Essays on Regulation of Platform Markets

Gaurav Jakhu

A dissertation submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

(Economics)

at the

Economics and Planning Unit
Indian Statistical Institute
New Delhi

2019

Supervisor : Prof. Prabal Roy Chowdhury

Contents

List of Figures

Abstract

The digital transformation and rise of online platforms have improved consumer welfare and business opportunities through better consumer choice and increased efficiency in trade. In this light, the authorities must weigh the concerns regarding the concentration of market power in the hands of a few large firms against the substantial benefits that they offer. In particular, user data is seen as a competitive resource in digital markets. In addition to better matches and information learning, activities of online firms in data collection and exploitation have raised concerns about erosion in user privacy and abuse of market power in digital markets. This raises the question of whether an intervention to reduce the data advantage of firms will improve or reduce social gains. In this light, this thesis analyzes the competitive and welfare effects of a specific form of data advantage arising from sharing consumer data among affiliated and unaffiliated firms. In addition, it also examines the competitive framework that should be part of antitrust intervention in such markets.

In doing so, Chapter 2 analyses the strategic and welfare impact of voluntary data sharing in platform markets. It is shown that, under data sharing, the upstream firm can invest higher in data collection, especially in markets with lower improvement in its advertising targeting rates. However, social welfare rises. Moreover, the exclusive technology sharing regime paradoxically improves the welfare of all users. Chapter 3 examines the competitive and welfare implications of alternative regulatory approaches to protect privacy namely i) restricting access to data owned by the firm's subsidiaries, and ii) empowering users to control data collection activities. It is shown that the former is always welfare reducing in the absence of any change in privacy level. Whereas, the latter can enhance user and social welfare in markets with large advertising targeting rates. Chapter 4 examines the private and social incentives to bundle when one firm can collect data regarding users from another market. It is shown that bundling is not profitable when investment in data collection and adver-

tising targeting rate are small. Moreover, user welfare and social welfare can move in opposite direction when both data collection and targeting rate are large, leading to a policy dilemma. Chapter 5 is a policy paper that discusses the nature of antitrust enforcement required in platform markets. It argues that the authorities should not pursue an ex-ante agenda and exclusively target objectionable activities that hurt consumers (not protecting some competitors) leaving other pro-competitive conducts that benefit consumers unregulated.

As a result, the theoretical models developed in the thesis contribute to the literature on platform markets by analyzing data sharing among firms. In the literature, so far, the current focus is on understanding the data collection on a platform. However, in our models, one firm can benefit from the functioning of another firm in a different market through data sharing. In addition, we consider privacy implications when data is shared by highlighting the mechanism that affects the privacy choice of a firm in an interconnected system. On the policy front, the thesis contributes to the ongoing debate on protecting user privacy and establishing a fair and competitive environment in digital markets.

Acknowledgements

My first and foremost thanks goes to my advisor Prabal Roy Chowdhury for guiding me throughout the Ph.D. journey. His continuous support and encouragement to the new ideas that I wanted to work upon as well as helpful comments and suggestions were invaluable and gave this dissertation the form as it has today. I am truly grateful to Payal Malik, Economics advisor at the Competition Commission of India (CCI), for introducing me to the concept of platforms which has formed the basis for my doctoral work. Also, our joint work on chapter 4 helped me learn a lot about the link between economic theory and practice.

I am thankful to other faculty members and my friends at the Economics and Planning Unit for their encouragement and support. Also, I have benefitted a lot from participating and presenting my work in various seminars. I appreciate and thank the participants and discussants for their insightful comments on my papers. I sincerely acknowledge the support received from administrative staff and workers, scholarship and travel support from ISI.

Last but not least. a heartful thanks to my parents for always believing in me and their unconditional support and love throughout my life.

Chapter 1

Introduction

1.1 Motivation

The Internet revolution has transformed the way economic activities are organized and goods or services are exchanged. Platforms are at the centre of this reorganization and play an important role in today's economy and society.[1] A platform is defined as an intermediary between two or more clearly defined groups exhibiting network effects, i.e., an agent in each group derives value (positive or negative) from the participation of agents in the same group or other group(s). Platforms can and do act as marketplaces, as search engines, as well as provide various services like social networking. Another common characteristic of these markets is that they are data-intensive, and the firms collect a large amount of information, also known as data, from their users. However, the presence of network effects, compounded by data exploitation, in such markets has led to tipping, establishing dominant platforms in the most pertinent online sectors. The table below provides a few examples of dominant platforms and their market shares. These dominant platforms rely on data as an economic in-

[1] According to the European Commission (Communication, A Digital Single Market Strategy for Europe, COM(2015) 192 final, 6 May 2015): "platforms have proven to be innovators in the digital economy, helping smaller businesses to move online and reach new markets. The rise of the sharing economy also offers opportunities for increased efficiency, growth and jobs, through improved consumer choice."

Sector	Major Player	Market Share
Social Networking	Facebook	Over 68 percent
Search Engine	Google	Over 89 percent
E-commerce	Amazon	Over 49 percent
Application Stores	Android	Over 74 percent

Table 1.1: Concentration of Market Power

put for their effective functioning,[2] which helps them achieve better targeting rates.[3] Besides the commodity characteristic of data, it can also impose a cost on users as they give away data when they join these platforms. Thus, data affects indirect network effects (targeting rate) on the advertising side, and mirror as a price on the user side. This raises ambiguity about how to regulate platforms and data collection and exploitation. In fact, there is an ongoing heated debate in academic and policy circles[4] on how to regulate pricing and non-pricing strategies of online platforms. Given the dual role of data and the presence of network effects, the authorities have to balance the trade-offs when considering intervention in digital markets. While it is important to establish a competitive environment in platform markets, it is likewise important to recognize the substantial benefits associated with them. They must follow a cautious approach and avoid over-intervention to preserve the efficiencies. Similarly, while consolidation of data is a source of competitive advantage, regulating user data need to balance the benefits from information advantage such as better matching and cost reduction (Goldfarb and Tucker (2019)) against potential harms from data collection and exploitation. [5]

[2]Data is also monetised through personalized services, better offers to users and/or sale to third parties.

[3]To get an insight into how user data raises the profitability of advertisements on these platforms, we can look at the rise in average advertising revenue per user (AARPU). AARPU is an indicator of how well internet firms can monetise their user base. Empirically, Facebook's AARPU increased from $ 2.3 in 2014 to $ 4.73 in 2017, whereas Google's AARPU also increased from $ 5 in 2014 to around $ 8 in 2017. Simlarly, AARPU has increased globally over other major platforms like Instagram, Youtube, Twitter, etc. This trend shows that advertisers are willing to spend more to reach each user as the targeting ability of these platforms has improved over time with greater data collection and exploitation.

[4]See, for e.g., Lasserre and Mundt (2017), a joint article by the German and Italian antitrust enforcers expressing concerns about the insurmountable advantage of platforms due to the large amount of user information that they can collect.

[5]Many high profile antitrust cases have acknowledged that standard economic principles are not applicable to data-driven platform markets and have incorporated privacy as a dimension of competi-

As a result, there has been a growth in research on the economic modelling of user data in platform markets. These studies have focussed on the impact of advertisement targeting on consumers (Bergemann and Bonatti (2011)), the effects of competition on privacy (Casadesus-Masanell and Hervas-Drane (2015)), the impact of taxation on data collection (Bloch and Demange (2018); Bourreau et al. (2018)), effects of privacy regulation on market structure (Campbell et al. (2015)), and a firm's investment in quality (Lefouili and Toh (2017)), etc. However, limited attention has been paid to the collection and sharing of data among firms. Other than direct data collection and exploitation, as analysed in previous studies, an important aspect is that platforms can share data among each other, whether affiliated or not.

Data sharing as a source of competitive advantage is widely practised in platform markets. It helps online firms to combine data sets from various sources, gain a better insight into users' behaviour, and create users' profiles related to culture, political orientation, browsing activities, etc. For instance, data on physical activity can be merged with geolocation data or restaurant check-ins and used to target advertisements regarding fitness centres to the health-conscious people. The table below gives examples of data collection by dominant firms from their affiliated partners.

Firm	**Affiliates**
Facebook	WhatsApp, Instagram, etc.
Google	Android, Google Maps, etc.

Table 1.2: Data Sharing between Affiliated Firms

In addition to data sharing between affiliated firms, as illustrated above, the technological developments and agreements between non-affiliated firms can facilitate data sharing. For instance, by using technologies such as APIs (application programming interfaces), online firms can collect data from third parties. Facebook, for example, stipulated contracts with third parties, for the use of its services such as "Like"

tion in its competition analysis. For example, in the Google/Doubleclick merger, Facebook/WhatsApp merger, Microsoft/Linkedln merger, Google Android Case, etc., the antitrust authorities considered the role of data as a source of competitive advantage in its final decisions. However, the trade-off from the use of data is not clear cut. In fact, there is a competitive ambiguity about the impact of data on competition and welfare among competition authorities.

button or Facebook login. Any third-party application adopting this technology voluntarily agrees to share data generated on its platforms, and, in turn, can receive user activity data from Facebook. The table below gives a few more examples of voluntary data sharing between non-affiliated firms.

Firm	Partnership with non-affiliates	Examples
Facebook	Contracts with third-party websites	Login through Facebook, etc.
	Contracts with mobile device manufacturers	eg: Apple, Samsung, etc.
Google	Contracts with third-party websites	Login through Gmail,
		Accelerated Mobile Pages (AMP), etc.

Table 1.3: Data Sharing between Unaffiliated Firms

These data sharing practices can have implications for privacy and welfare. A recent and infamous Cambridge Analytica debacle highlights the possibility of data leakages and misuses.[6] Moreover, data sharing can bring new distortions into the markets such as own content bias by a vertically integrated platform,[7] altering market structures, leveraging market dominance from one market to another market,[8] etc. And, it can also affect the antitrust analysis of exploitative and exclusionary conduct by online firms.[9]

This thesis fills this research gap and examines the role of data sharing among firms, affiliated or not. In doing so, the thesis analyses two areas of public policy that

[6] See an article by Kevin Granville, Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens, N.Y. TIMES (March 19, 2018), available at https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

[7] Edelman and Lai (2016) provide evidence that Google's prominent placement of its own search engine services drives traffic to its affiliated search flight listing.

[8] Newman (2014) argues that Google can use data regarding Android users to improve targeting in Google search through bundling of Android operating system with Google search app.

[9] A recent case demonstrates that competition authorities are concerned about data sharing practices in online platform markets. The German Competition Commission, recently, investigated whether Facebook abused its dominant position and collected user data against their consent. The proceeding aimed at examining the user data which Facebook collected from third-party apps via APIs and combined it with the user data from its social platform. In the final decision, it prohibited Facebook from collecting user data from third-party websites on its platform. Summary of the case is available at https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Fallberichte/Missbrauchsaufsicht/2019/B6-22-16.pdf

would require modification to promote efficiency and avoid user harm in platform markets - privacy regulation and competition policy.

In this context, the rest of the thesis is divided into four chapters. Chapter 2 studies the strategic and welfare effects of data sharing between unaffiliated firms. It analyses how data sharing affects the level of investment in data exploitation and welfare of users and society. Chapter 3[10] examines the effect of privacy regulation when there is data sharing between two affiliated firms. It looks at how i) restricting access to data owned by the firm's subsidiaries, and ii) empowering users to control data collection for targeted advertising affects the user and social welfare. Chapter 4 employs the framework developed in the preceding chapter to examine the competition policy and welfare implications of a non-price instrument to leverage market power, i.e., bundling when data collection and exploitation are present. Finally, chapter 5[11] is a policy paper that discusses how competition law and regulation should be modified to adjust for the challenges arising from the working of platform markets.

In the game theoretic models set up in the chapters, there is a monopolist firm in one of the markets. This competitive setting is relevant as it helps to identify how regulation would affect the market power of firms in such markets. Consequently, the thesis contributes to the ongoing debate on establishing a fair and competitive environment in platform markets. It specifically improves our understanding about i) the effects of data sharing among affiliated and unaffiliated firms on privacy and competition, ii) when a dominant platform should be regulated, and iii) how antitrust tools can be revamped for competition assessment.

Below, we briefly discuss the major chapters outlining the motivation and main results.

[10]Joint work with Prof. Prabal Roy Chowdhury, Indian Statistical Institute, New Delhi.
[11]Joint work with Payal Malik, Economics Advisor, Competition Commission of India (Jakhu and Malik (2017)).

1.2 Chapter 2: Data Sharing in Platform Markets

Information sharing between firms is an interesting area of research in industrial organization (e.g., Gal-or (1986), Gal-Or (1987), Sinha (2013)). However, the literature on data sharing between firms in platform markets is scarce. This chapter examines the competitive and welfare effects of data sharing between unaffiliated firms. As discussed in the previous section, this kind of voluntary data sharing can happen through the adoption of technology offered by a dominant firm to other small firms. This data sharing takes place in a vertical market structure in which a technology is offered by an upstream online platform like Google, Facebook, Amazon, etc., to downstream platforms such as content websites that require these platforms to access users. Examples of such technology include Google AMP and Facebook Instant Articles. The data sharing can circumvent the principles and rules formulated in privacy regulation which mainly concentrates on informing people about privacy issues and restrictions on direct data collection and exploitation. As a result, this chapter investigates how voluntary data sharing affects the level of investment in data exploitation and the resulting welfare effects.

To answer this research question, a game theoretic model is developed to understand the strategic interaction among the main players in the markets and derive market outcomes. It's a vertical market structure in which there are three firms: a single upstream firm and two downstream firms. Users must join the upstream firm to access one of the downstream firms. Thus, the upstream firm acts as the gatekeeper. All firms in the market operate as intermediaries and compete in the advertising market. The advertisers view placing advertisements in different firms as imperfect substitutes. All three firms obtain revenue through advertisements. In addition, the upstream dominant firm has the technological capability to invest in data collection and exploitation for commercial purposes, and obtain additional revenue through the sale of data, collected on its platform, to third parties. Further, the data sharing helps all three firms to improve advertising targeting rate on their platforms.

The model is investigated to understand the research question raised above. In doing so, the market conditions are characterized under which data collection on the upstream firm increases or decreases. It depends, among other things, on the extent of improvement in targeting ability in both markets. One of the key insights of the model is that in markets with a "small" improvement in advertising targeting rate in the upstream market, data exploitation by the upstream firm under no sharing regime is lower than under data sharing, whereas for a "large" improvement, it is the opposite. Intuitively, for a small improvement in the upstream market targeting rate, advertising competition intensifies and advertising prices fall under data sharing, which in turn increases data exploitation to reap profits. Whereas, for a large improvement in the upstream market targeting rate, advertising competition softens and advertising prices rise, which reduces data exploitation under data sharing. Next, the welfare effects of data sharing are examined. The net change in welfare depends on the following effects: improvement in the probability of match over the users, change in the level of data exploitation, net user benefit of data sharing, change in privacy costs, change in nuisance cost of advertisements, and change in transportation cost. It is shown that data sharing improves the targeting ability of the upstream firm which dominates any other effect and the advertising revenue goes up. Therefore, social welfare rises with data sharing because of a better probability of match over the user set. The model is extended in several directions to show the robustness of results and more nuanced results. The most interesting among them is that user welfare is highest under an exclusive data sharing regime, whereas social welfare is highest under a non-exclusive regime.

1.3 Chapter 3: Regulation of Consumer Data: Privacy and Welfare

An interesting feature of platform markets is that the platforms constantly revise boundaries of their activities and enter new areas. This is typically the case with major platforms such as Google and Facebook. Google has expanded into sectors such as application stores (Google Play), mobile operating systems (Android), content aggregation (YouTube, Google News, etc.). Facebook has entered into the mobile communication (WhatsApp), photo and video sharing (Instagram), etc. Linked to this rapid diversification of platforms is the ability to collect data from their affiliated partners and use it to make detailed profiles of users. For example, user location data from mobile devices can be used by Google to recommend specific restaurants in its vicinity. This gives these platforms a competitive advantage over other firms.

This form of data sharing, if harmful to users and society, cannot be mitigated under the privacy framework that has been formulated under recent General Data Protection Regulation (GDPR). It has tried to improve upon shortcomings of the informed consent approach. Under it, among other things, the data collector must specify the purpose of data collection and whether it will be shared with other parties. The regulation states the *lawful purposes* for which data can be collected and used including the informed consent of the data subject to the processing of data. Although it's an improvement, it has placed too much emphasis on individual action and cannot address the question of data as a competitive advantage for the tech giants like Google and Facebook. In fact, several pitfalls can reduce its effectiveness. Among them, the most serious is that a data controller can collect and use data for *legitimate interest or necessity* even when there is no user consent. For instance, Google can override the consent requirement by collecting location data from other apps even when we put on the do-not-track option in Android phones. So, this chapter examines alternative and complementary approaches to regulation to deal with data sharing in platform markets: i) restricting access to data owned by the subsidiaries, and ii) user control of data collection.[12] These two alternative regulatory proposals can provide some new insights and contribute to the ongoing privacy debate.

To investigate the competitive and welfare effects, a game theoretic model is set up in which one firm has a presence in multiple markets and can collect data regard-

[12]Informing user about the adverse effects of data collection is similiar to approaches in other regulatory areas, e.g., lablelling GM foods (Bansal et al. (2013)).

ing the users from these markets. In the model, there are two markets - market 1 and market 2, two firms G and S, a set of users, and a set of advertisers. Firm G is a monopolist in market 1 and competes with firm S in market 2 over users and advertisers, where users dislike advertisements, and advertisers obtain profit from placing advertisements. An interesting feature of the model is that firm G can collect data regarding the users who join it in market 1, which enables firm G to offer improved targeting of advertisements on its platform. However, data collection imposes privacy costs on users in market 1. In order to explain the competitive effects, two important parameters considered are - targeting rate of an advertisement (β) defined as the probability of informing a user by an advertisement and investment in data collection (k). An interesting insight from the model is that there can exist both symmetric and asymmetric business model equilibrium with firm G advertising financed and firm S user financed. This has been called as strategic differentiation in previous work (e.g., Calvano and Polo (2016)). Restricting access to data owned by subsidiaries reduces both user and social welfare. Whereas, in markets with a large advertising targeting rate, user control of data collection can improve both user and social welfare. This can provide policymakers new evidence to understand welfare distortions that can result from data usage restrictions. In the light of GDPR which has emphasized the right to data portability for users from one online firm to another, this alternative way of data collection and exploitation is worth investigating.

1.4 Chapter 4: Bundling in Platform Markets in the Presence of Data Advantage

Competition authorities across different jurisdictions have been investigating the anti-competitive practices of Google on the mobile application market. To put it succinctly, Google has imposed unfair conditions on mobile manufacturers under "Mobile Application Development Agreement". Under it, they have to pre-install Google Apps (notably Youtube and Google Maps) if they want to install Google Play on mobile devices. Since Google Play has the largest market share in the mobile operating system market, the manufacturers must offer it to their customers. This tends to prevent alternative competing apps from being installed by manufacturers. The fundamental question for an antitrust intervention is when such a bundling practice could be anti-competitive and reduce welfare. This has broader implications for understanding the nature of anti-competitive bundling in platform markets. Although there has been some research on this issue in economic theory[13] and policy[14], a particular issue that has remained unaddressed is the collection of user information by a dominant platform across markets to sustain its monopoly in the main market. This issue is interlinked with bundling practices of many platforms operating in the digital economy. So, the nature of antitrust intervention would depend on the ability of a platform to collect user information from various markets.

In this chapter, the game theoretic model developed in the preceding chapter is employed to understand the bundling incentives when an online firm can collect user data across markets and the welfare implications. In the model, the profitability of bundling depends on two opposite forces - i) bundling makes additional data available for targeted advertisements in market 2, and ii) bundled discount reduces firm G's profit in market 1. The interplay of the two effects depends on the targeting rate of an advertisement and investment in data exploitation. It is shown that bundling is profitable when, everything else equal, i) investment in data collection is sufficiently large, and/or ii) advertising targeting rate is large. This focus on the strength of network effects (targeting rate in this chapter) and its impact on digital products and regulatory design has been studied in the platform literature in other contexts, such as taxation in platform markets (e.g., Bloch and Demange (2018), and piracy in platform markets (e.g.,Lu and Poddar (2018)). Welfare analysis shows that there are regions where private and social incentives diverge esp. in markets with a small investment in data collection. Whereas, in markets with a large investment in data collection, bundling

[13] See for e.g., Amelio and Jullien (2012); and Choi and Jeon (2016).

[14] See for e.g., Bork and Sidak (2012); Manne and D. Wright (2010); Newman (2014); and Edelman and Geradin (2016).

is profitable and improves social welfare. So, including the role of data in bundling decisions affects the policy implications that can be drawn. Antitrust action should be focussed on markets with a large data collection and exploitation. Moreover, the welfare standard based on which antitrust action is taken can provide different suggestions. Markets with large data collection and large advertising targeting rates can have a rise in user welfare but a fall in social welfare because of bundling. In addition, the model is extended in several directions to show the robustness of baseline results and more nuanced results.

1.5 Chapter 5: Dilemma in Antitrust Enforcement: How Use of Economics can Guide Enforcement Rules in Multi-Sided Markets

This chapter[15] is a policy paper that discusses how existing areas of competition law should be applied to platform markets and the adequacy of laws to deal with challenges presented by platforms. In particular, it examines the common mistakes that regulators, policymakers and competition enforcers can make when evaluating the platform markets. It discusses antitrust intervention in the light of three pertinent issues raised in platform markets, i.e., i) pricing strategies, ii) search and net neutrality and iii) ex-ante vs ex-post regulation.

First, the exercise of market power is constrained by the risk of losing sales on the other side. This is based on the fundamental economic nature of platforms that different sides are interdependent. The competition authorities still work under the established economic paradigm that too low prices are a sign of predatory intent and a price above marginal cost signals market power. It is shown, through a case study analysis, that charging too low a price on one side is not exclusionary and is motivated by the efficiency arguments. In doing so, a competition assessment of the National Stock Exchange case under Indian competition law is done and it is shown that how

[15] See Jakhu and Malik (2017)

skewed pricing helped it to attain critical mass required to attract users on all sides and was not anti-competitive.

Second, it discusses how competition in online platform markets is significantly guided by a continuous research and development process, business model innovation and expansion into new sectors. As a result, new improved products can enter in a short time, market boundaries are blurred and competition is often for the market, i.e., new innovative firms can instantly come to the market and establish itself the market leader. In the light of this, a platform which looks dominant based on market share analysis conveys little about its actual market power.

Third, there is a divergence between the static price-oriented approach to antitrust analysis and dynamic efficiency which calls for a more restrained approach to regulation of online platforms. The ex-ante approach to regulation can do more harm than good. The shortcomings of the ex-ante approach are discussed in light of the continuing debate on the search and net neutrality. In the absence of any compelling evidence on substantial user harm, intervention should be cautious. In a nutshell, the antitrust enforcement should consider an effects-based approach in which consumers are protected when there is significant evidence of harm while promoting efficiency and not indulging in over-enforcement.

In conclusion, the thesis contributes to the current state of research in platform markets. On the theoretical front, the models developed in the thesis consider data sharing among firms. Whereas, in the literature so far, the current focus is on understanding the data collection on a platform. However, in our framework, one firm can benefit from the functioning of another firm in a distinct market through data sharing. In addition, we consider privacy implications when data is shared, highlighting the mechanism that affects the privacy choice of a firm in an interconnected system. Further, the thesis introduces a framework to model advertising competition in platform markets when there is data sharing between firms. This can serve as the starting point for a much more nuanced economic modelling of advertising targeting in future studies. The insights on competitive and welfare effects of data sharing are new and enhance the economic understanding of data-based competition. On the policy front, it provides clear cut recommendations to regulators about the market conditions under which regulation can be beneficial to users and society.

