_c) > L,$$

substituting optimal (t_m, t_c), we have

$$pn \min\{\varphi(x^*), \varphi(\hat{x})\} \geqslant L. \tag{1.22}$$

Therefore, a settlement with type $\theta_A \in [0, \overline{\theta}]$ is feasible if

$$pn \min\{\varphi(x^*), \varphi(\hat{x})\} \geqslant (1-p)(1+\theta_A n)\varphi(x^*),$$

or equivalently

$$\theta_A \leqslant \frac{p}{1-p} \min\{\frac{\varphi(\hat{x})}{\varphi(x^*)}, 1\}. \tag{1.23}$$

If $pn(t_m - t_c) \leqslant L$ such that litigation is not credible, the creator of any type θ_A will not respond to the settlement offer. If $\theta_A > \frac{p}{1-p} \min\{\frac{\varphi(\hat{x})}{\varphi(x^*)}, 1\}$, the creator will choose not to settle even if litigation is credible. □

Proof. Suppose $\mathcal{N}^*$ is such that $\varphi(\hat{x}) \leqslant \varphi(x^*)$. For user $i \in \mathcal{N}^*$, he will keep using open access. For user $i \notin \mathcal{N}^*$, he has a profitable deviation towards open access. There does not exist a $t > 0$ such that any user of any type will accept the offer. It follows that $\mathcal{N}^* = \mathcal{N}$.

Suppose $\mathcal{N}^*$ is such that $\varphi(\hat{x}) > \varphi(x^*)$. The owner can set $t = \min\{\varphi(\hat{x}) - \varphi(x^*), 0\}$. For user i who chose to pirate, he is weakly better off by accepting the offer. For user $i \in \mathcal{N}^*$, he has a profitable deviation towards the proprietary bundle. It follows that $\mathcal{N}^* = \varnothing$. □

Proof. Suppose $x^* > \hat{x}$. $x^* = \bar{x}(\overline{\theta}, 1; p, 1)$, and equation 1.10 can be simplified as

$$(1-p)[1+\overline{\theta}]\varphi'(x^*) = 1. \tag{1.24}$$

The optimal output of the owner satisfies

$$p\varphi'(\hat{x}) = 1 \tag{1.25}$$

Given $p < \frac{1+\overline{\theta}}{2+\overline{\theta}}$, we have

$$\begin{aligned}
& p < (1-p)(1+\overline{\theta}) \\
\Leftrightarrow & \frac{1}{(1-p)(1+\overline{\theta})} < \frac{1}{p} \\
\Leftrightarrow & \varphi'(x^*) < \varphi'(\hat{x}) \\
\Leftrightarrow & x^* > \hat{x}.
\end{aligned}$$

Suppose $\hat{x} \geqslant x^*$. $x^* = \bar{x}(\overline{\theta}, 0; p, q)$, and equation 1.11 is

$$(1-p)(1-\frac{p}{(1-p)\overline{\theta}})\varphi'(x^*) = q. \tag{1.26}$$

The optimal output of the owner satisfies

$$\varphi'(\hat{x}) = 1 \tag{1.27}$$

Given $p \geqslant \frac{1-q}{1+\frac{1}{\overline{\theta}}}$, we have

$$\begin{aligned}
& 1-(1+\frac{1}{\overline{\theta}})p \leqslant q \\
\Leftrightarrow & 1 \leqslant \frac{q}{(1-p)(1-\frac{p}{(1-p)\overline{\theta}})} \\
\Leftrightarrow & \varphi'(\hat{x}) \leqslant \varphi'(x^*) \\
\Leftrightarrow & \hat{x} \geqslant x^*.
\end{aligned}$$

□

1.C Appendix C. Figures and Tables

Appendix C. Figures and Tables

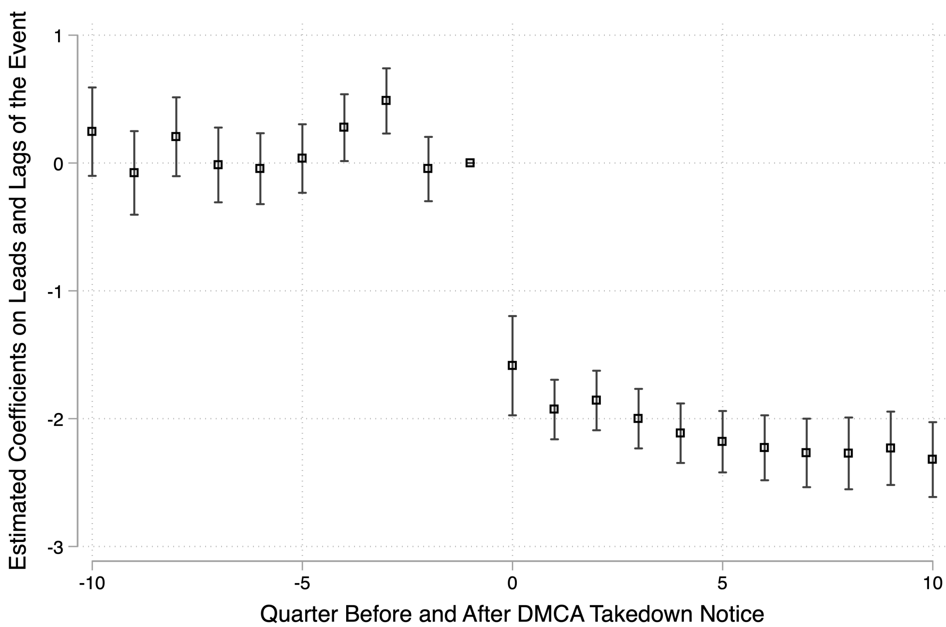


Figure 1.16: The Effect of Takedown Notice on the Number of Shared Repositories

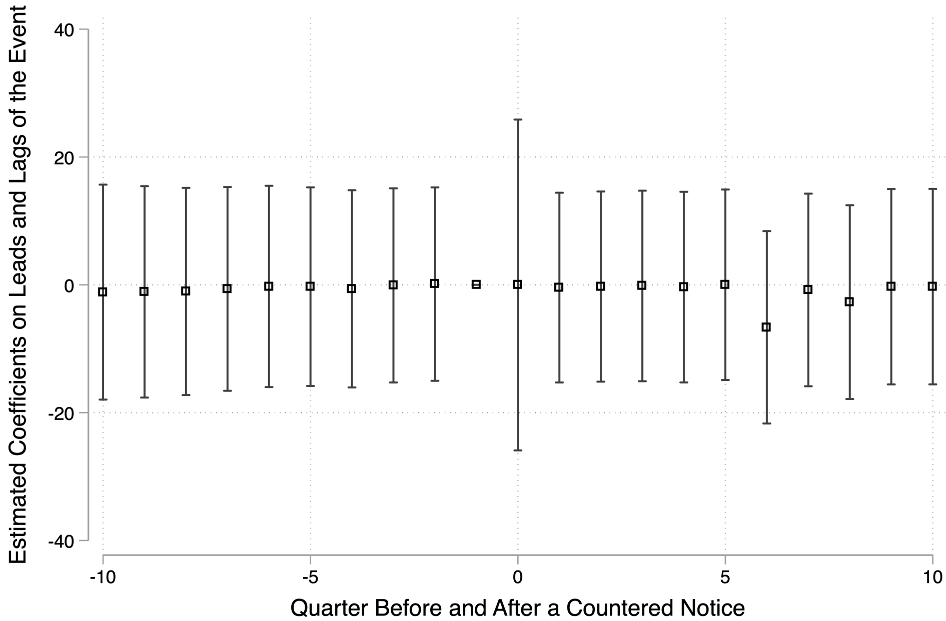


Figure 1.17: The Effect of Countered Notice on the Number of Shared Repositories

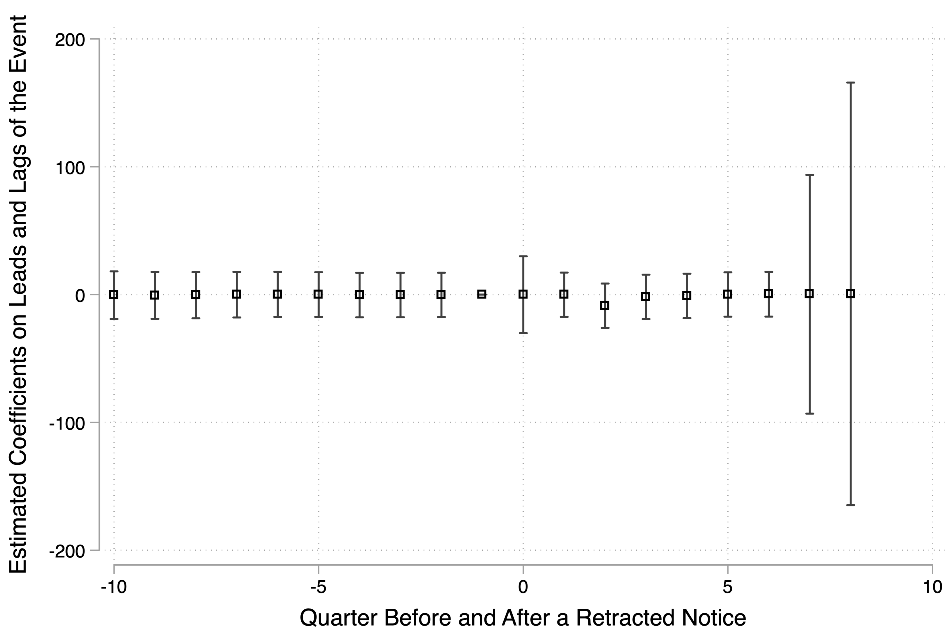


Figure 1.18: The Effect of Retracted Notice on the Number of Shared Repositories

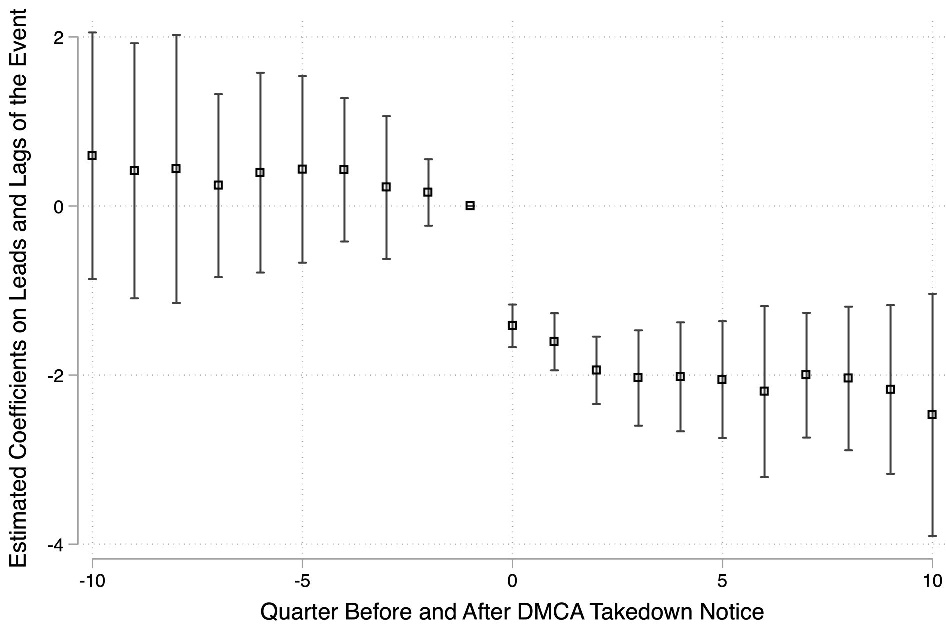


Figure 1.19: The Effect of Takedown Notice on the Number of Contributors to the Fork Repositories

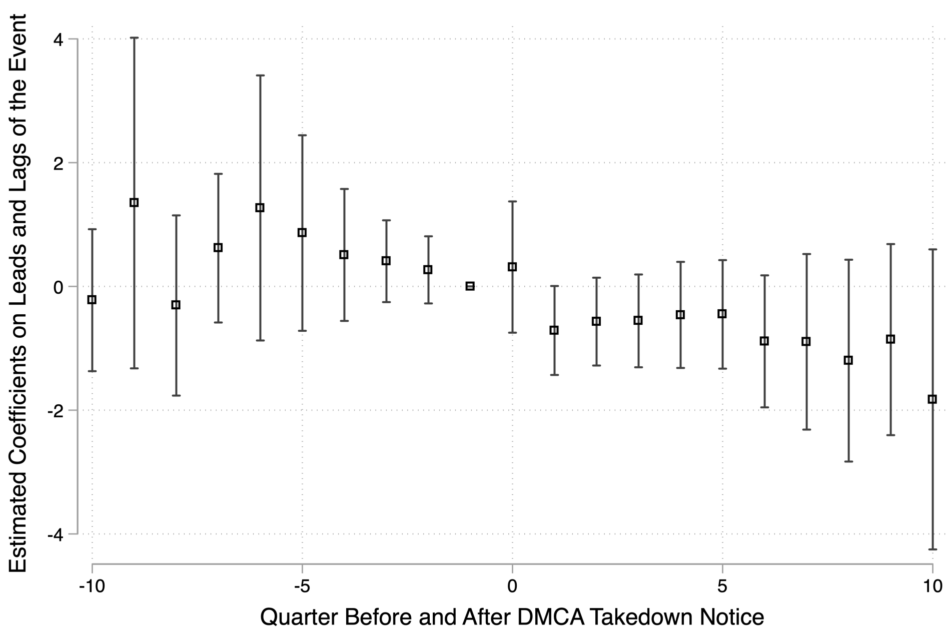
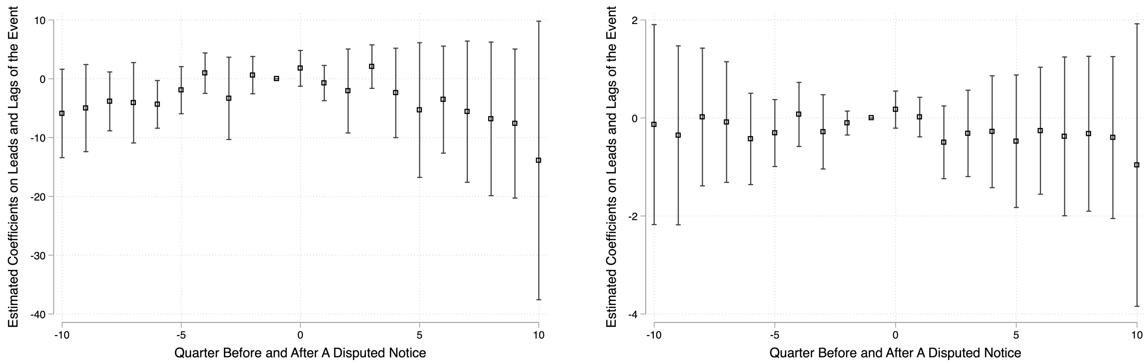


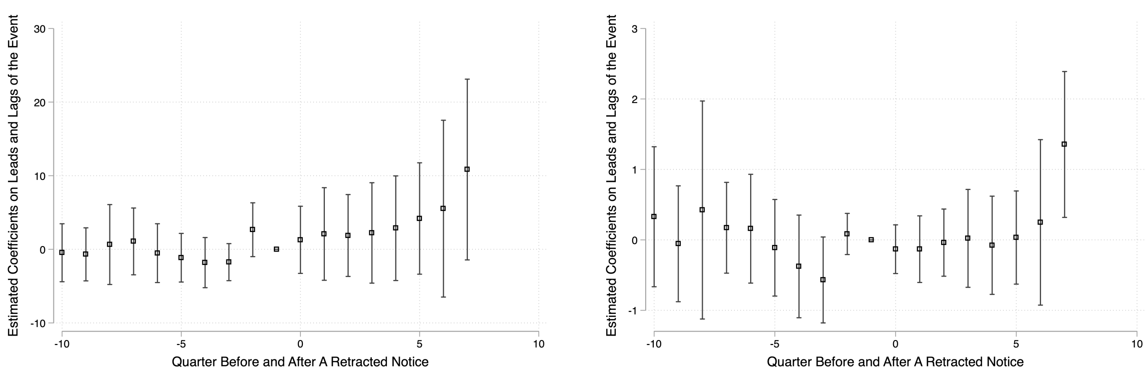
Figure 1.20: The Effect of Takedown Notice on the Number of Watchers to the Fork Repositories



(a) Number of Commits

(b) Number of Contributors

Figure 1.21: The Effect of a Countered Notice on Open Source Contributions



(a) Number of Commits

(b) Number of Contributors

Figure 1.22: The Effect of a Retracted Notice on Open Source Contributions

	(1) Average Commits	(2) Average Commits	(3) Total Commits	(4) Total Commits
$Enforcement$	-0.0054***	-0.0042**	-0.0064***	-0.0057**
	(0.0021)	(0.0022)	(0.0027)	(0.0029)
$Experience$		0.0256***		0.3926***
		(0.0058)		(0.0078)
$Experience^2$		-0.0072***		-0.0258***
		(0.0005)		(0.0007)
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	538,770	538,770	538,770	538,770
adj. R^2	0.027	0.035	0.129	0.154

Note: $commit_{ict}$ is the logarithm of the number of commits contributed to all the repositories shared by user i from country c in year t, and we use both average commits and total commits for robustness. $Enforcement_{ct}$ is the software compliance rate of country c in year t. $Experience$ is the number of years since the users' registration. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.2: Estimation of the effect of copyright enforcement on open source contributions

CHAPTER 2

MEDIATION AND COSTLY EVIDENCE

2.1 Introduction

Information asymmetry is at the heart of many inefficient resource allocations, leading to bargaining impasse, and many times failure of dispute resolution. Mediation, a *self-enforcing* procedure, is increasingly adopted to alleviate such informational problems. Traditionally, mediation is commonly used to resolve disputes in tort liability between the injured and the injury (e.g., accidents, medical malpractice, etc.), in wrongful termination between the employer and the employee, in bankruptcy and debt relief between the creditor and the debtor, and international conflicts where the two parties are sovereignty. With the growth of the digital technology and multi-sided platforms, mediation of online disputes becomes widespread (Rabinovich-Einy and Katsh (2012)). For instance, the e-commerce giant Taobao has handled approximately 2000 disputes of online transactions per day in 2012 (Erickson (2014)). Significant amount of mediated disputes has also been reported on other online platforms, including crowdfunding (Raymond and Stemler (2014)), and on-demand labor platform (DeVault et al. (2019)). The design of an efficient mediation procedure is thus of great importance.

Previous economic models of mediation has exclusively focused on facilitation where the mediator structures a communication process including holding private meetings with each party, understanding their concerns and passing on useful information to the other party. The role of the mediator is to transmit information. However, equally important in the practice of mediation is evidence-based recommendations. In an evaluative mediation, the mediator assists parties in reaching an agreement by checking the facts of each party's claims, pointing out the strength and weakness of the case, and predicting what a likely outcome would be if the two parties fail to settle (Brown and Ayres (1994), Roberts (2007), Nolan-Haley (2012)). An evaluative mediator would make evidence-based recommendations to the parties in relation to their outside option, in the hope that both parties can find a common ground for their agreement. Therefore the mediator also has a role to verify information.

In this paper we study the efficient mediation procedure where disputants are asymmetrically informed and hard evidence can be acquired or presented with a cost. We allow for a broad set of feasible mediation plans: the mediator can choose any combination of facilitative mediation and evaluative mediation. In addition, we compare two modes of evaluation: the adversarial procedure where the disputants present evidence, and the inquisitorial procedure where the mediator acquires evidence. [1] More precisely, our formulation builds on Spier (1994) and Hörner et al. (2015) where we adopt mechanism design and information design techniques to a canonical bilateral bargaining model that determines the division of a pie. We incorporate costly verifiable information, known as evidence, into the design problem.

There are three players: The mediator, the informed party which we call the plaintiff, the uninformed party which we call the defendant. The plaintiff and the defendant find themselves in a bargaining situation where the mediator tries to reach an agreement for the two parties. In the main model, we work with the baseline evidence structure: evidence can fully reveal the type, so once evidence is disclosed or acquired, the plaintiff's private information becomes known to the mediator (in Section 2.5, we relax this assump-

[1]The terminology is borrowed from the legal literature on the trial procedure in common-law countries where partisan advocates present their cases to an impartial jury, and the trial procedure in civil-law countries where judges take a much more active role in investigating the circumstances of the case. We use them more broadly here such that the adversarial model is a synonym for evidence disclosure from the disputants, and the inquisitorial model is a synonym for evidence acquisition of the mediator. In reality, evaluative techniques may lie on a spectrum between these two ideal models.

tion and extend our analysis to more general evidence structures). Under the inquisitorial procedure, the mediator commits whether to pay a cost to acquire evidence upon receiving a report from the plaintiff. Under the adversarial procedure, the mediator commits whether to ask the plaintiff to present evidence together with her report after the plaintiff decides whether to pay a cost to acquire evidence. In both cases, the mediator then converts the reports and the evidence to a recommended allocation and announces it to both parties, according to a pre-committed random mapping. The procedure is self-enforcing such that any party can opt out anytime. Rejection leads to the outside option that can mean going to trial, going on strike, or even a war. Efficiency requires the mediator to maximize the total payoffs of the two parties using transmitted and verified information subject to incentive and feasibility constraints.

Using this model as a vehicle of inquiry, we are able to investigate a rich set of questions on mediation: is the focus on facilitation without loss of generality? When should we use evaluation? What is the settlement rate of an efficient mediation? What is the amount of settlement associated with facilitation and evaluation respectively? Is the adversarial procedure more efficient than the inquisitorial procedure? What are the sources of inefficiency for the profit-maximizing mediation?

Any optimal mediation plan can be characterized by a simple threshold partitioning the ordered state space into truth set (where truth is revealed by the evidence) and pooling set (where no evidence is acquired or presented). Below the threshold, the mediator facilitates, all types are settled, and the settlement is the same across types. Above the threshold, the mediator evaluates, all types are settled, and a range of settlements can be implemented but the lower bound and the upper bound both increases with the type. The efficient mediation plan and the profit-maximizing mediation plan under both the adversarial and the inquisitorial are all such threshold mechanisms. As a result, our first theorem states that Full settlement of disputes is always the outcome irrespective of the procedures or the objectives. Therefore, partial resolution of disputes results from the literature's exclusive focus on facilitation, and full resolution of disputes is implementable once we incorporate evaluation.