Chapter 2

Data Sharing in Platform Markets

2.1 Introduction

This paper examines a recent phenomenon in platform markets, that of voluntary data sharing. Firms can gain a large pool of data on user profiles and activity through voluntary sharing of data among themselves. For instance, dominant platforms like Google, Facebook, etc., offer technology products to firms such as third party content providers which help these firms to improve their access to user data. In return, these dominant platforms also get access to user information from these third party firms which improve their targeting abilities. Hence, data sharing is a result of technology sharing between these firms. One example of such voluntary data sharing is Google AMP (accelerated mobile pages) project. Under this, publishers adopting AMP can improve their placement in Google search results and loading times on the mobile web. It also provides Google access to user data through these websites. Another example is Facebook offering Instant Articles technology under which content publishers get priority placement in the Facebook news feed and Facebook keeps detailed user data collected from these websites. And social logins that are being offered by online intermediaries also facilitate data sharing among unaffiliated digital firms.

Data is an important competitive resource for firms in digital markets. Access to a large amount of data can give an online firm a competitive advantage over others. One way to collect data is through accumulating information over users who access the platform for content. This direct collection of data by the platforms and its implications for the level of privacy and user welfare has been studied in the previous literature. However, little attention has been paid to the more recent voluntary data sharing among unaffiliated online firms and its strategic and welfare implications. In this paper, we develop a theoretical framework to understand how voluntary data sharing affects the equilibrium level of investment in data collection and exploitation in an upstream firm. This question is important as it contributes to our understanding of data as a source of market power for online intermediaries. It further helps us understand how a better and safe environment for user privacy can be attained.

We develop a game theoretic model in which there is a single upstream firm that users must join to access the downstream firms. All firms derive revenue through selling advertising space to the advertisers, who view advertisements on different firms as imperfect substitutes. The firms choose the total quantity of advertisements to be displayed on their platforms. In addition, the upstream firm can choose the level of investment in data exploitation to commercially exploit the personal data it can collect regarding the users on its platform. The downstream firms, on the other hand, do not have the technological capability to exploit the data left on its own platform, at a scale, such that it can be of commercial value and sold to third parties. However, information about users' profiles obtained from the dominant upstream firm can help increase their targeting rates. The data collected on the upstream firm includes browsing histories, location information, personal information and so on. However, data collection imposes privacy costs on users which depends on the strength of privacy preferences and the upstream firm's investment level. This modelling assumption is in line with the empirical observation about the concentration of user data and its exploitation in the hands of a few large firms, notably Google, Facebook, etc., which have raised serious privacy concerns. These big tech firms have the capability to collect, process and analyse user data across sectors and combine all the data sets to create valuable user profiles.[1]

In addition, there can be data sharing between the upstream firm and the downstream firms. It is possible only if there is an underlying technology offered by firm 0 that was adopted by the downstream firms. It helps each firm to obtain some useful information about its users which it could not have collected from its own platform. As a result, data sharing improves the targeting rate of advertisements on all the platforms. In the baseline model, it means that the targeting rate of both upstream and downstream firms improve with data sharing. The technology, if offered, is invariably adopted by the downstream firms. This helps us to focus on two different regimes - no data sharing and data sharing. However, in an extension, we consider a game analysing technology offer and adoption decisions. Thus, in this game, data sharing between firms would be endogenous and would depend on whether one or both downstream firms adopt the technology. We consider the business model of no advertisement targeting in which the upstream firm collects and uses user data for direct sale to third parties only.[2] So, improvement in the targeting rate of a firm is possible only through data sharing.

Using this model, we examine the impact of data sharing on the level of data exploitation by the upstream firm. Whether data collection on the upstream firm increases or decreases depends, among other things, on the extent of improvement in targeting ability in both markets. For a "small" improvement in advertising targeting rate in the upstream market, data exploitation by the upstream firm under no data sharing regime is lower than under data sharing, whereas for a "large" improvement, it is the opposite. The intuition for this result stems from how advertising competition and hence advertising prices are affected by data sharing. For a small improvement in the upstream market targeting rate, advertising competition intensifies and adver-

[1] Lerner (2014) and Newman (2013) argued that user data is an essential input in the working of online firms and the information is controlled by a handful of dominant firms with both the concentrated data processing power and supply of user data to dominate a particular sector. This has formed the basis for regulatory intervention as it can give rise to privacy risks. Similarly, Prüfer and Schottmuller (2017) developed a model of competition in data-driven markets in which the marginal cost of quality production is decreasing in the market size and the amount of user-generated data.

[2] In an extension, we also look at a mixed business model that can exist with data collection. In this model, the usptream firm use data for advertisement targeting and selling it to third parties.

tising prices fall under data sharing which increases data exploitation to reap profits. Whereas, for a large improvement in the upstream market targeting rate, advertising competition softens and advertising prices rise which reduces data exploitation under data sharing.

Turning to the welfare analysis, the net change in social welfare depends on the interplay of the following effects: change in advertising competition, improvement in the probability of match over the users, change in the level of data exploitation, change in nuisance cost of advertisements, and change in transportation cost. It is shown that data sharing improves the targeting ability of the upstream firm which dominates any other effect and the advertising revenue goes up. Therefore, social welfare rises with data sharing because of a better probability of match over the user set.

Then, we extend the baseline model in several directions. First, we provide an intuitive explanation for how the results would be affected when there is competition in the upstream market. Next, we analyze the equilibrium relations under mixed business model - advertisement targeting with sale of data to third parties. Interestingly, the level of data exploitation is always higher under data sharing relative to the no data sharing case. In another extension, we consider the possibility of exclusive data sharing along with non-exclusive data sharing in the baseline model. It is shown that for a very small improvement in targeting rate in the upstream market, data exploitation is the highest under non-exclusive sharing, whereas for a large improvement in targeting rate in the upstream market, data exploitation is the lowest under it. Turning to the welfare implications, it is shown that user welfare is the highest under an exclusive regime, whereas social welfare is the highest under a non-exclusive regime. The divergence in welfare analysis stems from the changes in transportation cost, privacy costs, nuisance cost of advertisements, improvement in user utility and better targeting for the advertiser. There is a net improvement in user utility despite excluding some users from the data sharing regime. So, user welfare increases with exclusivity. However, better advertisement targeting effect dominates other changes in social welfare which is highest under non-exclusive sharing. In the fourth extension, we consider the game that endogenizes the data sharing between firms. This is done by endogenizing the technology adoption decision of the downstream firms in which the upstream firm can make a non-exclusive offer to the downstream firms and they can decide whether to accept or reject the offer. The main result here is that it gives the upstream firm, offering technology, incentives to strategically over-invest in data exploitation or reduce the privacy level. The data sharing improves the revenue of the firm from advertisements through better targeting. A higher investment in data exploitation can increase revenue from the sale of data, but it also reduces advertising levels and raises privacy costs for the users. The net effect determines whether or not over-investment takes place. It is shown that in markets with intermediate to a large improvement in targeting rate in the upstream firm, strategic over-investment can be profitable to enforce technology adoption and thus data sharing. Finally, we briefly discuss extensions with endogenous location decisions and symmetric improvement in advertising efficiency.

2.1.1 Related literature

This study contributes to the growing strand of literature on platform markets which considers the data collection activities of firms in such markets. Firms collect data from their users and create value by selling it to third parties or using it for targeted advertisements (e.g., Bloch and Demange (2018); Casadesus-Masanell and Hervas-Drane (2015); Bourreau et al. (2018), etc.). This paper differs from the preceding literature in that it focuses on examining the interaction between privacy and data sharing. Here we study how data sharing between firms affects the privacy choice of an upstream firm. This issue is important as data sharing can affect the market power of the upstream firm through unfair data collection from the users. This paper differs from the previous studies which have focussed on how competition affects the data disclosure levels, impact of taxation on privacy levels, etc.

Data sharing can also be seen as a form of a contractual arrangement between firms. In two sided market literature, few papers have studied the platform contractual arrangement with content providers. Chou and Shy (1990, 1993, 1996) and Church and Gandal (1992, 1993, 2000) made significant contributions to the platform-component literature. They analyzed how indirect network effects influence the number of components on each platform. Hagiu and Lee (2011), Stennek (2014), and Weeds (2016) analyze the exclusionary effects of exclusive contracts between TV channels and distributors. Hogerdon and Ka Yat Yuen (2009) and Armstrong and Wright (2007) are few more papers that discuss and shows the existence of exclusive contracts under some parametric conditions.

However, none of the above mentioned studies consider data sharing agreements. Though information sharing between firms and the resulting strategic and welfare effects has been an interesting of research in industrial organization (e.g., Gal-or (1986); Gal-Or (1987), Vives (1990), Sinha (2013), etc.), the literature on data sharing in platform markets is scarce. The paper closest to this research is a recent study by Krämer et al. (2019). They study the competitive effects of social login adoption and find that social login can serve as an exploitative tool for the dominant firm. The content providers' profit may reduce with voluntary adoption of social logins yielding a prisoner dilemma outcome for them. Our paper differs from this study. First, the model set up is different from theirs. We consider advertising as a nuisance cost to the users and take into account the competition between upstream and downstream firms in the advertising market. Second, their focus is on the exploitation of downstream firms. Whereas, we consider the privacy implications of data sharing and how it affects the market outcomes.

The rest of the paper is organized as follows. Section 2.2 discusses the nature of data sharing in greater detail. In Section 2.3, the baseline model is set up where platforms finance themselves through advertising and sale of data to third parties. Section 2.4 studies the equilibrium relations and welfare effects. Section 2.6 studies some extensions and variations of the baseline model and draw new insights from that. Section 2.7 concludes. All proofs are in the appendix.

2.2 Nature of Data Sharing

Data sharing between platforms has been growing in popularity. A common feature is that sharing takes place between an upstream platform like Google, Facebook, etc., which offer generalized content like general search and social networking services and specialized platforms like content specific websites and mobile apps, with the upstream firm offering technology that is adopted by these websites.

As discussed in the introduction, this data sharing between unaffiliated platforms can improve the targeting rates of both platforms. This is because of the access to data that these platforms obtain from the other platforms. For instance, under social login through Facebook, websites and mobile apps can get access to basic information such as profile photo, demographic data, gender, networks, user ID and a list of friends. In addition, these websites can access other details such as users' likes, political and religious preferences, relationship status, location, photos, and even personal messages. Likewise, apps and websites using Google+ sign-in can access information of users' public profiles and their friend lists to optimize their service. On the other hand, Facebook and Google can get access to information on user activities on these websites. This two way sharing of user data can improve the targeting rate of advertisements on these platforms. Another instance is the adoption of Instant Articles or AMP technology by the websites which get faster loading speed on mobile apps. Under it, the general service platform gets access to audience data from these websites about content that the users read, like, location data, etc. All this data that can be obtained from the other unaffiliated platform can be used to personalize advertisements and improve the click-through rate on the platform.

This *indirect* way of data sharing is recent and has affected internet users through quality and privacy changes. On the quality front, more data has meant better customization and personalization of services. The data can be used to improve search services by the search engine and news feed on the social networking sites. It can also provide ease of login for users across multiple websites which reduces the need for transaction cost of registration. This drives up user welfare from joining these platforms. However, it can have negative implications for privacy. Tracking users across multiple platforms and sharing data with third parties can put user privacy at risk. For instance, although users benefit from the convenience of using mobile apps and websites through Facebook Login, they have to sacrifice their basic information and other details on Facebook. So, third party services can access a lot of information about their users. These vulnerabilities trigger concerns among users. Another source of concern can be how this data sharing affects data collection practices on the platform itself. As discussed earlier, the dominant platforms have the capabilities to collect and process data about user activities and use it for personalized advertisements, selling it to third parties, etc. The rate of data collection and exploitation on these dominant platforms can be affected by these exogenous data sharing regimes and lead to increased privacy risks. The social desirability of the data sharing regime remains unclear. This paper examines the effect of data sharing on user privacy and welfare.

2.3 The Model

Internet Users

There is an upstream firm 0 and two downstream firms 1 and 2 in the model. All three firms act as intermediaries connecting advertisers with users. To fix ideas, we can think of firm 0 as a social networking site (Facebook, LinkedIn, etc.) and the downstream firms as publishers. In this setting, a user has to use the upstream firm 0 to access the downstream market, connecting to *one* of the two downstream firms. Competition between downstream firms is modelled in the Hotelling framework, with firm 1 and 2 being located at two endpoints of the Hotelling line, firm 1 at point 0 and firm 2 at point 1.

Users are uniformly distributed on the Hotelling line and the utility they derive from joining firm $i = 1, 2$, is a decreasing function of the distance between firm i's

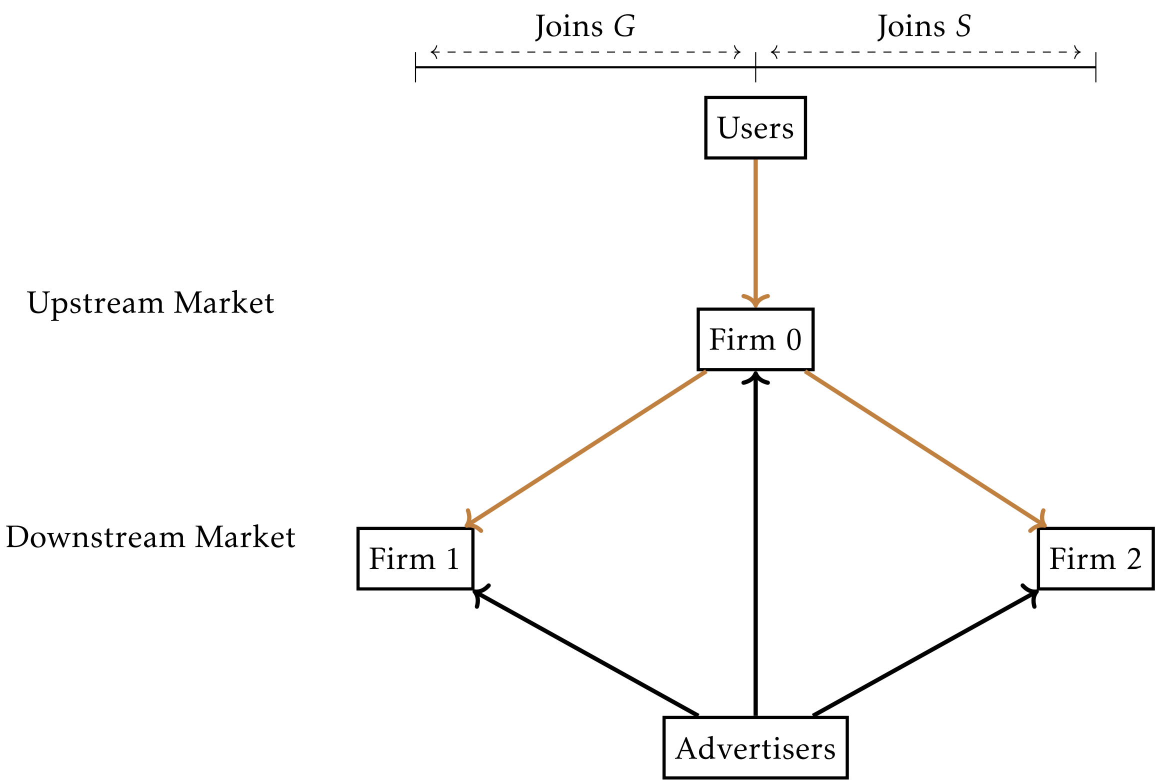


Figure 2.1: Market Structure

position and the user's location $x \in [0,1]$. There is no intrinsic difference between the qualities of the two downstream firms.

When a user located at x joins downstream firm i, her payoff is

$$U_i(x) = \begin{cases} V + I * \theta - \gamma_m(m_0 + m_1) - \gamma_p q_0 - tx, & \text{if } i = 1 \\ V + I * \theta - \gamma_m(m_0 + m_2) - \gamma_p q_0 - t(1-x), & \text{if } i = 2 \end{cases} \tag{2.1}$$

where V measures utility from accessing the content; θ measures improvement in user utility from data sharing; I is an indicator function that takes value one if the technology is shared between firm 0 and firm i and equals zero if there is no such sharing; m_i is the quantity of advertisements on firm i = 0, 1 and 2; $\gamma_m > 0$ measures the per unit nuisance cost of advertisements; q_0 is the level of investment in data exploitation by firm 0; $\gamma_p > 0$ measures the per unit user sensitivity to privacy; and $t > 0$ measures disutility from the discrepancy between the user's location x and firm i's location.

Users dislike advertisements that are bundled with the content on a firm and suffer a nuisance cost of $\gamma_m m_i$. This has been empirically validated in some media studies which found that advertising reduces users' utility (for e.g., Wilbur (2008); Depken and Wilson (2004)). Theoretical work has also characterized advertising as a nuisance to users (e.g., Anderson and Coate (2005)). On the demand side, nuisance and transportation costs also capture the interdependence of consumer demand among the firms in the downstream market through their effect on the equilibrium advertising levels. Besides that, a user is concerned about how much data the dominant upstream firm collects about her and sells to third parties, modelled via a disutility of $\gamma_p q_0$, when q_0 amount of data is collected from her. In this setup, the data is collected from the user when she joins firm 0. For instance, it might be the personal information that the user provides to register on a platform or behavioural data like search history. The reservation utility of the users is taken to be zero.

Advertisers

Advertisers want to place advertisements on firms to reach users. There is a continuum of identical advertisers whose mass is normalized to 1. The return from informing a user is normalized to 1 and the entire surplus is appropriated by the advertiser (see for e.g., Anderson and Coate (2005); Crampes et al. (2009)).

The probability that advertisers inform a single-homing user on firm i is given by $\rho^I(m_i)$ such that

$$\rho^I(m_i) = \alpha(1 + I * \beta_i)m_i, \tag{2.2}$$

where I is the indicator function as defined before; $\alpha > 0$ measures the effectiveness of a unit of an advertisement; m_i is the quantity of advertisements on firm i; and β_i measures the increase in probability due to data sharing. So, $\rho^I(m_i)$ depends on both m_i as well as whether firm i and 0 share data ($I = 1$) or not ($I = 0$). This gives

$$\rho^I(m_i) = \begin{cases} \alpha m_i, & \text{if firm i has no data over the user set,} \\ \alpha(1+\beta_i)m_i, & \text{if firm i has data over the user set.} \end{cases}$$

From the preceding equation, we can see that the network effect in our model is captured through the marginal return from placing an advertisement in a firm. It measures the value that an advertiser gets from an additional unit of participation on the user side. In firm i, $i = 0,1,2$, it can either be α, if there is no data sharing or $\alpha(1+\beta_i)$, if data is shared between firms.

Moreover, the preceding equation also implies that the improvement in advertising efficiency due to data sharing is asymmetric across firms. This is more likely to hold when firms are in different markets. In our model, the upstream firm and the downstream firms can have a different magnitude of improvement in advertising efficacy due to technological differences as well as differences in the nature of user data collected. For instance, Facebook has a huge resource base to track users across websites, whereas a small content provider may not have the same capabilities. More interestingly, the nature and relevance of data collected also differs across the firms. So, Facebook has access to behavioural data regarding users on other platforms, whereas the third-party firms on the Facebook platform get access to the basic user information about age, gender, likes, dislikes, etc. So, for the rest of the analysis, we simply assume that $\beta_1 = \beta_2 = \beta$ and $\beta_0 \neq \beta$. However, in the extensions section, we explore the possibility when the improvement in advertising efficacy is symmetric across firms, i.e., $\beta_0 = \beta$.

The advertising market is modelled in a way to capture an essential feature of the online advertising market, i.e., placing advertisements on two different firms are imperfect substitutes. This means that the marginal value of an advertisement on firm $i \in \{0,1,2\}$ decreases with an increase in the number of advertisements on the other firm $j \neq i$. This assumption is in line with earlier research work on platform markets (see for e.g., Cornière and Taylor (2014); Hahn and Singer (2008)).

Formally, if a user joins two firms 0 and $i = 1, 2$, then the probability is

$$\Pi(m_0, m_i) = 1 - (1 - \rho^I(m_0))(1 - \rho^I(m_i));\ i = 1, 2, \tag{2.3}$$

that the user is informed on atleast one of the firms. As can be seen from equation (2.3), $\frac{\partial \Pi^2}{\partial m_i \partial m_0} < 0$. This essentially captures the important assumption that the advertising quantities on the two different firms are imperfect substitutes. In other words, the demand for advertisements on a downstream firm depends on the quantity of advertisements placed on the upstream firm, which captures the interdependence of demand on the advertising side.

Note that there are two sets of users - i) who have joined firm 1 (N_1), and ii) who have joined firm 2 (N_2), where $N_1 + N_2 = 1$. All users join firm 0 by assumption. So, using equation (2.3), the probability of reaching a user who has joined firm 1 is $1 - (1 - \rho^I(m_0))(1 - \rho^I(m_1))$ and the probability of reaching a user who has joined firm 2 is $1 - (1 - \rho^I(m_0))(1 - \rho^I(m_2))$.[3]

Now, using these probability functions, we can find the revenue that the advertisers receive from purchasing advertisement quantity m_0, m_1 and m_2. Expected revenue on either firm i = 1,2, is $\rho^I(m_0)N_i + [1 - \rho^I(m_0)]\rho^I(m_i)N_i$. Thus, it equals the expected revenue from a set of users N_i from joining firm 0, $\rho(m_0)N_i$, plus the additional revenue from the same set of users on the downstream firm i, i.e., $[1-\rho^I(m_0)]\rho^I(m_i)N_i$. Let P_i denote the advertising price paid for a unit of an advertisement on firm $i \in \{0, 1, 2\}$. Thus, the expected revenue of advertisers (AW) is

$$AW = \rho^I(m_0) + [1 - \rho^I(m_0)][\rho^I(m_1)N_1] + [1 - \rho^I(m_0)][\rho^I(m_2)N_2], \tag{2.4}$$

which, in turn, gives the expected profit of advertisers as

$$\pi_a = AW - P_0 m_0 - P_1 m_1 - P_2 m_2, \tag{2.5}$$

[3]The user on each firm i = 1,2, is always multi-homing between it and firm 0.

where the first term in (2.4), i.e., $\rho^I(m_0)$ represents the revenue from reaching users on firm 0 and the second and third term, i.e., $[1-\rho^I(m_0)][\rho^I(m_i)N_i]$ for i =1,2, represents additional revenue obtained from reaching users that are not informed on firm 0. P_i, the advertising price, can also be interpreted as the marginal cost of a unit of an advertisement on firm i. The advertisers are price takers in the model and firms decide the quantities of advertisements to be displayed on their platforms, i.e., the choice of $m_i's$. So, the advertiser will participate in the advertising market as long as the marginal benefit from a unit of an advertisement is equal to its marginal cost P_i. This implies that the prices are determined to equate the demand for advertising slots by advertisers and the supply of advertising slots by firms, i.e., the choice of $m_i's$.

Firms

The profit of the firm 0 when it sells data directly to third parties is written as

$$\pi_0(m_0, q_0) = P_0 m_0 + Rq_0 - \frac{1}{2}q_0^2, \tag{2.6}$$

where the first term in the preceding equation is revenue from advertisements on firm 0, the second term captures revenue generation from the sale of data to third parties, and the last term is the cost of investment in data exploitation. Whereas, the profit of the dowsntream firm $i = 1,2$, is

$$\pi_i = P_i m_i, \tag{2.7}$$

So, firm 0 can sell advertising space and make investments in data collection and exploitation for sale to third parties, whereas downstream firms are only advertising financed. This distinction in the firms' business models emerges from the difference in the data processing abilities of the firms. We model a situation where data extraction costs are higher for upstream firms, and such costs being zero for the downstream firms is just an extreme form of that assumption. The downstream firms are small publishers or third-party websites which do not have sufficient resources, market base, and know-how to collect and exploit data at a scale as empirically observed on the dominant platforms like Facebook, Google, etc. Thus, downstream firms do not add significant value to the data collected, and face negligible costs of data extraction. Whereas, the upstream firm with significant value addition to the data collected faces a larger and significant data extraction cost.[4] Also, we do not model any additional cost of data sharing between firms. If any, this cost would be of fixed nature and hence, most of our results would remain qualitatively unchanged. A similar assumption was taken by Krämer et al. (2019), where they model the adoption of social login technology by third party content websites.