To gain some intuition on why the efficient mediation is a threshold mechanism, consider the objective of the mediator. Efficiency requires the mediator to settle as much as possible, while minimizing the cost of evidence. Since facilitation incurs no cost, the mediator would push it to the limit. That limit is the threshold that determines whether the mediator would base his recommendation on evidence. Without evidence, all types beyond the threshold would have to proceed to trial. Full resolution of disputes can be achieved because the mediator's evidence-based recommendation makes the private type common knowledge, and the plaintiff and the defendant realize they have a common interest to settle. And this mediation plan is indeed feasible. To satisfy individual rationality, the mediator gets both parties at least their outside option, such that settlement is always weakly better off. Incentive compatibility has two components, depending on whether the evidence is available. With hard evidence, the mediator can punish lying by refusing to mediate. Without evidence, the mediator has to ensure that a higher type has a weekly higher payoff so that the plaintiff finds truth-telling optimal. The mediator would optimally garble information to keep a grain of salt in the truthful reports, which in turn deters the plaintiff from misrepresentation. In effect, the noise in the communication channel gives the lower type some protection from exposure when he does not imitate the higher type, so it reduces the incentive of lower type to imitate higher type. Under the adversarial procedure, the mediator has to offer a premium in settlement for the evidence such that the plaintiff finds it optimal to acquire the costly evidence whenever needed. Under the inquisitorial procedure, the mediator would rely on the fees to make budget balanced.

Our second theorem concerns the comparison between the welfare maximization and the profit maximization. We identify two sources of inefficiency. The first source of inefficiency is the excessive use of evaluation, given procedure fixed. The second source of inefficiency is the discrepancy in the choice of procedure. For a welfare-maximizing mediator, the efficient adversarial mediation generates higher ex-ante total payoffs than the efficient inquisitorial mediation. The adversarial procedure is superior because it is more flexible in allocating the burden of proof, such that an adversarial mediator can use evaluation more efficiently. The adversarial mediation has a higher threshold so that an adversarial mediator also use facilitation more frequently. However, for a profit-maximizing mediator, the optimal inquisitorial mediation generates higher expected profits than the optimal adversarial mediation. The inquisitorial procedure is preferred because it is more flexible in extracting rents from the parties.

Incorporating costly evidence into the optimal design also increases the predictive power of the model. The derived mediation design matches the practice of professional mediators on mediation style, settlement rate, settlement terms, and satisfaction ratings. For example, facilitation associates with lower and narrower settlement, evaluation associates with higher and wider settlement, and settlement rate is nearly 100%.

We compare our efficient mediation plan with alternative forms of dispute resolution - arbitration and negotiation. We show that arbitration in general yields higher total payoffs by relaxing the self-enforcement constraint of mediation. Unmediated negotiation procedures cannot achieve the same mediated result, at least for simple protocols where disputants meet once and communicate simultaneously. We show that our efficient mediation easily extends to the Grossman-Milgrom evidence structure where evidence partially reveals the truth. As a final extension, we consider the Dye evidence structure where there is some probability that no evidence is available. The full settlement result no longer holds under the Dye evidence structure.

The rest of this paper proceeds as follows. The remainder of this section reviews the related literature. Section 2.2 formally presents the model. Section 2.3 characterizes the efficient adversarial procedure and the efficient inquisitorial procedure, and presents the key theorems on the full settlement result and the superiority of the adversarial procedure. Section 2.4 discusses the empirical implications of the theory. Section 2.5 discusses alternative forms of dispute resolution and extensions to general evidence structure. Section 2.6 concludes. All proofs are delegated to Appendix 2.A.

2.1.1 Related Literature

This paper contributes to the nascent literature on mediation in the Bayesian communication paradigm (Myerson (1991), Bergemann and Morris (2016)). Myerson (1991) pioneered in showing that mediation between parties with misaligned interests can improve the efficiency of the communication even though parties are restricted to unverifiable messages (i.e., Crawford and Sobel (1982)). Myerson's insight is that by adding noise, mediator actually makes the communication more informative. In a similar sender-receiver game, Goltsman et al. (2009) compares the ability of mediation, arbitration, and unmediated negotiation (finite rounds of cheap talk) to maximize the receiver's payoff. The optimal mediation filters the sender's private information and adds noise. Arbitration is (generically) more effective than mediation, while mediation is only sometimes more effective than unmediated negotiation. Bergemann and Morris (2016) classifies mediation as an important form of information design where the information designer has no informational advantages over the players. The main insight of this literature (Myerson (1991), Blume et al. (2007), Goltsman et al. (2009)) is that the mediator can improve welfare by introducing noise in the communication channel such that the incentives for misrepresentation is reduced. In this literature, it is usually assumed that information is unverifiable, while in our model, a pool of objective evidence is available but to access that pool players have to bear the burden of proof. Building upon their framework, we investigate how hard evidence changes the optimal mediation procedure and who should bear the burden of proof.

This paper contributes to the growing literature on alternative dispute resolution. This paper relates to the burgeoning literature on litigation and pretrial bargaining (see Spier (2007) and Daughety and Reinganum (2017) for excellent overviews of the literature). Most models follow earlier bilateral bargaining models by Bebchuk (1984), Reinganum and Wilde (1986), and Spier (1992). Spier (1994) and Klement & Neeman (2005) study dispute resolution from a mechanism design perspective. In both models, information revealed determines incentives to (re-)negotiate. Farmer & Pecorino (2005) study how costly evidence affects pretrial bargaining.

Works on mediation and alternative dispute resolution have been relatively little. In settings similar to bilateral bargaining when the preferences of the two parties are completely conflicting (i.e., split a pie), the mediator recommends a split of the pie based on agents' reports of their types, and if either party opts out, a default division of a reduced pie is implemented. In this setting, Fey and Ramsay (2010) shows that mediation cannot improve on unmediated communication if uncertainty only concerns agents' private costs of fighting. When symmetric agents' share of the (reduced) pie from conflict is determined by privately known strengths, however, Hörner, Morelli, and Squintani (2015) shows that arbitration and mediation are equally effective at minimizing conflict. Both outperform unmediated communication when the intensity of conflict is high, or asymmetric information is substantial. Extending this game, Meirowitz et al. (2019) shows that unmediated peace talks increase the incentive to militarize and so increase eventual conflict, but mediated peace talks reduce militarization and conflict. As mentioned in the introduction, previous theoretical works has exclusively focused on facilitative mediation (Myerson, 1991; Goltsman et al., 2009; Horner et al., 2015; Fanning, 2021). Exciting empirical research on mediation is emerging, highlighting the importance of evaluation (McDermott & Obar, 2004; Klerman & Klerman, 2015). In a very interesting paper, Balzer & Schneider (2020) also studies mediation and evidence, where evidentiary hearing is an outside option to the mediation process. By contrast, this paper highlights the use of evidence *in* the mediation process.

Our adversarial model relates to the literature on mechanism design with hard evidence. This paper contributes to the recent literature on evidence and mechanism design (Hart, Kremer & Perry, 2017; Ben-Porath, Dekel & Lipman, 2020). There are two approach to incorporate evidence: (i) verifiable disclosure of the informed party, and (ii) costly verification of the uninformed party. The early insight of this literature is that skepticism of the receiver can force the sender to voluntarily disclose and unravels to full revelation of truth (Grossman and Hart (1980), Grossman (1981), Milgrom (1981)). A large body of literature followed and can be categorized into two strands: one maintains the GM assumption and extends the conclusion, the other questions its robustness by bringing in other elements. In the first category, Okuno-Fujiwara et al. (1990) extends the unraveling result to more general games. Recently, Hart et al. (2017) shows that the uninformed party's commitment to transfer policy makes no difference to the outcome in a verifiable disclosure setting. Ben-Porath et al. (2019) further generalizes that result to multi-player mechanism design setting. In the second category, Jovanovic (1982) and **?** show that some information will be withheld if there is cost associated with disclosure [2]. Following their

[2]Other variations that support the withholding of information includes seller's lack of information and information acquisition by Dye (1986), Matthews and Postlewaite (1985), Farrel (1986), and Shavell (1994), alternative market structure including one-sided market by Fishman and Hagerty (1995), and oligopoly by Board (2009), and Hotz and Xiao (2013), and alternative category of information by Li and Madarasz (2008),

approach, we pay particular attention to the cost of evidence, and we consider all possible communication protocols by studying the information design problem. Moreover, we compare the disclosure setting to our main setup where it is the receiver who can request evidence. As such, we relax the assumption that off equilibrium message has to be truthful.

Our inquisitorial model relates to the literature on principle-agent model with costly state verification (Townsend (1979), Border and Sobel (1987), Mookherjee and Png (1989), Glazer and Rubinstein (2004), Ben-Porath et al. (2014)). The main differences are three-folds: i) In a model of costly verification, the uninformed party can directly learn the true type at the cost. We instead explicitly specify the process of evaluation on top of the evidence structure, and thus allow for risk consideration of information acquisition. ii) our central concern in this paper is the communication process instead of the transfer policy, iii) verification is committed in this literature, while evaluation of evidence has to be obedient for the receiver in our paper.

2.2 Model

Primitives. There is a plaintiff, a defendant, and a mediator. Let $\Omega := [0, 1]$ be the state space, where ω denotes its typical member. Nature randomly selects a state $\omega \in \Omega$ according to some non-degenerate commonly known distribution $\mu^0(\omega)$, and secretly reveals it to the plaintiff only.

Adversarial Mediation Plan. An adversarial mediation plan (AMP) is $\{\tilde{\pi}(x|m_1, e), t_1, t_2\}$.

and Board (2012).

Upon observing ω, the plaintiff decides whether to pay a cost $C \geqslant 0$ to acquire an evidence $e \in E$. The mediator asks the plaintiff to submit the pair of the message and the evidence (m_1, e) *privately* to the mediator. The mediator then converts the message m_1 and the evidence e to a public message $x \in \mathbb{R}_{\geqslant 0}$ according to a committed random mapping $\tilde{\pi}(x|m_1, e)$, and announces x to both parties. Based on $\{x, t_1, t_2\}$, the two parties decide upon whether to accept the recommended allocation $(x - t_1, -x - t_2)$, and the mediator collects $\{t_1, t_2\}$. Rejection leads to default allocation $(\omega - L_1, -\omega - L_2)$ (e.g. trial, strike, war).

Inquisitorial Mediation Plan. An inquisitorial mediation plan (IMP) is $\{\tilde{\pi}(x|m_1, e), t_1, t_2\}$. The mediator asks the plaintiff to send a *private* message $m_1 \in \mathcal{M}_1$ to the mediator. Based on m_1, the mediator commits whether to pay a cost $C \geqslant 0$ to acquire an evidence $e \in E$. The mediator then converts the message m_1 and the evidence e to a public message $x \in \mathbb{R}_{\geqslant 0}$ according to a committed random mapping $\tilde{\pi}(x|m_1, e)$, and announces x to both parties. Based on $\{x, t_1, t_2\}$, the two parties decide upon whether to accept the recommended allocation $(x - t_1, -x - t_2)$, and the mediator collects $\{t_1, t_2\}$. Rejection leads to default allocation $(\omega - L_1, -\omega - L_2)$.

Payoffs. The plaintiff's payoff under the adversarial mediation plan is

$$u_1 = \begin{cases} x - t_1 - CI\{e \neq \varnothing\} & \text{if an agreement is reached,} \\ \omega - L_1 - CI\{e \neq \varnothing\} & \text{otherwise.} \end{cases}$$

The defendant's payoff under the adversarial mediation plan is

$$u_2 = \begin{cases} -x - t_2 & \text{if an agreement is reached,} \\ -\omega - L_2 & \text{otherwise.} \end{cases}$$

The mediator's profit under the adversarial mediation plan is

$$\pi_0 = \begin{cases} t_1 + t_2 & \text{if an agreement is reached,} \\ 0 & \text{otherwise.} \end{cases}$$

The plaintiff's payoff under the inquisitorial mediation plan is

$$u_1 = \begin{cases} x - t_1 & \text{if an agreement is reached,} \\ \omega - L_1 & \text{otherwise.} \end{cases}$$

The defendant's payoff under the inquisitorial mediation plan is

$$u_2 = \begin{cases} -x - t_2 & \text{if an agreement is reached,} \\ -\omega - L_2 & \text{otherwise.} \end{cases}$$

The mediator's profit under the inquisitorial mediation plan is

$$\pi_0 = \begin{cases} t_1 + t_2 - CI\{e \neq \emptyset\} & \text{if an agreement is reached,} \\ -CI\{e \neq \emptyset\} & \text{otherwise.} \end{cases}$$

Strategy and Solution Concept. Every mediation plan induces a game between the plaintiff and the defendant. In this game, denote the plaintiff's mixed strategy under the adversarial mediation plan as $\tilde{\sigma}_1 : \Omega \to \Delta(\mathcal{M}_1 \times E)$, the plaintiff's mixed strategy under the inquisitorial mediation plan as $\tilde{\sigma}_1 : \Omega \to \Delta(\mathcal{M}_1)$, and the defendant's mixed strategy as $\tilde{\sigma}_2 : \mathcal{M}_2 \to \Delta(\mathbb{R})$. To interpret, a pure strategy of the plaintiff under AMP is to pick a pair of message m_1 and evidence e given the state ω. A pure strategy of the plaintiff under IMP is to pick a message m_1 given the state ω. A pure strategy of the defendant is to pick a real number as the amount of compensation given the public message x. We use perfect Bayesian equilibrium as the solution concept.

Efficiency. A mediation plan is efficient if it maximizes the *ex-ante* total payoffs, i.e., $\mathbb{E}[u_1 + u_2]$ subject to the plaintiff's and defendant's incentive compatibility and obedience constraints, and the mediator's budget constraint.

Profit Maximization. A mediation plan is profit-maximizing if it maximizes the expected profit of the mediator, i.e., $\mathbb{E}[\pi_0]$ subject to the plaintiff's and defendant's incentive compatibility and obedience constraints, and the mediator's budget constraint.

2.3 Analysis: Optimal Mediation Plan

We restrict our attention to direct mechanism w.l.o.g. where the report is a type and the message is an allocation. Given a direct mediation plan $\{\pi_0(x|\cdot), \pi_1(y|\cdot), I(\cdot), t_i\}$, we define some quantities that prove to be useful later. Define

$$p_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} \pi_0(x|\omega)dx, \quad p_\pi(\omega)x_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} x\pi_0(x|\omega)dx$$

where $p_\pi(\omega)$ is the probability of reaching an agreement if no evidence is presented or requested, and $p_\pi(\omega)x_\pi(\omega)$ is the expected settlement if no evidence is presented or requested. Similarly, define

$$q_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} \pi_1(y|\omega)dy, \quad q_\pi(\omega)y_\pi(\omega) = \int_{\underline{\omega}-L_1}^{\overline{\omega}+L_2} y\pi_1(y|\omega)dy$$

where $q_\pi(\omega)$ is the probability of reaching an agreement if an evidence is presented or requested, and $q_\pi(\omega)y_\pi(\omega)$ is the expected settlement if an evidence is presented or requested.

Define $X_1(\omega,\hat{\omega})$ and $X_2(\omega,\hat{\omega})$ as the expected payoff of the plaintiff and the defendant respectively if no evidence is presented or requested. Define $Y_1(\omega,\hat{\omega})$ and $Y_2(\omega,\hat{\omega})$ as the expected payoff of the plaintiff and the defendant respectively if an evidence is presented or requested.

The expected payoffs for the plaintiff are

$$X_1(\omega,\hat{\omega}) = p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) - t_1] + (1 - p_\pi(\hat{\omega}))(\omega - L_1)$$

$$Y_1(\omega,\hat{\omega}) = \begin{cases} q_\pi(\omega)[y_\pi(\omega) - t_1] + (1 - q_\pi(\omega))(\omega - L_1) & \text{if } \hat{\omega} = \omega, \\ \omega - L_1 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

Similarly, the expected payoffs for the defendant are

$$X_2(\omega,\hat{\omega}) = -p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) + t_2] - (1 - p_\pi(\hat{\omega}))(\omega + L_2)$$

$$Y_2(\omega,\hat{\omega}) = \begin{cases} -q_\pi(\omega)[y_\pi(\omega) + t_2] - (1 - q_\pi(\omega))(\omega + L_2) & \text{if } \hat{\omega} = \omega, \\ -\omega + L_2 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

The mediator needs to specify a different payoff when $\hat{\omega} \neq \omega$ as a threat to provide incentives for the plaintiff to tell the truth. Since mediation is self-enforcing, the harshest punishment in the off equilibrium Y_1 is to ask the two parties to proceed to outside options. Under such a threat there is no benefit from falsely reporting a $\hat{\omega}$ that triggers the verification for sure, conditioning on the participation constraints. Consequently, reporting under verification is always truthful, and we an denote the expected payoff under verification simply as $Y_1(\omega, \omega)$ and $Y_2(\omega, \omega)$.

2.3.1 Welfare Maximization: Inquisitorial Model

The efficient mediation plan determines $\{p_\pi(\cdot), q_\pi(\cdot), x_\pi(\cdot), y_\pi(\cdot), I(\cdot), t_i\}$ to maximize the total payoffs:

$$\max_{\substack{p_\pi(\cdot),q_\pi(\cdot),x_\pi(\cdot),y_\pi(\cdot),\\ I(\cdot),t_i}} \int_0^1 \Big(I(\omega)\sum_i X_i(\omega) + (1-I(\omega))\sum_i Y_i(\omega)\Big)\mu^0(\omega)\mathrm{d}\omega$$

$$\text{s.t.} \sum_i t_i \int_0^1 \Big(I(\omega)q_\pi(\omega) + (1-I(\omega))\, p_\pi(\omega)\Big)\mu^0(\omega)\mathrm{d}\omega \geqslant$$

$$C\int_0^1 I(\omega)\mu^0(\omega)\mathrm{d}\omega$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \;\; \forall\omega$$

$$X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_\mu[\omega|x] - L_2, \;\; \forall\omega$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant$$

$$\max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\; \omega - L_1\}, \;\; \forall\omega,\hat{\omega}$$

where the first constraint is *the budget constraint for evidence acquisition*, the second set of constraints is individual rationality for the plaintiff, the third set of constraints is individual rationality for the defendant, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

Note that $x_\pi(\omega), p_\pi(\omega)$ are defined on $\{\hat\omega | I(\hat\omega) = 0\}$, and are free to choose on the complement $\{\hat\omega | I(\hat\omega) = 1\}$. Likewise, $y_\pi(\omega), q_\pi(\omega)$ are defined on $\{\hat\omega | I(\hat\omega) = 1\}$, and are free to choose on $\{\hat\omega | I(\hat\omega) = 0\}$. The incentive compatibility for truth-telling can be broken down into the following four cases:

$$Y_1(\omega, \omega) \geqslant Y_1(\omega, \hat\omega) \quad \text{if } I(\omega) = I(\hat\omega) = 1. \tag{2.1}$$

This is equivalent to $q(\omega)y(\omega) \geqslant q_\pi(\omega)(\omega - L)$, which is implied by IR for plaintiff.

$$X_1(\omega, \omega) \geqslant Y_1(\omega, \hat\omega) \quad \text{if } I(\omega) = 0, I(\hat\omega) = 1. \tag{2.2}$$

This is equivalent to $p(\omega)x(\omega) \geqslant p_\pi(\omega)(\omega - L)$, which is implied by IR for plaintiff.

$$X_1(\omega, \omega) \geqslant X_1(\omega, \hat\omega) \quad \text{if } I(\omega) = I(\hat\omega) = 0. \tag{IC-M}$$

This is similar to a screening problem, and we can simplifies them as Myerson (1981).

$$Y_1(\omega, \omega) \geqslant X_1(\omega, \hat\omega) \quad \text{if } I(\omega) = 1, I(\hat\omega) = 0. \tag{IC-T}$$

This is similar to an auditing problem, and we can simplifies them as Townsend (1979). We start with (IC-M).

Lemma 2.1. *$\{x_\pi(\omega), p_\pi(\omega)\}$ is incentive compatible, if and only if (i) $p_\pi(\omega)$ is non-increasing in ω, and (ii) for any $\omega \in [\underline{\omega}, \overline{\omega}]$, $X_1(\omega) = \int_0^\omega [1 - p_\pi(\tilde\omega)] d\tilde\omega + X_1(0)$.*

Lemma 2.1 is a variant of the celebrated Myerson lemma in the mediation problem that establishes the necessary and sufficient conditions for a direct mechanism to be incentive compatible. The first condition ensures the weak monotonicity of the function p_π such that a higher type is always rewarded with a weakly higher probability of reaching an agreement. The second condition shows that the expected payoff of different types of the plaintiff are pinned down by the function p_π and by the expected payoff of the lowest type of the plaintiff $X_1(0)$. Any two indirect mechanism, which give rise to the same p_π and $X_1(0)$ once the plaintiff optimizes, therefore imply the same expected payoff for all types of the plaintiff. It is easy to see that $Y_1(\omega) \geqslant X_1(\omega)$ plus IC-M imply IC-T; thus given IC-M we only have to deal with $Y_1(\omega) \geqslant X_1(\omega)$. Combining it with IR for plaintiff, we have

$$q_\pi(\omega)[y_\pi(\omega) - t_1] \geqslant \max\{X_1(\omega) - (1 - q_\pi(\omega)(\omega - L_1)), q_\pi(\omega)(\omega - L_1)\}, \tag{2.3}$$

which states that the expected payoff of the plaintiff whose type is verified by the evidence should be weakly larger than both the expected payoff without verification (i.e., $X_1(\omega)$) and the outside option (i.e., $\omega - L_1$). Otherwise, the plaintiff would either misreport his type to avoid verification or to opt out mediation. The next lemma tells us that this constraint is binding.

Lemma 2.2. *For any* $\omega \in [0,1]$, $q_\pi(\omega)[y_\pi(\omega) - t_1] = \max\{X_1(\omega) - (1 - q_\pi(\omega))(\omega - L), q_\pi(\omega)(\omega - L_1)\}$ *if* $q_\pi(\omega) < 1$, $y_\pi(\omega) - t_1 \geqslant \max\{X_1(\omega), \omega - L_1\}$ *if* $q_\pi(\omega) = 1$.

Lemma 2.2 is an extension of lemma 3.1 in Townsend (1979) to the mediation problem, and it follows from the optimality of the mediation plan. Suppose this constraint is non-binding, then it would always be possible to reduce $y_\pi(\omega) - t_1$ without violating any incentive or budget constraint. Doing so would relax the defendant's participation constraint and thus make it feasible to raise the agreement probability $q_\pi(\omega)$. Thus a mediation plan with this constraint being non-binding would involve inefficiently low settlement rate.

We know $X_1(\omega)$ is determined by $p_\pi(\omega)$ and $X_1(0)$. Now consider the following program

$$\begin{aligned} \min_{p_\pi(\omega)} \quad & (L_1 + L_2) \int_0^1 [1 - I(\omega)][1 - p_\pi(\omega)]\mu^0(\omega) d\omega \\ s.t. \quad & p_\pi(\omega) \text{is non-increasing} \end{aligned} \tag{2.4}$$

Lemma 2.3. *$p_\pi(\omega)$ solves the above program if there exists a ω_I^* such that*

$$p_\pi(\omega) = \begin{cases} 1 & \textit{if } \omega \leqslant \omega_I^*, \\ 0 & \textit{if } \omega > \omega_I^*. \end{cases} \tag{2.5}$$

Lemma 2.3 is an application of the extreme point theorem which states that a function p_π that is an extreme point and that maximizes the total payoffs among all extreme points also maximizes the total payoffs among all functions. We may thus simplify the mediation problem further: instead of considering all weakly monotone functions, it is sufficient to restrict our attention to the set of extreme points which requires $p_\pi \in \{0, 1\}$ for all ω. And an extreme point is monotone if and only if it is a step function. Thus p_π is characterized by a single threshold type ω_I^*. The mediator always recommend an agreement for reports below ω_I^*, and always recommend the outside option for reports above ω_I^*. Combining lemma 2.3 with lemma 2.1, we have

$$X_1(\omega) = \begin{cases} X_1(0) & \text{if } \omega \leqslant \omega_I^*, \\ \omega - \omega_I^* + X_1(0) & \text{if } \omega > \omega_I^*. \end{cases} \tag{2.6}$$

By definition, this implies $x_\pi(\omega) - t_1 = x_\pi(0) - t_1 = \omega_I^* - L_1$ for $\omega \in [0, \omega_I^*]$. Therefore,

$$x_\pi(\omega) = \begin{cases} \omega_I^* - L_1 + t_1 & \text{if } \omega \leqslant \omega_I^*, \\ \omega - L_1 + t_1 & \text{if } \omega > \omega_I^*. \end{cases} \tag{2.7}$$

It is straightforward to check that this satisfies IR for plaintiff, i.e., $x_\pi(0) - t_1 \geqslant \omega - L_1$ for $\omega \in [0, \omega_I^*]$ and $\omega - L_1 \geqslant \omega - L_1$ for $\omega \in (\omega_I^*, 1]$. IR for the defendant requires whenever $\omega \in [0, \omega_I^*]$, $x_\pi(0) + t_2 \leqslant \mathbb{E}[\omega|\omega \leqslant \omega_I^*] + L_2$. That is, for a ω_I^* to be feasible, it has to satisfy

$$\omega_I^* - \mathbb{E}[\omega|\omega \leqslant \omega_I^*] \leqslant L_1 + L_2 - (t_1 + t_2). \tag{2.8}$$

Corollary 2.1. $\omega_I^* \in \{\omega' | \omega' - \mathbb{E}[\omega|\omega \leqslant \omega'] \leqslant \sum_i (L_i - t_i)\}$.

Notice that so long as $L_1 + L_2 \geqslant t_1 + t_2$, such a ω_I^* always exist. As either L_1 or L_2 becomes larger, this set possibly grows larger. As either t_1 or t_2 becomes larger, this set shrinks. In general, there could be multiple ω_I^* satisfying this condition. To have a unique ω_I^*, we need $\omega' - \mathbb{E}[\omega|\omega \leqslant \omega']$ to be a monotone increasing function of ω'. Figure 2.1 depicts a unique threshold ω_I^* when $\omega' - \mathbb{E}[\omega|\omega \leqslant \omega']$ to be a monotone increasing function of ω'. As shown by Burdett (1996), the necessary and sufficient condition would be a restriction on $\mu^0(\omega)$ such that $\int_{-\infty}^{\omega} F(x)\mathrm{d}x$ is log-concave. This can be satisfied if the distribution $F(\omega)$ is log-concave (but not necessarily). Several well-known distributions are log-concave, e.g., uniform, normal, and exponential distributions.

Lemma 2.4. *$q_\pi(\omega) = 1$ for any $\omega \in \{\omega | I(\omega) = 1\}$ if $t_1 + t_2 \leqslant L_1 + L_2$.*

The first best $q_\pi(\omega) = 1$ can be obtained because evidence fully reveals ω. As long as the mediation fee is less than the loss from the outside option, both parties prefer to settle in this ex-post complete information environment.

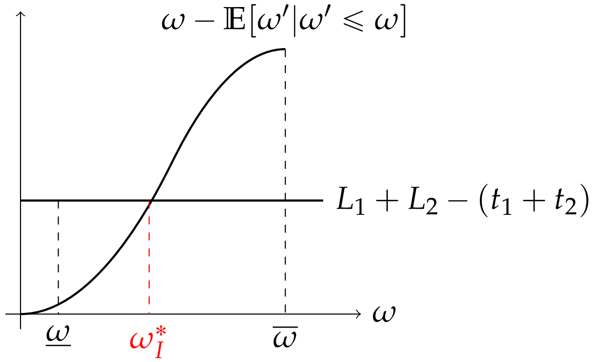


Figure 2.1: A Unique Threshold

Now the only task remaining is to choose which set of types to verify. The goal of the mediation is to help the parties make the most informed decision with the least cost of evidence. As such, the optimality consists of two intuitive requirements: provide more information and optimize the use of evidence. Let λ_T be the shadow price of the budget.

Lemma 2.5. $\{\omega|I(\omega) = 1\}$ *is a connected set, and is nonempty if* $\sum_i L_i \geqslant \lambda_T C + (1 - \lambda_T)\sum_i t_i$.

To understand lemma 2.5, rewrite the condition as $\sum_i L_i - \sum_i t_i \geqslant \lambda_T(C - \sum_i t_i)$. The left-hand side is the increased size of pie due to a higher settlement rate, whereas the right-hand side is the utility value of a deficit in the budget. The mediator will acquire evidence on type ω if the benefit justifies the cost. The benefit is the improvement in the settlement rates $q_\pi(\omega) - p_\pi(\omega)$. The cost is the mediation fees that balance the budget for evidence acquisition. Thus $I(\omega) = 1$ if either the increase in the size of pie is relatively large or the cost of evidence is relatively small. Equivalently, this condition requires the shadow price does not exceed the fraction $\frac{\sum_i L_i - \sum_i t_i}{C - \sum_i t_i}$ such that evidence acquisition is worth the dollar. Notice that in the special case $\sum_i t_i = 0$, the fraction becomes simply the loss-evidence cost ratio, i.e., $(L_1 + L_2)/C$.

A disconnected auditing set would be suboptimal, since then an improvement is available by shifting to a connected set with the same probability mass. Such a shift would make it possible to relax the participation constraint and raise the settlement probability. By lemma 2.5, $\{\omega | I(\omega) = 1\} = [\omega_I^*, 1]$. The budget constraint implies

$$t_1 + t_2 \geqslant C[1 - F^0(\omega_I^*)], \tag{2.9}$$

where $F^0(\omega)$ is the CDF of prior belief. The two other constraints on transfers are

$$\begin{aligned} t_1 + t_2 &\leqslant L_1 + L_2 - (\omega_I^* - \mathbb{E}[\omega | \omega \leqslant \omega_I^*]) && \text{if } \omega \leqslant \omega_I^*, \\ t_1 + t_2 &\leqslant L_1 + L_2 && \text{if } \omega > \omega_I^*. \end{aligned} \tag{2.10}$$

Lemma 2.6. *In any efficient inquisitorial mediation plan,* $t_1 + t_2 = C[1 - F(\omega_I^*)]$.

In an efficient mediation, both the budget constraint and the participation constraint for the defendant are binding. Given a threshold mechanism, the budget constraint becomes $t_1 + t_2 = C[1 - F(\omega_I^*)]$. Note that the distribution of the fees is indeterminate.

Proposition 2.1. *An efficient inquisitorial mediation plan is characterized by a threshold* ω_I^* *such that* $\omega_I^* - \mathbb{E}[\omega | \omega \leqslant \omega_I^*] = L_1 + L_2 - C[1 - F^0(\omega_I^*)]$,
for any $\omega \leqslant \omega_I^*$, $I(\omega) = 0$, $p_\pi(\omega) = 1$, *and* $x_\pi(\omega) = \omega_I^* - L_1 + t_1$,
for any $\omega > \omega_I^*$, $I(\omega) = 1$, $q_\pi(\omega) = 1$, *and* $y_\pi(\omega) \in [\omega - L_1 + t_1, \omega + L_2 - t_2]$,
where $t_1 + t_2 = C[1 - F(\omega_I^*)]$.

Figure 2.2 depicts the efficient inquisitorial mediation plan shown in Proposition 2.1. Individual rationality constraints says that the motivator wants both players to get at least their outside option, such that settlement is always weakly better off. Incentive compatibility has two components, depending on whether the mediator audits. Because the mediator can directly learn the truth, he knew whoever is lying. So he can punish lying by refusing to mediate. He also has to ensure when he does not audit, the plaintiff

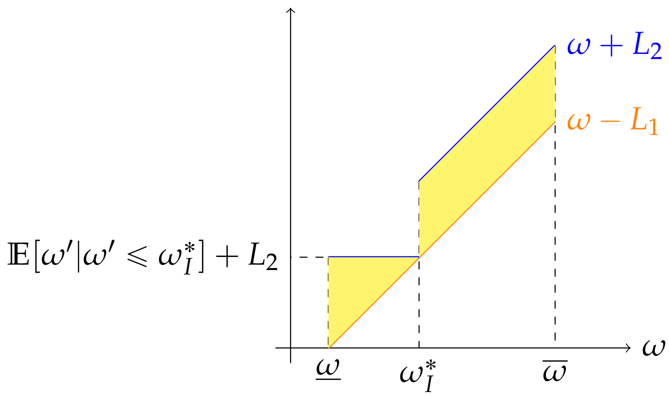


Figure 2.2: Efficient Inquisitorial Mediation Plan

is also telling the truth that requires the p function to be monotone. So that a higher type always have a weekly higher payoff.

Now, let's say we have a candidate pool of all the feasible mediation plans. How should we choose among them? Think about the mediators' objective. The mediator wants the two parties to settle as much as possible. In the meantime, he wants to save auditing cost and facilitative mediation does not cost anything. He will push that to the limit. That's where you find ω_I^* in figure 2.1. Below ω_I^*, all cases are settled, they have exactly the same allocation, and no transfer is ever being paid. But above ω_I^*, without evaluation all cases go to trial. If the cost of auditing is less or equal to the total loss of going to court, then it is worthwhile to audit. Once the mediator audits, omega becomes common knowledge. The plaintiff and the defendant will realize they have a common interest to avoid the loss for both of them. So they are willing to settle, and they are willing to pay for the auditing cost. How do I know it's a threshold but not any other partitions? Weaker cases are always easier to settle. You can tell that from the monotonicity of the p function.

2.3.2 Welfare Maximization: Adversarial Model

The expected payoffs for the plaintiff are

$$X_1(\omega,\hat{\omega}) = p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) - t_1] + (1 - p_\pi(\hat{\omega}))(\omega - L_1)$$

$$Y_1(\omega,\hat{\omega}) = \begin{cases} q_\pi(\omega)[y_\pi(\omega) - t_1] + (1 - q_\pi(\omega))(\omega - L_1) & \text{if } \hat{\omega} = \omega, \\ \omega - L_1 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

Similarly, the expected payoffs for the defendant are

$$X_2(\omega,\hat{\omega}) = -p_\pi(\hat{\omega})[x_\pi(\hat{\omega}) + t_2] - (1 - p_\pi(\hat{\omega}))(\omega + L_2)$$

$$Y_2(\omega,\hat{\omega}) = \begin{cases} -q_\pi(\omega)[y_\pi(\omega) + t_2] - (1 - q_\pi(\omega))(\omega + L_2) & \text{if } \hat{\omega} = \omega, \\ -\omega + L_2 & \text{if } \hat{\omega} \neq \omega. \end{cases}$$

The efficient mediation plan determines $\{p_\pi(\cdot), q_\pi(\cdot), x_\pi(\cdot), y_\pi(\cdot), I(\cdot), t_i\}$ to maximize the total payoffs:

$$\begin{aligned} \max_{\substack{p_\pi(\cdot),q_\pi(\cdot),x_\pi(\cdot),y_\pi(\cdot),\\ I(\cdot),t_i}} & \int_0^1 \Big(I(\omega)\sum_i X_i(\omega) + (1 - I(\omega))\sum_i Y_i(\omega)\Big)\mu^0(\omega)\mathrm{d}\omega \\ \text{s.t. } & Y_1(\omega,\omega) - X_1(\omega,\omega) \geqslant C, \ \ \forall\omega \\ & X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \ \ \forall\omega \\ & X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_\mu[\omega|x] - L_2, \ \ \forall\omega \\ & X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \\ & \qquad \max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\ \omega - L_1\}, \ \ \forall\omega,\hat{\omega} \end{aligned}$$

where the first constraint is *the obedience constraint for evidence disclosure*, the second set of constraints is individual rationality for the plaintiff, the third set of constraints is individual rationality for the defendant, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

We can proceed as before. Lemma 2.1 holds. The obedience constraint in effect strength (IC-T). Because of that, we need a revised version of lemma 2.2 for the adversarial mediation plan.

Lemma 2.7. *For any $\omega \in [0,1]$, $q_\pi(\omega)[y_\pi(\omega) - t_1] = \max\{X_1(\omega) - (1 - q_\pi(\omega))(\omega - L_1) + C, q_\pi(\omega)(\omega - L_1)\}$ if $q_\pi(\omega) < 1$, $y_\pi(\omega) - t_1 \geqslant \max\{X_1(\omega) + C, \omega - L_1\}$.*

Lemma 2.3 holds such that there exists a ω_A^* where $p_\pi(\omega) = 1$ for $\omega \in [0, \omega_A^*]$ and $p_\pi(\omega) = 0$ for $\omega \in (\omega_A^*, 1]$. It follows that for $\omega \in [0, \omega_A^*]$, $x_\pi(\omega) - t_1 = \omega_A^* - L$. Combined with IR for the defendant, we have

$$\omega_A^* - \mathbb{E}[\omega | \omega \leqslant \omega_A^*] \leqslant L_1 + L_2 - (t_1 + t_2). \tag{2.11}$$

The revised version of Lemma 2.4 in an adversarial mediation plan has to take the burden of proof for the plaintiff into account when comparing with her outside option, and it is as follows:

Lemma 2.8. *$q_\pi(\omega) = 1$ for any $\omega \in \{\omega | I(\omega) = 1\}$ if $C + t_1 + t_2 \leqslant L_1 + L_2$.*

Proposition 2.2. *An efficient inquisitorial mediation plan is characterized by a threshold ω_A^* such that $\omega_A^* - \mathbb{E}[\omega | \omega \leqslant \omega_A^*] = L_1 + L_2$,*
for any $\omega \leqslant \omega_A^*$, $I(\omega) = 0$, $p_\pi(\omega) = 1$, and $x_\pi(\omega) = \omega_A^* - L_1$,
for any $\omega > \omega_A^*$, $I(\omega) = 1$, $q_\pi(\omega) = 1$, and $y_\pi(\omega) \in [\omega - L_1 + C, \omega + L_2]$,
where $t_1 = t_2 = 0$.

Theorem 2.1. *The* ex-ante *total payoffs under the efficient adversarial mediation plan is* higher *than that under the efficient inquisitorial mediation plan.*

The superiority of the adversarial procedure comes from its flexibility in allocating the burden of proof. Under the inquisitorial procedure, the cost of evidence is evenly shared by all types. By contrast, the adversarial model allows the parties to pay the cost only when they need the evidence to settle. The adversarial model use evaluation more efficiently. In fact, if the mediator can charge conditional fee in the inquisitorial model, same welfare can be achieved. Theorem 2.1 is silent about any inefficient mediation plans. It could be an inefficient adversarial plan generates lower payoffs than an inquisitorial plan.

2.3.3 Profit Maximization: Inquisitorial Model

The profit-maximizing mediation plan determines $\{p_\pi(\cdot), q_\pi(\cdot), x_\pi(\cdot), y_\pi(\cdot), I(\cdot), t_i\}$ to maximize the expected profits of the mediator:

$$\max_{\substack{p_\pi(\cdot),q_\pi(\cdot),x_\pi(\cdot),y_\pi(\cdot),\\ I(\cdot),t_i}} [t_1 + t_2] \int_0^1 \Big(I(\omega) q_\pi(\omega) + (1 - I(\omega))\, p_\pi(\omega) \Big) \mu^0(\omega) d\omega$$

$$\text{s.t.} \sum_i t_i \int_0^1 \Big(I(\omega) q_\pi(\omega) + (1 - I(\omega))\, p_\pi(\omega) \Big) \mu^0(\omega) \mathrm{d}\omega \geqslant$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \ \ \forall \omega$$

$$X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_{\mu^1}[\omega] - L_2, \ \ \forall \omega$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant$$

$$\max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\ \omega - L_1\}, \ \ \forall \omega, \hat{\omega}$$

where the first constraint is *the budget constraint for evidence acquisition*, the second set of constraints is individual rationality for the plaintiff, the third constraint is individual rationality for the defendant, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

Proposition 2.3. *A profit-maximizing inquisitorial mediation plan is characterized by a threshold $\hat{\omega}_I$ such that $\hat{\omega}_I = \operatorname{argmin}_{\omega} \omega - \mathbb{E}[\omega'|\omega' \leqslant \omega] + C[1 - F(\omega)]$,*
for any $\omega \leqslant \hat{\omega}_I$, $I(\omega) = 0$, $p_\pi(\omega) = 1$, and $x_\pi(\omega) = \hat{\omega}_I - L_1 + t_1$,
for any $\omega > \hat{\omega}_I$, $I(\omega) = 1$, $q_\pi(\omega) = 1$, and $y_\pi(\omega) \in [\omega - L_1 + t_1, \omega + L_2 - t_2]$,
where $t_1 + t_2 = L_1 + L_2 - (\hat{\omega}_I - \mathbb{E}[\omega|\omega \leqslant \hat{\omega}_I])$.

Profit maximization requires him to use evaluation for every single type ω. The inefficiency lies in the excessive use of evaluation.

2.3.4 Profit Maximization: Adversarial Model

The profit-maximizing mediation plan determines $\{p_\pi(\cdot), q_\pi(\cdot), x_\pi(\cdot), y_\pi(\cdot), I(\cdot), t_i\}$ to maximize the expected profits of the mediator:

$$\max_{\substack{p_\pi(\cdot), q_\pi(\cdot), x_\pi(\cdot), y_\pi(\cdot), \\ I(\cdot), t_i}} [t_1 + t_2] \int_0^1 \Big(I(\omega) q_\pi(\omega) + (1 - I(\omega))\, p_\pi(\omega) \Big) \mu^0(\omega) d\omega$$

$$\text{s.t. } Y_1(\omega,\omega) - X_1(\omega,\omega) \geqslant C, \ \ \forall \omega$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1, \ \ \forall \omega$$

$$X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_\mu [\omega|x] - L_2, \ \ \forall \omega$$

$$X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant$$
$$\max\{X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})],\ \omega - L_1\}, \ \ \forall \omega, \hat{\omega}$$

where the first constraint is *the obedience constraint for evidence disclosure*, the second set of constraints is individual rationality for the plaintiff, the third set of constraints is individual rationality for the defendant, the last set of constraints is incentive compatibility for double deviation of truth-telling and opting out.

Proposition 2.4. *A profit-maximizing adversarial mediation plan is characterized by a threshold* $\hat{\omega}_A$ *such that* $\hat{\omega}_A - \mathbb{E}[\omega|\omega \leqslant \hat{\omega}_A] = C$*,*
for any $\omega \leqslant \hat{\omega}_A$*,* $I(\omega) = 0$*,* $p_\pi(\omega) = 1$*, and* $x_\pi(\omega) = \hat{\omega}_A - L_1 + t_1$*,*
for any $\omega > \hat{\omega}_A$*,* $I(\omega) = 1$*,* $q_\pi(\omega) = 1$*, and* $y_\pi(\omega) \in [\omega - L_1 + t_1 + C, \omega + L_2 - t_2]$*,*
where $t_1 + t_2 = L_1 + L_2 - C$*.*

Theorem 2.2. *The* ex-ante *profits under the profit-maximizing inquisitorial mediation plan is* higher *than that under the profit-maximizing adversarial mediation plan.*

The inquisitorial procedure is more flexible in extracting rents from the parties. Under the inquisitorial procedure, the mediator can always choose to acquire evidence, ensure a settlement with the evidence, and collect mediation fee conditional on the settlement. The inquisitorial procedure allows the mediator to extract the maximal rents from the disputants, i.e., $t_1 + t_2 = L_1 + L_2 - C$. By contrast, the mediator has to satisfy the obedience constraint for evidence disclosure under the adversarial procedure.

Comparing Theorems 2.2 and 2.1, we see a discrepancy in the choice of procedures: a profit-maximizing mediator prefers the inquisitorial procedure, whereas a welfare-maximizing mediator prefers the adversarial procedure.

Theorem 2.3. *Full settlement of disputes is the outcome irrespective of the mediation procedures or the mediator's objective.*

In all the optimal mediation plans, outside option is never triggered. This is easily understandable for the efficient mediation because triggering the outside option means a loss to the total payoffs. It is perhaps surprising that this continues to be the case for profit maximization. The fact that mediation fees can be collected only if the mediation is successful partially align the objective of a self-interested mediator with that of a social planner.