In the next section, we will examine the effect of data sharing on user privacy and welfare under an alternate business model. To begin with, the business model is one with "no" advertisement targeting with the direct sale of data and then extend the analysis to targeted advertisements in the next section.

Timing of the game

We consider a dynamic multi-stage game, where the timing is as follows:

Stage 1: Firm 0 chooses the level of investment in data exploitation q_0.

Stage 2: Firm 0 chooses the quantity of advertisements m_0.[5]

Stage 3: Firm $i = 1, 2$, chooses the quantity of advertising slots, m_1 and m_2 respectively.

Stage 4: Given m_0, m_1 and m_2, the prices P_0, P_1 and P_2 adjust so that the advertising market clears.

[4]See, for e.g., Newman (2013), Lerner (2014) and Prüfer and Schottmuller (2017) for a discussion on the link between a dominant platform's marginal cost of production and the amount of data collected. In a different setting, de Cornière and Sarvary (2018) analyse the impact of content bundling by a dominant platform when it could provide pesonalized user services based on data collected and provide access to downstream third-party content providers.

[5]One could argue in favour of a different timing in which firm 0 chooses q_0 and m_0 simultaneously. This is formally equivalent to the timing considered in the model. Since it can be argued that the investment in data exploitation is a longer term decision, we stick to the timing in which firm 0 chooses q_0 first and then an advertising decision is made.

Stage 5: Users decide i) whether or not to join firm 0, and ii) which downstream firm to join or not join either.

We look for a subgame perfect Nash equilibrium (henceforth equilibrium) of the game. When solving the game, two different regimes are considered - i) no data sharing, when firm 0 does not share the technology with the downstream firms , and ii) data sharing, when firm 0 shares the technology with downstream firms, which facilitates data sharing.

2.4 Equilibrium Analysis under "Sale of Data" Model

The outcome of the game depends on the data sharing regime, i.e., whether there is data sharing or not. We make the following assumption for the rest of the analysis.

Assumption 2.1. $V \geq V^{'}$.

The threshold value $V^{'}$ is derived in appendix I. It ensures that (i) fixed utility V is high enough so that all users obtain a non-negative net utility from joining firm $i = 1$, 2, and (ii) it is not profitable for firm 0 to exclude some users from the market.

2.4.1 No Data Sharing

Consider the no data sharing regime, indicated by a superscript "*nt*".

Efficiency Benchmark: Second Best Outcome

First, the "efficient" level of data collection and advertising levels are obtained that sets our benchmark. Note that the focus is on the "second best" outcome, where the social planner sets q_0, m_0, m_1 and m_2 to maximize social welfare taking the users' participation decisions and the advertisers' decisions as given. Since advertising prices are just transfers, social welfare will be

$$SW^{nt} = \int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx + \pi_a + \pi_0 + \pi_1 + \pi_2$$
$$= \underbrace{\int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx}_{\text{Users' Surplus}} + \underbrace{AW^{nt}}_{\text{Advertisers' Revenue}} + \underbrace{Rq_0 - \frac{1}{2}q_0^2}_{\text{Profit from the Sale of Data.}} \quad . \tag{2.8}$$

Thus, social welfare is composed of three components. First, the surplus from users' participation in the market. Second, advertisers' revenue from placing advertisements on three firms. Last, the net profit from the sale of data to third parties. A user will purchase from the nearest downstream firm on the Hotelling line. The participation decisions of the users gives us the demand function for each firm $i = 1,2$. The indifferent user is located at point $\hat{x}$ obtained by solving $U_1(x) = U_2(x)$ (see equation (2.1)) and $\hat{x}$ equals

$$\hat{x} = \frac{1}{2} + \frac{\gamma_m(m_2 - m_1)}{2t}. \tag{2.9}$$

From this, the demand for firm 1 is $\hat{x}$ and for firm 2 is $1 - \hat{x}$. Thus, the socially optimal solution can be summarized as follows:

Proposition 2.1. *The efficient solution is characterized by* $m_1 = m_2 = m_0 = \widetilde{m}$, *and* $q_0 = \widetilde{q}_0$, *where*

$$\widetilde{m} = \frac{\alpha - \gamma_m}{\alpha^2}, \text{ and } \widetilde{q}_0 = R - \gamma_p. \tag{2.10}$$

The socially optimal level of data collection is such that marginal social cost, i.e., marginal cost of privacy on user side γ_p plus the marginal cost of data exploitation q_0 equals marginal social benefit of data collection R. Similarly, socially optimal advertising level is such that the marginal social cost of advertisements, i.e., $\gamma_m + \alpha^2\widetilde{m}$ equals marginal social benefit α. It can be seen from the preceding equation that $\widetilde{m} > 0$ if $\alpha > \gamma_m$ and $\widetilde{q}_0 > 0$ if $R > \gamma_p$, i.e., only if the marginal social benefit is sufficiently large. Otherwise, the optimal value is 0. Moreover, the socially optimal advertising levels do not depend on the transportation cost. This is because transportation cost affects a user's utility linearly and there are no quality differences between the two downstream firms. So, socially optimal advertising levels are symmetric and at the margin, transportation cost across the two firms is equal and cancels out. Hence, the advertising levels are independent of the transportation cost parameter.

Equilibrium

We now solve for equilibrium when there is no technology sharing.

Stage 5: The demand for each downstream $i = 1,2$, denoted by N_i, is given by

$$N_i = \begin{cases} [t + \gamma_m(m_j - m_i)]/2t, & \text{if } 2V - 2\gamma_p q_0 - \gamma_m(2m_0 + m_i + m_j) - t \geq 0, \\ [V - \gamma_m(m_0 + m_i) - \gamma_p q_0]/t, & \text{otherwise.} \end{cases}$$

Stage 4: The advertising prices will be determined. As discussed earlier, advertisers are price takers in all the markets and will place advertisements such that price equals the marginal benefit of an advertisement. The advertiser's profit is given by equation (2.5). Substituting (2.3) into (2.5), we have

$$\pi_a = \alpha m_0 + [1 - \alpha m_0][\alpha m_1 N_1 + \alpha m_2 N_2] - P_0 m_0 - P_1 m_1 - P_2 m_2. \tag{2.11}$$

Thus, inverse advertising demand functions for each firm can be written as

$$P_0^{nt} = \alpha[1 - \alpha m_1 N_1 - \alpha m_2 N_2], \tag{2.12}$$

$$P_i^{nt} = [1 - \alpha m_0]\alpha N_i, \; i = 1,2. \tag{2.13}$$

Note from the preceding equations (2.12) and (2.13) that, given a linear probability function, the price P_i is independent of the choice of m_i.

Stage 3: Each firm i =1,2, maximizes its profit function π_i using the inverse demand function (equation (2.9)) and advertising prices (equations (2.12) and (2.13)). The maximization problem of firm i is

$$\underset{m_i}{Max}\ \pi_i \text{ subject to } U_i(\hat{x}(m_i, m_j); m_0) \geq 0. \tag{2.14}$$

This gives us the best response of firm i, m_i^{nt}, as a function of m_0 and q_0 as

$$m_i^{nt} = \begin{cases} t/\gamma_m, & \text{if } V - \gamma_p q_0 - \gamma_m m_0 > 3t/2, \\ [V - t/2 - \gamma_p q_0 - \gamma_m m_0]/\gamma_m, & \text{if } t < V - \gamma_p q_0 - \gamma_m m_0 \leq 3t/2, \\ [V - \gamma_p q_0 - \gamma_m m_0]/2\gamma_m, & \text{if } 0 \leq V - \gamma_p q_0 - \gamma_m m_0 \leq t. \end{cases}$$

Stage 2: Firm 0 operates as the stackelberg leader. It can set the advertising level m_0 given the subsequent best response of downstream firms and users. So, it can set the optimal m_0^{nt} such that the indifferent user gets zero utility in equilibrium. Under assumption 2.1, this yields the solution

$$m_0^{nt} = \frac{V - t - \gamma_p q_0}{\gamma_m}. \tag{2.15}$$

Under assumption 2.1, the optimal value of m_0^{nt} is such that the market remains covered for the subsequent stages and the indifferent user gets zero utility. This gives the optimal advertising levels at downstream firms as

$$m_1^{nt} = m_2^{nt} = \frac{t}{2\gamma_m} \tag{2.16}$$

This choice of m_0^{nt} is intuitive. If firm 0 increases its advertising to $m_0^{nt} + \epsilon$, then the market becomes uncovered and under assumption 2.1, it is not profitable for firm 0 to do that. Whereas, if it reduces the advertising level then the market remains covered and it suffers a loss.

Stage 1: It can be seen that the optimal m_0^{nt} falls with a higher q_0. Also, a reduction in privacy, i.e., a higher q_0, has two opposite effects on firm 0's profits: i) it reduces firm 0's optimal advertising revenues, and ii) it increases the revenue from the sale of data to third parties. Now, substituting (2.15) in (2.6) gives firm 0's profit as

$$\pi_0^{nt} = \alpha m_0^{nt}\left[1 - \frac{\alpha t}{2\gamma_m}\right] + Rq_0^{nt} - \frac{1}{2}(q_0^{nt})^2. \tag{2.17}$$

Maximizing (2.17) with respect to q_0 gives the equilibrium level of data exploitation as

$$q_0^{nt} = R - \frac{\alpha\gamma_p}{\gamma_m}\left[1 - \frac{\alpha t}{2\gamma_m}\right]. \tag{2.18}$$

An interesting question to ask in the model is can there be too few or too many advertisements in the market equilibrium? Provided that $\alpha > \gamma_m$, it is possible that the equilibrium advertising level may be bigger or smaller than the social optimum depending on the transportation cost "t". A lower transport cost means that users can easily substitute between the two firms. From equation (2.15), it can be seen that equilibrium advertising on the upstream firm increases as t decreases. This holds because t affects both the final value of good to users $\bar{V}$ and the rate of data collection q_0, both of which are a decreasing function of t. On the contrary, a lower t reduces equilibrium advertising on downstream firms. Intuitively, there is more competition for users with lower t and thus users can easily switch between two firms. Thus, if t is sufficiently small then equilibrium advertising on the downstream market is lower and on the upstream market is higher than the social optimum. Next, comparing the level of data collection under market outcome with efficient level $\widetilde{q}_0$, it can be seen that marginal user cost γ_p is additionally weighted by $\frac{\alpha\gamma_p}{\gamma_m}$. Since advertising price is never greater than one in the model, there is always over-exploitation of data under the private outcome. These results are summarized in the following proposition.

Proposition 2.2. *If competition on the user side is strong, i.e., small t and net network benefit of advertisement is large, i.e., $\alpha > \gamma_m$, then market outcome leads to i) an over-exploitation of user data, ii) too many advertisements in the upstream market, and iii) too few advertisements in the downstream markets.*

2.4.2 Data Sharing

Consider the case when firm 0 shares technology with both the downstream firms, which, in turn, facilitates data sharing between them. This case is denoted by a superscript "t". Since both downstream firms' quality goes up by θ, there is no vertical quality difference that can occur.

Efficiency Benchmark: Second Best Outcome

The indifferent user's location is still at $\hat{x}$ such that

$$\hat{x} = \frac{1}{2} + \frac{\gamma_m(m_2 - m_1)}{2t}. \tag{2.19}$$

Using this, social welfare under technology sharing is

$$\begin{aligned} SW^{nt} &= \int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx + \pi_a + \pi_0 + \pi_1 + \pi_2 \\ &= \underbrace{\int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx}_{\text{Users' Surplus}} + \underbrace{AW^t}_{\text{Advertisers' Revenue}} + \underbrace{Rq_0 - \frac{1}{2}q_0^2}_{\text{Profit from Sale of Data.}} . \end{aligned} \tag{2.20}$$

Substituting the values for $U_1(x), U_2(x)$ in the preceding equation, the socially optimal values can be found.

Proposition 2.3. *The efficient solution is characterized by $m_1 = m_2 = \widetilde{m}$, $m_0 = \widetilde{m}_0$ and $q_0 = \widetilde{q}_0$, where*

$$\widetilde{m} = \frac{\alpha(1+\beta_0)-\gamma_m}{\alpha^2(1+\beta_0)(1+\beta)},\ \widetilde{m}_0 = \frac{\alpha(1+\beta)-\gamma_m}{\alpha^2(1+\beta_0)(1+\beta)},\ and\ \widetilde{q}_0 = R-\gamma_p. \tag{2.21}$$

The efficient level of investment in data collection and exploitation by firm 0 remains unaffected by data sharing. This is due to the way data is monetized here. The only source of monetisation is the sale of data to third parties. However, efficient advertising level depends on the targeting rates of both markets. At the socially efficient level, improvement in advertising efficiency β_i of each firm i, $i = 0,1,2$ affects its advertising level negatively. The intuition for this is as follows. For firm 0, from a social point of view, a larger value of β_0 improves the social return to placing advertisements in firm 0, but reduces the social return from placing advertisements in downstream firms. The latter effect dominates the former which implies that an increase in β_0 reduces $\widetilde{m}_0$. A similar line of reasoning can explain the negative effect of β on $\widetilde{m}$.

Equilibrium

In this section, we solve for equilibrium under technology sharing. At *Stage 5*, the demand for each downstream $i = 1,2$, denoted by N_i, is given by

$$N_i = \begin{cases} [t+\gamma_m(m_j-m_i)]/2t, & \text{if } 2V-2\gamma_p q_0-\gamma_m(2m_0+m_i+m_j)-t \geq 0, \\ [V+\theta-\gamma_m(m_0+m_i)-\gamma_p q_0]/t, & \text{otherwise.} \end{cases}$$

Turning to *stage 4*, the advertiser now has a higher probability of match on all the firms. Substituting (2.3) in (2.5), its profit function is given as

$$\pi_a = \alpha(1+\beta_0)m_0 + [1-\alpha(1+\beta_0)m_0][\alpha(1+\beta)m_1N_1 + \alpha(1+\beta)m_2N_2] - P_0m_0 - P_1m_1 - P_2m_2. \tag{2.22}$$

From the preceding equation, it can be seen that the targeting rate increases by β_0 on firm 0 and by β on downstream firms 1 and 2. Using this, the inverse advertising demand functions can be written as

$$P_0^t = \alpha(1+\beta_0) - \alpha(1+\beta_0)[\alpha(1+\beta)m_1N_1 + \alpha(1+\beta)m_2N_2], \tag{2.23}$$

$$P_i^t = [1 - \alpha(1+\beta_0)m_0]\alpha(1+\beta)N_i, i = 1, 2. \tag{2.24}$$

At *stage 3*, each firm i =1,2, maximizes its profit function π_i using the inverse demand function (equation (2.19)) and the advertising prices (equations (2.23) and (2.24)). This gives us the best response function of firm i as

$$m_i^t = \begin{cases} t/\gamma_m, & \text{if } V + \theta - \gamma_p q_0 - \gamma_m m_0 > 3t/2, \\ [V - t/2 + \theta - \gamma_p q_0 - \gamma_m m_0]/\gamma_m, & \text{if } t < V + \theta - \gamma_p q_0 - \gamma_m m_0 \leq 3t/2, \\ [V + \theta - \gamma_p q_0 - \gamma_m m_0]/2\gamma_m, & \text{if } 0 \leq V + \theta - \gamma_p q_0 - \gamma_m m_0 \leq t, \end{cases}$$

Next, at *stage 2*, firm 0 sets its advertising level at m_0^t such that the market remains covered and the indifferent user gets zero utility. This gives

$$m_0^t = \frac{V - t + \theta - \gamma_p q_0}{\gamma_m}, \tag{2.25}$$

and the optimal advertising levels at downstream firms' as

$$m_1^t = m_2^t = \frac{t}{2\gamma_m}. \tag{2.26}$$

At *stage 1*, firm 0 sets the value of q_0 to maximize its profit. The profit function is given as

$$\pi_0^t = \alpha(1+\beta_0)m_0^t\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right] + Rq_0^t - \frac{1}{2}(q_0^t)^2. \tag{2.27}$$

Maximizing (2.27) w.r.t. q_0 gives the equilibrium level of q_0 as

$$q_0^t = R - \alpha(1+\beta_0)\frac{\gamma_p}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]. \tag{2.28}$$

Under data sharing, a new source of difference in advertising levels under the social optimum and the market equilibrium can result from the parameter value of improvement in targeting rates in the two markets. From (2.21), it can be seen that parameter β_0 reduces the socially optimal advertising level in firm 0 and parameter β reduces the socially optimal advertising level in firm $i = 1,2$. However, under market equilibrium, the equilibrium advertising level in the upstream market, i.e., m_0^t increases with β_0, and the equilibrium advertising level in the downstream market, i.e., m_i^t, $i = 1,2$, is independent of the advertising targeting rates. So, for a sufficiently large β_0 and β, there can be an overprovision of advertisements in the market equilibrium. In the downstream market, the net effect would also depend upon the value of the transportation cost. If transportation cost parameter t is sufficiently small such that $t < 2\gamma_m\widetilde{m}$, where $\widetilde{m}$ is as defined in equation (2.21), then there is an underprovision of advertisements in market equilibrium by the downstream firms 1 and 2. As argued before, there is always an over-exploitation of data in the equilibrium. These results are summarized in the following proposition.

Proposition 2.4. *If competition on the user side is sufficiently strong, i.e., small t, improvement in upstream advertising targeting rate, i.e., β_0 is large, improvement in downstream advertising targeting rate, i.e., β is small, and the net network benefit of advertisements is large, i.e., $\alpha > \gamma_m$, then market outcome leads to i) an over-exploitation of user data, and ii) too many advertisements in the upstream market, and ii) too few advertisements in the downstream market.*

2.5 Data Sharing versus No Data Sharing

2.5.1 Advertising and Data Exploitation Comparison

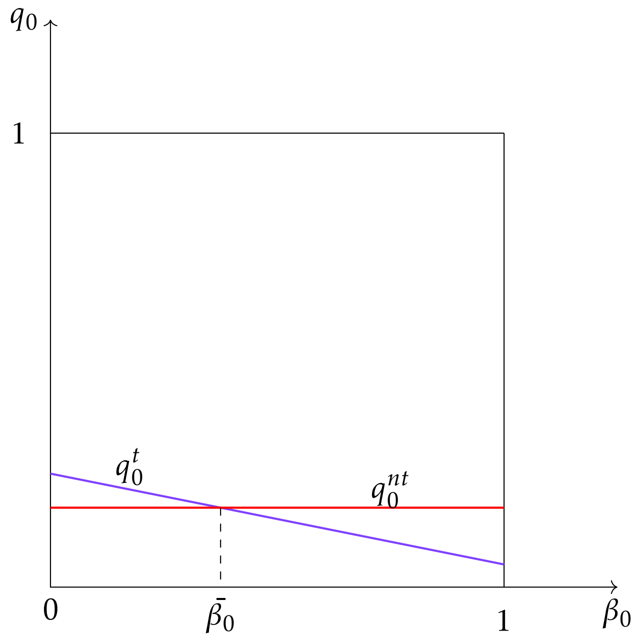
We start by comparing the two regimes for the advertising levels. Under no data sharing, the advertising quantity at firm 0 is given by (2.15) and under data sharing, it is given by (2.25). Thus, it can be seen that the *advertising level of firm 0 is higher under data sharing relative to no data sharing regime.*

A comparison of optimal data exploitation levels under the two regimes yields the following result:

Proposition 2.5. *For a sufficiently small value of β, there exists $\bar{\beta}_0$ such that*

(i) For $\beta_0 \leq \bar{\beta}_0$, $q_0^t > q_0^{nt}$,

(ii) For $\beta_0 > \bar{\beta}_0$, $q_0^t \leq q_0^{nt}$.



q_0^t is the unconstrained level of investment in data exploitation under data sharing and q_0^{nt} is the unconstrained level under no data sharing regime. The figure is drawn for parameter values $V = 2$, $\theta = 0.25$, $\gamma_p = 0.8$, $\gamma_m = 0.5$, $t = 0.5$, $\alpha = 0.2$ and $\beta = 0.8$.

Figure 2.2: Data Exploitation under Different Regimes

Figure 2.2 depicts the optimal choice of the investment in data exploitation under a specific range of parameter values. The above proposition suggests that data sharing can affect data exploitation on the usptream firm 0 either way. In markets where targeting rate on the upstream firm can rise significantly (large β_0), data sharing can reduce the level of data exploitation. In order to interpret the result in the preceding proposition, we need to look at how advertising price is affected by data sharing. Advertising price charged to the advertisers in the upstream market and investment in the data exploitation are substitutes in the model. So, a rise in advertising price on firm 0 reduces data exploitation. There are two effects of data sharing: improvement in probability of match over the users and intensified dvertising competition, the interplay of which determines the advertising price and change in data exploitation. For low values of β_0, net advertising competition intensifies, lowering advertising price, and the investment in data exploitation increases under data sharing. Whereas, for larger values of β_0, improvement in probability of match is sufficient to offset intensified advertising competition, raising advertising price and reducing the investments in data exploitation under data sharing.

2.5.2 Welfare Analysis

Since prices are just transfers from advertisers to firms, social welfare is the sum of users' surplus, advertisers' revenue, revenue from the sale of data and total cost of data exploitation.

$$SW^k = UW^k + AW^k + Rq_0^k - \frac{1}{2}(q_0^k)^2, \tag{2.29}$$

where k denotes regime nt or t; UW^k is user welfare under regime k; AW^k is advertisers' revenue under regime k; Rq_0^k is the revenue of firm 0 from the sale of data to third parties; and the last term is the total cost of data exploitation.

User Welfare

First, we investigate how overall user surplus changes under different scenarios. It is defined as the integral over the utility of all purchasing users. Hence, user welfare under regime k is

$$UW^k = \int_0^{\hat{x}^k} [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_1^k) - tx]dx \\ + \int_{\hat{x}^k}^1 [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_2^k) - t(1-x)]dx, \quad (2.30)$$

where $k = t$ or nt.

Proposition 2.6. *User welfare is the same under both scenarios, i.e.,* $UW^t = UW^{nt}$.

To understand this result, we need to look into the effects that come into play. Data sharing improves the quality of services for the firms. So, users, who join firm i receive a higher utility. Also, it raises the equilibrium advertising level of firm 0 which raises the nuisance cost of these users. The users are exposed to total advertisements $m_0^t + t/\gamma_m$ under data sharing, whereas they are exposed to $m_0^{nt} + t/\gamma_m$ advertisements under no data sharing. Since the advertising level on firm 0 is such that the market remains covered in the model, any surplus utility from data sharing cancels out. Also, sharing does not distort the distribution of users across the two downstream firms. Taken together, the overall loss from higher nuisance cost and gain from an improvement in service balance out each other and the user welfare remains the same.

Advertisers' Revenue

Here, we look at how advertisers' revenue change with data sharing. Substituting the values for $\rho^I(m_i)$ in equation (2.4) under the two regimes and taking the difference of two, the change in advertisers' revenue $\Delta AW = AW^t - AW^{nt}$ can be written as

$$\Delta AW = \alpha(1+\beta_0)m_0^t\left[1 - \frac{\alpha(1+\beta)t}{\gamma_m}\right] + \frac{\alpha\beta t}{\gamma_m} - \alpha m_0^{nt}\left[1 - \frac{\alpha t}{\gamma_m}\right]. \quad (2.31)$$

The following can be stated about how advertisers' revenue change with technology sharing.

Proposition 2.7. *For a sufficiently large* θ

(i) advertisers' revenue rise with data sharing, and

(ii) the difference in advertisers' revenue expands with an increase in the value of parameters β_0, and β, and contracts with an increase in γ_p.