2.4 Empirical Implications

Our theory has two types of empirical implications. First, for a given mediation context, it provides predictions on the mediator's and the disputants' behaviors, the settlement rate and the settlement terms. Second, comparing across mediation contexts, it has implications for how the mediation practice should vary with the underlying parameters. The model has two groups of underlying parameters: institutional parameters, which include the distribution of the default division $\mu(\omega)$; technological parameters, which consist of the cost of evidence C and the cost of trial L_1 and L_2.

Mediator's behavior. In the optimal mediation plan, $I(\omega) > 0$ for some ω. Evaluation is popular among professional mediators. McDermott and Obar (2004) studies employment mediation programs under the Equal Employment Opportunity Commission. They find that 48.5% of the mediators used evaluative techniques, and 24.9% of the mediators used hybrid techniques. Even in a mediation program labeled as facilitative, a substantial number of mediators engage in evaluative behavior, mainly reality checking.

A key insight of the model is that a profit-maximizing mediator is biased towards inquisition. Unfortunately there is no direct test of this hypothesis. However, the empirical literature provides clues on how this might be true.

Settlement Rate. In the optimal mediation plan, The combination of evaluation and facilitation yields much higher settlement rate than pure facilitation, i.e., $\int_0^1 \left(I(\omega)q_\pi(\omega) + (1 - I(\omega))\, p_\pi(\omega)\right) dF(\omega) \geqslant \int_0^1 p_\pi(\omega) dF(\omega)$. Empirical studies on mediation in international conflicts find that although facilitation can be effective at increasing the probability of settlement and de-escalation of disputes (Dixon, 1996), evaluation appear more effective in ending the disputes (Fey and Ramsay, 2010). Building a measure of mediator information using the mediator's diplomatic representation, its trading relationships, and its military alliance with the countries in conflict, Savun (2008) finds that mediators with access to independent information from these sources are much more likely to resolve conflicts successfully.

In the optimal mediation plan, the settlement rate is 100%, i.e., $\int_0^1 \left(I(\omega)q_\pi(\omega) + (1 - I(\omega))\, p_\pi(\omega)\right) dF(\omega) = 1$. Empirically mediation with hybrid techniques can lead to very high settlement rate. Klerman and Klerman (2015) finds that a mediator's proposal, when used, leads to very high settlement rate (over 99 percent). Therefore, partial resolution of disputes results from the literature's exclusive focus on facilitation, and full resolution of disputes is the outcome once we incorporate evaluation.

Settlement Terms. In the optimal mediation plan, facilitation leads to a singleton of settlement, i.e., $x_\pi(\omega) - t_1 = \hat{\omega}_I - L_1$, and evaluation has a range of settlement, i.e., $y_\pi(\omega) - t_1 \in [\omega - L_1, \omega + L_2 - (t_1 + t_2)]$. Evaluation leads to higher amount of settlement, i.e., $\min y_\pi(\omega) \geqslant x_\pi(\omega)$, or average dollar benefits $\frac{\int_{\hat{\omega}_I}^1 y_\pi(\omega) dF(\omega)}{1 - F(\hat{\omega}_I)} \geqslant \frac{\int_0^{\hat{\omega}_I} x_\pi(\omega) dF(\omega)}{F(\hat{\omega}_I)}$. McDermott and Obar (2004) finds that facilitation associates with lower and narrower settlement, evaluation associates with higher and wider settlement.

McDermott and Obar (2004) compares evaluation with counsel representation versus without representation. They find that when evaluative mediation is combined with representation, the settlement is far more than if the evaluation was conducted without representation (\$31,276 vs. \$5,988). When parties are represented by counsels, it is most likely that the adversarial procedure would prevail. Without representation, evaluative mediation is mostly close to the inquisitorial model. Our theory indeed predicts that the average settlement $\int_0^1 \left(I(\omega)y_\pi(\omega) + (1 - I(\omega))\, x_\pi(\omega)\right) dF(\omega)$ is lower in the inquisitorial model as it allows the profit-maximizing mediator to extract more rent.

McDermott and Obar (2004) also finds a puzzling result: when asking about the disputants' satisfaction with the mediated outcome, facilitation was on average rated higher than evaluation, even though evaluation results in higher amount of settlement. Our model can explain this if we define satisfaction as the difference between the settlement and the outside option, i.e., $x_\pi(\omega) - t_1 - (\omega - L_1)$ and $y_\pi(\omega) - t_1 - (\omega - L_1)$. It could be that settlement under evaluation ends up in the lower end such that satisfaction under evaluation is zero, i.e., $\min y_\pi(\omega) - t_1 - (\omega - L_1) = 0$. In fact, for a profit-maximizing mediator, even the higher end generates very little satisfaction as the mediator extracts almost the entire rent, i.e., $\max y_\pi(\omega) - t_1 - (\omega - L_1) = \hat{\omega}_I - \mathbb{E}[\omega|\omega \leqslant \hat{\omega}_I]$. Satisfaction under facilitation, however, is positive for all $\omega < \hat{\omega}_I$ and is higher for lower types. Consequently, $\frac{\int_0^{\hat{\omega}_I} [x_\pi(\omega) - \omega] dF(\omega)}{F(\hat{\omega}_I)} \geqslant \frac{\int_{\hat{\omega}_I}^1 [y_\pi(\omega) - \omega] dF(\omega)}{1 - F(\hat{\omega}_I)}$, the average ratings is higher for facilitation than evaluation even though evaluation generates higher dollar benefit on the book.

2.5 Extensions and Robustness

2.5.1 Alternative Forms of Dispute Resolution: Arbitration

The arbitration problem with ex-ante IR is

$$\begin{aligned}
\min_{I(\omega),x(\omega),p(\omega),y(\omega),q(\omega)} \quad & [L_1 + L_2] \int_0^1 [I(\omega)(1-q(\omega)) + (1-I(\omega))(1-p(\omega))] \mu^0(\omega) d\omega \\
s.t. \quad & C \int_0^1 I(\omega)\mu^0(\omega) d\omega \leqslant T \\
& X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \omega - L_1 \quad \forall\omega \\
& X_2(\omega,\omega) + I(\omega)[Y_2(\omega,\omega) - X_2(\omega,\omega)] \geqslant -\mathbb{E}_{\mu^0}[\omega] - L_2 \quad \forall\omega \\
& X_1(\omega,\omega) + I(\omega)[Y_1(\omega,\omega) - X_1(\omega,\omega)] \geqslant \\
& X_1(\omega,\hat{\omega}) + I(\hat{\omega})[Y_1(\omega,\hat{\omega}) - X_1(\omega,\hat{\omega})] \quad \forall\omega,\hat{\omega}
\end{aligned}$$

The first constraint is the budget constraint for evidence acquisition, the second set of constraints is individual rationality for plaintiff, the third constraint is individual rationality for player 2, the last set of constraints is incentive compatibility for truth-telling.

The only difference between arbitration and mediation is that two parties have to agree ex-ante to the arbitration decision. So the self-enforcement constraint of mediation is relaxed, and the ex-ante IR replaces the ex-interim IR in the mediation program. The efficient arbitration plan is again a threshold mechanism where $\omega^* = \min\{\overline{\omega}, L_1 + L_2 - \mathbb{E}_{\mu^0}[\omega]\}$.

2.5.2 Alternative Forms of Dispute Resolution: Negotiation

Shavell (1989) considers a screening game where the informed plaintiff can costless disclose verifiable private information and the uninformed defendant makes the offer. He finds an equilibrium under which there are no trials: plaintiffs with strong cases reveal their type, while plaintiffs with weak cases remain silent and receive a pooling offer that all accept.

Sobel (1989), however, shows that costly voluntary disclosure will not take place if the opposing party makes the final offer. If the plaintiff reveals her type but the defendant makes the final offer, the defendant will gain all the benefits from settlement through this final offer. Thus, there is no benefit to the plaintiff from revealing her type, and with positive cost of disclosure, she strictly prefers to remain silent. As a consequence, no trial equilibrium in Shavell (1989) will take place if and only if disclosure is costless.

2.5.3 Grossman-Milgrom Evidence Structure

In this subsection, we consider more general evidence structure under which our results extend. We start with describing the availability of the evidence. It is common knowledge that the pool of evidence is an exogenous mapping $\mathcal{E}: \Omega \to 2^{2^{\Omega}}$ that satisfies the following two property:

a. Authenticity property: $E \in \mathcal{E}(\omega)$ implies $\omega \in E$.

b. Consistency property: $\omega' \in E \in \mathcal{E}(\omega)$ implies $E \in \mathcal{E}(\omega')$.

Interpretation of $\mathcal{E}(\omega)$ is the set of events that can be proved conclusively by some documents or other forms of tangible evidence when the true state is ω. The authenticity property says that any evidence must contain truth. The consistency property states that if an evidence available for ω does not rule out ω', that evidence is also available when the true state is ω'.

We assume the evidence structure is complete such that $\mathcal{E}(\omega) = \{E \in 2^{\Omega} | \omega \in E\}$. If the true state is ω_2, the pool of evidence is $\{\omega_2, \{\omega_0, \omega_2\}, \{\omega_1, \omega_2\}, \Omega\}$. Likewise for ω_1 and ω_0. Following Bull and Watson (2007), in the adversarial model, plaintiff's disclosure strategy is to choose a pair of message and evidentiary action (m, E). Given the prior $\mu^0(\omega)$, the true state ω, and an evidence E, evidence disclosure updates belief as follows:

$$\mu^1(\omega'|E) = \begin{cases} 0 & \text{if } \omega' \notin E, \\ \frac{\mu(\omega'|m)}{\sum_{\tilde{\omega}\in E} \mu(\tilde{\omega}|m)} & \text{if } \omega' \in E. \end{cases} \tag{2.12}$$

The Grossman-Milgrom evidence disqualifies some states of the world instead of directly representing the truth. Any type outside the evidence cannot be true state. Belief regarding types inside the evidence are Bayesian updated.

To study the inquisitorial model under this evidence structure, we have to go beyond what is currently available in the literature and formulate the technology of evidence acquisition. Our formulation of the evidence acquisition is based on this evidence structure. The mediator specifies an intended evidence E, and observes the realization of an indicator variable $\rho \in \{T, F\}$ where $\rho = T$ means the intended evidence is available, i.e., $E \in \mathcal{E}(\omega)$ and $\rho = F$ means the intended evidence is not available, i.e., $E \notin \mathcal{E}(\omega)$. Given the prior μ^0, the true state ω, the intended evidence E, and the realization ρ, evidence acquisition updates belief as follows: if $\rho = T$ such that $\omega \in E$,

$$\mu^1(\omega'|E,T) = \begin{cases} 0 & \text{if } \omega' \notin E, \\ \frac{\mu(\omega'|m)}{\sum_{\tilde{\omega} \in E} \mu(\tilde{\omega}|m)} & \text{if } \omega' \in E, \end{cases} \tag{2.13}$$

if $\rho = F$ such that $\omega \notin E$,

$$\mu^1(\omega'|E,F) = \begin{cases} 0 & \text{if } \omega' \in E, \\ \frac{\mu(\omega'|m)}{\sum_{\tilde{\omega} \notin E} \mu(\tilde{\omega}|m)} & \text{if } \omega' \notin E. \end{cases} \tag{2.14}$$

If the intended evidence is acquired, evaluation informs the mediator that any type outside the evidence cannot be truth. If the intended evidence is not acquired, evaluation is not uninformative but instead tells the mediator that any type within the evidence cannot be truth.

We now consider the implications of the Grossman-Milgrom evidence structure. Once we move from costly auditing to costly evidence, two complications arise. The first one is the evidence acquired will be a set instead of the true state. But that itself doesn't stop us from learning the truth, if E is chosen properly. In particular, when the plaintiff reports omega, the mediator will check whether the true state is at least omega. If that's true, he will treat the type of the plaintiff as the smallest element in that set. So a higher type will be assigned a finer set. No one wants to misreport. Otherwise, he will either be found lying or he will receive a lower payoff that's actually consistent with the classic unraveling result in voluntary disclosure.

Let $\underline{\omega}_E = \inf E(\omega)$ be the smallest element of the acquired evidence $E(\omega)$. The expected payoffs for the plaintiff and the defendant are

$$
\begin{aligned}
Y_1(\omega, \hat{\omega}, E(\hat{\omega})) &= \begin{cases} [y_\pi(\underline{\omega_E}) - q_\pi(\underline{\omega_E})\tau_1(\underline{\omega_E})] + (1 - q_\pi(\underline{\omega_E}))(\omega - L) & \text{if } \hat{\omega} \in E(\hat{\omega}), \\ \omega - L & \text{if } \hat{\omega} \notin E(\hat{\omega}). \end{cases} \\
Y_2(\omega, \hat{\omega}, E(\hat{\omega})) &= \begin{cases} -[y_\pi(\underline{\omega_E}) + q_\pi(\underline{\omega_E})\tau_2(\underline{\omega_E})] - (1 - q_\pi(\underline{\omega_E}))(\omega + L) & \text{if } \hat{\omega} \in E(\hat{\omega}), \\ -\omega + L & \text{if } \hat{\omega} \notin E(\hat{\omega}). \end{cases}
\end{aligned} \tag{2.15}
$$

Proposition 2.5. *$\{E(\omega), \pi_E(y|\omega)\}$ solves the mediation problem with costly evidence if $E(\omega) = [\omega, 1]$ for any ω such that $I(\omega) = 1$, $\pi_E(y|\omega) = \pi_0(x|\omega)$ for $E = \varnothing$, $\pi_E(y|\omega) = \pi_1(y|\omega)$ for $E \neq \varnothing$.*

2.5.4 Dye Evidence Structure

The complication added by the Dye evidence structure is that evaluation may be inconclusive. And that gives rise to the second threshold. That problem is similar to facilitative mediation with one critical difference: the defendant, in his posterior belief, thinks being inconclusive higher types are more likely, because the evidence assigned to them is harder to find. In fact, we prove in the paper that if η satisfies a monotonicity condition, then the second threshold can even reach the highest type, such that all cases can be settled if and only if evaluative mediation is used.

Theorem 2.4. *An efficient mediation plan with Dye evidence structure is characterized by two thresholds $\{\omega^*, \omega_I^*\}$ such that:*

$$
\begin{aligned}
\omega^* &= \sup\{\omega | \omega - \mathbb{E}[\omega' | \omega' \leqslant \omega] = L_1 + L_2\}, \\
\omega_I^* &= \min\left\{\overline{\omega},\ \sup\{\omega | \omega - \mathbb{E}_\eta[\omega' | \omega' \in [\omega^*, \omega]] = L_1 + L_2 - C\}\right\}.
\end{aligned}
$$

(i) For any $\omega \leqslant \omega^*$, $E(\omega) = \varnothing$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 1$, and $x_\pi(\omega) = \omega^* - L_1$;

(ii) For any $\omega > \omega^*$:

if $C > L_1 + L_2$, then $E(\omega) = \varnothing$, $t_1 = t_2 = 0$, $p_\pi(\omega) = 0$; if $C \leqslant L_1 + L_2$, then $E(\omega) = [\omega, \overline{\omega}]$, $t_1 + t_2 = C$, $q_\pi(\omega) = 1$, $y_\pi(\omega) \in [\omega - L_1 + t_1, \omega + L_2 - t_2]$. (iii) For any $\omega \in [\omega^*, \omega_I^*]$, $r_\pi(\omega) = 1$, $z_\pi(\omega) = \omega_I^* - L_1$,

(iv) For any $\omega > \omega_I^*$, $r_\pi(\omega) = 0$.

2.6 Concluding Remarks

This paper studies the efficient mediation procedure where disputants are asymmetrically informed and hard evidence can be acquired probabilistically at a cost. A mediator commits ex-ante to a mediation plan that generates stochastic messages for the uninformed party, based on the informed party's reports, and acquired costly evidence. The model encompasses both facilitative mediation where mediator only transmits information, and evaluative mediation where mediator bases recommendation on evidence, thus has to acquire information.

The efficient mediation plan features a simple threshold partitioning the ordered state space into truth set (where truth is revealed by the evidence) and pooling set (where no evidence is acquired or presented). Below the threshold, the mediator facilitates, all types are settled, and the settlement is the same across types. Above the threshold, the mediator evaluates, all types are settled, and a range of settlements can be implemented but the lower bound and the upper bound both increases with the type. Both the efficient adversarial mediation plan and the efficient inquisitorial mediation plan are such threshold mechanisms. As a result, our first theorem states that irrespective of the procedures, full resolution of disputes is the efficient outcome. Therefore, partial resolution of disputes results from the literature's exclusive focus on facilitation, and full resolution of disputes is implementable once we incorporate evaluation.

While facilitation is the exclusive focus of previous literature, our findings suggest that evaluation is equally important for efficiency. Our results speak directly to the evaluative-facilitative debate central in mediation. Our findings highlight mediation default, which bears implications for the design of online platform, dispute resolution, ratings, and international relations.

Our efficient mediation appears consistent with several empirical findings (McDermott & Obar, 2004; Klerman & Klerman, 2015). Substantial percent of mediators use both evaluative and facilitative techniques, and/or "hybrid" techniques. A mediator's proposal, when used, leads to very high settlement rate (over 99 percent). Pure facilitation has a comparatively narrow range of settlement. Evaluation results in higher amount of settlement.

We conclude this paper by relating it to the literature that no unmediated negotiation procedures can achieve the same mediated result. Mediation has a strict benefit. We also advocate the policy of mediation default to resolve costly disputes, especially in developing countries where the legal costs are high, and with the advancement in information technology, the cost of evaluation is getting lower and lower.

2.A Appendix A. Proofs

Appendix A. Proofs

Direct Mechanism

The following revelation principle is in the spirit of Myerson (1991). The following lemma tells us that we can focus only on direct, truthful, and obedient mediation plans.

Lemma 2.9. *Given an arbitrary mediation plan $\tilde{\pi}(m_r|m_s)$ which implements a random mapping from states to joint distributions of (D, a) as an outcome of a perfect Bayesian equilibrium, there exists a direct, truthful, and obedient mediation plan $\pi(D, a_T, a_F|\omega)$ that implements the same random mapping as an outcome of a perfect Bayesian equilibrium.*

Proof. Construct a direct mechanism π using type report ω as input and a recommended set D and recommended action a_z contingent on $\rho = \{T, F\}$ as the output:

$$\pi(D, a_T, a_F|\omega) \equiv \sum_{m_r}\sum_{m_s}\left(\prod_z \tilde{\sigma}_a(a_z|D, m_r, \rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega).$$

In the following, we verify that this mechanism implements the same distribution of outcomes by means of a truthful and obedient PBE, where $\sigma_s(\omega|\omega) = 1$, $\sigma_r(D|D, a_T, a_F) = 1$, and $\sigma_a(a_z|D, a_T, a_F, \rho) = 1$.

First, the sender finds truth-telling optimal:

$$\begin{aligned}
&U_s(\omega|\omega) \equiv \sum_{D,\rho,a_z} a_z \Pr(\rho|D,\omega)\pi(D,a_T,a_F|\omega) \\
= &\sum_{m_s,m_r,D,\rho,a_z} a_z\tilde{\sigma}_a(a_T|D,m_r,T)\tilde{\sigma}_a(a_F|D,m_r,F)\Pr(\rho|D,\omega)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega) \\
\geqslant &\sum_{m_s,m_r,D,\rho,a_z} a_z\tilde{\sigma}_a(a_T|D,m_r,T)\tilde{\sigma}_a(a_F|D,m_r,F)\Pr(\rho|D,\omega)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega') \\
= &\sum_{D,\rho,a_z} a_z \Pr(\rho|D,\omega)\pi(D,a_T,a_F|\omega') = U_s(\omega'|\omega),
\end{aligned}$$

where the inequality follows because in the original PBE, $\tilde{\sigma}_s(\cdot|\omega)$ is a better strategy than $\tilde{\sigma}_s(\cdot|\omega')$ when the true state is ω.

Second, the receiver's beliefs are consistent with the equilibrium. Upon receiving recommendation (D, a_T, a_F),

$$\begin{aligned}
&\mu^1(\omega|D,a_T,a_F) = \frac{\pi(D,a_T,a_F|\omega)\mu^0(\omega)}{\sum_\omega \pi(D,a_T,a_F|\omega)\mu^0(\omega)} \\
= &\frac{\mu^0(\omega)\sum_{m_r}\sum_{m_s}\left(\prod_z \tilde{\sigma}_a(a_z|D,m_r,\rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega)}{\sum_\omega \mu^0(\omega)\sum_{m_r}\sum_{m_s}\left(\prod_z \tilde{\sigma}_a(a_z|D,m_r,\rho)\right)\tilde{\sigma}_r(D|m_r)\tilde{\pi}(m_r|m_s)\tilde{\sigma}_s(m_s|\omega)}.
\end{aligned}$$

Upon receiving recommendation (D, a_T, a_F) and seeing the test outcome ρ,

$$\begin{aligned}
\mu^2(\omega|D,a_T,a_F,T) &= \begin{cases} 0 & \text{if } \omega \notin D, \\ \frac{\mu^1(\omega|D,a_T,a_F)}{\sum_{\omega'\in D}\mu^1(\omega|D,a_T,a_F)} & \text{if } \omega \in D. \end{cases} \\
\mu^2(\omega|D,a_T,a_F,F) &= \begin{cases} 0 & \text{if } \omega \in D, \\ \frac{\mu^1(\omega|D,a_T,a_F)}{\sum_{\omega'\in D^c}\mu^1(\omega|D,a_T,a_F)} & \text{if } \omega \notin D. \end{cases}
\end{aligned} \tag{2.16}$$

Third, the receiver finds obedience optimal. Notice that by Bayes' rule,

$$\mu^1(\omega|D, a_T, a_F) = \sum_{m_r} \Pr(m_r|D, a_T, a_F)\tilde{\mu}^1(\omega|m_r),$$

where $\tilde{\mu}^1(\omega|m_r)$ is the receiver's belief upon receiving message m_r in the original PBE. Therefore, for every m_r such that the receiver's strategy selects plan (D, a_T, a_F) with positive probability in the original PBE, we have:

$$\begin{aligned} & \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D, \omega)\tilde{\mu}^1(\omega|m_r) - c(D) \\ \geqslant \; & \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D', \omega)\tilde{\mu}^1(\omega|m_r) - c(D'), \end{aligned}$$

for any (D', a'_T, a'_F). Averaging over all such m_r, we know that in the direct mediation plan:

$$\begin{aligned} & U_r(D, a_T, a_F|D, a_T, a_F) \equiv \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D, \omega)\mu^1(\omega|D, a_T, a_F) - c(D) \\ = \; & \sum_{m_r} \Pr(m_r|D, a_T, a_F) \sum_{\rho,\omega} v(a_z, w) \Pr(\rho|D, \omega)\tilde{\mu}^1(\omega|m_r) - c(D) \\ \geqslant \; & \sum_{m_r} \Pr(m_r|D, a_T, a_F) \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D', \omega)\tilde{\mu}^1(\omega|m_r) - c(D') \\ = \; & \sum_{\rho,\omega} v(a'_z, w) \Pr(\rho|D', \omega)\mu^1(\omega|D, a_T, a_F) - c(D') \\ = \; & U_r(D', a'_T, a'_F|D, a_T, a_F). \end{aligned}$$

□

Proof of Lemma 2.1

Proof. Consider two types ω and $\hat{\omega}$ where $\omega > \hat{\omega}$. IC-M requires

$$\begin{aligned} p(\omega)x(\omega) + (1 - p(\omega))(\omega - L) &\geqslant p(\hat{\omega})x(\hat{\omega}) + (1 - p(\hat{\omega}))(\omega - L) \\ p(\hat{\omega})x(\hat{\omega}) + (1 - p(\hat{\omega}))(\hat{\omega} - L) &\geqslant p(\omega)x(\omega) + (1 - p(\omega))(\hat{\omega} - L) \end{aligned} \tag{2.17}$$

Adding the two inequalities and rearranging, we get

$$(p(\hat{\omega}) - p_{\pi}(\omega))(\omega - \hat{\omega}) \geqslant 0. \tag{2.18}$$

Since $\omega > \hat{\omega}$, we have $p_{\pi}(\omega) \leqslant p(\hat{\omega})$.