Sale of Data

The last component of social welfare is the profit from the sale of data to third parties. The difference under the two regimes is

$$\Delta R = Rq_0^t - Rq_0^{nt} + \frac{1}{2}(q_0^{nt})^2 - \frac{1}{2}(q_0^t)^2. \tag{2.32}$$

Substituting the values for q_0^{nt} (equation (2.18)) and q_0^t (equation (2.28)) in the preceding equation, it can be written as

$$\Delta R = \left\{\frac{\alpha\gamma_p}{\gamma_m}\left[1 - \frac{\alpha t}{\gamma_m}\right]\right\}^2 - \left\{\frac{\alpha(1+\beta_0)\gamma_p}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{\gamma_m}\right]\right\}^2.$$

The result for the sale of data is summarized as follows:

Proposition 2.8. *There exists β_0^r such that when $0 \leq \beta_0^r \leq 1$, we have that*

(i) for $\beta_0 < \beta_0^r$, profit from the sale of data rises with data sharing,

(ii) for $\beta \geq \beta_0^r$, profit from the sale of data falls with data sharing, and

(iii) the difference in profit from the sale of data contracts with an increase in β_0, expands with an increase in β, and contracts with an increase in γ_p.

Taken together the components of social welfare can move in different directions. The major advantage of data sharing is that it improves the probability of match over the users for the advertisers. On the other hand, the two main concerns that can be raised are of loss of privacy and increased nuisance cost of advertisements. Privacy loss can occur since data exploitation can go up with data sharing. Another concern, from the firm 0's point of view, is that data sharing reduces the profit from the sale of data. However, data sharing improves the probability of match and this effect might outweigh the other concerns, i.e., loss of privacy, increased nuisance cost of advertisements, and the fall in profit from the sale of data to third parties.

Proposition 2.9. *Social welfare is higher under data sharing, i.e.,* $SW^t > SW^{nt}$.

2.6 Extensions

2.6.1 Competition in the Upstream Market

In the presence of a competing firm E, in equilibrium there will be a reduction of advertisements on firm 0. An increase in q_0 would now affect firm 0's profit through three channels. Two of these are present in the current model, namely i) it will improve the revenue from selling data to third parties, and ii) it will raise the cost of data exploitation. There will be a third and novel effect, in that this may make firm 0 more or less attractive for consumers, as well as advertisers, relative to firm E. What happens would depend on how consumers trade-off targeting gains θ against nuisance costs of advertising. For instance, we conjecture that in the presence of data sharing the marginal cost of raising q_0 goes up, so that q_0 would be lower under data sharing. The intuition is as follows. The advertising efficiency is higher with data sharing and thus, increases the marginal cost of increasing q_0, i.e., the fall in advertising revenue. Moreover, if competition in the upstream market is sufficiently strong, then the loss of revenue from advertisements would be exacerbated due to users switching, and this competition effect would dominate the positive effect of increasing q_0 on the revenue from the sale of data. Hence, for sufficiently strong competition between upstream firms, the level of investment in data exploitation, q_0, would always be lower with data sharing relative to the no data sharing scenario. A full analysis is beyond the scope of this chapter though.[6]

[6] I would like to thank an anonymous reviewer for raising this issue.

2.6.2 Alternate Business Model

In this section, we examine the general case when firm 0 uses user data both for advertisement targeting and selling it directly to third parties. This requires advertising targeting at firm 0 to be a function of q_0, $\alpha + \delta q_0$ where $\alpha, \delta > 0$. The parameter $\delta > 0$ captures the effect of data exploitation on the advertisement targeting rate. If $\delta = 0$, we are back to the baseline model of no advertisement targeting. Thus, the advertiser's profit under no data sharing is

$$(\alpha + \delta q_0)m_0 + [1 - (\alpha + \delta q_0)m_0][\alpha(m_1 N_1 + m_2 N_2)]] - P_0 m_0 - P_1 m_1 - P_2 m_2, \tag{2.33}$$

whereas under data sharing it is

$$(\alpha + \delta q_0)(1 + \beta_0)m_0 + [1 - (\alpha + \delta q_0)(1 + \beta_0)m_0][(\alpha + \delta q_0)(1 + \beta)(m_1 N_1 + m_2 N_2)] - P_0 m_0 - P_1 m_1 - P_2 m_2. \tag{2.34}$$

From the preceding equation, it can be seen that, under data sharing, the advertising targeting rate of downstream firms is a function of q_0. This is because firm 0 shares the data collected through its investment in the data exploitation with the downstream firms. As earlier, the profit function of firm 0 is $P_0(q_0)m_0 + Rq_0 - \frac{1}{2}q_0^2$, whereas profit of the downstream firm $i = 1, 2$, is $P_i m_i$. It is difficult to obtain general, tractable results for the advertising targeting case, but equilibria under the two regimes can be computed for specific parameter values. To do so, we consider an example with the parameter values: $V = 2$, $\alpha = 0.2$. $t = 0.5$, $\gamma_m = 0.5$, $\gamma_p = 0.8$, $\theta = 0.25$, and $\beta = 0.8$. In this example, we find that *the investment in data exploitation is larger under data sharing relative to the no data sharing regime.*

While this result is of course true for a particular example, one can make the intuitive case that this should hold for a larger parameter space. This can be explained as follows. Under data sharing, the variable q_0 affects firm 0's profit through various channels. Three of them are already present in the baseline model, namely i) it reduces firm 0's advertising levels, ii) it increases the revenue from the sale of data, and iii) it increases the cost of data exploitation. In addition, there will be two more novel effects, that affects the attractiveness of firm 0 for the advertisers, namely iv) it improves the advertising targeting rate on firm 0, and v) it intensifies advertising competition with the downstream firms. The net impact on firm 0's choice of data exploitation depends on how change in advertising competition affects the profitability of firm 0. The advertising competition intensifies because downstream firms' targeting rate is also affected by q_0 with data sharing. Moreover, it tends to limit the fall in advertising quantitiy of firm 0. When the impact of data exploitation on targeting rate of advertisers δ is large, then the loss from a fall in advertising quantity would be small, and would be dominated by a larger advertising targeting rate and better revenue from the sale of data. Hence, for a sufficiently large δ the net effect is that the intensfied advertising competition increases the investment in data exploitation under the data sharing regime vis-à-vis no data sharing regime.[7]

2.6.3 Exclusive Data Sharing

In the baseline model, we assumed that there was data sharing between the upstrem firm 0 and *both* downstream firms. Suppose, now in addition to that, there can also be a regime"*et*" in which there is exclusive data sharing between firm 0 and one downstream firm i. Then firm i and firm j, $i \neq j$, $i, j = 1, 2$, will have a vertical quality difference measured by θ. At *Stage 5*, the demand for downstream firm i is given by

$$N_i = \begin{cases} [t + \theta + \gamma_m(m_j - m_i)]/2t, & \text{if } 2V + \theta - 2\gamma_p q_0 - \gamma_m(2m_0 + m_i + m_j) - t \geq 0, \\ [V + \theta - \gamma_m(m_0 + m_i) - \gamma_p q_0]/t, & \text{otherwise,} \end{cases} \tag{2.35}$$

whereas the demand for firm j is given as

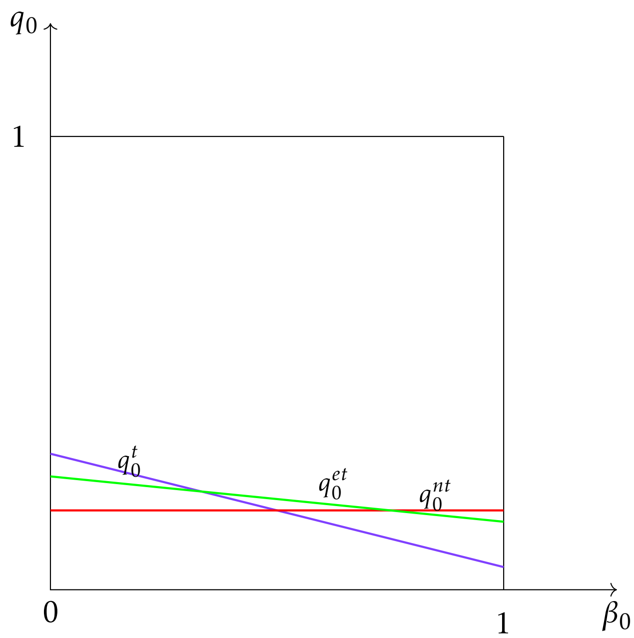
[7] I would like to thank an anonymous reviewer for raising this issue.

$$N_j = \begin{cases} [t - \theta + \gamma_m(m_i - m_j)]/2t, & \text{if } 2V + \theta - 2\gamma_p q_0 - \gamma_m(2m_0 + m_i + m_j) - t \geq 0, \\ [V - \gamma_m(m_0 + m_j) - \gamma_p q_0]/t, & \text{otherwise.} \end{cases} \tag{2.36}$$

On the advertising side, firm 0 and firm i will have higher advertising effectiveness. The profit function of the advertiser is

$$\alpha(1+\beta_0)m_0N_i + \alpha m_0 N_j + [1 - \alpha(1+\beta_0)m_0]\alpha(1+\beta)m_iN_i + [1 - \alpha m_0]\alpha m_j N_j - P_0m_0 - P_im_i - P_jm_j. \tag{2.37}$$

Figure 2.3 below gives us a comparison of the level of data exploitation under different regimes.



q_0^t (blue line) is the unconstrained level of investment in data exploitation under data sharing; q_0^{nt} (red line) is the unconstrained level under no data sharing regime; and q_0^{et} (green line) is the unconstrained level under exclusive data sharing regime. The figure is drawn for parameter values $V = 2$, $\theta = 0.25$, $\gamma_p = 0.5$, $\gamma_m = 0.5$, $t = 0.5$, $\alpha = 0.2$ and $\beta = 0.8$.

Figure 2.3: Exclusive Sharing and Data Exploitation under Different Regimes

The equilibrium level of data exploitation under regime "t" and "et" decreases monotonically with an increase in β_0. This is because β_0 affects the advertising competition. The sensitivity to increase in β_0 is higher under regime "t". A higher β_0 gives a higher probability of match over all users under regime t and thus advertisers are willing to pay more under it vis-à-vis other regimes. As a result, advertising prices increase faster under regime t. So, when β_0 increases advertising prices rise more under regime"t" than under regime"et", the optimal q_0^t falls faster. This shows that there exists a threshold β_0 below which $q_0^t > q_0^{et}$ and above which $q_0^t < q_0^{et}$. Second, the optimal q_0^{nt} remains fixed. This implies that for some values of β_0, q_0^t and q_0^{et} will equal q_0^{nt}. So, again for small values of β_0, q_0^t and q_0^{et} are greater than q_0^{nt}. Whereas, for higher values of β_0, q_0^t and q_0^{et} are smaller than q_0^{nt}. Based on this, the following can be summarised about the level of data exploitation under different regimes.

Proposition 2.10. *If improvement in the targeting rate of downstream firms is sufficiently large, then for small values of β_0, data exploitation is highest under data sharing, whereas for large values of β_0 it is highest under no data sharing regime.*

Next, a normative analysis of the equilibrium relations is done. Based on the analysis, following can be concluded.

Proposition 2.11. *User welfare is highest under exclusive sharing, i.e., $UW^{et} > UW^t = UW^{nt}$. Whereas, social welfare is highest under non-exclusive sharing, i.e., $SW^t > SW^{et} > SW^{nt}$.*

This gives us a paradoxical result that excluding some users from data sharing can raise overall user welfare. To understand this result, we need to look into the effects that come into play. Exclusive sharing improves the services for the firm which adopts it. So, users, who join the patronized firm i receive a higher utility. Also, exclusivity raises the quantity of advertisements by exclusive firm i which raises the nuisance cost of these users. On the other hand, users, who join excluded firm j are exposed to fewer advertisements by firm j. The practice also distorts the distribution of users across the two downstream firms. Some users who were initially at firm j will join the patronized firm i. This leads to a rise in overall transportation cost and nuisance cost of advertisements and leads to distortion. However, due to the redistribution of users across two firms, there will be a net improvement in user welfare from better quality available on firm i exclusively. This dominates the rise in transportation cost and nuisance cost. Hence, user welfare rises.

2.6.4 Data Sharing Decision

In the baseline model, data sharing was taken as exogenous. In this section, we extend the model and introduce a new stage at which firm 0 offers the technology to the downstream firms. Since, it is the technology sharing between firms that facilitate data sharing, we call it the technology adoption stage. However, the only way this adoption affects the economic environment is through facilitating data sharing. Since data exploitation is a longer term choice variable in the model, a new stage is introduced after firm 0 chooses the level of investment in data exploitation. The new timing of the game is as follows:

Stage 1: Firm 0 chooses the level of investment in data exploitation q_0.

Stage 2a: Firm 0 decides whether or not to share technology.

Stage 2b: If it makes an offer, then firm 1 and firm 2 decide sequentially whether to accept or reject the offer.

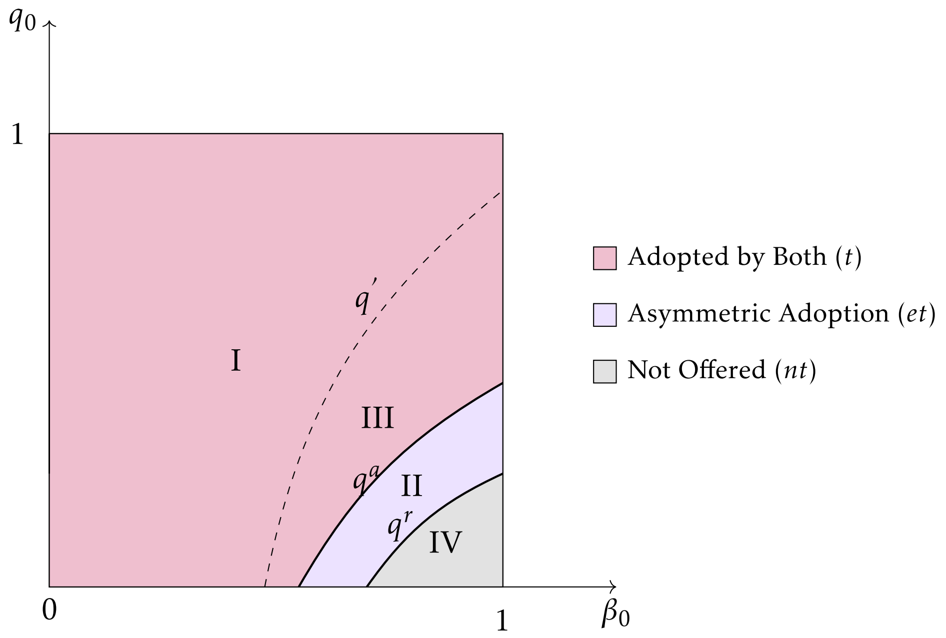
Stage 3: Firm 0 chooses the quantity of advertisements m_0.

Stage 4: Firm 1 and 2 choose the quantity of advertisements, m_1 and m_2 respectively.

Stage 5: Advertisers observe m_0, m_1 and m_2. Advertising market clears: P_0, P_1 and P_2 adjust to equalize the supply and demand for advertisements on each platform.

Stage 6: Users decide i) whether or not to join firm 0, and ii) which downstream firm to join or not join either.

Figure 2.4 below describes the market outcome of technology offer and adoption game. As can be seen from the figure, there are thresholds q^a and q^r that suffice to delineate the market outcomes. There are four sub-regions that have different properties in terms of either the nature of offer, adoption or the impact of technology sharing on downstream firms' profit. The technology is offered and adopted by both firms in



The technology is offered and adopted by both firms in region I, offered and adopted by a single firm in region II and not adopted in region IV. In region III, both downstream firms are worse off because of adoption of the technology, i.e., there is prisoner dilemma outcome. The figure is drawn for parameter values $V = 1.5$, $\theta = 0.25$, $\gamma_p = 0.5$, $\gamma_m = 0.5$, $t = 0.6$, $\alpha = 0.2$ and $\beta = 0.8$. For other values of the parameters, one or the other region may not exist but the properties of each region remains the same.

Figure 2.4: Technology Sharing: Offer and Adoption Decisions

region I. In other regions, technology is either offered and adopted by a single firm, i.e., asymmetric adoption (region II) or not offered (region IV). In addition to that, both firms accept technology but would have been better off under a no offer scenario in region III. This result is similar to the prisoner dilemma situation which has been established in a previous study by Krämer et al. (2019). They find that firms can be in a prisoner dilemma situation when social login is adopted by both special interest content providers.

Next, it would be interesting to do some comparative statics to analyse how these regions would change with a change in parameter values. The two parameters of interest are - user sensitivity to privacy measured by γ_p and improvement in the targeting rate of downstream firms measured by β. The former affects how much surplus users are left with when it joins a platform, while the latter affects the competition in the advertising market. The details of comparative statics are relegated to the appendix.

Here, the main results are summarised.

It should be noted that the parameter γ_p relates to vertical competition in the market for users. It affects all thresholds in the same direction. The other parameter β affects the advertising market competition. It determines the competitive advantage of downstream firms which affects the profitability of technology sharing for firm 0.

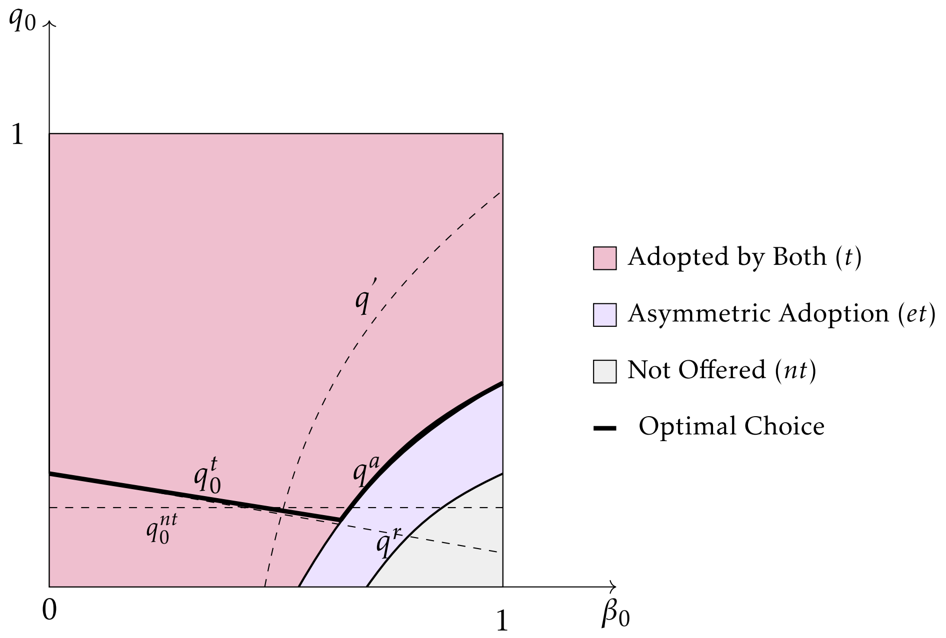
With an increase in γ_p, the offer and adoption thresholds shift downward. When downstream firms' targeting rate β increases then, paradoxically, thresholds q^a and q^r decrease. This can be understood from the likely effect of β on advertising competition. As β increases, the advertising market becomes more competitive under technology sharing. So, each firm's profit from asymmetric adoption would be higher. Hence, in equilibrium, the prisoner dilemma outcome becomes more likely as each firm ends up adopting the technology. The above analysis can help in characterizing the market conditions under which the technology is offered and adopted by both firms, offered and adopted by a single firm or not offered.

Proposition 2.12. *When the technology offer and adoption decisions are endogenous, then*

(i) the technology is a) offered and adopted by both firms when the level of data exploitation by firm 0 is intermediate to large, b) offered and adopted by a single firm when firm 0's data exploitation is intermediate and improvement in its targeting rate is large, and c) not offered when improvement in its targeting rate is very large and data exploitation is very low;

(ii) the likelihood of technology sharing increases, everything else equal, a) with an increase in user sensitivity to privacy (γ_p), and b) with an increase in advertising competition, i.e., an increase in the targeting rate of downstream firms (β).

At stage 1, firm 0 chooses the optimal investment in data exploitation q_0. Now, it has to compare the change in its profits from choosing an unconstrained level of investment in data exploitation q_0^i and strategically choosing a lower or higher q_0 to enforce a particular regime in subsequent stages. The regions and their corresponding investment schedules (dashed lines q_0^t and q_0^{nt}) are shown in figure 2.5. Firm 0 can opt to choose the unconstrained level of data exploitation or strategically reduce or



The figure is drawn for parameter values V = 1.5, θ = 0.25, γ_p = 0.5, γ_m = 0.5, t = 0.6, α = 0.2 and β = 0.8. q_0^t and q_0^{nt} are the unconstrained levels of investment in data exploitation under no data (technology) sharing and data (technology) sharing regime.

Figure 2.5: Strategic Choice of Data Exploitation

increase the investment to enforce a particular regime. From figure 2.5, it can be seen that for low values of β_0, firm 0 can choose unconstrained levels q_0^{nt} or q_0^t. For high values of β_0, firm 0 can again choose either q_0^{nt} and do not offer the technology or strategically increase q_0 to q^a or q^r to enforce sharing.

For the relevant parameter choice, the optimal choice of data exploitation is shown by the bold line in figure 2.5. For low to an intermediate value of β_0, firm 0 optimally chooses unconstrained q_0^t.[8] For a high value of β_0, firm 0 strategically increases q_0 to q^a to enforce technology adoption by both firms.[9] This important insight is summarized in the following proposition.

Proposition 2.13. *When firm 0 can choose its level of data exploitation then, for high values of β_0, it can strategically increase the level of data exploitation on its platform to enforce technology sharing.*

Intuitively, parameter β_0 affects advertising competition. So, when β_0 is large, then

[8]If we take a value of β sufficiently high, it might be the case that firm 0 chooses q_0^{nt} instead of q_0^t to avoid adverse advertising competition.

[9]For higher values of β, firm 0 might choose q^r to enforce asymmetric adoption.

firm 0 can increase q_0 to a level to induce technology sharing and gain from the higher probability of matches despite increased competition in the advertising market.[10]

The two main conditions which affect the level of data exploitation by firm 0 are:

i) Improvement in advertising efficiency, measured by β_0. This can also be interpreted as the investment in third-party user tracking by the firm.

ii) Benefits to users from data collection, measured by θ.

Empirically, it is observed that user tracking capability is concentrated in the hands of a small set of firms such as Facebook, Google, Twitter, and Microsoft. These firms have made significant investment in tracking users across third-party content websites.[11] Hence, it is more likely that β_0 would be higher and based on the results, there will be excessive privacy intrusion on these dominant platforms. Infact, data collection by them is a source of antitrust concern across many jurisdictions. In a recent antitrust case, German competition authority ruled that Facebook has engaged in excessive data collection, amounting to violation of users' privacy terms.

2.6.5 Endogenous Location

In the baseline model, the location of the downstream firms was taken as exogenously given. However, in reality, we see that content providers choose the nature of content before joining an aggregator platform like Facebook, and competing in the advertising market. For instance, news providers can choose the political inclination of their news content before deciding about the advertising quantities. Our model can be used to analyze the endogenous location choice by the downstream firms. The complete analysis is beyond the scope of the paper. However, we can provide an intuition about the equilibrium location decisions. To do so, we can add a new stage at the beginning of the game at which the downstream firms choose their locations on the Hotelling line before the firms compete in the advertising market. For firm i, $i = 1, 2$, choosing a location closer to the location of the other firm j, $j \neq i$, on the Hotelling line, intensifies

[10]But if β was large then, for a very large value of β_0, advertising competition is intense and hence, firm 0 can gain only through an asymmetric adoption.