Equation IC-M means that for all ω, we have

$$X_1(\omega) = \max_{\hat{\omega} \in \{\hat{\omega} | I(\hat{\omega}) = 0\}} p(\hat{\omega})x(\hat{\omega}) - (1 - p(\hat{\omega}))L + (1 - p(\hat{\omega}))\omega \tag{2.19}$$

By envelop theorem, we have $X_1'(\omega) = 1 - p_{\pi}(\omega)$ whenever differentiable.

By the second fundamental theorem of calculus (Royden and Fitzpatrick (2010) shows it extends to piece-wise differentiability), we obtain

$$X_1(\omega) = \int_0^{\omega} [1 - p(\tilde{\omega})] d\tilde{\omega} + X_1(0) \quad \forall \omega. \tag{2.20}$$

□

Proof of Lemma 2.2

Proof. If $q_\pi(\omega) = 0$, since $X_1(\omega) - (1 - q_\pi(\omega))(\omega - L) \leqslant 0$, the equation holds.

Assume $q_\pi(\omega) \in (0, 1)$, suppose $y_\pi(\omega) > \max\{\frac{X_1(\omega)}{q_\pi(\omega)} - \frac{1-q_\pi(\omega)}{q_\pi(\omega)}(\omega - L), (\omega - L)\}$. There exists $\tilde{y}(\omega) \geqslant \omega - L_1$ and $\tilde{y}(\omega) < y_\pi(\omega)$. Observe that $\tilde{y}(\omega)$ is feasible but would relax the IR constraint for the defendant and thus make it feasible to raise the agreement probability $\tilde{q}(\omega) > q_\pi(\omega)$. This contradicts the fact that $q_\pi(\omega)$ is a maximizer. Therefore, $q_\pi(\omega)y_\pi(\omega) = \max\{X_1(\omega) - (1 - q_\pi(\omega))(\omega - L), q_\pi(\omega)(\omega - L)\}$ for all ω where $q_\pi(\omega) < 1$.

If $q_\pi(\omega) = 1$, then IC is not binding and $y_\pi(\omega) \geqslant \max\{X_1(\omega), (\omega - L)\}$. □

Proof of Lemma 2.3

Proof. Let $\mathcal{P}$ be the set of all bounded non-increasing functions such that $p_\pi(\omega) \in [0, 1] \forall \omega \in [0, 1]$. This is the set the designer can choose. We endow $\mathcal{P}$ with the linear structure and the metrics induced by L^1-norm.

Notice that $\mathcal{P}$ is compact and convex. Denote L_0 as the objective function and notice further that L_0 is continuous and linear in $p_\pi(\omega)$. By the Extreme Point Theorem (Ok, 2007, p.658), the set of extreme points of $\mathcal{P}$ is nonempty which includes a $p_\pi(\omega)$ such that

$$L_0(p_\pi(\omega)) \geqslant L_0(\tilde{p}(\omega)) \quad \forall p \in \mathcal{P}. \tag{2.21}$$

Therefore a function $p_\pi(\omega)$ that is an extreme point of $\mathcal{P}$ and that minimizes L_0 among all extreme points of $\mathcal{P}$ also minimizes L_0 among all functions in $\mathcal{P}$. A function $p_\pi(\omega)$ is an extreme point of $\mathcal{P}$ if $p_\pi(\omega) \in \{0,1\}$ for almost all $\omega \in [0,1]$. The designer can thus restrict attention to non-stochastic $p_\pi(\omega)$.

An extreme point $p_\pi(\omega)$ is non-increasing if and only if there exists a ω^* such that

$$p_\pi(\omega) = \begin{cases} 1 & \text{if } \omega \leqslant \omega^*, \\ 0 & \text{if } \omega > \omega^*. \end{cases} \tag{2.22}$$

□

Proof of Lemma 2.4

Proof. $q_\pi(\omega) = 1$ for $\omega \in \{\omega | I(\omega) = 1\}$ obviously solves the unconstrained program. We show it is also feasible in the constrained program if $t_1 + t_2 \leqslant L_1 + L_2$.

By lemma 2,

$$y_\pi(\omega) - t_1 = \begin{cases} \max\{\omega^* - L_1, \omega - L_1\} = \omega^* - L_1 & \text{if } I(\omega) = 1 \& \omega \leqslant \omega^*, \\ \omega - L_1 & \text{if } I(\omega) = 1 \& \omega > \omega^*. \end{cases} \tag{2.23}$$

It remains to check whether IR constraint for the defendant is satisfied. By Bayes rule,

$\mathbb{E}[\omega|I(\omega)=1]=\omega$. If $I(\omega)=1$ and $\omega \leqslant \omega^*$, we have

$$\begin{aligned} y_\pi(\omega)+t_2 &= \omega^* - L_1 + t_1 + t_2 \\ &\leqslant \omega^* + L_2 - (\omega^* - \omega) \\ &= \omega + L_2 \end{aligned} \tag{2.24}$$

If $I(\omega)=1$ and $\omega > \omega^*$, we have

$$\begin{aligned} y_\pi(\omega)+t_2 &= \omega - L_1 + t_1 + t_2 \\ &\leqslant \omega + L_2 \end{aligned} \tag{2.25}$$

□

Proof of Lemma 2.5

Proof. Suppose the function $I^*(\omega)$ is a solution to the arbitration problem. Define $h(\omega)$ to be the deviation between $I^*(\omega)$ and some other feasible function $I(\omega)$. For any constant a, the function $I(\omega) = I^*(\omega) + ah(\omega)$ is also feasible. With both $I^*(\omega)$ and $h(\omega)$ held fixed, consider the Lagrangian as a function of a,

$$\begin{aligned} \mathcal{L}(a) = &\int_0^1 \left[q_\pi(\omega) + [I^*(\omega)+ah(\omega)]\,[p_\pi(\omega)-q_\pi(\omega)]\,\right](\sum_i L_i - \sum_i t_i)\mu^0(\omega)d\omega \\ &+ \lambda_T \int_0^1 \{\,[I^*(\omega)+ah(\omega)]\,[C-(p_\pi(\omega)-q_\pi(\omega))\sum_i t_i] - p_\pi(\omega)\sum_i t_i\}\mu^0(\omega)d\omega \end{aligned} \tag{2.26}$$

By assumption, the function $\mathcal{L}(a)$ obtains its optimum at $a = 0$. This implies that for all $h(\omega)$, the first order derivative

$$\mathcal{L}'(a) = \int_0^1 \{[p_\pi(\omega) - q_\pi(\omega)][\sum_i L_i - \sum_i t_i] + \lambda_T[C - (p_\pi(\omega) - q_\pi(\omega)) \sum_i t_i]\} h(\omega)\mu^0(\omega) d\omega \leqslant 0 \quad \forall a > 0 \tag{2.27}$$

Therefore, if $h(\omega) > 0$, $[p_\pi(\omega) - q_\pi(\omega)][\sum_i L_i - \sum_i t_i] \leqslant \lambda_T[(p_\pi(\omega) - q_\pi(\omega)) \sum_i t_i - C]$; if $h(\omega) < 0$, $[p_\pi(\omega) - q_\pi(\omega)][\sum_i L_i - \sum_i t_i] \geqslant \lambda_T[(p_\pi(\omega) - q_\pi(\omega)) \sum_i t_i - C]$. Since $I^*(\omega) + ah(\omega)$ is feasible, it follows that

$$I^*(\omega) = \begin{cases} 1 & \text{if } q_\pi(\omega) - p_\pi(\omega) \geqslant \frac{\lambda_T C}{\sum_i L_i - (1+\lambda_T)\sum_i t_i}, \\ 0 & \text{if } q_\pi(\omega) - p_\pi(\omega) < \frac{\lambda_T C}{\sum_i L_i - (1+\lambda_T)\sum_i t_i}. \end{cases} \tag{2.28}$$

Since $p_\pi(\omega)$ is non-increasing, $\{\omega | I(\omega) = 1\}$ is a nonempty connected set if (i) there exists $\underline{\omega}$ such that $q_\pi(\underline{\omega}) - p_\pi(\underline{\omega}) \geqslant \frac{\lambda_T C}{\sum_i L_i - (1+\lambda_T)\sum_i t_i}$, (ii) $q_\pi(\omega)$ is non-decreasing in $[0, 1]$. Set $q_\pi(\omega) = 0$ for $\omega \in [0, \omega_I^*)$, then $q_\pi(\omega)$ is non-decreasing. It follows that (i) becomes $\sum_i L_i \geqslant \lambda_T C + (1 + \lambda_T) \sum_i t_i$, the stated condition. □

Proof of Lemma 2.6

Proof. In an efficient mediation, the budget constraint has to be binding, otherwise we can always increase the total payoffs by reducing $t_1 + t_2$. The optimality conditions for

$I(\omega)$ becomes

$$I(\omega) = \begin{cases} 1 & \text{if } q_\pi(\omega) - p_\pi(\omega) \geqslant \frac{\lambda_T C}{L_1+L_2}, \\ 0 & \text{if } q_\pi(\omega) - p_\pi(\omega) < \frac{\lambda_T C}{L_1+L_2}. \end{cases} \tag{2.29}$$

$$= \begin{cases} 1 & \text{if } \lambda_T C \leqslant L_1 + L_2, \\ 0 & \text{otherwise.} \end{cases} \tag{2.30}$$

□

CHAPTER 3

PROFIT-CAP POLICY: THEORY AND APPLICATIONS

3.1 Introduction

The large profits earned by some industries, such as the Big Tech or Big Pharma, have been a source of much debate and acrimony in recent times. The quandary is created by the fact that their production technologies are characterized by economies of scale that make it difficult to solve the problem by the traditional method of using antitrust law to break up these large corporations into several small firms, and additionally by the fact that these industries provide goods and services that are considered essential in today's world, and so we cannot risk damaging their incentive to produce by our zealousness to curtail their profits. The problem has been heightened by the recent controversy concerning the global availability of COVID-19 vaccines, which have generated large profits but at the same time are in short supply in most developing economies.

How we reconcile these conflicting objectives and get these firms to supply more, and to reach more consumers, without making demands on them in ways so unmindful of incentives that the policy backfires, resulting in the corporations cutting down production or even closing down, is the big challenge. It is for this reason that many nations, such as Germany and Switzerland, have resisted the popular demand to revoke patents. In the long run this may damage production by hurting the incentive of firms to undertake costly research and produce new drugs.

This is an important debate that will, no doubt, play out for a while. The aim of this paper is to draw on economic theory to make a small contribution to a segment of this debate. It is a simple theoretical result that we prove here but we believe it can have large, real-life policy implications.

In most industrial settings, the policy of curtailing profit tends to go hand in hand with damaging incentives and the efficiency of markets, which in turn causes supply to drop (Katz and Rosen, 1985, Levin, 1985, Vives, 1999). In many industries, such as for those producing luxury goods, we may decide that the price of some efficiency loss is worth it in order to curtail mega profits accruing in a few hands. But for essential goods and services such as vaccines and telecommunication we do not have that latitude. This is one reason why, despite the large profits earned by the Big Tech and the Big Pharma economists hesitate to recommend the blunt policy of curtailing their profits.

What is missed out in this traditional argument is the fact that a restriction on profits, when placed on a group of two or more firms works very differently from when the restrictions are firm-specific. What we show in this paper is that if a profit cap is set, for instance, on the entire vaccine producing sector, that is on the total profit of the industry, that may increase, rather than decrease, the supply of vaccines and enhance overall efficiency. If an aggregate profit cap is placed on all the big pharmaceutical firms, in a specific way (to be clarified later), that can heighten competition among these firms, making them produce more and driving the oligopoly towards the competitive equilibrium. In other words, a profit cap, done right, can be a virtual substitute for antitrust law, with the advantage that this is achieved without breaking up firms (which hurts consumers buying from sectors like the Big Tech and the pharmaceutical industry, where there are large, natural economies of scale). One interesting by-product of this analysis is the 'duopoly rule,' which shows that, in some circumstances, the ideal market structure is the duopoly.

These abstract theoretical insights can then be used to design other kinds of interventions that are being talked about nowadays. There is, for instance, talk of mandatory donations of vaccines by corporations, which would then be used by the governments of advanced economies to distribute vaccines to developing countries that are unable to compete and buy enough vaccines for their citizenry. But how should we distribute the mandatory donation responsibility across the firms without damaging their incentive to produce? The main result in the paper gives a surprising insight into the design of mandatory donation. The main result also has implications for e-commerce and digital platforms' antitrust problem.

3.1.1 Related Literature

This paper contributes to the literature on regulations in oligopoly markets (Armstrong and Sappington, 2007). The two main policy instruments studied extensively in the literature are price caps, and taxation. The goal of the policies is to increase welfare by curbing market power. The literature on price cap shows that under deterministic demand, imposing a profit cap expands total output and increases consumer welfare and total welfare (Earle et al., 2007). [1] The literature on taxation shows that imposition of the tax unambiguously reduce total output, but may increase industry profits in some cases (Katz and Rosen, 1985, Levin, 1985, Myles, 1987). To our knowledge, this is the first paper in the literature that studies the welfare improving role of a profit-cap policy. In doing so, we consider the design of ex-post liability to enforce this policy, and allow for a range of policy instruments, including taxation and donation requirement. The desirable effect of this policy cannot be replicated by taxation alone; in fact, in our equilibrium the tax rate is zero. The effect might be reached by a price cap. However, the price-cap literature is

[1] See Grimm and Zöttl (2010) for a general characterization of the case under stochastic demand, Reynolds and Rietzke (2018) for the case of endogenous entry.

silent on the enforcement problem (i.e., what happens when firms do not respect the price cap) and cannot be easily translated as a donation requirement. In this sense, our paper is complementary to the literature on price cap.

The industry-level profit cap this paper proposes is reminiscent of joint liability where if one individual deviates from the prescribed behavior, the entire group is held liable. The idea that joint liability has the potential to improve social welfare has been well established in the microfinance lending literature (Ghatak and Guinnane, 1999, Besley and Coate, 1995) and recently extended to the sovereign debt literature (Basu and Stiglitz, 2015). In this paper, the firms have no contractual relationship or common collateral, but the fact that tax liability is applied universally if any firm disregard the profit cap gives every firm an incentive to expand output prescribed by the profit cap. To our knowledge, this is the first paper that applies the idea of joint liability to support a regulation in the product market.

The rest of the paper proceeds as follows. Section 3.2 presents a profit cap model with linear demand and states the main result on profit cap and efficiency in such context. Section 3.3 generalizes and analyzes the profit cap model by allowing general demand functions and endogenous entry. Section 3.4 discusses the two applications of the theory to vaccine supply and e-commerce. All formal proofs are relegated to Appendix 3.A.

3.2 A Simple Model on How Profit Caps Increase Efficiency

To understand the intuition which is crucial to extend the theoretical model to actual domains of policy, we shall begin with a special case, with strong assumptions, which enables us to explain the gist of the analysis and one of the main results. Let $N = \{1, \ldots, n\}$ be a set of $n \geqslant 2$ identical firms with a constant marginal cost of production $c > 0$. The firms engage in a Cournot competition with homogeneous good and a linear inverse market demand function given by:

$$p(X) = a - bX,$$

for some $a, b > 0$, and with X denoting the aggregate demand for the good. Further, let us for now assume that firms do not incur any fixed cost of production.

The game happens as follows. In period 1, the government announces a profit cap $\hat{\Pi} \geqslant 0$ for the industry. This means that government has made a commitment to adjust t to make sure that the industry's profit does not exceed $\hat{\Pi}$. After seeing how much the firms produce, government will set a per unit tax t to ensure that total profit is less than $\hat{\Pi}$. Government is a player that mechanically does this. The game that we are about to describe may be called the "$\hat{\Pi}$–oligopoly". In period 2, all firms simultaneously announce how much they will produce:

$$\mathbf{x} \equiv (x_1, x_2, \ldots, x_n) \geqslant 0.$$

In period 3, government chooses t to make sure that total profit does not exceed $\hat{\Pi}$.

Given $\mathbf{x}$ and t, firm i's profit is

$$\pi_i(\mathbf{x}, t) = (a - bX)x_i - (c + t)x_i.$$

Define the industry profit

$$\Pi(\mathbf{x}, t) = \sum_{i \in N} \pi_i(\mathbf{x}, t).$$

Formally, given the profit cap rule, this is how government selects t in the $\hat{\Pi}$–oligopoly. If $\mathbf{x}$ is such that $\Pi(\mathbf{x}, 0) \leqslant \hat{\Pi}$, then $t(\mathbf{x}, \hat{\Pi}) = 0$. Otherwise, $t = t(\mathbf{x}, \hat{\Pi})$ is chosen such that $\Pi(\mathbf{x}, t) = \hat{\Pi}$. Let $X(\hat{\Pi})$ denotes the aggregate output of the industry in equilibrium. The following theorem characterizes the equilibrium outcome of $\hat{\Pi}$–oligopoly.

Theorem 3.1. *In the subgame perfect equilibrium of a $\hat{\Pi}$–oligopoly, $t = 0$. As the profit cap decreases, total output increases, i.e., $\hat{\Pi}' < \hat{\Pi}$ implies $X(\hat{\Pi}') \geqslant X(\hat{\Pi})$. As $\hat{\Pi} \to 0$, the equilibrium output of the $\hat{\Pi}$–oligopoly converges to the competitive equilibrium.*

We shall sketch the intuition behind the theorem. The formal proof is easy to create out of that. Figure 3.1 shows the inverse demand function, marked DD, and also shows the marginal cost curve, which is the horizontal line drawn through the point marked c.

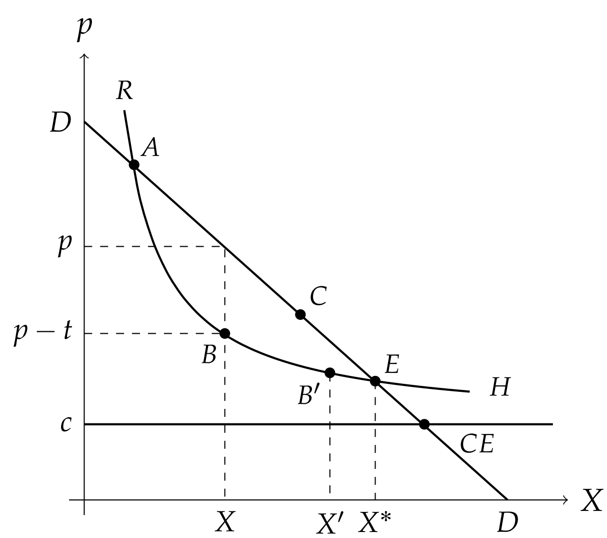


Figure 3.1: Linear Demand Example

We have also marked on the figure, a point C where the standard Cournot equilibrium occurs (that is, with no taxation), and the competitive equilibrium is of course at the point of intersection between the demand curve and marginal cost curve, denoted by CE.

Now, suppose the game will be played as a $\hat{\Pi}$–oligopoly ($\hat{\Pi} > 0$). To see what the equilibrium will be like draw the line defined by $(p - c)X = \hat{\Pi}$, or $p = \hat{\Pi}/X + c$. This will be the rectangular hyperbola, with c being treated as the origin. The rectangular hyperbola is shown by the curved line marked RH. It follows that given the profit cap at $\hat{\Pi}$ and the inverse demand function the final equilibrium (p, X), that is, price and aggregate production will be on the line $DABB'ED$. The segment of the demand curve to the right of this results in aggregate profit that exceed $\hat{\Pi}$.

Assume that the industry produces $(x_1, \ldots, x_n)$ such that $X = \sum_{i=1}^{n} x_i$ is as shown in Figure 3.1. If there was no tax, price would be at p as shown in the figure. It follows that in the $\hat{\Pi}$–oligopoly government will step in, in period 2, with a tax of t as shown in the figure. This will ensure post-tax profit is $\hat{\Pi}$.

Could this be a subgame perfect equilibrium? Clearly not. Take any firm i. If this firm unilaterally expands production, X will rise, say to X'. Government will then adjust t so that the industry settles at B'. Clearly, the post-tax price at B' is lower than at B. Hence all the other firms with positive production will see their profit decline. Since total profit at B and at B' is the same, firm i's profit would be higher at B'. hence, B could not have been an equilibrium.

By this reasoning equilibrium can only be at E. Note, if they were anywhere on the demand curve DD, to the right of E, that cannot be an equilibrium, since on the stretch ED, $t = 0$, and we know that with zero tax, the Cournot equilibrium is at point C. And given our assumptions this is a stable equilibrium. So some firm would want to unilaterally cut output.

The only thing that remains to be checked is whether E is indeed an equilibrium. Mark the point on the horizontal axis below E by X^*, as shown. Define $x^* \equiv X^*/n$. Hence if every firm produces x^*, total output will be X^* (where the regular demand curve meets rectangular hyperbola).

Let $\phi(x_i, x)$ be the profit earned by firm i when i produces x_i and every other firm produces x. Hence

$$\phi(x_i, x) = [a - b(x_i + x(n-1))]x_i - cx_i.$$

Hence,

$$\frac{\partial \phi(x_i, x)}{\partial x_i} = a - bx(n-1) - 2bx_i - c. \tag{3.1}$$

Let us denote the standard Cournot equilibrium output (with no profit cap) by X^C. Thus, X^C (not shown in the figure) is the point vertically below C in Figure 3.1. Since all firms are identical, in the standard Cournot equilibrium each firm produces $x^C = X^C/n$. From the fact that this is a Cournot equilibrium it follows that

$$\left.\frac{\partial \phi(x_i, x^C)}{\partial x_i}\right|_{x_i = x^C} = 0.$$

Using 3.1, we therefore have

$$a - b(n+1)x^C - c = 0.$$

It follows that,

$$\left.\frac{\partial \phi(x_i, x^*)}{\partial x_i}\right|_{x_i = x^*} = a - b(n+1)x^* - c < 0$$

since $x^* > x^C$. Hence, if all firms are producing x^* (that is, we are at point E in Figure 3.1), no firm will want to produce more.

However, as we have already seen, in the ÎI–oligopoly no firm will want to cut production (since that will result in government raising t in stage 2 to hold industry-wide profit constant at $\hat{\Pi}$). It follows that every firm producing x^* with total output being X^* and equilibrium tax rate, $t^* = 0$ (all represented by point E in Figure 3.1) is part of the subgame perfect equilibrium of $\hat{\Pi}$–oligopoly. This establishes the claim in Theorem 1 that, in equilibrium, $t = 0$. Further as the profit cap become severe, that is, as $\hat{\Pi}$ falls the rectangular hyperbola RH will shift to the left. Hence, the point of intersection with the demand, that is, E, will shift to the right along the demand curve DD. This establishes that as $\hat{\Pi}$ converges to 0, the equilibrium output of the $\hat{\Pi}$–oligopoly will converge to the competitive equilibrium. This completes the proof of Theorem 1.

Before closing this section, it may be worth asking an interesting question. We just saw that point E in Figure 3.1 with all firms producing the same amount of vaccines (namely, x^*) depicts a subgame perfect equilibrium of the $\hat{\Pi}$–oligopoly. This however does not tell us whether there are other equilibria where, of course, the same total amount, X^*, is produced but different firms produce different amounts.

The answer to this question is yes. For all $\hat{\Pi}$–oligopoly where $\hat{\Pi}$ is less than the standard Cournot profit (that is, the profit cap is of some consequence, there will be subgame perfect equilibria where the outcome is asymmetric across firms.

We will illustrate this with the case of duopoly, that is, $n = 2$. Let us assume Figure 3.1 illustrates the case of a duopoly. Thus C depicts the Cournot equilibrium of a standard Cournot duopoly (that is, with no government intervention). As is well-known, this equilibrium can also be depicted on a graph where the horizontal axis shows firm 1's output, x_1, and the vertical axis shows firm 2's output, x_2. If in this space we draw the reaction functions or best-response functions of the two firms, as shown by R_1 and R_2 in Figure 3.2, the point of intersection of the two reaction functions depicts the Cournot equilibrium.

Thus, point C in Figure 3.2 depicts the Cournot equilibrium and happens to be the counterpart of C in Figure 3.1. Through C draw a (negatively-sloped) 45°-line as depicted by the line AA. Clearly, the sum of x_1 and x_2 at each point on line AA adds up to $X^C = x^C + x^C$.

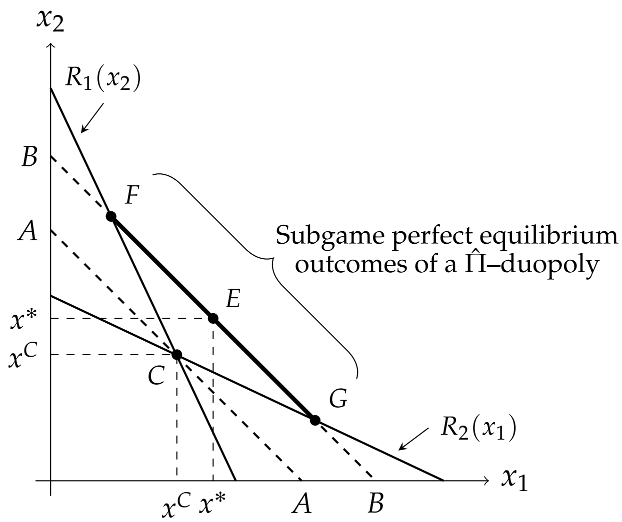


Figure 3.2: Best-Response

Now consider the aggregate output depicted by X^* in Figure 3.1. This is greater than X^C (which is the point vertically below C in Figure 3.1 on the horizontal axis). So if we want to depict a line of (x_1, x_2) such that $x_1 + x_2 = X^C$, this will be a (negatively-sloped) 45°–line to the right of AA. This is shown as BB in Figure 3.2. The mid-point of this line (x^*, x^*) is the counterpart of the equilibrium point E in Figure 3.1. This is shown as point E in Figure 3.2. In short, E in Figure 3.2 is also a depiction of the subgame perfect equilibrium of the ÎI–oligopoly.

To see if there are other combinations of x_1 and x_2 that add up to X^* and constitute part of an equilibrium, we need to locate points (x_1, x_2) on the line BB such that neither will want to unilaterally produce more. As we have already seen, producing less will cause the tax rate to rise and this is never worthwhile, unilaterally.

From the two reaction functions it is immediately obvious that all points (x_1, x_2) that lie on the line BB and are above both reaction functions R_1 and R_2, constitute subgame perfect equilibrium outcomes. Graphically, all points on the line segment FG and only those points constitute equilibria of the $\hat{\Pi}$–oligopoly.

To be more concrete, the reaction functions in Figure 3.2 are given by

$$R_1(x_2) = \frac{a - c - bx_2}{2b} \qquad \text{and} \qquad R_2(x_1) = \frac{a - c - bx_1}{2b}.$$

One can also think of a best-response of the whole $\hat{\Pi}$–oligopoly game as:

$$\hat{R}_i(x_j) = \begin{cases} R_i(x_j) & \text{for} \quad R_i(x_j) + x_j \notin [\underline{X}, \overline{X}] \\ \overline{X} - x_j & \text{for} \quad R_i(x_j) + x_j \in [\underline{X}, \overline{X}] \end{cases},$$

where

$$\overline{X} \equiv \frac{(a-c) + \sqrt{(a-c)^2 - 4b\hat{\Pi}}}{2b}$$
$$\underline{X} \equiv \frac{(a-c) - \sqrt{(a-c)^2 - 4b\hat{\Pi}}}{2b}.$$

Equilibrium outcomes of the $\hat{\Pi}$–oligopoly is given by (x_1, x_2) that solves $x_i = \hat{R}_i(x_j)$ for $i, j = 1, 2$. One can see from Figure 3.2 that the equilibrium output X^* is indeed $\overline{X}$.

In general, $\frac{a-c}{2b} \leqslant \overline{X} \leqslant \frac{a-c}{b}$. The upper bound occurs when $\hat{\Pi} = 0$, while the lower bound occurs at the monopoly profit $\hat{\Pi} = \frac{(a-c)^2}{4b}$. If $\hat{\Pi}$ were to be any higher $\overline{X}$ would not exist.

For the profit cap to be effective, however, we must have $X^C = \frac{2(a-c)}{3b} < \overline{X} < \frac{a-c}{b}$.

The line BB in Figure 3.2 shows $x_1 + x_2 = \bar{X}$, which demonstrates this case. The reason for this is because in the linear model, the profit cap curve, $\hat{\Pi}/X + c$, touches the demand at the point where marginal revenue is c. This is at the monopoly quantity, $\frac{a-c}{2b}$, and it happens at $\hat{\Pi} = \frac{(a-c)^2}{4b}$. At this level, the game is still not affected since the profit cap curve does not yet "catch" the original Cournot. For the profit cap to be in effect, it has to go down to the Cournot profit, $\hat{\Pi} = \frac{2(a-c)^2}{9b}$, or below. This corresponds to $\overline{X} = \frac{2(a-c)}{3b}$ or above.

3.3 General Model and Analysis

This section generalizes the $\hat{\Pi}$–oligopoly model in two ways. First, we allow for general inverse demand functions. Second, firms incur fixed costs when entering the market. Only firms that chose to enter will be able to compete in the production stage.

3.3.1 General Demand Function

Consider the market for vaccines. Let us suppose the inverse demand function is given by $p(X)$, where X is the industry output and $p'(X) < 0$. The properties on this function are specified later. As in standard oligopoly analysis, the demand comes from a multitude of consumers who are price takers, whereas the supply is from a limited number of firms, each of whose choice can affect price.

The global (inverse) demand $p(X)$ is assumed to satisfy the following standard assumptions:

Assumption 3.1. *$p(0) \in (c, \infty)$ and there exists $Z > 0$ such that $p(X) > 0$ for $X \in [0, Z)$ and $p(X) = 0$ for $X \geqslant Z$.*

Assumption 3.2. *$p(X)$ is twice-continuously differentiable and $p'(X) < 0$ for $X \in [0, Z)$.*

Assumption 3.3. *For all $X \in [0, Z)$, $p'(X) + p''(X)x_i < 0$, for $x_i \in [0, X]$.*

Assumptions 1, 2, and 3 ensure that a unique Nash Equilibrium exists in the standard Cournot competition among the producing firms (Gaudet and Salant, 1991). We denote the standard Cournot equilibrium by $\mathbf{x}^{NE}$ and let $X^{NE} = \sum_i x_i^{NE}$. A useful observation is that Assumptions 1-3 imply a single-crossing condition of the marginal revenue and such crossing occurs at the *monopoly output* in this market, X^M. Further, define X^{PC} to be the *perfectly competitive output*, $p(X^{PC}) = c$.

Lemma 3.1. *$MR(X) = p(X) + p'(X)X$ is strictly decreasing for $X \in [0, Z)$ and there exists a unique $X^M \in (0, Z)$ such that $MR(X^M) = c$.*

We analyze the profit cap subgame with fixed number of firms n. Given the output profile $\mathbf{x}$, the government sets the tax rate t such that $\Pi(\mathbf{x}, t) = \hat{\Pi}$ if the unregulated profit $\Pi(\mathbf{x}, 0)$ exceeds the profit cap $\hat{\Pi}$.

The proposed implementation of the profit cap defines a function $t = t(X, \hat{\Pi})$ for the tax rate, which must satisfy $t(X, \hat{\Pi})X = \max\{\Pi(X) - \hat{\Pi}, 0\}$ since the tax revenue must cover the excess industry profit, if any. This means that the profit cap tax satisfies:

$$t(X, \hat{\Pi}) = \max\left\{p(X) - \hat{p}(X, \hat{\Pi}), 0\right\}, \tag{3.2}$$

where q is the *profit cap curve* defined as

$$\hat{p}(X, \hat{\Pi}) \equiv \frac{\hat{\Pi}}{X} + c.$$

We can think of $\hat{p}(\hat{\Pi}, X)$ as an inverse demand curve. Its corresponding the marginal revenue curve has the property that it is constant at c for all X.

Since the government commits to this tax scheme, we can solve for the equilibrium output in stage 3 by deriving the firms' reaction functions. This is given by

$$\hat{R}_i(\mathbf{x}_{-i}) \equiv \underset{x_i \geqslant 0}{\text{argmax}} \, [p(x_i + X_{-i}) - c - t(x_i + X_{-i}, \hat{\Pi})]x_i$$

Define $\underline{X}(\hat{\Pi})$ and $\overline{X}(\hat{\Pi})$ as

$$\underline{X}(\hat{\Pi}) \equiv \inf \{X : p(X) = \hat{p}(X, \hat{\Pi})\}$$
$$\overline{X}(\hat{\Pi}) \equiv \sup \{X : p(X) = \hat{p}(X, \hat{\Pi})\}$$

with the convention that $\sup \varnothing = -\infty$ and $\inf \varnothing = \infty$, so that $[\underline{X}(\hat{\Pi}), \overline{X}(\hat{\Pi})] = \varnothing$ when $p(X) = \hat{p}(X, \hat{\Pi})$ has no solution. An important lemma is that the profit cap curve does not cross the inverse demand curve more than twice.

Lemma 3.2. *There exists at most two solutions to $p(X) = \hat{p}(X, \hat{\Pi})$. If the solution is unique, then it is X^M.*

Another related lemma characterizes the behavior of $\underline{X}(\hat{\Pi})$ and $\overline{X}(\hat{\Pi})$.

Lemma 3.3. *If $\hat{\Pi} > (p(X^M) - c)X^M$, then $[\underline{X}(\hat{\Pi}), \overline{X}(\hat{\Pi})] = \varnothing$. If $\hat{\Pi} = (p(X^M) - c)X^M$, then $\underline{X}(\hat{\Pi}) = \overline{X}(\hat{\Pi}) = X^M$. As $\hat{\Pi} \to 0$, $\underline{X}(\hat{\Pi}) \to 0$ and $\overline{X}(\hat{\Pi}) \to X^{PC}$.*

The above lemmas ensure that the model with a general $p(X)$ behaves the same way as a linear demand when it comes to the profit cap result. Thus, the general analogue of Theorem 3.1 holds.

Theorem 3.2. *The equilibrium industry output with profit cap is*

$$X^*(\hat{\Pi}) = \max\{X^{NE}, \overline{X}(\hat{\Pi})\}.$$

Moreover, the equilibrium profit is $\Pi^* = \min\{\Pi^{NE}, \hat{\Pi}\}$ *and the equilibrium tax rate is zero,* $t(X^*, \hat{\Pi}) = 0$.

As a sanity check, if $\hat{\Pi} \to \infty$, then the profit cap is not binding and there is no solution to $p(X) = \hat{p}(X, \hat{\Pi})$. Therefore, $[\underline{X}, \overline{X}]$ is empty and $\hat{R}_i = R_i$ as expected.

3.3.2 Fixed Cost and Entry

Denote by $\mathcal{M} = \{1, \ldots, m\}$, $m \geqslant 1$, the set of firms that have the capability to innovate and produce. Each of these firms, should it decide to get into vaccine production, has to incur a fixed cost of f_i in terms of R&D and other start 'entry costs'. We shall follow the convention of indexing firms by increasingly higher R& D fixed costs. Hence:

$$0 \leqslant f_1 \leqslant f_2 \leqslant f_3 \leqslant \ldots \leqslant f_m.$$

Each firm has to decide, as in standard oligopoly, whether or not to incur the fixed cost and enter the market, and if it does enter, it has to decide how much to produce in the next period. However, in this analysis the firms have to be mindful of the fact that there is a government or regulator that sets a per unit tax, t, that all firms that produce have to pay; and, further, the tax rate chosen by the government is dependant on the total amount produced by the industry. [2] Hence, unlike in a standard Cournot oligopoly, producers will now be aware that the amount produced by them will effect the value of t. We shall model this by treating the government as a player in the game, who uses t to place a profit cap, $\hat{\Pi}$, but an unusual one. It is a cap on the entire industry's *variable* profit. We shall call the $(m+1)$-player game with m firms and the government, a "$\hat{\Pi}$–oligopoly."

Our aim is to show that placing a variable profit cap is a strategy that can be a surprisingly powerful instrument not just to curb excessive profit, but to, at the same time, encourage greater production of vaccines and higher overall efficiency.

The firms engage in a two-stage game. In the first stage, each firm chooses whether or not to innovate, $b_i \in \{0, 1\}$, by paying the fixed cost $f_i \geqslant 0$. Firms that choose to innovate, $b_i = 1$, engage in a Cournot competition with the profit cap policy.[3] Otherwise, i chooses not to innovate, $b_i = 0$, and does not enter the competition. Let $\mathcal{N} = \{i \in \mathcal{M} : b_i = 1\}$ be the set of firms that developed the vaccine. Denote by $\mathbf{x} = (x_1, \ldots, x_n) \geqslant 0$ the vector of their outputs, where firm i produces output x_i at a constant marginal cost $c > 0$, identical across all firms. Let $X = \sum_i x_i$ denote aggregate output—the supply of vaccines in the global market, and define $X_{-i} = X - x_i$ to be the aggregate output of all firms except firm i. The industry's fixed cost is denoted $F = \sum_{i \in \mathcal{N}} f_i$.

Profit-cap Policy The tax rate t is chosen after production decision and it is propor-

[2] We think of t abstractly as the tax rate in our analysis. It can be thought of as output tax, lowering the market price of the good to, or as input tariff, increasing the marginal cost of production. Practically and most interestingly, the tax can be thought of as requiring the firms to give away some of its output. This is a *donation requirement* and we will discuss how to implement it in the next section.

[3] We assume that there is property rights. In absence of property rights, all firms can produce as long as at least one firm chooses to innovate, assuming that there is no technological barrier once the firm has the knowledge. For simplicity, we further assume that innovation leads to a guaranteed success and that vaccines produced by different firms are identical.

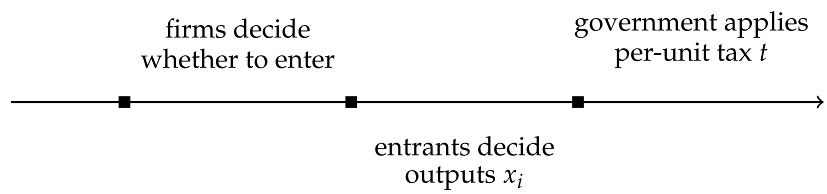


Figure 3.3: The Timing of the Model

tional to how much the industry exceeds a predetermined industry variable profit cap $\hat{\Pi}$. That is, $t = t(X, \hat{\Pi})$ is set such that given the output X, the industry variable profit never exceeds $\hat{\Pi}$. We denote firm i's variable profit with tax level t and output profile $\mathbf{x} = (x_1, \ldots, x_n)$ as

$$\pi_i(\mathbf{x}, t) = p(X)x_i - cx_i - tx_i \tag{3.3}$$

The industry variable profit depends only on the aggregate output:

$$\Pi(X, t) = \sum_{i \in \mathcal{N}} \pi_i(\mathbf{x}, t) = [p(X) - c - t]X \tag{3.4}$$

We write $\Pi(X, 0) = \Pi(X)$ for simplicity. It is important to note that the profit cap is on the industry's variable profit, $\Pi(X)$, and not on the industry's total profit $\Pi(X) - F$.

Timing. The timing of the $\hat{\Pi}$–oligopoly game, depicted in Figure 3.3, is summarized as follows:

(i) Firms in $\mathcal{M}$ choose whether or not to innovate, $b_i \in \{0, 1\}$, paying a fixed cost $f_i \geqslant 0$.

(ii) Firms in $\mathcal{N} = \{i \in \mathcal{M} : b_i = 1\}$ engage in a Cournot competition, accruing industry variable profit $\Pi(X)$.

(iii) Government levies a per-unit tax $t \geqslant 0$ such that $\Pi(X) - tX \leqslant \hat{\Pi}$.

Solution Concept. The equilibrium concept we use is that of Subgame Perfect Equilibrium. The strategy of each firm includes two decisions: whether or not to innovate and how much to produce given innovation. Firm i's strategy is a tuple (b_i, x_i), where $b_i : [0, \infty) \times [0, \infty) \to \{0, 1\}$, maps the announced profit cap and the firm's cost of innovation to an entry decision, and $x_i : 2^{\mathcal{M}} \to [0, \infty)$ maps the set of producers into a production decision. For ease of exposition, we require that $x_i(\mathcal{N}) = 0$ if $i \notin \mathcal{N}$. We say that a profile of strategies $(b_i^*, x_i^*)_{i \in \mathcal{M}}$ constitutes a Subgame Perfect Equilibrium if the profile of strategies constitutes a Nash Equilibrium of every subgames and write $X^* \equiv \sum_i x_i^*$.

Welfare. Government announces the profit cap $\hat{\Pi} \geqslant 0$ and commits to a tax scheme $t = t(\hat{\Pi}, X)$ to be imposed in the last period. For welfare comparison, define the total surplus in the usual way as

$$TS(X) = \int_0^X [p(\xi) - c] \, \mathrm{d}\xi - F = \underbrace{\int_0^X [p(\xi) - p(X)] \mathrm{d}\xi}_{\text{consumer surplus}} + \overbrace{[p(X) - c] \, X - F}^{\text{producer surplus}},$$

where X is the industry output and $F = \sum_{i \in \mathcal{N}} f_i$ is the sum of the development costs of the producers. The surplus $TS(X)$ is increasing in X. By nature of the standard Cournot competition, more firms results in higher industry output, so the first term of expression goes up. More firms, however, means higher total development costs. This is indeed what happens if $\hat{\Pi}$ is high so as to not be binding. If $\hat{\Pi}$ is low, then firms may not find it profitable to develop the product. This may lead to $\mathcal{N} = \varnothing$, which means zero output, and consequently, zero social surplus.

As will be shown, an appropriately chosen profit cap can potentially increase output beyond what would have occurred in its absence, while at the same time reducing the number of firms that entered.

Solving the game backwards, we split our analysis into two steps. First, we consider the last two stages of the game. This is an n-firm Cournot competition with profit cap. Second, we consider the entry decisions of m potential producers given the announced profit cap. Moreover, we comment on how the government should choose the variable profit cap.

Non-binding Profit Cap

Before analyzing the full game, let us first analyze the equilibrium entry decision of the firms in the standard Cournot competition, that is, without imposing a profit cap. For simplicity, assume that $m = 3$ and that the firms are indexed by non-decreasing fixed costs: $0 \leqslant f_1 \leqslant f_2 \leqslant f_3$. Generalization to general m can be done straightforwardly.