[11]Binns et al. (2018)

advertising competition. This makes possible larger gains for the upstream firm in the advertising market. Hence, we expect that it would raise both advertising level, m_0, and the investment in data exploitation, q_0. This would further reduce firm i's profit. Hence, the net effect of locating close to the other firm j in the downstream market would be a reduction in the firm i's profit. Thus, we expect firms in the downstream market to choose extreme locations as we have assumed in the baseline model. Moreover, with data sharing, this equilibrium would exist for a larger parameter space.[12]

2.6.6 Case: When Advertising Efficiency is Symmetric across Firms

In the baseline model, we examined the case when an improvement in advertising efficiency is asymmetric across firms and derived our main results. This is the salient case and is based on the empirical observation that the investment in third-party user tracking is asymmetric across dominant platforms like Facebook, Google, etc. and small websites. However, for the sake of completeness, in this section, we explore the case when it could be symmetric, i.e., $\beta_0 = \beta$. We summarize the main result on the market equilibrium in the following proposition.

Proposition 2.14. *When $\beta_0 = \beta$, then for sufficiently strong competition in the downstream market, i.e., small t,*

(i) the equilibrium level of advertising levels in the downstream market remains unchanged, i.e., $m_i^{nt} = m_i^t, i = 1, 2$,

(ii) equilibrium advertising level in the upstream market increases with data sharing, i.e., $m_0^{nt} < m_0^t$, and

(iii) the level of investment in data exploitation decreases with data sharing, i.e., $q_0^{nt} > q_0^t$.

The new insight that can be obtained on the market equilibrium is that the equilibrium level of data exploitation always decreases with data sharing. Intuitively, an increase in β raises advertising efficiency in both markets. Hence, for sufficiently strong competition on the user side in the downstream market, an increase in advertising

[12] I would like to thank an anonymous reviewer for raising this issue.

efficiency of the upstream firm offset the positive effect of the intensified advertising competition on q_0. The net effect is to reduce the investment in data exploitation.

Next, we examine the welfare implications under the constraint, $\beta_0 = \beta$. The main result is summarized as follows.

Proposition 2.15. *When $\beta_0 = \beta$, then for a sufficiently large θ, a large β and a small t, we have*

(i) user welfare remains unchanged with data sharing, i.e., $UW^{nt} = UW^{t}$,

(ii) advertising revenue increases with data sharing, i.e., $AW^{t} > AW^{nt}$,

(iii) profit from the sale of data increases with data sharing, and

(iv) social welfare increases with data sharing, i.e., $SW^{nt} < SW^{t}$.

The user welfare remains unchanged for the same reason as explained in the baseline model. Interestingly, both advertising revenue and profit from the sale of data rises for a sufficiently large β. This is an important result as it shows, contrary to the intuition, that firm 0's profit increases both from advertisements and sale of data with data sharing. The reason is as follows. Advertising revenue increases because, for a large θ, equilibrium advertising levels in firm 0 increases sufficiently with data sharing. Also, for a sufficiently large β, the net profit from the sale of data increases with data sharing because of a reduction in the cost of data exploitation . The social welfare always increases with data sharing because an increase in advertising revenue offset any distortion that arises with data sharing.[13]

2.7 Conclusion

This chapter analysed the strategic and welfare impact of voluntary data sharing through technology adoption in platform markets. It is shown that, under data sharing, the upstream firm can invest higher in data collection, especially in markets with a lower improvement in its advertising targeting rates. However, social welfare rises, irrespective of the business model. So, the focus of the intervention should not be on

[13] I would like to thank an anonymous reviewer for raising this issue.

how data is monetized, but whether or not firms share data. The analysis suggests that data sharing can be beneficial to society.

In an extension, it is shown that exclusive sharing can raise user welfare. Exclusive offer benefits the users of the winning firm. It also benefits excluded users because of the lower nuisance cost of advertisements. So, a ban on exclusive offer may reduce the welfare of all users in the model. Also, a regulation prohibiting discrimination in technology offers can lead to fewer offers of technology. In figure 2.4, the area where downstream firms themselves adopted technology asymmetrically, a non-discrimination rule can reduce technology adoption, and thus data sharing.

2.8 Appendices

Appendix I: Derivation of Covered Market Condition

In this appendix, we derive the covered market condition that is stated in assumption 2.1. There are two different regimes - no data sharing and data sharing, that we have to consider.

No Data Sharing

Suppose firm 0 sets m_0 such that some users do not join any of the two downstream firms. In this case, the demand for firm $i = 1,2$, is given by

$$N_i = \frac{V - \gamma_p q_0 - \gamma_m(m_0 + m_i)}{t} \tag{2.38}$$

Using the preceding equation, each firm $i = 1,2$, maximizes its profit π_i w.r.t. m_i and then at stage 2, firm 0 chooses the advertising quantity m_0. This gives

$$m_i^u = \frac{V - \gamma_p q_0 - \gamma_m m_0^u}{2\gamma_m}, \tag{2.39}$$

$$m_0^u = \frac{-2(\gamma_m - \alpha\widetilde{V}) + \sqrt{4(\alpha\widetilde{V} - \gamma_m)^2 + 3\alpha(2\gamma_m\widetilde{V} - \alpha\widetilde{V}^2)}}{3\alpha\gamma_m}, \tag{2.40}$$

where $\widetilde{V} = V - \gamma_p q_0$. Using the value of m_i^u from the preceding equation in (2.38), it can be seen that the market is covered if $\widetilde{V} - \gamma_m m_0^u > t$. Now, using the value of m_0^u from the preceding equation gives a threshold $\bar{V}$ which equals

$$\bar{V} = \frac{t(4\gamma_m - 3\alpha t)}{2(\gamma_m - \alpha t)} + \gamma_p q_0, \tag{2.41}$$

such that the market is covered if $V \geq \bar{V}$.

Data Sharing

Consider the Data sharing regime. If firm 0 sets m_0 such that some users do not join any of the two downstream firms, then demand for firm $i = 1, 2$, is given by

$$N_i = \frac{V + \theta - \gamma_p q_0 - \gamma_m(m_0 + m_i)}{t}. \tag{2.42}$$

Using the preceding equation, each firm $i = 1, 2$, maximizes its profit π_i w.r.t. m_i and then at stage 2, firm 0 chooses the advertising quantity m_0. This gives

$$m_i^u = \frac{\widetilde{V} + \theta - \gamma_m m_0^u}{2\gamma_m}, \tag{2.43}$$

$$m_0^u = \frac{-2[\gamma_m - \alpha(1+\beta)(\widetilde{V} + \theta)]}{3\alpha(1+\beta)\gamma_m} + \frac{\sqrt{4[\alpha(1+\beta)(\widetilde{V} + \theta) - \gamma_m]^2 + 3\alpha(1+\beta)[2\gamma_m(\widetilde{V} + \theta) - \alpha(1+\beta)(\widetilde{V} + \theta)^2]}}{3\alpha(1+\beta)\gamma_m}, \tag{2.44}$$

where $\widetilde{V} = V - \gamma_p q_0$. Similar to previous case it can be shown that the market is covered if $\widetilde{V} + \theta - \gamma_m m_0^u > t$. Using the value of m_0^u from the preceding equation gives a threshold $V^{'}$ which equals

$$V^{'} = \frac{t[4\gamma_m - 3\alpha(1+\beta)t]}{2[\gamma_m - \alpha(1+\beta)t]} - \theta + \gamma_p q_0, \tag{2.45}$$

such that the market is covered if $V \geq V^{'}$. It can be shown that $\bar{V} < V^{'}$, where $\bar{V}$ is as given in equation (2.41). So, $V^{'}$ is binding. Hence, if market is covered under data sharing, then it remains covered under no data sharing.

Appendix II: Existence of Nash Equilibrium

We need to show that the equilibrium advertising levels and data exploitation under the two regimes are a subgame perfect nash equilibrium.

No Data Sharing

We need to show that the equilibrium advertising quantities and data exploitation specified in equations (2.15), (2.16) and (2.18) constitute an equilibrium. The proof proceeds in two parts. First, given q_0^{nt}, m_0^{nt} and m_j^{nt}, a downstream firm i will have no incentive to deviate. Second, given firm 1 and 2's best response functions, firm 0 cannot do any better by deviating from m_0^{nt} and q_0^{nt}.

Step 1: No downstream firm will deviate: Substitute the value for m_0^{nt}, q_0^{nt} and m_j^{nt} in firm i's profit function given in equation (2.7). This gives an unconstrained advertising level for firm i equal to $\frac{3t}{4\gamma_m}$. At this level, the user located at point $\frac{1}{2}$ gets a negative utility. Hence, the market becomes uncovered. Reducing its advertising level is not profitable as the profit function is concave. So, when deviating, firm i cannot do an unconstrained optimization with a covered market. It will make the market uncovered and user demand for it will be as given in equation (2.38). This will give the best deviation advertising quantity as $m_i = \frac{t}{2\gamma_m}$. So, firm i cannot set a higher advertising quantity by deviating. Similarly, it can be shown that firm j cannot do any better by deviating.

Step 2: Firm 0 will not deviate: We need to show that firm 0 will not deviate from m_0^{nt} and q_0^{nt}. Since the profit function is concave, it suffices to show that given q_0^{nt}, firm 0 will not deviate to a higher advertising level. Suppose firm 0 sets a higher advertising level at $m_0^d = m_0^{nt} + \epsilon$, for some $\epsilon > 0$. First, we need to find the best response of downstream firms. Given other firms' choices, if the downstream firm i keeps the market covered, then it sets $m_i^{'} = \frac{t}{2\gamma_m} - \epsilon$. This $m_i^{'}$ is less than the unconstrained level it could have set under full market coverage. Since at an unconstrained level, the user located at $\frac{1}{2}$ gets negative utility and profit function is concave, $m_i^{'}$ is the best firm i

can do. However, at this level, it can be shown that $\pi_i(m_i^{'}) < \pi_i(m_i^{nt})$. So, firm i's best response is not to the keep market covered. Under partial market coverage, firm i sets its advertising level at $m_i^d = m_i^u$ given in equation (2.39) and it equals

$$m_i^d = \frac{t - \gamma_m \epsilon}{2\gamma_m}. \tag{2.46}$$

It can be shown that

$$\pi_i(m_i^d) = [1 - \alpha m_0^{nt}]\alpha \left[\frac{t}{4\gamma_m} - \frac{\epsilon}{2} + \frac{\epsilon^2}{4\gamma_m t}\right] > \pi_i(m_i^{'}) = [1 - \alpha m_0^{nt}]\alpha \left[\frac{t}{4\gamma_m} - \frac{\epsilon}{2}\right].$$

Hence, best response is to keep partial market coverage. Now given partial market coverage, we need to show that for firm 0, $\pi_0(m_0^d) - \pi_0(m_0^{nt}) \leq 0$. After some algebra, this can be written as

$$\pi_0(m_0^d) - \pi_0(m_0^{nt}) = \alpha m_0^{nt}\left[\alpha\epsilon - \frac{\alpha\epsilon^2\gamma_m}{2t} - \frac{\gamma_m \epsilon}{t}\right] + \alpha\epsilon\left[1 - \frac{\alpha t}{2\gamma_m} - \frac{\gamma_m \epsilon}{t} + \alpha\epsilon - \frac{\alpha\epsilon^2\gamma_m}{2t}\right] \tag{2.47}$$

Putting in the value for m_0^{nt} from equation (2.15), the expression in preceding equation is less than 0 if

$$V \geq t\left[1 - \frac{\alpha/2}{\gamma_m/t + \alpha\epsilon\gamma_m/2t - \alpha}\right] + \gamma_p q_0^{nt}. \tag{2.48}$$

Given assumption 2.1, this will hold. So, firm 0 will not deviate to m_0^d. Since profit function is concave in its arguments, firm 0 can do no better by deviating from q_0^{nt}. Hence proved.

Data Sharing

The proof will follow the same steps as under no data sharing regime. Since the proof is very similar, we omit the details here. Firm 0 will have no incentive to deviate to $m_0^t + \epsilon$ if

$$V \geq t\left[1 - \frac{\alpha(1+\beta)/2}{\gamma_m/t + \alpha(1+\beta)\epsilon\gamma_m/2t - \alpha(1+\beta)}\right] - \theta + \gamma_p q_0^t. \quad (2.49)$$

Under assumption 2.1, this condition will be satisfied. Hence proved.

Appendix III: Proofs of Baseline Model

Proof of proposition 2.1

Social welfare under no data sharing regime can be written as

$$SW^{nt} = V - \gamma_m m_0 - \gamma_p q_0 - \gamma_m m_1 \hat{x} - \gamma_m m_2(1-\hat{x}) - \frac{t}{2} + t\hat{x} - t\hat{x}^2 + \alpha m_0 + [1-\alpha m_0][\alpha m_1 \hat{x} + \alpha m_2(1-\hat{x})] + Rq_0 - \frac{1}{2}q_0^2. \quad (2.50)$$

The efficient solution is found by maximizing the preceding equation w.r.t. q_0, m_0, m_1, and m_2. The first-order conditions yield

1. $$\frac{\partial SW}{\partial q_0} = -\gamma_p + R - q_0 = 0, \quad (2.51)$$

2. $$\frac{\partial SW}{\partial m_0} = -\gamma_m + \alpha - \alpha[\alpha m_1 \hat{x} + \alpha m_2(1-\hat{x})] = 0, \quad (2.52)$$

3. $$\frac{\partial SW}{\partial m_1} = -\gamma_m\left[\frac{1}{2} + \frac{\gamma_m(m_2 - 2m_1)}{2t}\right] - \gamma_m m_2\left[\frac{\gamma_m}{2t}\right] - t\left[\frac{\gamma_m}{2t}\right] + 2t\hat{x}\left[\frac{\gamma_m}{2t}\right] + [1-\alpha m_0]\left\{\alpha\left[\frac{1}{2} + \frac{\gamma_m(m_2 - 2m_1)}{2t}\right] + \alpha m_2\left[\frac{\gamma_m}{2t}\right]\right\} = 0, \quad (2.53)$$

4. $$\frac{\partial SW}{\partial m_2} = -\gamma_m\left[\frac{1}{2}+\frac{\gamma_m(m_1-2m_2)}{2t}\right]-\gamma_m m_1\left[\frac{\gamma_m}{2t}\right]+t\left[\frac{\gamma_m}{2t}\right]-2t\hat{x}\left[\frac{\gamma_m}{2t}\right]$$
$$+[1-\alpha m_0]\left\{\alpha\left[\frac{1}{2}+\frac{\gamma_m(m_1-2m_2)}{2t}\right]+\alpha m_1\left[\frac{\gamma_m}{2t}\right]\right\}=0. \quad (2.54)$$

Solving the F.O.Cs (2.51) - (2.54) simultaneously gives us the required solution.

Proof of proposition 2.3

Social welfare under data sharing is

$$SW^t = V+\theta-\gamma_m m_0-\gamma_p q_0-\gamma_m m_1\hat{x}-\gamma_m m_2(1-\hat{x})-\frac{t}{2}+t\hat{x}-t\hat{x}^2+\alpha(1+\beta_0)m_0$$
$$+[1-\alpha(1+\beta_0)m_0][\alpha(1+\beta)m_1\hat{x}+\alpha(1+\beta)m_2(1-\hat{x})]+Rq_0-\frac{1}{2}q_0^2.$$

The solution can be found by maximizing the preceding equation w.r.t. q_0, m_0, m_1 and m_2. Then using first-order conditions, we can get the required solutions given in the main text.

Proof of proposition 2.5

The difference in the optimal choice under the two regimes is

$$q_0^t-q_0^{nt}=\frac{\alpha\gamma_p}{\gamma_m}\left\{1-\frac{\alpha t}{2\gamma_m}-(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\},$$
$$\text{Now, } q_0^t-q_0^{nt}\geq 0 \text{ if } \beta_0\leq\frac{1-\alpha t/2\gamma_m}{1-\alpha(1+\beta)t/2\gamma_m}-1. \quad (2.55)$$

The R.H.S in equation (2.55) is defined as the threshold $\bar{\beta_0}$. Since $\beta\in[0,1]$, it can be seen that $0\leq\bar{\beta_0}\leq 1$. Hence proved.

Proof of proposition 2.6

User welfare under regime k can be written as

$$UW^k = \int_0^{\hat{x}^k} [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_1^k) - tx]dx + \int_{\hat{x}^k}^1 [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_2^k) - t(1-x)]dx, \quad (2.56)$$

where $k = t, nt$; $\hat{x}^k$ is the market share under regime k; and m_j^k is the advertising level on firm $j \in \{0, 1, 2\}$ under regime k. Equation (2.56) can be rewritten as

$$UW^k = V + I * \theta - \gamma_m[m_0^k + m_1^k \hat{x}^k + m_2^k(1 - \hat{x}^k)] - \gamma_p q_0^k - t\left[\frac{(\hat{x}^k)^2}{2}\right] - t(1 - \hat{x}^k) + t\left[\frac{1}{2} - \frac{(\hat{x}^k)^2}{2}\right].$$

Substituting the values for m_0, m_1, m_2, and q_0 under different regimes and after some calculations, we get $UW^t = UW^{nt} = t/4$. Thus, user welfare remains the same under the two regimes. Hence proved.

Proof of proposition 2.7

Change in the aggregate advertiser revenue due to data sharing is

$$\Delta AW = \left\{\alpha(1+\beta_0)m_0^t + [1 - \alpha(1+\beta_0)m_0^t]\left[\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\} - \left\{\alpha m_0^{nt} + [1 - \alpha m_0^{nt}]\left[\frac{\alpha t}{2\gamma_m}\right]\right\}.$$

After some algebra, it can be written as

$$\Delta AW = \alpha(1+\beta_0)m_0^t\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right] + \frac{\alpha\beta t}{\gamma_m} - \alpha m_0^{nt}\left[1 - \frac{\alpha t}{2\gamma_m}\right].$$

Now, putting in the values for m_0^t and m_0^{nt} as given in (2.25) and (2.15), we get

$$\Delta AW = \left\{\frac{V-t-\gamma_p R}{\gamma_m}\right\} * \left\{\alpha(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right] - \alpha\left[1-\frac{\alpha t}{2\gamma_m}\right]\right\} + \frac{\alpha\beta t}{2\gamma_m} +$$
$$\frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right] - \left\{\frac{\alpha\gamma_p}{\gamma_m}\left[1-\frac{\alpha t}{2\gamma_m}\right]\right\}^2 + \left\{\frac{\alpha(1+\beta_0)\gamma_p}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\}^2. \tag{2.57}$$

It can be seen that if θ is sufficiently large, then the preceding equation is greater than 0. In order to prove the second part of the proposition, we need to calculate the derivative of ΔAW w.r.t. β_0, β, and γ_p. After some calculations, they are

1. $$\frac{\partial \Delta AW}{\partial \beta_0} = \alpha\left\{\frac{V-t+\theta-\gamma_p R}{\gamma_m} + \frac{\alpha\gamma_p^2(1+\beta_0)}{\gamma_m^2}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right],$$

2. $$\frac{\partial \Delta AW}{\partial \beta} = 1 - \frac{\alpha(1+\beta_0)}{\gamma_m}\left[V-t+\theta-\gamma_p R\right] - \frac{2\alpha^2(1+\beta_0)^2\gamma_p^2}{\gamma_m^2}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right],$$

3. $$\frac{\partial \Delta AW}{\partial \gamma_p} = \alpha\left[1-\frac{\alpha t}{2\gamma_m}\right]\left[\frac{R}{\gamma_m} - \frac{2\alpha\gamma_p}{\gamma_m^2}\left(1-\frac{\alpha t}{2\gamma_m}\right)\right]$$
$$- \alpha(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\left[\frac{R}{\gamma_m} - \frac{2\alpha\gamma_p}{\gamma_m^2}\left(1-\frac{\alpha t}{2\gamma_m}\right)\right].$$

Since $\beta \in [0,1]$, it can be shown that the derivative of ΔAW w.r.t. β_0 and β is greater than 0, whereas w.r.t γ_p, it is less than 0. Hence proved.

Proof of proposition 2.8

The change in profit from the sale of data to third parties is

$$\Delta R = [Rq_0^t - \frac{1}{2}(q_0^t)^2] - [Rq_0^{nt} - \frac{1}{2}(q_0^{nt})^2].$$

Substituting the values for q_0^t and q_0^{nt} as defined in equation (2.18) and (2.28), ΔR can be written as

$$\Delta R = \left\{ \frac{\alpha \gamma_p}{\gamma_m} \left[1 - \frac{\alpha t}{2\gamma_m} \right] \right\}^2 - \left\{ \frac{\alpha(1+\beta_0)\gamma_p}{\gamma_m} \left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m} \right] \right\}^2. \tag{2.58}$$

Setting $\Delta R = 0$ gives a threshold β_0^r such that

$$\beta_0^r = \frac{1 - \alpha t / 2\gamma_m}{1 - \alpha(1+\beta)t/2\gamma_m} - 1. \tag{2.59}$$

So, if $\beta_0 \leq \beta_0^r$, then $\Delta R > 0$, otherwise it is less than 0. Next, the partial derivative of ΔR w.r.t. β_0, β, and γ_p are

1. $\dfrac{\partial \Delta R}{\partial \beta_0} = -\dfrac{2\alpha^2(1+\beta_0)\gamma_p^2}{\gamma_m^2} \left[1 - \dfrac{\alpha(1+\beta)t}{2\gamma_m} \right]^2,$

2. $\dfrac{\partial \Delta R}{\partial \beta} = \dfrac{2\alpha t}{\gamma_m} \left\{ \dfrac{\alpha(1+\beta_0)\gamma_p}{\gamma_m} \left[1 - \dfrac{\alpha(1+\beta)t}{2\gamma_m} \right] \right\},$

3. $\dfrac{\partial \Delta R}{\partial \gamma_p} = \dfrac{2\alpha\gamma_p}{\gamma_m} \left\{ \left[1 - \dfrac{\alpha t}{2\gamma_m} \right]^2 - (1+\beta_0)^2 \left[1 - \dfrac{\alpha(1+\beta)t}{\gamma_m} \right]^2 \right\}.$

It can be seen that $\frac{\partial \Delta R}{\partial \beta_0} < 0$, and $\frac{\partial \Delta R}{\partial \beta} > 0$. There exists β_0^r as defined in (2.59) such that if $\beta_0 \leq \beta_0^r$, then $\frac{\partial \Delta R}{\partial \gamma_p} \geq 0$, and for $\beta_0 > \beta_0^r$, it is less than 0. Hence proved.

Proof of proposition 2.9

The change in social welfare under regime "*t*" and regime"*nt*" is

$$\Delta SW = \Delta UW + \Delta AW + \Delta R.$$

ΔAW and ΔR are as defined in (2.57) and (2.58). User welfare under regime "*t*" and "*nt*" are equal. So, $\Delta UW = 0$. This implies that

$$\Delta SW = \left\{\frac{V - t - \gamma_p R}{\gamma_m}\right\}\left\{\alpha(1+\beta_0)\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right] - \alpha\left[1 - \frac{\alpha t}{2\gamma_m}\right]\right\} + \frac{\alpha\beta t}{2\gamma_m} + \frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]. \quad (2.60)$$

For sufficiently high V as we have assumed, (2.60) is greater than 0. Hence proved.