Without profit cap,[4] the game consists of two periods, corresponding to the development stage and the Cournot stage as shown in Figure 3.4.

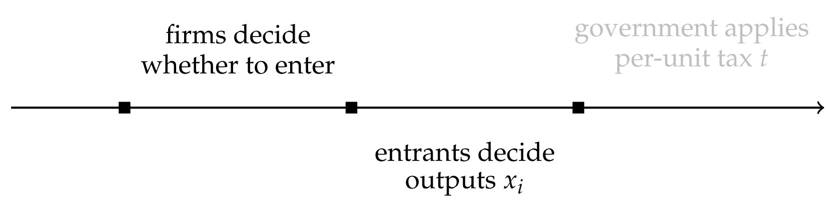


Figure 3.4: The Timing of the Game without Profit Cap

Let Π^n denote the industry profit in the standard Cournot equilibrium with n firms. With Assumptions 1-3, equilibrium is symmetric and each firm gets

$$\pi^n = \frac{\Pi^n}{n}.$$

Moreover, both Π^n and π^n are decreasing in n.

[4]This is also the case when profit cap is too high. The game and its equilibrium behave as if there is no profit cap. We go further into detail below.

We start by noting that in the Cournot stage, firms that entered compete without taking into account the fixed cost which was incurred prior to this stage. That is, once entered, they behave as in the standard Cournot with n firms. Now, in the development stage, firms play a simultaneous-move game, where each firm chooses whether or not to develop the product.[5] Hence, the subgame perfect equilibrium of this two-stage game (without profit cap) can be found by analyzing the normal form game in Table 3.1, where firm 3 chooses either the left or the right matrix, firm 2 chooses the column, and firm 1 chooses the row.

In

	In	Out
In	$\pi^3 - f_1, \pi^3 - f_2, \pi^3 - f_3$	$\pi^2 - f_1, 0, \pi^2 - f_3$
Out	$0, \pi^2 - f_2, \pi^2 - f_3$	$0, 0, \pi^1 - f_3$

Out

	In	Out
In	$\pi^2 - f_1, \pi^2 - f_2, 0$	$\pi^1 - f_1, 0, 0$
Out	$0, \pi^1 - f_2, 0$	$0, 0, 0$

Table 3.1: Normal-Form Game of Entry without Profit Cap

If all three firms choose to develop, profit to each firm from the Cournot stage would be π^3. The net profit to each firm would then be $\pi^3 - f_i$. For this to be an equilibrium, no firm would want to unilaterally deviate. This holds if $\pi^3 \geqslant f_i$ for all i. An equilibrium with two firms, say firm 1 and firm 2, developing the product exists if firms 1 and 2 do not want to unilaterally deviate, $\pi^2 \geqslant f_1$ and $\pi^2 \geqslant f_2$, and firm 3 does not want to develop, $\pi^3 < f_3$.

From this, one can see that the equilibrium entry decisions depend on the distribution of f_i's relative to π^n's. For example, one can check that if $f_1 \leqslant \pi^3 \leqslant f_2 \leqslant \pi^2 \leqslant f_3$, then $\{1, 2\}$ is a unique equilibrium. However, if $f_1 \leqslant \pi^3 \leqslant f_2 \leqslant f_3 \leqslant \pi^2$, then the equilibria are $\{1, 2\}$ and $\{1, 3\}$. Two important special cases are of interest.

[5] This is not the same game as when firms choose both to develop (pay the fixed cost) *and* the quantity of output simultaneously. In such a game, a unilateral deviation from no production to positive production does not allow incumbent firms to re-optimize their choices. Consider a two-firm equilibrium with firms 1 and 2 producing $x_1 + x_2 = X^2$ and enjoying profits $\pi^2 - f_1$ and $\pi^2 - f_2$, respectively. A unilateral deviation by an inactive firm, say firm 3, considers whether $\max_{x_3} p(X^2 + x_3)x_3 \geqslant f_3$, not whether $\pi^3 \geqslant f_3$ as in the game we are studying.

Proposition 3.1. *If $f_i = f$ for all i, then equilibrium entry is*

$$\mathcal{M}^* = \begin{cases} \{\{1,2,3\}\} & \text{if} \quad f \leqslant \pi^3 \\ \{\{1,2\},\{1,3\},\{2,3\}\} & \text{if} \quad \pi^3 \leqslant f \leqslant \pi^2 \\ \{\{1\},\{2\},\{3\}\} & \text{if} \quad \pi^2 \leqslant f \leqslant \pi^1 \\ \{\varnothing\} & \text{if} \quad \pi^1 \leqslant f \end{cases}$$

If $f_1 \leqslant f_2 \leqslant \pi^2 \leqslant f_3$, then $\mathcal{M}^* = \{\{1,2\}\}$.

Binding Profit Cap

With profit cap, we impose an assumption of symmetric equilibrium. Thus, when the profit cap is binding, each firm produces the same output in equilibrium and receive a profit of $\hat{\Pi}/n$.

First, note that the analysis without profit cap is indeed what also happens when profit cap $\hat{\Pi}$ is too high, so that it does not affect any equilibrium behavior. This is shown in Figure 3.5 as the rectangular hyperbola $\hat{\Pi} > \Pi^1$ does no intersect the demand DD. As $\hat{\Pi}$ decreases to Π^1, it touches the demand at C^1. This is the monopoly output, where marginal revenue equals marginal cost. Standard Cournot equilibrium is still not affected at this profit cap.

Now, consider a profit cap such that $\Pi^3 < \hat{\Pi} < \Pi^2$. At this point, if $n = 2$, the standard Cournot profit exceeds the profit cap, so each firm will increase production until the point marked $\hat{C}^2$. At this point, each firm gets $\hat{\Pi}/2$. If, however, three firms were to come in, then their standard Cournot equilibrium does not exceed the profit. Equilibrium

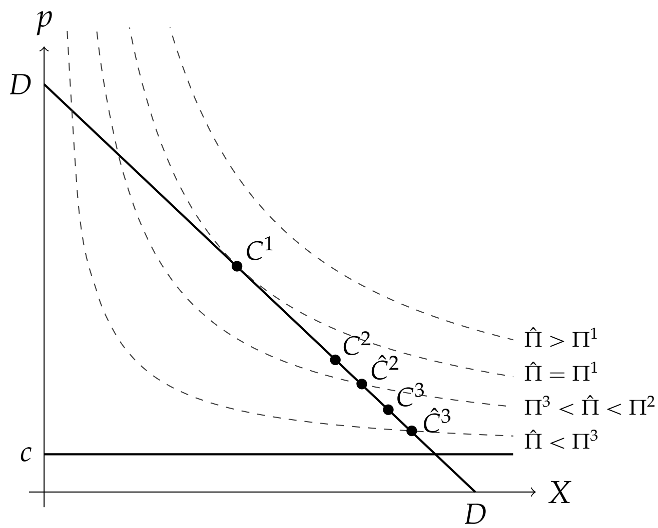


Figure 3.5: Profit Cap with Entry

will be at C^3 and each firm gets $\Pi^3/3$. Thus, in the Cournot competition with profit cap $\hat{\Pi}$, the equilibrium variable profit of each firm when there are n firms in the market is:

$$\hat{\pi}^n \equiv \frac{1}{n}\min\{\Pi^n, \hat{\Pi}\}.$$

The subgame perfect equilibrium of the game with endogenous entry and profit cap $\hat{\Pi}$ can then be found by analyzing the normal-form game in Table 3.2. This is the same game as above, but with $\hat{\pi}^n$ instead of π^n.

In

	In	Out
In	$\hat{\pi}^3 - f_1, \hat{\pi}^3 - f_2, \hat{\pi}^3 - f_3$	$\hat{\pi}^2 - f_1, 0, \hat{\pi}^2 - f_3$
Out	$0, \hat{\pi}^2 - f_2, \hat{\pi}^2 - f_3$	$0, 0, \hat{\pi}^1 - f_3$

Out

	In	Out
In	$\hat{\pi}^2 - f_1, \hat{\pi}^2 - f_2, 0$	$\hat{\pi}^1 - f_1, 0, 0$
Out	$0, \hat{\pi}^1 - f_2, 0$	$0, 0, 0$

Table 3.2: Normal-Form Game of Entry with Profit Cap

3.3.3 Optimal Profit Cap and the Duopoly Rule

The above analysis leads us to a rather interesting and attractive property of a duopoly. Suppose the government cares about efficiency, that is, it wants to maximize the total social surplus that is generated by the industry. How should government set the profit cap? Ideally, the government would want a duopoly, with the two most efficient firms (that is, with the lowest fixed costs) entering.

Why is the equilibrium $\{1,2\}$ socially desirable? To see this, note that a profit cap increases efficiency in two ways. First, if the number of firms is fixed, profit cap means lower price and higher quantity output, thus higher total surplus. Second, it lowers the number of firms in equilibrium, thereby lowering total fixed costs incurred by the industry without affecting consumer surplus and variable profit.

This holds as long as there is competition among firms, so efficiency is guaranteed to increase up to the point where only two firms choose to enter the market and the two firms that have the lowest development costs are indeed the most efficient ones. This idea is what we term the "duopoly rule".

There are however reasons why the government may not always be able to direct the industry to this outcome. Note that proposition 3.1 implies that if we can choose $\hat{\Pi}$ such that $f_1 \leqslant f_2 \leqslant \hat{\pi}^2 \leqslant f_3$, then we can ensure that the unique equilibrium is $\{1,2\}$. Can we always do this? The answer is no. First, note that since $\hat{\pi}^n \leqslant \pi^n$, the equilibrium set of entrants is always smaller with profit cap. So if $f_1 \leqslant \pi^2 \leqslant f_2 \leqslant f_3$ and only firm 1 enters the market without profit cap, then with profit cap either only firm 1 still enters the market or no one enters at all. Second, even if we can choose $\hat{\Pi}$ to ensure a 2-firm equilibrium, we may not be able to ensure the uniqueness of $\{1,2\}$. If, say $\pi^3 \leqslant f_1 < f_2 = f_3 \leqslant \pi^2$, then any $\hat{\Pi}$ either result in $\mathcal{M}^* = \{\{1\}\}$ or $\mathcal{M}^* = \{\{1,2\},\{1,3\},\{2,3\}\}$.

The duopoly rule is however interesting and worth keeping in mind for future work on designing more sophisticated government interventions.

3.4 Two Applications of the Theory

The theory developed has policy significance. It can address some of the current discussions on vaccination distribution.

Patent Waiver. There has been a lot of discussion on waiving intellectual property protection on the COVID-19 vaccines. Originally proposed by countries such as India and South Africa, the World Health Organization (WHO) as well as the Biden administration have supported the waiver, while the European Union has been against it. In an official statement, the US Trade Representative Katherine Tai said "the Administration believes strongly in intellectual property protections, but in service of ending this pandemic, supports the waiver of those protections for COVID-19 vaccines." The European Commission, on the other hand, has proposed alternative policy[6] focusing on minimizing trade restrictions and granting manufacturing licenses.

Opponents of the patent waiver argue that it is "the wrong tool for the right goal"[7] because the problems we face are shortages in vaccines manufacturing, which is not solved by lifting the patent. Many argue that upholding patent right should in fact be

[6] www.reuters.com/world/europe/eu-executive-submits-vaccine-access-proposal-wto-2021-06-04 (accessed June 2021).

[7] https://blog.petrieflom.law.harvard.edu/2021/05/05/covid-vaccine-patent-waiver (accessed June 2021).

a priority[8]: "the proposal undermines the very system that produced the life-saving science in the first place. And it destroys the incentive for companies to take risks to find solutions for the next health emergency." Waiver the intellectual property right would be "a bad precedent that would do no immediate good and substantial long-term harm."[9]

In terms of the current model, if the patent is waived, there will be free entry and, over time, the market structure would move towards being perfectly competitive. In such a competitive market with free entry, the price is equal to the marginal cost of production and in equilibrium all firms earn zero profit. However, foreseeing such an outcome would deter firms in the future to incur the fixed cost of innovation. Our policy proposal of profit cap either via output tax or donation requirement discussed below can be viewed as an alternative to patent waiver. It can be fine-tuned to retain the incentive for innovation.

Donation. In light of the recent announcement[10] that the US would donate 500 million doses of the Pfizer-BioNTech vaccines over the next year, it is interesting to see how the profit cap scheme described above can be implemented via donation requirement instead of imposing a per unit tax on vaccines.

Note that instead of the government purchasing the vaccines at cost from the firms and then donating the vaccines as a "humanitarian obligation," the government can require an industry-level donation of αX, where $\alpha \geqslant 0$ works to cap industry profit the same way as a unit tax does. Correspondingly, firm i then would donate αx_i and its total

[8]https://www.economist.com/by-invitation/2021/04/20/michelle-mcmurry-heath-on-maintaining-intellectual-property-amid-covid-19 (accessed June 2021)

[9]https://www.wsj.com/articles/patent-busting-wont-help-vaccinate-the-world-faster-11620591133 (accessed May 2021).

[10]https://www.nytimes.com/2021/06/10/world/europe/biden-vaccine-500-million.html (accessed June 2021).

production is $(1+\alpha)x_i$. Firm i's profit is then given by

$$\pi_i(\mathbf{x}, \alpha) = p(X)x_i - c(1+\alpha)x_i,$$

while the industry profit is

$$\Pi(X, \alpha) = \sum_i \pi_i(\mathbf{x}, \alpha) = (p(X) - c(1+\alpha))X.$$

Theorem 3.2 goes through in this implementation by defining the donation requirement as

$$\alpha(X, \hat{\Pi}) = t(X, \hat{\Pi})/c.$$

It is interesting to note that the donation requirement results in no donation in equilibrium. Yet, the requirement constraint is binding and it would work to broaden supply of COVID-19 vaccines by driving up the production via competition, and driving down the market price. In the limit, the donation requirement pushes the industry to produce at the competitive level and lowering the market price.

3.4.1 E-commerce and Platforms' Antitrust Problem

Suppose that there are n platforms for the product in question. The stage game of this model is the simultaneous quote of a price to the buyers p^b and a price to the sellers p^s by each platform. Let platform i's pair of price be denoted as (p_i^b, p_i^s). The best price to the buyers is the lowest; that is, it is the quote of the dealer who is asking for the smallest amount of money in order to sell the product to a customer. Let $\underline{p}^b = \min_i p_i^b$. Similarly, let $\overline{p}^s = \max_i p_i^s$, the highest that any platform is willing to pay for a product. The price spread or the markup is therefore is $\underline{p}^b - \overline{p}^s$.

In the current period, when these quotes are binding, all buy orders are executed at a price equal to a. The profit to a platform from participating in this transaction depends on the markup and the volume of sales. The profit that each dealer makes depends on the fraction of the aggregate order flow that he receives. We assume the platforms hold no inventory and cannot oversell. Each platform's profits of the stage game is $(\underline{p}^b - \overline{p}^s)\frac{X(\underline{p}^b)}{n}$. If the platform does not post the best price, it makes zero profit. In case there is a tie, every platform with $\underline{p}^b$ gets an equal fraction of the sales. We maintain the assumption of constant marginal cost, so the supply curve is perfectly elastic, which makes $\overline{p}^s = c$.

Let us now turn to Nash equilibrium. It is straightforward to see that $p_i^b = c$ for all i is a Nash equilibrium of the stage game. If every other platform is posting a competitive price, a single platform has no profitable deviation. Every platform makes zero profits in this equilibrium; consequently, this will constitute the benchmark punishment regime in a repeated game.

Let us look at purely collusive pricing in the stage game. If the platforms set prices in order to maximize their collective profits, they would set the price to solve the following maximization problem: $(\underline{p}^b - \overline{p}^s)X(\underline{p}^b)$. The profit-maximizing collusive price is $p^m = c - \frac{X'(p^m)}{X(p^m)}$. Note that the collusive markup is therefore $-\frac{X'(p^m)}{X(p^m)}$, whereas the competitive markup is 0. The quantity traded at the collusive price is $X(p^m) < X(c)$, leading to the classical monopoly inefficiency of underproduction.

The collusive price can be sustained in a subgame perfect Nash equilibrium by several strategies. One strategy is the grim trigger strategy: each platform begins with the collusive price and maintains them as long as the others have done so in the past. If any platform undercuts, from that period onward all platforms go to pricing at $\underline{p}^b = c$ forever thereafter. Clearly, once the punishment regime has been initiated, no platform can profitably deviate by pricing at any price other than $p_i^b = c$. If the platform never deviates, it expects a continued payoff of $\frac{(p^m-c)X(p^m)}{n(1-\delta)}$, where δ is the discount factor. By undercutting the other platforms by ε, a single platform makes a profit of $\frac{(p^m-c)X(p^m)}{n(1-\delta)}$ in the current period but earns zero profit thereafter. Sustaining collusion is the better option if $\frac{(p^m-c)X(p^m)}{n(1-\delta)} \geqslant (p^m - c)X(p^m)$. Equivalently, we have $n(1-\delta) \leqslant 1$. Either a lower number of platforms or a higher discount factor will make the collusion easier. There are other more-benign punishment might also possibly sustain collusive pricing. One punishment, in the event of a departure from collusive pricing by some platform, is to price forever thereafter at the alternative stage game Nash equilibrium. Another punishment is a forgiving trigger in which markups narrow for some number of rounds to punish deviation but then go back to the collusive pricing.

Now consider the effect of a profit cap $\hat{\Pi}$. The profit cap $\hat{\Pi}$ is enforced by t as follows:

$$t = \begin{cases} \tau(\mathbf{p}^b, \hat{\Pi}) & \text{if } \Pi(\mathbf{p}^b, 0) > \hat{\Pi}, \\ 0 & \text{if } \Pi(\mathbf{p}^b, 0) \leqslant \hat{\Pi}. \end{cases} \tag{3.5}$$

where the present discounted profits of the cartel is

$$\Pi(\mathbf{p}^b, \tau) = \sum_{t=0}^{\infty} \delta^t (p_t^b - c - \tau) X(p_t^b). \tag{3.6}$$

The differences from the main model are (i) the cartel chooses price instead of quantity and (ii) the cartel makes decision repeatedly. Theorem 3.2 applies by redefining the control variable to be $X(p_t^b)$ and viewing every period as a separate firm. In equilibrium no tax liability is triggered, the platform cartel preserves the economies of scale, and the binding profit cap increases trades between more sellers and buyers.

3.A Appendix A. Proofs

Proof of Lemma 3.1. The marginal revenue $MR(X)$ is strictly decreasing for $0 \leqslant X < Z$ because

$$MR'(X) = p'(X) + p''(X)X + p'(X) < 0$$

by Assumptions 2 and 3. Now, $MR(0) > c$ by Assumption 1. Note that $X^{PC} < Z$ because the inverse demand is strictly decreasing and $c > 0$. At X^{PC}, $MR(X^{PC}) = p(X^{PC}) + p'(X^{PC})X^{PC} < c$. Thus, we have that $MR(X^{PC}) < c < MR(0)$, so by the Intermediate Value Theorem and the fact that $MR(X)$ is strictly decreasing on such domain, there exists a unique $X^M < Z$ such that $MR(X^M) = c$. □

Proof of Lemma 3.2. We write $\hat{p}(X)$ for $\hat{p}(X, \hat{\Pi})$ when $\hat{\Pi}$ is fixed. The proof proceeds in three steps. First, we claim a relationship between the nature if the solution and the marginal revenue at such solution. Second, we show that if the solution is unique then it X^M. Third, we show that if the solution is not unique there can only be two solutions.

Step 1. Consider a solution X to $p(X) = \hat{p}(X)$. There are three possibilities at X: (i) q crosses p from above, (ii) q crosses p from below, and (iii) q is tangent to p. We claim that (i) holds if and only if $MR(X) > c$. If q crosses p from above, then because both derivatives are negative, $p'(X) > q'(X)$. Note that $q'(X) = -\frac{\hat{p}(X)-c}{X}$, so $p'(X) > -\frac{\hat{p}(X)-c}{X}$. Substituting $p(X) = \hat{p}(X)$ and rearranging yields $MR(X) > c$. Analogously, at any solution, (ii) holds if and only if $MR(X) < c$ and (iii) holds if and only if $MR(X) = c$.

Step 2. If the solution is unique then only (iii) can hold at such solution and by Lemma 1 it is at X^M. Suppose not and (i) holds, thus q crosses p from above at such solution, then because $\lim_{X\to\infty} \hat{p}(X) = c$ and p eventually crosses c there must exist at least another solution, a contradiction. A similar contradiction occurs if (ii) holds because $p(X)$ is bounded but q grows unboundedly as X approach 0.

Step 3. Suppose the solution to $p(X) = \hat{p}(X)$ exists and is not unique. Let $\underline{X}$ be the smallest solution. If $MR(\underline{X}) < c$, then q crosses p from below at $\underline{X}$. Because MR, p, and q are all decreasing, there can be no larger crossings, a contradiction. If $MR(\underline{X}) = c$, then the two curves are tangent at $\underline{X}$. For any crossing at a larger X, q must cross p from above, which cannot happen since at such point $MR(X)$ would have been below c. Thus, $MR(\underline{X}) > c$ and q crosses p from above at $\underline{X}$. A similar argument shows that $MR(\overline{X}) < c$, where $\overline{X}$ is the largest solution, and thus q crosses p from below at $\overline{X}$. Now if there are more solutions in between $\underline{X}$ and $\overline{X}$, then it would mean that MR crosses c more than once, which contradicts the assumptions on the inverse demand function. □

Proof of Lemma 3.3. If $\hat{\Pi} > (p(X^M) - c)X^M$, then $\hat{\Pi} > (p(X) - c)X$ for all X. That is, the profit cap is too large and $p(X)$ and $\hat{p}(X, \hat{\Pi})$ do not intersect. We show this by contradiction. Suppose there exists an X such that $\hat{\Pi} \leqslant (p(X) - c)X$. If such X is unique then it must be that $\hat{\Pi} = (p(X) - c)X$ and so at this unique X, $p(X) = \hat{p}(X)$. From Lemma 2, X must be X^M, a contradiction. If such X is not unique, then define the smallest and largest such X as $\underline{X}$ and $\overline{X}$, respectively. It must be the case that $MR(\underline{X}) > c$ and $MR(\overline{X}) < c$. By Intermediate Value Theorem, there is an X between $\underline{X}$ and $\overline{X}$ such that $MR(X) = c$, but that X must be X^M, a contradiction.