Appendix IV: Extensions

Exclusive Data Sharing

From equation (2.37), it is clear that the targeting rate on firm 0 increases by β_0 and on firm i by β. Using this, the advertising prices are

$$P_0^{et} = \alpha(1+\beta_0)N_i + \alpha N_i - \alpha^2(1+\beta_0)(1+\beta)m_i N_i - \alpha^2 m_j N_j,$$
$$P_i^{et} = [1 - \alpha(1+\beta_0)m_0]\alpha(1+\beta)N_i,$$
$$P_j^{et} = [1 - \alpha m_0]\alpha N_j.$$

The demand functions of the downstream firms are given in equations (2.35) and (2.36). Using the advertising prices and demand functions, we can solve for the equilibrium advertising quantities. Downstream firms' best response functions are

$$m_i^{et} = \begin{cases} (3t+\theta)/3\gamma_m, & \text{if } V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 > 3t/2, \\ [V + 3\theta/4 - \gamma_p q_0 - \gamma_m m_0 - t/2]/\gamma_m, & \text{if } t < V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 \le 3t/2, \\ [V + \theta - \gamma_p q_0 - \gamma_m m_0]/2\gamma_m, & \text{if } 0 \le V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 \le t. \end{cases} \quad (2.61)$$

$$m_j^{et} = \begin{cases} (3t-\theta)/3\gamma_m, & \text{if } V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 > 3t/2, \\ [V + \theta/4 - \gamma_p q_0 - \gamma_m m_0 - t/2]/\gamma_m, & \text{if } t < V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 \le 3t/2, \\ [V - \gamma_p q_0 - \gamma_m m_0]/2\gamma_m, & \text{if } 0 \le V + \theta/2 - \gamma_p q_0 - \gamma_m m_0 \le t. \end{cases} \quad (2.62)$$

At *stage 2*, using the best response functions given in (2.61) and (2.62), firm 0 sets the optimal advertising level as[14]

$$m_0^{et} = \frac{V - t + \theta/2 - \gamma_p q_0}{\gamma_m}, \tag{2.63}$$

and the optimal advertising levels of the downstream firms are

$$m_1^{et} = \frac{2t+\theta}{4t}, \text{ and } m_2^{et} = \frac{2t-\theta}{4t}. \tag{2.64}$$

At *stage 1*, firm 0 will choose the level of data exploitation given that it shares the technology exclusively with one downstream firm i. This gives the choice of q_0^{et} as

$$q_0^{et} = R - \alpha\frac{\gamma_p}{\gamma_m}\left\{1 + \frac{\beta_0(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right\}. \tag{2.65}$$

Proof of Proposition 2.10

The level of investment in data exploitation under different regimes are

$$q_0^{nt} = R - \alpha\frac{\gamma_p}{\gamma_m}\left[1 - \frac{\alpha t}{2\gamma_m}\right],$$

$$q_0^{t} = R - \alpha(1+\beta_0)\frac{\gamma_p}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right],$$

$$q_0^{et} = R - \alpha\frac{\gamma_p}{\gamma_m}\left\{1 + \frac{\beta_0(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right\}.$$

The difference $q_0^{et} - q_0^{nt}$ equals

$$q_0^{et} - q_0^{nt} = \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} - \beta_0\left[\frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m}\right]. \tag{2.66}$$

[14]Under assumption 2.1, it will be optimal for firm 0 to set m_0 such that the market remains covered in equilibrium.

Setting $q_0^{et} - q_0^{nt} = 0$ gives a threshold $\beta_0^{'}$ such that

$$\beta_0^{'} = \left\{ \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} \right\} \left\{ \frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} \right\}^{-1}. \tag{2.67}$$

So, if $\beta_0 \leq \beta_0^{'}$ then $q_0^{et} \geq q_0^{nt}$, otherwise $q_0^{et} < q_0^{nt}$. Similarly, the difference $q_0^{et} - q_0^{t}$ equals

$$q_0^{et} - q_0^{t} = \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} - \frac{\alpha\beta t}{2\gamma_m} + \beta_0 \left[\frac{1}{2} - \frac{\theta}{4t} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(1+\beta)t}{2\gamma_m} \right]. \tag{2.68}$$

Setting $q_0^{et} - q_0^{t} = 0$ gives a threshold $\beta_0^{''}$ such that

$$\beta_0^{''} = \left\{ \frac{\alpha\beta t}{2\gamma_m} - \frac{\alpha\theta^2}{8t\gamma_m} - \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} \right\} \left\{ \frac{1}{2} - \frac{\theta}{4t} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(1+\beta)t}{2\gamma_m} \right\}^{-1}. \tag{2.69}$$

So, if $\beta_0 \geq \beta_0^{''}$, then $q_0^{et} \geq q_0^{t}$, otherwise $q_0^{et} < q_0^{t}$. It remains to be shown that $\beta_0^{'}$, $\beta_0^{''} \in [0,1]$. If α is sufficiently small, then this will hold. Hence proved.

Proof of proposition 2.11

User welfare is written as

$$UW^k = \int_0^{\hat{x}^k} [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_1^k) - tx]dx + \int_{\hat{x}^k}^1 [V + I * \theta - \gamma_p q_0^k - \gamma_m(m_0^k + m_2^k) - t(1-x)]dx, \tag{2.70}$$

where $k = t, et$ or nt; $\hat{x}^k$ is the market share under regime k; m_j^k is the advertising level on firm $j \in \{0,1,2\}$ under regime k. This can be rewritten as

$$UW^k = V + I * \theta - \gamma_m[m_0^k + m_1^k\hat{x}^k + m_2(1-\hat{x}^k)] - \gamma_p q_0^k - t\left(\frac{(\hat{x}^k)^2}{2}\right) - t(1-\hat{x}_k) + t\left(\frac{1}{2} - \frac{(\hat{x}^k)^2}{2}\right). \tag{2.71}$$

The optimal value of m_1 and m_2 is $t/2\gamma_m$ under regime t and nt. Suppose, under regime et, there is exclusive data sharing between firm 0 and firm i. Then the optimal $m_i = (2t+\theta)/4\gamma_m$ and optimal $m_j = (2t-\theta)/4\gamma_m$. Using this, we can find the demand for each firm i and j. Under the exclusive regime, the demand for firm i ($\hat{x}^{et}$) is $(2t+\theta)/4t$ and demand for firm j ($1-\hat{x}^{et}$) is $(2t-\theta)/4t$. Using these market shares, the advertising levels given in equations (2.63) and (2.64), and the level of data exploitation given in equation (2.65), the user welfare under the regime et (UW^{et}) is $t/4+\theta^2/16t$. Whereas, user welfare under regime t and nt is $t/4$. Thus, user welfare is the highest under exclusive data sharing.

Social welfare under different regimes are

$$SW^{nt} = UW^{nt} + \alpha m_0^{nt} + [1-\alpha m_0^{nt}]\alpha[m_1^{nt}N_1^{nt} + m_2^{nt}N_2^{nt}] + Rq_0^{nt} - \frac{1}{2}(q_0^{nt})^2,$$

$$SW^{t} = UW^{t} + \alpha(1+\beta_0)m_0^{t} + [1-\alpha(1+\beta_0)m_0^{t}]\alpha(1+\beta)[m_1^{t}N_1^{t} + m_2^{t}N_2^{t}] + Rq_0^{t} - \frac{1}{2}(q_0^{t})^2,$$

$$SW^{et} = UW^{et} + \alpha m_0^{et}[(1+\beta_0)N_1^{et} + N_2^{et}] + [1-\alpha(1+\beta_0)m_0^{et}][\alpha(1+\beta)m_1^{et}N_1^{et}] + [1-\alpha m_0^{et}]\alpha m_2^{et}N_2^{et} + Rq_0^{et} - \frac{1}{2}(q_0^{et})^2.$$

The difference in social welfare under regime t and et is

$$
\begin{aligned}
SW^t - SW^{et} = & \frac{\alpha[V-t-\gamma_p R]}{\gamma_m}\left\{(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\} \\
& - \frac{\alpha[V-t-\gamma_p R]}{\gamma_m}\left\{\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2+(2t-\theta)^2]\right]\right\} \\
& + \frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right] - \frac{\alpha\theta}{2\gamma_m}\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2+(2t-\theta)^2]\right] \\
& + \frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m} - \frac{\theta^2}{16t} \\
> & \frac{\alpha[V+\theta-t-\gamma_p R]}{\gamma_m}\left\{(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\} \\
& - \frac{\alpha[V+\theta-t-\gamma_p R]}{\gamma_m}\left\{\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2+(2t-\theta)^2]\right]\right\} \\
& + \frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m} - \frac{\theta^2}{16t}.
\end{aligned}
$$

The R.H.S in the preceding equation is greater than 0 if $\beta_0 > \tilde{\beta_0}$, where

$$
\begin{aligned}
\tilde{\beta_0} = & \left\{\left[\frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m}\right]\left[1-\left(\frac{\alpha(V+\theta-t-\gamma_p R)}{\gamma_m}\right)^{-1}\right] - \frac{\theta^2}{16t}\right\} \\
& \left\{\frac{1}{2} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\theta}{4t} - \frac{\alpha(1+\beta)t}{2\gamma_m}\right\}^{-1}.
\end{aligned}
$$

Given the assumption that $\alpha < \gamma_m/(1+\beta)t$, it can be shown that $\tilde{\beta_0} < 0$. Hence, $SW^t > SW^{et}$.

Similarly, the difference in social welfare under regime *et* and *nt* is

$$SW^{nt} - SW^{et} = \frac{\alpha[V - t - \gamma_p R]}{\gamma_m} \left\{ \left[1 - \frac{\alpha t}{2\gamma_m} \right] - \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha\theta^2}{8t\gamma_m} \right\}$$

$$- \frac{\alpha[V - t - \gamma_p R]}{\gamma_m} \left[1 + \beta_0 \frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m} [(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2] \right]$$

$$- \frac{\theta^2}{16t} - \frac{\alpha\theta}{2\gamma_m} \left[1 + \beta_0 \frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m} [(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2] \right].$$

This can be rewritten as

$$SW^{nt} - SW^{et} = \left[\frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} \right] \left[\frac{\alpha[V - t - \gamma_p R]}{\gamma_m} - 1 \right]$$

$$- (\beta_0) \left[\frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} \right] - \frac{\theta^2}{16t}$$

$$- \left[\frac{\alpha\theta}{2\gamma_m} \right] \left\{ 1 + \beta_0 \frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m} [(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2] \right\}.$$

Since $\alpha < \gamma_m/(1+\beta)t$, the preceding equation is less than 0. Hence, $SW^{nt} < SW^{et}$.

Technology Adoption Game

Adoption decisions can lead to four different scenarios under technology sharing: both firms adopt the technology (aa); both firms reject the technology (rr); only firm i accepts the technology (ar); only firm j accepts the technology (ra). Firm i's profit under each scenario is

$\pi_i^{aa} = [1 - \alpha(1+\beta_0)m_0^{aa}]\alpha(1+\beta)m_i^{aa}N_i^{aa}$: Both firms accept the technology,

$\pi_i^{ar} = [1 - \alpha(1+\beta_0)m_0^{ar}]\alpha(1+\beta)m_i^{ar}N_i^{ar}$: Firm i accepts and firm j rejects the technology,

$\pi_i^{ra} = [1 - \alpha m_0^{ra}]\alpha m_i^{ra}N_i^{ra}$: Firm i rejects and firm j accepts the technology,

$\pi_i^{rr} = [1 - \alpha m_0^{rr}]\alpha m_i^{rr}N_i^{rr}$: Both firms reject the tehcnology.

The market share of firm i depends on whether other firm j accepts or rejects the technology. If both firms accept or reject, then they are identical in the eyes of users and the market share of each downstream firm is $\frac{1}{2}$. If firm i accepts the technology, then it gains a competitive advantage over firm j and $N_i^{ar} > N_j^{ar}$. Moreover, the technology adoption increases the targeting rate of firm i by β. The technology adoption also increases the targeting rate of firm 0 by β_0 over the users joining firm i, i.e., N_i. The values for m_i and m_0 under different scenarios are the same as in previous sections. The profit functions as given above can be used to determine the thresholds q^a and q^r.

Proof of proposition 2.12

At *stage 2b*, firms 1 and 2 decide whether to accept or reject the technology offer if firm 0 makes an offer. The firms move sequentially in the model and since the firms are symmetric, the order of moves doesn't affect the results. Each firm considers the impact of its decision on the expected profit that it can get. This depends on i) improvement in user utility, measured by θ, affecting the competitive position in the user market, and ii) improvement in targeting rate in the downstream market, measured by β, affecting the competitive position in the advertising market.

Suppose firm 0 has made an offer. Then, there are two different scenarios that firm $j \neq i$, $i, j = 1, 2$, which moves at a later stage, has to consider, i.e., whether firm i accepts or rejects the offer in the previous stage. If firm i accepts the offer, then firm j will also accept the offer if $\pi_j^{aa} \geq \pi_j^{ar}$. This inequality implies that, given firm i accepts, firm j's profit from accepting (π_j^{aa}) should be at least as large as its profit from rejecting the offer (π_j^{ar}). Equality of the two profits gives a threshold q^a such that

$$q^a = \frac{\gamma_m}{\gamma_p} * \left\{ \frac{V-t}{\gamma_m} - \frac{[1+\beta-(2t-\theta)^2/4t^2]}{\alpha[(1+\beta_0)(1+\beta)-(2t-\theta)^2/4t^2]} + \frac{\theta[(1+\beta_0)(1+\beta)-(2t-\theta)^2/8t^2]}{\gamma_m[(1+\beta_0)(1+\beta)-(2t-\theta)^2/4t^2]} \right\}. \quad (2.72)$$

Firm j will accept the offer, given that firm i also accepts the offer, if and only if the level of data exploitation is at least as large as the threshold q^a, i.e., $q_0 \geq q^a$.

Similarly, if firm i rejects the offer, then firm j will accept if its profit from accepting (π_j^{ra}) is at least as large as its profit from rejecting (π_j^{rr}), i.e., $\pi_j^{ra} \geq \pi_j^{rr}$. At equality, there is a threshold q^r such that

$$q^r = \frac{\gamma_m}{\gamma_p} * \left\{ \frac{V-t}{\gamma_m} - \frac{[(1+\beta)(2t+\theta)^2/4t^2 - 1]}{\alpha[(1+\beta_0)(1+\beta)(2t+\theta)^2/4t^2 - 1]} + \frac{\theta[(1+\beta_0)(1+\beta)(2t+\theta)^2/8t^2]}{\gamma_m[(1+\beta_0)(1+\beta)(2t+\theta)^2/4t^2 - 1]} \right\}. \quad (2.73)$$

Firm j will accept the offer, given that firm i rejects, if and only if the level of data exploitation is greater than or equal to the threshold value q^r, i.e., $q_0 \geq q^r$. A comparison of the two thresholds will show that $q^a > q^r$.

At *stage 2a)*, when firm i moves, its adoption decision will depend on the level of data exploitation. Three subcases can be defined:

(i) $0 \leq q_0 \leq q^r$: Firm j always rejects the offer.

(ii) $q^r < q_0 \leq q^a$: Firm j rejects if firm i accepts and firm j accepts if firm i rejects.

(iii) $q^a < q_0 \leq 1$: Firm j always accepts the offer.

Since downstream firms' profit functions are symmetric, firm i's choice will yield the same thresholds as for firm j. Under case (i), firm j always rejects the offer. Firm i will accept the offer if its payoff from accepting the offer (π_i^{ar}) is at least as large as the payoff from rejecting the offer (π_i^{rr}). So, for the relevant range of q_0, firm i will also reject the offer. Under Case (ii), firm i will accept the offer if $\pi_i^{ar} \geq \pi_i^{ra}$, where π_i^{ar} is firm i's profit if it accepts (and firm j rejects) and π_i^{ra} is firm i's profit if it rejects (and firm j accepts). This will require

$$q_0 \geq \frac{\gamma_m}{\gamma_p} \left\{ \frac{V+\theta/2-t}{\gamma_m} - \frac{[(1+\beta)(2t+\theta)^2 - (2t-\theta)^2]}{\alpha[(1+\beta_0)(1+\beta)(2t+\theta)^2 - (2t-\theta)^2]} \right\}.$$

For a sufficiently small α, R.H.S in the preceding equation is less than zero. So, firm i always accepts the offer if firm j rejects it in the subsequent stage. Under case (iii), firm i will accept the offer, given firm j always accept, if its payoff from accepting (π_i^{aa}) is at least as large as the payoff from rejecting (π_i^{ra}), i.e., $\pi_i^{aa} \geq \pi_i^{ra}$. This holds for the relevant range of q_0. It follows from the above discussion that both firms reject the offer under case (i); firm i accepts and firm j rejects the offer under case (ii); and both firms accept the offer under case (iii).

Having derived the equilibrium adoption decisions, the next question to ask is how the profit of firm i=1,2, is compared across these adoption decisions. We need to compare the profits under different scenarios.

i) It can be seen that, for a sufficiently large θ, $m_0^{aa} > m_0^{ar}$ and $m_i^{ar} > m_i^{aa}$. Therefore, putting in the values for advertising levels gives $[1-\alpha(1+\beta_0)m_0^{aa}]\alpha(1+\beta)m_i^{aa}N_i^{aa} < [1-\alpha(1+\beta_0)m_0^{ar}]\alpha(1+\beta)m_i^{ar}N_i^{ar}$, which means $\pi_i^{aa} < \pi_i^{ar}$.

ii) $\pi_i^{aa} \geq \pi_i^{rr}$ gives $[1-\alpha(1+\beta_0)m_0^{aa}]\alpha(1+\beta)m_i^{aa}N_i^{aa} \geq [1-\alpha m_i^{rr}]\alpha m_i^{rr}N_i^{rr}$. Putting in the values for m_i and N_i under different regimes, it can be shown that the inequality holds if $q_0 \geq q^{'}$, where

$$q^{'} = \frac{V-t}{\gamma_p} + \frac{\theta[(1+\beta)(1+\beta_0)]}{\gamma_p[(1+\beta)(1+\beta_0)-1]} - \frac{\beta\gamma_m}{\gamma_p\alpha[(1+\beta)(1+\beta_0)-1]}.$$

A few more calculations will show that $q^{'} > q^a$. Hence, this implies that

a) $\pi_i^{aa} \leq \pi_i^{rr}$ for $q^a < q_0 \leq q^{'}$,

b) $\pi_i^{aa} > \pi_i^{rr}$ for $q^{'} < q_0 \leq 1$.

iii) $\pi_i^{ar} \geq \pi_i^{rr}$ if $[1-\alpha(1+\beta_0)m_i^{ar}]\alpha(1+\beta)m_i^{ar}N_i^{ar} \geq [1-\alpha m_0^{rr}]\alpha m_i^{rr}N_i^{rr}$. This holds if $q_0 \geq q^r$, where q^r is as defined in (2.73). So, it always hold.

iv) $\pi_i^{aa} \geq \pi_i^{ra}$ if $[1-\alpha(1+\beta_0)m_0^{aa}]\alpha(1+\beta)m_i^{aa}N_i^{aa} \geq [1-\alpha m_0^{ra}]\alpha m_i^{ra}N_i^{ra}$. This gives $q_0 \geq q^a$, where q^a is as defined in (2.72).

v) Since $m_0^{ra} > m_0^{rr}$ and $m_i^{ra} < m_i^{rr}$, this implies $[1-\alpha m_0^{ra}]\alpha m_i^{ra}N_i^{ra} < [1-\alpha m_0^{rr}]\alpha m_i^{rr}N_i^{rr}$. Therefore, $\pi_i^{ra} < \pi_i^{rr}$.

It can be concluded that i) profits for firm i is highest when firm j≠i rejects the technology, and ii) the two firms can be in an equilibrium where profits are lower when both accept the technology than under no adoption scenario, i.e., there exists $q^{'}$ such that for firm i

a) $\pi^{ar} \geq \pi^{rr} \geq \pi^{aa} \geq \pi^{ra}$; for $q^{a} < q_0 \leq q^{'}$,

b) $\pi^{ar} \geq \pi^{aa} \geq \pi^{rr} \geq \pi^{ra}$; for $q^{'} < q_0 \leq 1$.

At *stage 1*, anticipating the effect of its choice on adoption decision and advertising levels in the subsequent stages, firm 0 decides whether or not to make an offer. The choice of firm 0 will depend on the level of data exploitation q_0 and hence, on three subcases defined in the last subsection. As defined previously, π_0^{t} is firm 0's profit when it makes an offer, and the technology is adopted by both firms, π_0^{nt} is firm 0's profit when the offer is rejected by both downstream firms or if it doesn't make an offer. A comparison of firm 0's profit under the two regimes is done to find the offer threshold. It will prefer offering the technology over not offering at all if $\pi^{t} \geq \pi^{nt}$. This gives a threshold q^{o} such that

$$q^{o} = V - t - \theta(1+\beta_0)[1-\alpha(1+\beta)t/2\gamma_m]C^{-1}, \text{ where} \tag{2.74}$$

$$C = 1 - \alpha t/2\gamma_m - (1+\beta_0)(1-\alpha(1+\beta)t/2\gamma_m). \tag{2.75}$$

So, if $q_0 \geq q^{o}$ then firm 0 will offer the technology to both downstream firms. It can be shown that this offer threshold is a decreasing function of β_0. When both downstream firms reject the offer, firm 0 optimally makes no offer as offering the technology doesn't change its payoffs.

Next, we need to analyse the signs of partial derivatives of offer and adoption thresholds.

The partial derivative of offer threshold q^{o} with respect to γ_p is

$$\frac{\partial q^{o}}{\partial \gamma_p} = 0,$$

Similarly, the partial derivative of offer threshold w.r.t. β is

$$\frac{\partial q^o}{\partial \beta} = \theta(1+\beta_0)C^{-1}\left\{\frac{\alpha t}{2\gamma_m} + \frac{(1+\beta_0)\alpha t}{2\gamma_m}\left[1 - \frac{\alpha t(1+\beta)}{2\gamma_m}\right]C^{-1}\right\} > 0,$$

where C is as defined in equation (2.75). Now, we need to evaluate the signs of the partial derivatives of adoption thresholds. To begin, the partial of q^a w.r.t. γ_p is

$$\frac{\partial q^a}{\partial \gamma_p} = -\frac{V-t}{\gamma_p^2} + \frac{\gamma_m[1+\beta-(2t-\theta)^2/4t^2]}{\alpha\gamma_p^2[(1+\beta)(1+\beta_0)-(2t-\theta)/4t^2]} - \frac{\theta[(1+\beta)(1+\beta_0)-(2t-\theta^2)/8t^2]}{\gamma_p^2[(1+\beta_0)(1+\beta)-(2t-\theta)/4t^2]},$$

Setting the preceding equation equal to 0. There exists a threshold β_0^{ao} such that

$$\beta_0^{ao} = \frac{(\theta)(2t-\theta)^2/8t^2 + \gamma_m[1+\beta-(2t-\theta)^2/4t^2]/\alpha - [(2t-\theta)^2/4t^2](V-t)}{(1+\beta)(V-t-\theta)} - 1, \quad (2.76)$$

and if $\beta_0 > \beta_0^{ao}$ then $\partial q^a/\partial\gamma_p < 0$. Also, setting the value of q^a given in (2.72) equal to 0, there exists a threshold β_0^a such that

$$\beta_0^{a} = \frac{(\theta)(2t-\theta)^2/8t^2 + \gamma_m[1+\beta-(2t-\theta)^2/4t^2]/\alpha + [(2t-\theta)^2/4t^2](V-t)}{(1+\beta)(V-t-\theta)} - 1, \quad (2.77)$$

and $q^a > 0$ if $\beta_0 > \beta_a$. A comparison of the two thresholds will show that $\beta_0^a > \beta_0^{ao}$. So, when $q^a > 0$ then $\partial q_0/\partial\gamma_p < 0$ for all $\beta_0 > \beta_0^a$.