If $\hat{\Pi} = (p(X^M) - c)X^M$, then $\hat{p}(X^M) = p(X^M)$. By Lemma 1 and Lemma 2, $p(X)$ and $\hat{p}(X)$ must be tangent at X^M and thus is the only solution. Lastly, as $\hat{\Pi} \to 0$, either $p(X) - c \to 0$ or $X \to 0$, or both. The smallest solution thus goes to zero, and the largest goes to X^{PC}. □

Proof of Theorem 3.2. First, we characterize the reaction functions $\hat{R}_i(\mathbf{x}_{-i})$. We suppress dependence on $\hat{\Pi}$ in some of our notation. Suppose $p(X) = \hat{p}(X)$ has no solution, then it must be because $\hat{p}(X) > p(X)$ for all X. Hence $t(X, \hat{\Pi}) = 0$ for all X and $\hat{R}_i(\mathbf{x}_{-i}) = R_i(\mathbf{x}_{-i})$, where $R_i(\mathbf{x}_{-i})$ is the reaction function of firm i in the standard Cournot oligopoly. From now, suppose one or two solutions to $p(X) = \hat{p}(X)$ exist.

Given $\mathbf{x}_{-i}$ and hence X_{-i}, the domain of optimization can be split up into two sections. Either

$$x_i \in A \equiv [0, \underline{X} - X_{-i}) \cup (\overline{X} - X_{-i}, \infty)$$

or

$$x_i \in B \equiv [\underline{X} - X_{-i}, \overline{X} - X_{-i}]$$

In A, $X < \underline{X}$ or $X > \overline{X}$, which means that $t(X, \hat{\Pi}) = 0$ and we are back to standard Cournot, so $\hat{R}_i(\mathbf{x}_{-i}) = R_i(\mathbf{x}_{-i})$ for such X. In B, $t(X, \hat{\Pi}) = p(X) - \hat{p}(X, \hat{\Pi})$, so for such X, the objective function becomes

$$[p(x_i + X_{-i}) - c - t(x_i + X_{-i}, \hat{\Pi})]x_i = \frac{\hat{\Pi}}{X} x_i,$$

which means that in this range firms essentially engage in a contest for a fixed prize of $\hat{\Pi}$ with Tullock winning probability x_i/X, but without incurring any costs! This leads to firm i choosing x_i as high as possible as a best-response to any X_{-i}, and so $\hat{R}_i(\mathbf{x}_{-i}) = \overline{X} - X_{-i}$.

Thus, for a given $\hat{\Pi}$, the reaction function of firm i is

$$\hat{R}_i(\mathbf{x}_{-i}) = \begin{cases} R_i(\mathbf{x}_{-i}) & \text{if} \quad R_i(\mathbf{x}_{-i}) + X_{-i} \notin [\underline{X}, \overline{X}] \\ \overline{X} - X_{-i} & \text{if} \quad R_i(\mathbf{x}_{-i}) + X_{-i} \in [\underline{X}, \overline{X}] \end{cases}$$

The equilibrium then the profile $\mathbf{x}^*$ such that for all i, $x_i^* = \hat{R}_i(\mathbf{x}_{-i}^*)$.

If $X^{NE} \notin [\underline{X}, \overline{X}]$, then $X^* = X^{NE}$, simply the standard Cournot equilibrium. We can show that the only possibility is $X^{NE} > \overline{X}$. In this case, $p(X^{NE}) < \hat{p}(X^{NE}, \hat{\Pi})$, so $t(X^{NE}, \hat{\Pi}) = 0$. If not, then $X^* = \overline{X}$ and again $t(X^*, \hat{\Pi}) = 0$. □

BIBLIOGRAPHY

Andreoni, James, "Philanthropy," *Handbook of the economics of giving, altruism and reciprocity*, 2006, *2*, 1201–1269.

Armstrong, Mark and David EM Sappington, "Recent developments in the theory of regulation," *Handbook of industrial organization*, 2007, *3*, 1557–1700.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 06 1962, *29* (3), 155–173.

Arrow, KJ, "Economic welfare and the allocation of resource for inventions, in the rate and direction of inventive activity: economic and social factors," *N. Bureau*, 1962.

Athey, Susan and Glenn Ellison, "Dynamics of open source movements," *Journal of Economics & Management Strategy*, 2014, *23* (2), 294–316.

Balzer, Benjamin and Johannes Schneider, "Managing A Conflict: Optimal Alternative Dispute Resolution," 2020.

Basu, Kaushik and Joseph E Stiglitz, "Sovereign debt and joint liability: an economic theory model for amending the Treaty of Lisbon," *The Economic Journal*, 2015, *125* (586), F115–F130.

Belleflamme, Paul and Martin Peitz, "Digital Piracy: Theory," *The Oxford Handbook of the Digital Economy*, 2012.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L Lipman, "Optimal allocation with costly verification," *American Economic Review*, 2014, *104* (12), 3779–3813.

_, _, and _, "Mechanisms with evidence: Commitment and robustness," *Econometrica*, 2019, *87* (2), 529–566.

Bénabou, Roland and Jean Tirole, "Incentives and prosocial behavior," *American economic review*, 2006, *96* (5), 1652–1678.

Benkler, Yochai, *The wealth of networks*, Yale university press, 2008.

Bergemann, Dirk and Stephen Morris, "Information design, Bayesian persuasion, and Bayes correlated equilibrium," *American Economic Review*, 2016, *106* (5), 586–91.

Besley, Timothy and Stephen Coate, "Group lending, repayment incentives and social collateral," *Journal of Development Economics*, February 1995, *46* (1), 1–18.

Blume, Andreas, Oliver J Board, and Kohei Kawamura, "Noisy talk," *Theoretical Economics*, 2007, *2* (4), 395–440.

Bohnet, Iris, Bruno S Frey, and Steffen Huck, "More order with less law: On contract enforcement, trust, and crowding," *American Political Science Review*, 2001, *95* (1), 131–144.

Boldrin, Michele and David Levine, "The case against intellectual property," *American Economic Review*, 2002, *92* (2), 209–212.

Border, Kim C and Joel Sobel, "Samurai accountant: A theory of auditing and plunder," *The Review of economic studies*, 1987, *54* (4), 525–540.

Brown, Jennifer Gerarda and Ian Ayres, "Economic rationales for mediation," *Virginia Law Review*, 1994, pp. 323–402.

Bull, Jesse and Joel Watson, "Hard evidence and mechanism design," *Games and Economic Behavior*, 2007, *58* (1), 75–93.

Burdett, Kenneth, "Truncated means and variances," *Economics Letters*, 1996, *52* (3), 263–267.

Casadesus-Masanell, Ramon and Pankaj Ghemawat, "Dynamic mixed duopoly: A model motivated by Linux vs. Windows," *Management Science*, 2006, *52* (7), 1072–1084.

Charness, Gary and Matthew Rabin, "Understanding social preferences with simple tests," *The quarterly journal of economics*, 2002, *117* (3), 817–869.

Coase, R. H., "The Problem of Social Cost," *The Journal of Law Economics*, 1960, *3*, 1–44.

Cooper, David J and John H Kagel, "Other-Regarding Preferences," in "The Handbook of Experimental Economics, Volume 2," Princeton University Press, 2016, pp. 217–289.

Crawford, Vincent P and Joel Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1431–1451.

Demsetz, Harold, "Toward a Theory of Property Rights," *American Economic Review*, 1967, *57* (2), 347–359.

_ , "The private production of public goods," *The Journal of Law and Economics*, 1970, *13* (2), 293–306.

DeVault, Ileen A, Maria Figueroa, Fred B Kotler, Michael Maffie, and John Wu, "On-demand platform workers in New York State: The challenges for public policy," 2019.

Diamond, Peter, "Optimal tax treatment of private contributions for public goods with and without warm glow preferences," *Journal of Public Economics*, 2006, *90* (4-5), 897–919.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Courts," *The Quarterly Journal of Economics*, 2003, *118* (2), 453–517.

Dye, Ronald A, "Disclosure of nonproprietary information," *Journal of accounting research*, 1985, pp. 123–145.

Earle, Robert, Karl Schmedders, and Tymon Tatur, "On Price Caps Under Uncertainty," *The Review of Economic Studies*, January 2007, *74* (1), 93–111.

Ellickson, Robert C, *Order without law*, Harvard University Press, 1991.

Ellingsen, Tore and Magnus Johannesson, "Pride and prejudice: The human side of incentive theory," *American economic review*, 2008, *98* (3), 990–1008.

Erickson, Alizila, "How Taobao Is Crowdsourcing Justice in Online Shopping Disputes," 2014.

Fanning, Jack, "Mediation in reputational bargaining," *American Economic Review*, 2021, *111* (8), 2444–72.

Farrell, Joseph and Garth Saloner, "Standardization, compatibility, and innovation," *The RAND Journal of Economics*, 1985, pp. 70–83.

Fey, Mark and Kristopher W Ramsay, "When is shuttle diplomacy worth the commute? Information sharing through mediation," *World Politics*, 2010, *62* (4), 529–560.

Gaudet, Gérard and Stephen W. Salant, "Uniqueness of Cournot Equilibrium: New Results from Old Methods," *The Review of Economic Studies*, 1991, *58* (2), 399–404.

Gentzkow, Matthew and Emir Kamenica, "Costly persuasion," *American Economic Review*, 2014, *104* (5), 457–62.

Ghatak, Maitreesh and Timothy W. Guinnane, "The economics of lending with joint liability: theory and practice," *Journal of Development Economics*, October 1999, *60* (1), 195–228.

Glazer, Jacob and Ariel Rubinstein, "On optimal rules of persuasion," *Econometrica*, 2004, *72* (6), 1715–1736.

Gneezy, Uri and Aldo Rustichini, "A fine is a price," *The Journal of Legal Studies*, 2000, *29* (1), 1–17.

Goldfarb, Avi and Catherine Tucker, "Digital economics," *Journal of Economic Literature*, 2019, *57* (1), 3–43.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani, "Mediation, arbitration and negotiation," *Journal of Economic Theory*, 2009, *144* (4), 1397–1420.

Gonzalez-Barahona, Jesus M, Gregorio Robles, Roberto Andradas-Izquierdo, and Rishab Aiyer Ghosh, "Geographic origin of libre software developers," *Information Economics and Policy*, 2008, *20* (4), 356–363.

Gordon, H Scott et al., "The Economic Theory of a Common-Property Resource: The Fishery," *Journal of Political Economy*, 1954, *62*, 124–124.

Gould, David M, William C Gruben et al., "The role of intellectual property rights in economic growth," *Journal of Development Economics*, 1996, *48* (2), 323–350.

Gousios, Georgios, "The GHTorrent dataset and tool suite," in "Proceedings of the 10th Working Conference on Mining Software Repositories" MSR '13 IEEE Press Piscataway, NJ, USA 2013, pp. 233–236.

Grimm, Veronika and Gregor Zöttl, "Price regulation under demand uncertainty," *The BE Journal of Theoretical Economics*, 2010, *10* (1).

Grossman, Sanford J, "The informational role of warranties and private disclosure about product quality," *The Journal of Law and Economics*, 1981, *24* (3), 461–483.

_ **and Oliver D Hart**, "Disclosure laws and takeover bids," *The Journal of Finance*, 1980, *35* (2), 323–334.

Han, Lanshan and Andrew L. Liu, "On Nash–Cournot games with price caps," *Operations Research Letters*, January 2013, *41* (1), 92–97.

Hart, Sergiu, Ilan Kremer, and Motty Perry, "Evidence games: Truth and commitment," *American Economic Review*, 2017, *107* (3), 690–713.

Hörner, Johannes, Massimo Morelli, and Francesco Squintani, "Mediation and peace," *The Review of Economic Studies*, 2015, *82* (4), 1483–1501.

Johnson, Justin Pappas, "Open source software: Private provision of a public good," *Journal of Economics & Management Strategy*, 2002, *11* (4), 637–662.

Johnson, William R, "The economics of copying," *Journal of Political Economy*, 1985, *93* (1), 158–174.

Jovanovic, Boyan, "Truthful disclosure of information," *The Bell Journal of Economics*, 1982, pp. 36–44.

Kalliamvakou, Eirini, Georgios Gousios, Kelly Blincoe, Leif Singer, Daniel M German, and Daniela Damian, "The promises and perils of mining github," in "Proceedings of the 11th working conference on mining software repositories" 2014, pp. 92–101.

Kamenica, Emir and Matthew Gentzkow, "Bayesian persuasion," *American Economic Review*, 2011, *101* (6), 2590–2615.

Katz, Michael L and Carl Shapiro, "Network externalities, competition, and compatibility," *The American economic review*, 1985, *75* (3), 424–440.

_ **and Harvey S Rosen**, "Tax analysis in an oligopoly model," *Public Finance Quarterly*, 1985, *13* (1), 3–20.

Klein, Benjamin, Andres V Lerner, and Kevin M Murphy, "The economics of copyright" fair use" in a networked world," *American Economic Review*, 2002, *92* (2), 205–208.

Klement, Alon and Zvika Neeman, "Against compromise: A mechanism design approach," *Journal of Law, Economics, and Organization*, 2005, *21* (2), 285–314.

Klerman, Daniel and Lisa Klerman, "Inside the Caucus: An Empirical Analysis of Mediation from Within," *Journal of Empirical Legal Studies*, 2015, *12* (4), 686–715.

Lerner, Josh and Jean Tirole, "Some simple economics of open source," *The journal of industrial economics*, 2002, *50* (2), 197–234.

_ and _, "The economics of technology sharing: Open source and beyond," *Journal of Economic Perspectives*, 2005, *19* (2), 99–120.

_ and _, "The scope of open source licensing," *Journal of Law, Economics, and Organization*, 2005, *21* (1), 20–56.

_, Parag A Pathak, and Jean Tirole, "The dynamics of open-source contributors," *American Economic Review*, 2006, *96* (2), 114–118.

Lessig, Lawrence, "Free culture: how big media uses technology and the law to lock down culture and control creativity," 2004.

_, *Code: And other laws of cyberspace*, ReadHowYouWant. com, 2009.

Levin, Dan, "Taxation within Cournot oligopoly," *Journal of Public Economics*, August 1985, *27* (3), 281–290.

Liebowitz, Stan J, "File sharing: creative destruction or just plain destruction?," *The Journal of Law and Economics*, 2006, *49* (1), 1–28.

Marron, Donald B and David G Steel, "Which countries protect intellectual property? The case of software piracy," *Economic inquiry*, 2000, *38* (2), 159–174.

McDermott, E Patrick and Ruth Obar, "What's going on in mediation: An empirical analysis of the influence of a mediator's style on party satisfaction and monetary benefit," *Harv. Negot. L. Rev.*, 2004, *9*, 75.

Merrill, Thomas W, "Introduction: the Demsetz thesis and the evolution of property rights," *The Journal of Legal Studies*, 2002, *31* (S2), S331–S338.

Meurer, Michael J, "The settlement of patent litigation," *The RAND Journal of Economics*, 1989, pp. 77–91.

Miceli, Thomas J and Richard P Adelstein, "An economic model of fair use," *Information Economics and Policy*, 2006, *18* (4), 359–373.

Milgrom, Paul and John Roberts, "Relying on the information of interested parties," *The RAND Journal of Economics*, 1986, pp. 18–32.

Milgrom, Paul R, "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 1981, pp. 380–391.

Mookherjee, Dilip and Ivan Png, "Optimal auditing, insurance, and redistribution," *The Quarterly Journal of Economics*, 1989, *104* (2), 399–415.

Mustonen, Mikko, "Copyleft—the economics of Linux and other open source software," *Information Economics and Policy*, 2003, *15* (1), 99–121.

Myerson, R, "Game Theory: Analysis of Conflict Harvard Univ," *Press, Cambridge*, 1991.

Myerson, Roger B, "Optimal auction design," *Mathematics of operations research*, 1981, *6* (1), 58–73.

Myles, Gareth D, "Tax design in the presence of imperfect competition: An example," *Journal of Public Economics*, 1987, *34* (3), 367–378.

Nagaraj, Abhishek, "Does copyright affect reuse? Evidence from Google Books and Wikipedia," *Management Science*, 2018, *64* (7), 3091–3107.

Nalebuff, Barry, "Credible pretrial negotiation," *The RAND Journal of Economics*, 1987, pp. 198–210.

Nolan-Haley, Jacqueline, "Mediation: The new arbitration," *Harv. Negot. L. Rev.*, 2012, *17*, 61.

Novos, Ian E and Michael Waldman, "The effects of increased copyright protection: An analytic approach," *Journal of Political Economy*, 1984, *92* (2), 236–246.

Novshek, William, "On the Existence of Cournot Equilibrium," *The Review of Economic Studies*, January 1985, *52* (1), 85–98.

Oberholzer-Gee, Felix and Koleman Strumpf, "The effect of file sharing on record sales: An empirical analysis," *Journal of Political Economy*, 2007, *115* (1), 1–42.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura, "Strategic information revelation," *The Review of Economic Studies*, 1990, *57* (1), 25–47.

Ostrom, Elinor, *Governing the commons: The evolution of institutions for collective action*, Cambridge university press, 1990.

_ , "Beyond markets and states: polycentric governance of complex economic systems," *American Economic Review*, 2010, *100* (3), 641–72.

Peitz, Martin and Patrick Waelbroeck, "Piracy of digital products: A critical review of the theoretical literature," *Information Economics and Policy*, 2006, *18* (4), 449–476.

Penney, Jonathon W, "Privacy and Legal Automation: The DMCA as a Case Study," *Stan. Tech. L. Rev.*, 2019, *22*, 412.

Posner, Richard A, "Intellectual property: The law and economics approach," *Journal of Economic Perspectives*, 2005, *19* (2), 57–73.

Rabinovich-Einy, Orna and Ethan Katsh, "Technology and the future of dispute systems design," *Harv. Negot. L. Rev.*, 2012, *17*, 151.

Raymond, Anjanette H and Abbey Stemler, "Trusting strangers: Dispute resolution in the crowd," *Cardozo J. Conflict Resol.*, 2014, *16*, 357.

Raymond, Eric S. and Bob Young, *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*, USA: O'Reilly amp; Associates, Inc., 2001.

Reynolds, Stanley S. and David Rietzke, "Price caps, oligopoly, and entry," *Economic Theory*, October 2018, *66* (3), 707–745.

Rob, Rafael and Joel Waldfogel, "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*, 2006, *49* (1), 29–62.

Roberts, Kenneth M, "Mediating the evaluative-facilitative debate: Why both parties are wrong and a proposal for settlement," *Loy. U. Chi. LJ*, 2007, *39*, 187.

Romer, Paul M, "Two strategies for economic development: using ideas and producing ideas," *The World Bank Economic Review*, 1992, *6* (suppl_1), 63–91.

Rosen, J. B., "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games," *Econometrica*, 1965, *33* (3), 520–534.

Seng, Daniel, "Copyrighting Copywrongs: An Empirical Analysis of Errors with Automated DMCA Takedown Notices," *Santa Clara High Tech. LJ*, 2020, *37*, 119.

Shin, Hyun Song, "The burden of proof in a game of persuasion," *Journal of Economic Theory*, 1994, *64* (1), 253–264.

Spier, Kathryn E, "Pretrial bargaining and the design of fee-shifting rules," *The RAND Journal of Economics*, 1994, pp. 197–214.

Stern, Nicholas, "The effects of taxation, price control and government contracts in oligopoly and monopolistic competition," *Journal of Public Economics*, 1987, *32* (2), 133–158.

Stiglitz, Joseph E et al., "Knowledge as a global public good," *Global public goods: International cooperation in the 21st century*, 1999, *308*, 308–325.

Suber, Peter, "A very brief introduction to open access," 2010.

Szidarovszky, F. and S. Yakowitz, "A New Proof of the Existence and Uniqueness of the Cournot Equilibrium," *International Economic Review*, 1977, *18* (3), 787–789.

Takhteyev, Yuri and Andrew Hilts, "Investigating the Geography of Open Source Software through Github."

Torvalds, Linus and David Diamond, *Just for fun: The story of an accidental revolutionary*, Harper Audio, 2001.

Townsend, Robert M, "Optimal contracts and competitive markets with costly state verification," *Journal of Economic theory*, 1979, *21* (2), 265–293.

Urban, Jennifer M and Laura Quilter, "Efficient process or chilling effects-takedown notices under Section 512 of the Digital Millennium Copyright Act," *Santa Clara Computer & High Tech. LJ*, 2005, *22*, 621.

_ , Joe Karaganis, and Brianna Schofield, "Notice and takedown in everyday practice," *UC Berkeley Public Law Research Paper*, 2017, (2755628).

v Engelhardt, Sebastian and Andreas Freytag, "Institutions, culture, and open source," *Journal of Economic Behavior & Organization*, 2013, *95*, 90–110.

Varian, Hal R, "Copying and copyright," *Journal of economic perspectives*, 2005, *19* (2), 121–138.

Vives, Xavier, *Oligopoly pricing: old ideas and new tools*, MIT press, 1999.

Waldfogel, Joel, "Copyright protection, technological change, and the quality of new products: Evidence from recorded music since Napster," *The journal of law and economics*, 2012, *55* (4), 715–740.

_ , "Copyright research in the digital age: Moving from piracy to the supply of new products," *American Economic Review*, 2012, *102* (3), 337–42.

_ , "How digitization has created a golden age of music, movies, books, and television," *Journal of economic perspectives*, 2017, *31* (3), 195–214.

_ **and Imke Reimers**, "Storming the gatekeepers: Digital disintermediation in the market for books," *Information economics and policy*, 2015, *31*, 47–58.

Weesie, Jeroen, "Incomplete information and timing in the volunteer's dilemma: A comparison of four models," *Journal of conflict resolution*, 1994, *38* (3), 557–585.

World Bank, "World development report: Knowledge for development," 1998.

Zhang, Xiaoquan Michael and Feng Zhu, "Group size and incentives to contribute: A natural experiment at Chinese Wikipedia," *American Economic Review*, 2011, *101* (4), 1601–15.

Zittrain, Jonathan, *The future of the internet–and how to stop it*, Yale University Press, 2008.