The partial of threshold q^r w.r.t. γ_p is

$$\frac{\partial q^r}{\partial \gamma_p} = -\frac{V-t}{\gamma_p^2} + \frac{\gamma_m[(1+\beta)(2t+\theta)^2/4t^2-1]}{\alpha\gamma_p^2[(1+\beta)(1+\beta_0)(2t+\theta)^2/4t^2-1]} - \frac{\theta[(1+\beta)(1+\beta_0)(2t+\theta)^2/8t^2]}{\gamma_p^2[(1+\beta)(1+\beta)(2t+\theta)^2/4t^2-1]} < 0.$$

Setting the preceding equation equal to 0 gives a threshold β_0^{ro} such that

$$\beta_0^{ro} = \frac{\gamma_m[(1+\beta)(2t+\theta)^2/4t^2-1]/\alpha-(V-t)}{(1+\beta)[(V-t)(2t+\theta)^2/4t^2+\theta(2t+\theta)^2/8t^2]}-1, \tag{2.78}$$

and $\partial q^r/\partial\gamma_p < 0$ if $\beta_0 > \beta_0^{ro}$. Also, setting (2.73) equal to 0 gives a threshold β_0^r such that

$$\beta_0^{r} = \frac{\gamma_m[(1+\beta)(2t+\theta)^2/4t^2-1]/\alpha-(V-t)}{(1+\beta)[(V-t)(2t+\theta)^2/4t^2+\theta(2t+\theta)^2/8t^2]}-1, \tag{2.79}$$

and $q^r > 0$ if $\beta_0 > \beta_0^r$. It can be shown that $\beta_0^r > \beta_0^{ro}$. So, for all $\beta > \beta_0^r$, $\partial q^r/\partial\gamma_p < 0$.

The partial derivative of adoption thresholds with respect to β are

$$\frac{\partial q^a}{\partial\beta} = -\frac{\theta(1+\beta_0)(2t-\theta)^2}{8t^2[(1+\beta_0)(1+\beta)-(2t-\theta)^2/4t^2]} < 0,$$

$$\frac{\partial q^r}{\partial\beta} = -\frac{\theta(1+\beta_0)(2t+\theta)/8t^2+\beta_0\gamma_m(2t+\theta)^2/4\alpha t^2}{\gamma_p[(1+\beta_0)(1+\beta)(2t+\theta)^2/4t^2-1]} < 0.$$

These partials are less than zero. Hence proved.

Proof of proposition 2.13

In order to prove this proposition, we have to compare the profits that firm 0 can get from choosing either a) unconstrained investment levels or b) strategically increasing or decreasing the data exploitation to enforce technology sharing or no sharing. From figure 2.5, we can define four subcases:

Case i) When β_0 is low or intermediate and we are in the region where no offer is made (IV) or technology is adopted by both firms (I). In that case, firm 0 can choose the unconstrained level of data exploitation q_0^{nt} or q_0^t. In figure 2.5, firm 0 chooses q_0^t.

Case ii) When β_0 is intermediate, we are in region I. Then, firm 0 can either choose q_0^t or reduce it to 0 and do not offer the technology. Since for $q_0 > q''$, it is profitable for firm 0 to offer the technology, there is no incentive for it to deviate and choose q_0 equal to 0. So, it will choose the unconstrained level q_0^t.

Case iii) When β_0 is large, we are in region II, firm 0 can either choose q_0^t and enforce asymmetric adoption or it can increase q_0 to q^a and enforce adoption by both firms. It can be shown that it will choose q^a to enforce adoption by both firms.
Case iv) When β_0 is large, we are in region IV, firm 0 can either choose unconstrained level q_0^{nt} (provided it is less than q^r) or it can choose a higher q_0 to enforce technology sharing. In figure 2.5, it optimally chooses q^a to enforce adoption by both firms.

Hence, for large values of β_0, firm 0 can strategically choose a higher q_0 to enforce technology adoption, and thus data sharing.

Chapter 3

Regulation of Consumer Data: Privacy and Welfare

3.1 Introduction

Various online platforms can and do collect individualised user data whenever consumers go online for various purposes - to search content, view movies, listen to music, shop, do social networking, etc. - monetising this data in various ways. In particular such data can be used to boost the targeting efficacy of advertisements that appear on these platforms. To give just one example, user location data from mobile devices can be used to recommend specific restaurants in a user's vicinity. In fact, such data advantages can accrue to not just the platform collecting this data, but by all affiliated platforms owned by the firm that collects the data. It may be argued, for example, that data from Whatsapp and Instagram enhances the advertising efficacy of Facebook which owns these two platforms. Similarly, data from LinkedIn benefits Microsoft, and that from Android and Chrome is useful to Google.

This widespread collection and exploitation of user data is of serious concern from a regulatory viewpoint,[1] both because of privacy concerns,[2] as well as its potential anti-competitive abuses. The regulatory response in both the European Union (EU) and the US has focussed on the informed consent approach which emphasises that a firm cannot collect data without the users' consent. The recent General Data Protection Regulation (GDPR) by the EU extends this approach by specifying the *lawful purposes* for which data can be collected and used.

The literature, however, suggests that the informed consent approach by itself may not meet the twin challenges of privacy, and anti-competitive mis-uses. This can happen because of informational externalities (Choi et al. (2019)), data collection disproportionately affecting small firms (Campbell et al. (2015)), countervailing price adjustments (Preibusch et al. (2013) and Strahilevitz and Kugler (2016)), etc. The policy literature has also criticised the informed consent approach, e.g., Solove (2004, 2012), Lenard and Rubin (2010, 2015), etc. These papers have focussed on cognitive problems such as bounded rationality and structural problems such as assessment of harm at an individual level vis-à-vis social harm that reduces the efficacy of privacy self management. Moreover, even in the absence of such consent a data controller can sometimes collect and use data for *legitimate interest or necessity*, especially when it controls multiple apps. For instance, Google can override the consent requirement through collecting location data from other apps even if one selects the do-not-track option in Android phones.

Given the potential limitations of the informed consent approach,[3] in this study we examine two complementary regulatory policies. One possibility is to strengthen the informed consent approach by empowering the users, allowing them to control the potential uses of such personal data. Going back to our earlier example, under this

[1]The economic analysis of regulation and its impact on competition and welfare has been studied extensively, most notably environmental regulation (e.g., Jaffe and Palmer (1997), Bansal (2008), Goldschlag and Tabarrok (2018)). However, the debate about privacy regulation is still at a nascent stage, and information practices of online firms, data sharing in particular, requires further investigation.

[2]Recently, Acquisti et al. (2013) examines the value that people place on privacy. Brill (2011) and Kerber (2016) emphasize on the interaction of consumer protection and competition law in privacy protection regulation.

[3]Lerner (2014) discusses how data collection and generation affects the firms and users, in particular whether aggressive antitrust intervention is required.

approach Google cannot use the location data collected by it for any other purpose (other than the specific one for which it was collected), if the consumer so decides. This can be thought of as formalizing the recently adopted General Data Protection Regulation (GDPR) that mandates, for example, that in addition to giving consent for using their personal data, consumers have other rights, e.g. right to access data, right to restrict processing, right to data portability, right to have their data deleted, etc.[4] In a similar spirit, the German competition commission have ruled that Facebook cannot merge information obtained from Facebook accounts and its other services, like WhatsApp and Instagram, without user permission.[5,6]

The second policy is motivated by the fact that many regulatory authorities are contemplating whether to challenge previous mergers and acquisitions by large technology firms such as Google, Facebook, Amazon, and Apple, e.g., those between Whatsapp and Facebook, as well as LinkedIn and Microsoft. Many of these mergers were approved based on a price-oriented approach. However, some regulatory authorities now seem to be coming around to the view that such a price centric approach under-estimates the long-term competitive advantage such merged firms can obtain by combining their data sets. For example, the German Monopolkomission suggests that "it is conceivable that the concentration-related combination of data stocks on the platform of an acquirer enables its operator to prevail over competitors solely by virtue of permanently having superior knowledge, for e.g., of the user preferences."(see Monopolkomission (2015)). Motivated by such concerns, we analyse a scenario where merged firms do not have access to the data collected by its subsidiaries.

While it is clear that prevention of data-sharing, using either these two policies,

[4]Another aspect of GDPR is that explicit consent makes consumers aware of the ways in which their personal data is collected and used which they might find privacy intrusive. So, they tend to become increasingly familiar with the costs and benefits of data collection. In addition, GDPR can also make price discrimination based on personal data more difficult. Z.Borgesius and Poort (2017) argue that "personalised pricing generally entails the processing of personal data. Data protection law requires a company to inform people about the purpose of processing their personal data. Hence, companies must say so if they personalise prices; such transparency may impact the practice and outcome of personalised pricing."

[5]see https://www.nytimes.com/2019/02/07/technology/germany-facebook-data.html

[6]Informed user approach can have similarities with policy approaches in environmental regulation that focus on informing consumers about the environmental impact, e.g., lablelling GM foods (Bansal et al. (2013)).

can help in preventing the abuse of monopoly power arising from access to data, the welfare implications however are not clear given the various trade-offs involved. One well known and important trade-off arises from the fact that access to more personalized data may help improve the quality of services themselves. In our study, we introduce another possible trade-off arising from the fact that privacy regulations can affect the market structure itself, in particular the business models adopted by the affected firms. The issue of regulations affecting the business models is relevant given that various digital firms adopt different business models. In some markets, the competing platforms adopt business models that are similar. For example, in case of Netflix and Amazon prime videos, both adopt a subscription based business model, whereas in case of Google and Bing, both charge their advertisers, but not the users. Whereas in other markets, different firms adopt different business models. For instance, in the social networking space, while Facebook's business model relies on financing its operations through advertisements, Tencent (in China) has focussed on selling value added services to users.

In this study, we therefore start by setting up a two-sided framework that allows for endogenous choice of business models, as well as data collection in related markets by some firms, but not by others. We then use this framework to examine the two policy proposals of interest. Formally, we develop a game theoretic model in which there are two markets 1 and 2, two firms G and S, a set of users (who dislike advertisements) and a set of advertisers. Firm G is a monopolist in market 1 and collects data regarding the users who join it in this market. Firm G however competes with firm S in market 2 over users and advertisers, with the data it collects in market 1 enabling firm G to offer improved targeting of advertisements in market 2 vis-à-vis firm S. In order to explain the competitive effects, we focus on two important parameters - targeting rate of an advertisement (β), and the extent of data collection and exploitation (k).

Our analysis provides a theory of business models in platform markets, finding that it is intrinsically linked to the possibility of data collection, as well as the efficacy of advertising targeting, i.e., β. We find that when β is large, both firms choose to employ a advertising financed business model, charging the advertisers but not the consumers. Whereas, if β is small, they adopt an user financed model, charging only the users. For intermediate values of β, however, there will be strategic differentiation (Calvano and Polo (2016)) with firm G opting for advertisement financing, whereas firm S opts for user financing. These results are intuitive. For β large, the potential net surplus from advertising is large, so that the firms do not charge the consumers in market 2 so as to ensure that their market share is as large as possible, making their profits from the advertisers instead. Similarly, for β small. For β intermediate, the divergence in business strategies arise from the fact that G has a relative advantage in advertisement targeting because of its presence in market 1, and hence prefers advertisement financing.

The analysis suggests that advertising financing will be adopted by firms which collect consumer data, e.g., those in social media and search markets. Consistent with this proposition, we find that Google, which competes with Bing (owned by Microsoft), and Facebook, which competes with Twitter, both compete on the basis of advertising and provide free services to users. Whereas, in the market for office software applications, there is strategic differentiation in that while Microsoft, one of the key competitors, adopts user financing for its office products, Google's G suite is a free online software with financing coming from advertisements. Note that in line with our analysis, while Microsoft commits to privacy protection, Google can and does collect information about users. In online music and video streaming market, the major players such as Amazon Prime and Netflix have adopted user financing.

Next, we analyse the two policy proposals discussed above. We find that user control of data collection can be potentially beneficial, especially in markets where the targeting technology is very effective (large β). This is because with targeting being already quite effective, data collection adds little to targeting ability while imposing privacy costs on users. Consequently, user control reduces the aggregate amount of information collected and this need not be offset by higher prices to users, so that there is an increase in both user and social welfare. On the other hand, we find that restricting access to data reduces both user, as well as social welfare. Intuitively, restricting access to data reduces the efficacy of advertisements which, in turn, leads firms to charge higher user prices reducing both user and social welfare.

3.1.1 Related literature

The literature on two sided markets is vast,[7] and we briefly discuss the strands of literature that are most relevant for the present paper. First, one strand of the literature examines the competitive interaction between firms in two-sided markets where the competing firms have different business models, in particular in the context of the pay TV and free to air TV business models, e.g., Peitz and Valletti (2008), Armstrong and Weeds (2007), Anderson and Jullien (2016), and Thöne et al. (2016). In contrast, our study endogenously solves for the business model.

Second, there is another strand of literature where the data collection *technology* is itself endogenous, e.g., Bloch and Demange (2018), Casadesus-Masanell and Hervas-Drane (2015), Dimakopoulos and Sudaric (2018), and Lefouili and Toh (2017). By way of contrast, we assume that the data collection technology is exogenously given, though the aggregate amount of data collection is of course endogenous, and depends on the number of users joining firm G in market 1. Moreover, unlike this literature, we endogenize the business model chosen by the firms.

The papers closest to ours are Calvano and Polo (2016) and Prüfer and Schottmuller (2017), both of whom allow for endogenous determination of the market structure. Calvano and Polo (2016) was the first paper to show that *strategic differentiation*, with different firms adopting different business models, can emerge as an equilibrium phenomenon. The result in Calvano and Polo (2016) is, however, driven by their targeting technology which ensures that multi-homing users are easier to inform. In contrast, we show that strategic differentiation, if it arises in equilibrium, is driven by the fact that some firms have greater access to data. Moreover, unlike Calvano and Polo (2016), we do not allow for multi-homing.

[7]See, among others, Caillaud and Jullien (2003), and Hagiu (2006), for papers on two sided markets.

In a recent paper, Prüfer and Schottmuller (2017) study a dynamic model of R&D competition in which the authors make the important observation that firms can leverage user data to enter new markets - the *data driven domino effect*. However, our study differs from their study in several respects. First, the focus of the two papers is different. While Prüfer and Schottmuller (2017) examine the data driven domino effect, in the present paper, we investigate the competitive and welfare implications of data collection from a connected market so as to improve advertising targeting in the core market. Moreover, we examine whether asymmetric business model equilibria exist or not. Second, the modelling framework is different in the two papers. For one, Prüfer and Schottmuller (2017) set up a demand side model in which additional user information reduces the firm's cost of innovation, whereas we employ a two-sided market framework in which user information is used to improve targeting rate. For another, in JS the strategic variable is investment in quality, whereas in our model firms choose user prices and advertising quantities.

Turning to the policy literature, a related study is Mccan and Hall (2018) that also goes beyond the informed consent approach and explicitly considers the role of banning data-sharing as an alternate mechanism to protect consumer privacy. In a similar way, our study also argues that banning data-sharing improves consumer privacy. However, unlike their paper, we also bring out the trade-off resulting from the impact on user pricing which could be detrimental to overall user and social welfare.

The rest of the study is organized as follows. Section 3.2 sets up the baseline model. Section 3.3 discusses the market outcomes and welfare. Section 3.4 examines policy proposals to protect privacy. Section 3.5 discusses a few extensions of the baseline model. Finally, Section 3.6 concludes. All proofs are in the appendix.

3.2 The Model

We consider a framework with two markets 1 and 2, and three sets of agents, namely two firms, G and S, unit mass of users, and a unit mass of advertisers. Users decide which of the firms to join in the two markets, if at all, advertisers decide which firm(s) to advertise on in market 2, and firms decide on the prices they charge their users, as well as how much advertising to allow.

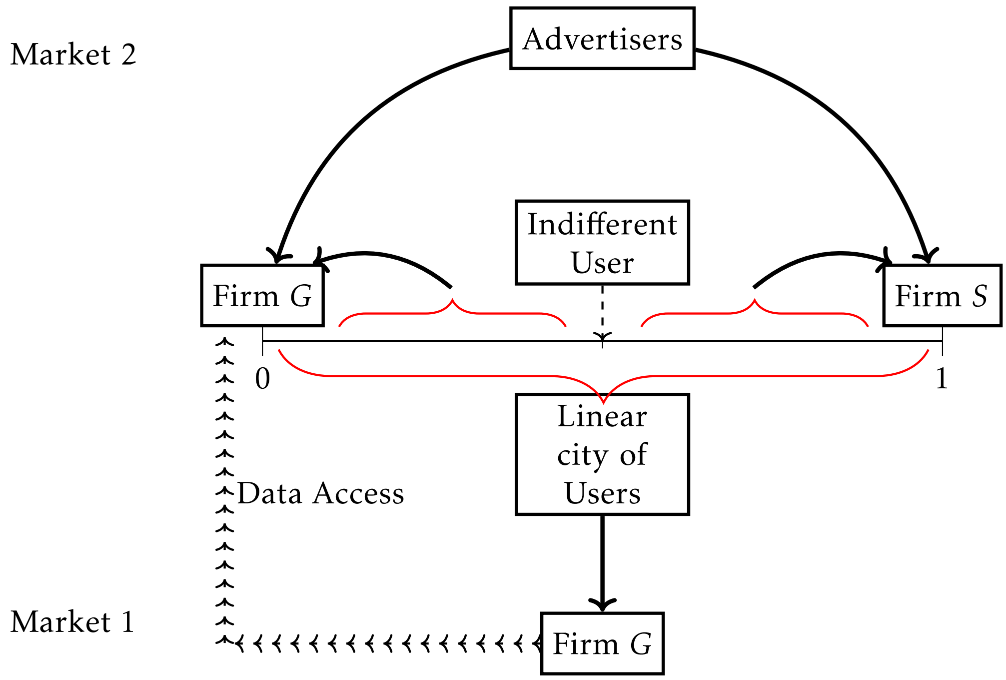


Figure 3.1: Market structure

■ **Firms.** Firm G is a monopolist in market 1, whereas firms G and S compete à la Hotelling in market 2 with firm G being located at point 0 and firm S at point 1 of the unit Hotelling interval. Firm G charges its users q_{Gi} in market i, $i = 1, 2$, and firm S charges its users q_{S2} in market 2. We assume that all prices are non-negative, i.e., q_{Gi}, $q_{S2} \geq 0$, $i = \{1, 2\}$. This is realistic since typically in these markets firms offer services for free and do not provide subsidies to the users (notably Google and Facebook).[8] We next define some notations for future use:

- N_{G1} denotes the number of users joining firm G in market 1.
- N_{i2}, $i = G, S$, denotes the number of users joining firm i in market 2.
- p_{i2}, $i = G, S$, denotes the price per unit of advertisement in firm i in market 2.

[8]If we relax this assumption then, in additon to the results that we obain in next section, there might exist market equilibrium in which firms adopt a mixed business model with advertisments and subsidies to the users.

On the advertising side, given p_{i2}, firms decide on their demand for advertisement, m_{i2} $i = \{1,2\}$, and advertisers decide on whether to advertise or not in any given firm. Firm G also collects data in market 1, where the amount of data collected per user is denoted k, $k > 0$ being exogenous.

Let π_i denote the profit of firm i, $i = G, S$, which includes its revenue collection on both consumers side, as well as from the advertisers. Thus, we have that:

$$\pi_G = q_{G1}N_{G1} + q_{G2}N_{G2} + p_{G2}m_{G2}, \tag{3.1}$$

$$\text{and, } \pi_S = q_{S2}N_{S2} + p_{S2}m_{S2}. \tag{3.2}$$

■ **Users.** There is a unit mass of users with each user being defined by a pair (v,c), where v denotes her gross utility from joining firm G in market 1, and c defines her location on the Hotelling line in market 2, where (v,c) is uniformly distributed over the unit square $[0,1]\times[0,1]$. Data collection by firm G imposes a privacy cost of γk on all consumers who join it in market 1, where $\gamma > 0$ is per unit privacy cost. A (v,c) type user gets a gross utility of $v\theta$ in market 1, where $\theta > 0$, so that her net utility in market 1 is

$$U_{G1}(v) = v\theta - q_{G1} - \gamma k. \tag{3.3}$$

Next, consider the user (v,c)'s utility in market 2. There is no multi-homing in the model, and thus she either joins firm G, or firm S.[9] There is no intrinsic quality difference between the two firms however and thus the user obtains a gross utility of X, irrespective of which firm she joins. However, her transportation cost is τc, if she joins firm G, and $\tau(1-c)$, if she joins firm S. Moreover, users have an aversion for advertisement, captured by a per unit disutility of t from advertising, where $t > 0$, so that her disutility from advertisement is tm_{G2}, if she joins firm G, and tm_{S2}, if she

[9]While in reality of course there is some degree of mult-ihoming, our assumption of single-homing is without loss of generality, since we find that our results remain qualitatively the same even if we allow for multi-homing.

joins firm S.[10,11] Thus, her net utility in market 2, $U_{i2},\ i=\{1,2\}$, equals

$$\begin{cases} X - tm_{G2} - q_{G2} - \tau c, & \text{if she joins firm } G, \\ X - tm_{S2} - q_{S2} - \tau(1-c), & \text{if she joins firm } S. \end{cases}$$

Note that nuisance and transportation costs, through their effect on the equilibrium advertising levels, affect the consumer demand for each platform in market 2. This brings in a strategic interdependence among firms while choosing their advertising levels. The reservation utility of all users is normalized to zero in *both* markets, 1 and 2.

■ **Advertisers.** There is a unit mass of identical advertisers all of whom want to generate "attention" for their product or service through placing advertisements in firm G and/or S in market 2. The advertising market is competitive, with the two firms, G and S, as well as the advertisers being price takers in this market. Thus, given the market prices the two firms decide on their demand for advertisement, and the advertisers decide on their supply. In equilibrium, prices adjust to ensure that demand equals supply. This entails that in equilibrium, p_{i2} equals the advertisers' marginal benefit from a unit of an advertisement in firm i, so that advertisers are indifferent between advertising and not advertising.

We next turn to deriving the advertisers' net payoff from placing advertisements in firms G and S in market 2. We normalize the advertisers' gross payoff from informing a single user to 1. Thus, the expected revenue per unit of an advertisement from any given user in firm i is

$$1 \times \text{"Probability of informing this single user"}.$$

The probability of informing a user in the model also captures the network effects on the advertiser side. It depicts the marginal increase in value from placing advertisements on a platform, when an additional user joins that platform. The larger the user

[10] A few media studies have found that advertising reduces a user's utility (see for e.g., Wilbur (2008)). Other theoretical work that has also characterised advertising as a nuisance to users include Anderson and Coate (2005).

[11] While the nuisance cost of advertising tm_{i2} is assumed to be linear, the results remain qualitatively the same for convex nuisance costs.

base, the higher the value from placing advertisements on the platform. We next note that the "probability of informing a single user" should naturally depend on whether the concerned firm has individualised data regarding the user or not. First, consider firm S. Recall that firm S has no such data, given that she has no presence in market 1 which is the source of such data in our framework. Thus, for firm S, the probability that an advertisement informs a single user is given by a baseline advertising efficiency parameter β, $0 < \beta \leq 1$. Thus, an advertiser's gross advertising revenue from placing m_{S2} advertisements in firm S equals

$$\beta m_{S2} N_{S2}. \tag{3.4}$$

Next, consider firm G which, however, does have access to such individualised data for the users who join it in both the markets, collecting k units of data per consumer in market 1. Thus, for these consumers, the probability that an advertisement in firm G informs them is higher at $\beta + k(1-\beta)$. Recalling that the number of users who join firm G in market 1 is N_{G1}, out of the N_{G2} users who join firm G in market 2, $N_{G1}N_{G2}$ are present in market 1 as well. Thus, the total revenue to an advertiser from these users is $[\beta + k(1-\beta)]m_{G2}N_{G1}N_{G2}$. Whereas, firm G has no data advantage from the users not joining firm G in market 1, i.e., a total of $(1-N_{G1})N_{G2}$ users. Thus, the total revenue from these users is $\beta m_{G2}(1-N_{G1})N_{G2}$ for an advertiser. Hence, the total gross advertising revenue from placing m_{G2} advertisements in firm G is[12]

$$[(\beta + k(1-\beta))N_{G1} + \beta(1-N_{G1})]m_{G2}N_{G2} = [\beta + k(1-\beta)N_{G1}]m_{G2}N_{G2}. \tag{3.5}$$

From the preceding equation, we can see that advertising revenue on firm G depends on the aggregate data collected in market 1, which in turn depends on the pricing regime in market 1. Thus, the cross market data advantage establishes the interdependence between firm G's revenue in the two markets.

■ **Timing of the game.** We solve for the subgame perfect Nash equilibrium (henceforth equilibrium) of the following two-stage game:

Stage 1: Firms G and S simultaneously choose user prices and the quantity of

[12] In the derivation of advertising revenue it is assumed that the probability of informing a single user remains between $[0,1]$, i.e. $\beta m_{S2} < 1$ and $[\beta + k(1-\beta)N_{G1}]m_{G2} < 1$. This will hold under the parametric condition that $t \geq \tau$, which is ensured by assumption 1 later.

advertising, i.e., firm G chooses q_{G1}, q_{G2}, and m_{G2}, and firm S chooses q_{S2} and m_{S2}. These choices become public knowledge.

Stage 2: Users decide whether to join firm G or not in market 1, and which firm to join in market 2. Simultaneously advertisers decide which platform to advertise on in market 2, G, or S, or both.

Let a market equilibrium be denoted by $\left\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\right\}$.

3.3 Equilibrium Analysis

3.3.1 Market Outcome

We assume that X, i.e., the users' gross utility in market 2 from joining either firm, is sufficiently large relative to the per unit transport costs (τ), so as to ensure full market coverage. This of course is the case of interest, since there is no strategic interaction between the two firms otherwise. In addition, we impose parametric restrictions such that the probability of match remains between [0,1] in equilibrium. This assumption is in line with the empirical observation that advertising technology is not efficient enough to ensure perfect targeting. So, the nuisance cost of advertisements t is sufficiently large relative to the per unit transport cost τ.

Assumption 1 (Full Market Coverage): $\min\{\frac{2X}{3}, t\} \geq \tau$.

Further, for ease of exposition we assume that joining firm G in market 1 is attractive. Formally, we have

Assumption 2: $\theta > k\left[\frac{\tau}{2t} + \gamma\right]$.

We relax this assumption in Appendix IV, showing that the results go through qualitatively.

■ **Stage 2.** Recall that in market 1, a user has the option of either joining firm G or not, whereas in market 2, she can decide which firm to join, G or S (there is no multi-homing, and, given Assumption 1, the user always joins either firm G, or firm S). Since, from a user's point of view, these two decisions are independent, the demand functions in each market can be derived separately.

In market 1, let the user $\hat{v}$ be indifferent between joining and not joining, so that

$$U_{G1}(\hat{v}) \equiv \hat{v}\theta - q_{G1} - \gamma k = 0 \Rightarrow \hat{v} = \frac{q_{G1}}{\theta} + \frac{\gamma k}{\theta}. \tag{3.6}$$

All users with a valuation of v that exceeds $\hat{v}$ purchases, so that the demand for firm G in market 1 is

$$N_{G1} = 1 - \frac{q_{G1}}{\theta} - \frac{\gamma k}{\theta}. \tag{3.7}$$

In market 2, let the user $\hat{c} \in [0,1]$ be indifferent between joining firm G and firm S, i.e., $U_{G2}(\hat{c}) = U_{S2}(\hat{c})$. Thus, the demand for firm G, denoted N_{G2}, and that for firm S, denoted N_{S2}, are

$$N_{G2} = \hat{c} = \frac{1}{2} + \frac{tm_{S2} - tm_{G2}}{2\tau} + \frac{q_{S2} - q_{G2}}{2\tau}, \tag{3.8}$$

$$\text{and, } N_{S2} = 1 - \hat{c} = \frac{1}{2} + \frac{tm_{G2} - tm_{S2}}{2\tau} + \frac{q_{G2} - q_{S2}}{2\tau}. \tag{3.9}$$

Next, consider the market for advertising. As discussed earlier, market clearing entails that the marginal benefit of a unit of advertisement in firm i is equal to its marginal cost p_{i2}. Thus, given (3.4) and (3.5), the inverse advertising demand functions are[13]

$$p_{G2} = [\beta + k(1-\beta)N_{G1}]N_{G2}, \tag{3.10}$$

$$\text{and, } p_{S2} = \beta N_{S2}. \tag{3.11}$$

■ **Stage 1.** At stage 1, using the inverse advertising demand functions defined in equations (3.10) and (3.11) and user demand functions defined in equations (3.7), (3.8) and (3.9), the two firms maximize their profits with respect to user prices and

[13] More generally, there can be diminishing returns to advertisements on each platform i.e. the marginal advertising revenue would fall with increase in advertisements. In this paper, however, we have simplified the setting by assuming constant returns to advertisements.

advertising quantities.

We then characterize the business model being followed by the two firms. For ease of exposition, we divide all possible business models into the following four classes depending on which side of the market is being charged in market 2:

(i) *Advertising financed* (A): both firms charge only the advertisers in market 2.

(ii) *User financed* (U): both firms charge only the users in market 2.

(iii) *Strategic differentiation* (S): firm i charges one side and firm $j \neq i$ charges the other side of the market in market 2.

(iv) *Mixed business model* (M): any other possible price configuration.

We can now characterize the optimal strategies of the two firms.

Proposition 3.1. *Let Assumptions 1 and 2 hold.*

(i) There does not exist an equilibrium where either firm charges both sides of the market.

(ii) There exists β_S and $\beta_G(k)$ where $0 \leq \beta_S, \beta_G(k) \leq 1$, $\beta_G(k)$ is decreasing in k, and $\beta_S = \beta_G(0)$, such that given a market characterized by (β, k), we have that:

(a) If $\beta_S < \beta \leq 1$, the equilibrium is advertising financed, i.e., both firms charge only the advertisers, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{ \left(Max\left\{ \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)\tau}{4t}, 0 \right\}, 0, \frac{\tau}{t} \right); \left(0, \frac{\tau}{t} \right) \right\}.$$

(b) If $\beta \leq \beta_G(k)$, the equilibrium is user financed, i.e., both firms charge only the users, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{ \left(Max\left\{ \frac{\theta}{2} - \frac{\gamma k}{2}, 0 \right\}, \tau, 0 \right); (\tau, 0) \right\}.$$

(c) If $\beta_G(k) < \beta \leq \beta_S$, the equilibrium involves strategic differentiation with firm G being advertiser financed, and firm S being user financed, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{ \left(Max\left\{ \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)\tau}{4t}, 0 \right\}, 0, \frac{\tau}{t} \right); (\tau, 0) \right\}.$$

Figure 3.2 describes the choice of business models in $k - \beta$ space. Note that the marginal return from advertising is β for firm S, whereas for firm G, it is $\beta + k(1-\beta)$ if the user joins it in market 1, and β otheriwse. Therefore, for β large, the marginal return from advertising is large for both firms, and both opt for a business model that is advertising financed. Whereas, for β small, the marginal revenue from advertising is low, and both firms opt for user financing. Interestingly, when β is at an intermediate level, firm G finds it optimal to opt for advertiser financing, while firm S finds it optimal to opt for user financing, since the marginal return from advertising on firm G is higher because of its greater access to data.

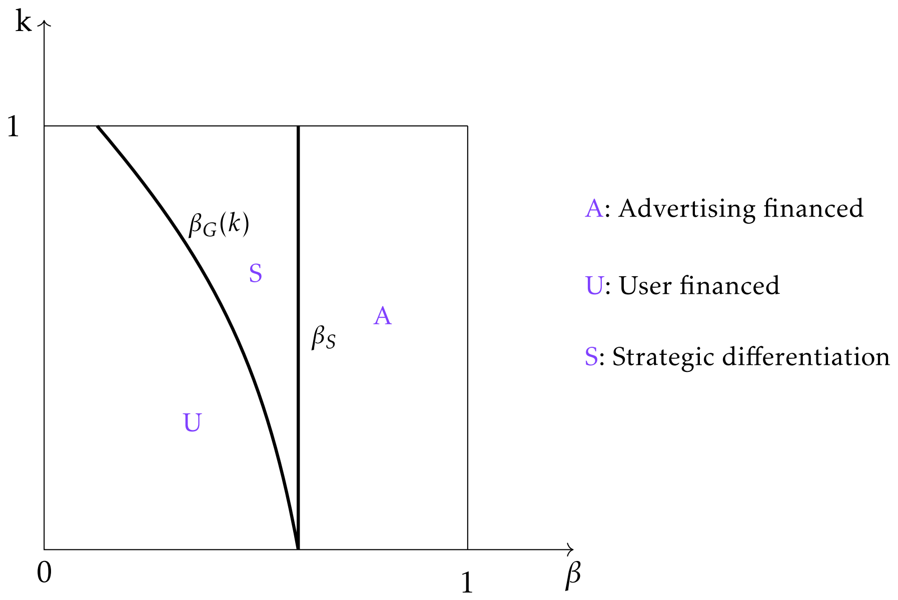


Figure 3.2: Choice of business model ($\gamma = 0.15$, $t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.6$)

As we discussed in the introduction, Proposition 3.1 is consistent with reality. Recall, for example, that Google and Bing (owned by Microsoft), both are advertising financed and do not charge the users. We next provide an example of strategic differentiation.

■ **Strategic differentiation in the market for professional social networking services.** Consider LinkedIn (now owned by Microsoft) and XING which compete in this market. In its competition analysis of Microsoft/ Linkedin merger (see Case M.8124 Microsoft / LinkedIn, para 350), the European Commission highlighted the differences in LinkedIn and XING's privacy policy. It says that "during the registration process, XING asks users to actively accept XING's privacy policy and Terms & Conditions by ticking a box, whereas LinkedIn users accept LinkedIn's privacy policy automatically when they press the button *join now*."[14] Consistent with our analysis, we find that while LinkedIn offers a lower degree of privacy protection than its competitor and relies primarily on advertiser financing, XING, which provides more consumer protection, relies mainly on premier subscription.

3.3.2 Welfare

We next turn to comparing the consumers' surplus and aggregate welfare under the various business models.

■ **Consumers' surplus.** The consumers' surplus is given by:

$$UW^* = \int_{\frac{q^*_{G1}}{\theta}+\frac{\gamma k}{\theta}}^{1} [v\theta - q^*_{G1} - \gamma k] dv + \int_0^{\hat{c}} \left[X - tm^*_{G2} - q^*_{G2} - \tau c\right] dc + \int_{\hat{c}}^{1} \left[X - tm^*_{S2} - q^*_{S2} - \tau(1-c)\right] dc, \quad (3.12)$$

where the optimal value of strategic variables depends on the equilibrium business model (see Proposition 3.1). Interestingly, we find that consumers' surplus is maximal when firm G adopts an advertising financed business model.

Proposition 3.2. *Let Assumptions 1 and 2 hold. Given any market* (β, k), *consumers' surplus under an advertising financed business model equals that under a strategic differentiation equilibrium, and both exceed that under a user financing equilibrium.*

This proposition is intuitive. Consider market 2. Under user-financing, while consumers have to pay a positive price in market 2, they are not subject to any nuisance

[14] It continues, "Moreover, when XING introduces new services which have an implication on how it collects and/or uses its members' data, it explicitly seeks active consent from the members. In addition, regardless of whether members give their consent in such specific cases or not, they will be able to continue to use XING as such without losing any of the functions to which they previously had access. In contrast, when LinkedIn makes changes to its collection, storing, processing or usage of personal data, LinkedIn only informs the members of those changes and considers that LinkedIn members agree with those changes, if they continue to use LinkedIn's services after they have been notified of the changes".

cost from advertising. Whereas, under other market forms, while consumers may not pay any price for joining, they may have to endure advertisements. Thus, all business models involve a trade-off. In market 1, however, under either advertising financing, or strategic differentiation, firm G adopts advertisement as a source of revenue. This incentivises firm G to lower its prices in market 1, since by doing so it can induce more users to join (thus gathering a greater amount of data which it can monetise in market 2). It is this lower price for users in market 1 in case of advertising financing by firm G that leads to a larger consumer surplus in this case.

■ **Aggregate welfare.** The aggregate welfare from adopting alternate business models is defined as the sum of consumers' surplus, advertisers' profit and firms' profit. Since prices are just transfers in the model we have that

$$SW^* = \int_{\frac{q^*_{G1}}{\theta}+\frac{\gamma k}{\theta}}^{1} [v\theta - \gamma k]\, dv + \int_0^{\hat{c}} \left[X - tm^*_{G2} - \tau c\right] dc + \int_{\hat{c}}^{1} \left[X - tm^*_{S2} - \tau(1-c)\right] dc$$
$$+ [\beta + k(1-\beta)N^*_{G1}]m^*_{G2}N^*_{G2} + \beta m^*_{S2}N^*_{S2}. \quad (3.13)$$

Proposition 3.3. *Let Assumptions 1 and 2 hold, and consider a market* (β, k)*.*

(i) For $\beta > \beta_S$*, social welfare is highest in case the equilibrium is advertising financed.*

(ii) For $\beta \leq \beta_S$*, there exist a threshold* $\beta_E(k)$*, where* $0 \leq \beta_E(k) \leq \beta_S$*, and* $\beta_E(k)$ *is decreasing in* k*, such that*

(i) for $\beta_E(k) < \beta \leq \beta_S$*, social welfare is highest under a strategic differentiation equilibrium, whereas*

(ii) if $0 \leq \beta \leq \beta_E(k)$*, then social welfare is maximal under a user financed equilibrium.*

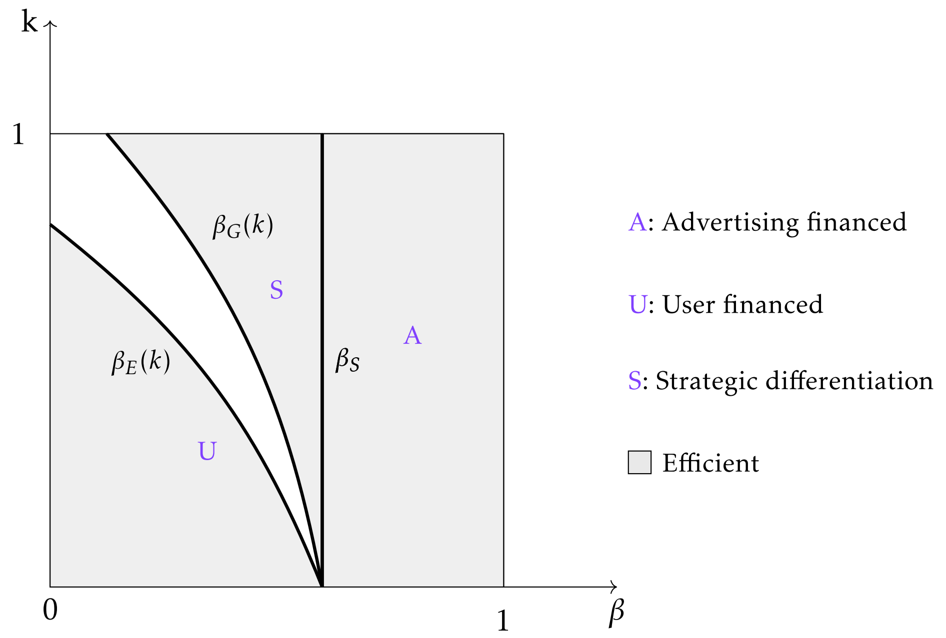


Figure 3.3: Social welfare comparison ($t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.6$)

Figure 3.3 illustrates the preceding proposition.. Why does advertising financing maximize social welfare for β large? In this case, the surplus generated from advertisement targeting is large, which is best exploited using an advertisement financed business model, since this induces firm G to follow pricing strategies that ensure that the number of consumers joining firm G in market 1 is large. Whereas, for β small, advertising surplus net of nuisance costs is low vis-à-vis the privacy costs. Under user financing there is no advertisement, which however is not a great loss to aggregate surplus since advertising surplus is low enough to begin with. On the other hand, user financing has the advantage that with no advertising revenue to be had, firm G has little incentive to induce a large number of users to join in market 1, thus lowering aggregate privacy costs. For intermediate values of β, aggregate surplus is highest under strategic differentiation. While β is large enough such that firm G doing advertisement is good, for firm S, it is best that they do user financing.

3.4 Privacy Regulations

With an explosion in data collection, privacy concerns have also exacerbated. We next use the baseline model to understand how various regulations to protect user privacy would affect the market structure, and consequently consumers' surplus and welfare. In particular, we consider two different regulatory approaches: (a) user control of data collection, and (b) restricting access to data owned by subsidiaries.

3.4.1 User Control of Data Collection

Recall from our earlier discussion, that user control of data is one of the aspects that the GDPR focuses on. In this sub-section, we examine how such user control affects welfare. In our framework, this can be formalized as the users having control over the level of k. We modify the existing game so that the game begins in an initial stage, call it stage 0, where the users simultaneously and non-cooperatively decide on their own choice of k. The rest of the game tree is identical to that in the baseline model. We begin by establishing that in equilibrium all users optimally set $k = 0$.

Proposition 3.4. *Let Assumptions 1 and 2 hold and consider the new game with users simultaneously deciding about the choice of k under a non discriminatory pricing regime. Then there exists a unique equilibrium in which all users set $k = 0$.*

Next, using (3.12), we have that the change in user welfare following a change in k, i.e.,

$$\frac{\partial UW}{\partial k} = -\left[1 - \frac{q^*_{G1}}{\theta} - \frac{\gamma k}{\theta}\right] * \left[\frac{\partial q^*_{G1}}{\partial k} + \gamma\right], \tag{3.14}$$

where recall q^*_{G1} is the optimal value of user price in market 1. Similarly, the change in social welfare is written as

$$\frac{\partial SW}{\partial k} = \left(\frac{1}{\theta}\right)\left[-\frac{q^*_{G1}}{\partial k} - \gamma\right] * \left[\frac{\theta}{2}\left(1 + \frac{q^*_{G1}}{\theta} + \frac{\gamma k}{\theta}\right) - \gamma k\right] + \left[1 - \left(\frac{q^*_{G1}}{\theta} + \frac{\gamma k}{\theta}\right)\right] * \left[\frac{1}{2}\left(\frac{\partial q^*_{G1}}{\partial k} + \gamma\right) - \gamma\right] + \left[k(1-\beta)\left(\frac{\partial N^*_{G1}}{\partial k}\right) + (1-\beta)N^*_{G1}\right] m^*_{G2} N^*_{G2}. \quad (3.15)$$

The following proposition characterizes the change in welfare caused by this policy.

Proposition 3.5. *Let Assumptions 1 and 2 hold, and consider a market* (β, k)*. There exists a threshold* $\bar{\beta} \in [0,1]$ *such that*

(i) for $\beta_G(k) < \beta \leq \bar{\beta}$*, enabling user control of data leads to a decrease in user and social welfare, whereas*

(ii) for i) $\beta > \bar{\beta}$*, and ii)* $0 \leq \beta \leq \beta_G(k)$*, user and social welfare increases if user control of data is enabled.*

The threshold $\bar{\beta}$ *is decreasing in* γ *and* t*.*

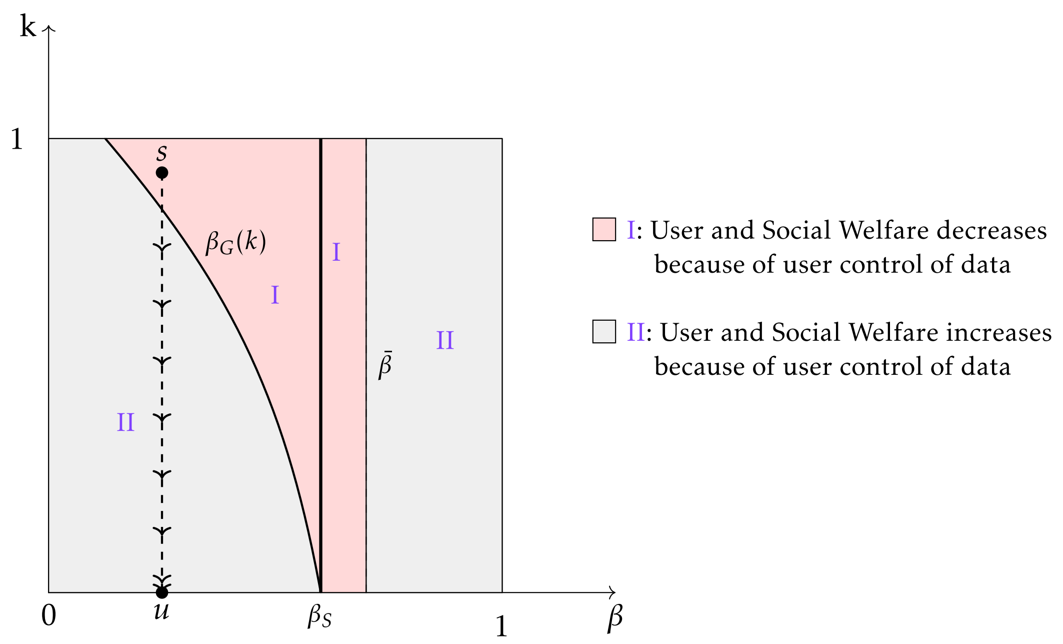


Figure 3.4: User control of data and welfare ($\gamma = .15$, $t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.6$)

Figure 3.4 illustrates the preceding proposition in $k-\beta$ space. Intuitively, β affects the extent of user discount that firm G can give to users and improvement in advertising surplus vis-à-vis privacy cost imposed on users as k increases. This, in addition to market structure, determines the welfare of user control on data collection. For $0 \leq \beta \leq \beta_S$, we have user financed equilibrium. In this case, firms do not use data for advertisement targeting and hence, do not offer user discounts to induce more data collection. The only effect of data collection is to raise user privacy costs. When $\beta_G(k) < \beta \leq \bar{\beta}$, strategic differentiation or advertising financed are the equilibrium outcomes. So, firm G offers user discounts to induce more user participation for data collection. In this case, user control and reduction in k leads to higher prices which offset better consumer privacy achieved, reducing user welfare. In addition, lower advertising surplus also reduces social welfare. Whereas, when $\beta > \bar{\beta}$, we have advertising financed as the equilibrium and the rise in user price is compensated by lower privacy costs, raising user as well as social welfare. From a policy perspective, the above discussion implies that for β small or large, user control is beneficial as reduction in privacy costs are sufficient to offset any other negative effect of regulation. However, for intermediate values of β. the regulation can counter-intuitively reduce welfare by raising user prices despite reduction in privacy costs.

Finally, from the above figure, we also note that the welfare effects of user control, among other channels as described, can also mediate through a business model shift. To emphasize this point, suppose the initial equilibirum is at point s with strategic differentiation as the outcome. Now, as a result of user control of k, the equilibrium moves to point u at which user financed model is the outcome. So, the net effect on social welfare can decomposed into two distinct effects. One, there is the continous effect of reduction in data collection that tend to raise user prices and reduce advertising surplus. However, there is another discontinous effect as we move from point s to u. We call it the business model effect, measured as "$\tau(t-\beta)/2t$". It tends to raise social welfare as firm G shifts to user financing, which is socially efficient for $\beta \leq t$. This new effect has not been considered in the previous literature and the present paper is the first one to explicitly consider the role of privacy regulation through business model changes.

3.4.2 Restricting Access to Data Owned by Subsidiary Firms

Recent thinking in antitrust analysis is now beginning to recognize that access to data can provide significant market power. For instance, in the context of the *WhatsApp/Facebook* and *Microsoft/LinkedIn* merger. the European Commission indicated that data can be a source of competitive advantage.[15]

In this sub-section, we analyse how curtailing a merged firm's ability to access and process data owned by its subsidiary affects welfare. In our framework, this would mean that firm G cannot use data across markets, however, while setting prices and advertising quantities, it can still act as a conglomerate. Under this scenario, the market shares for the two firms would remain the same as defined in equation (3.7), (3.8) and (3.9). On the advertising side, due to no data-sharing, the advertising prices are

$$p_{G2} = \beta N_{G2}, \text{ and } p_{S2} = \beta N_{S2}. \tag{3.16}$$

As in the baseline model, in equilibrium, firms would only charge one side of the market in market 2. Moreover, with neither firm having a relative advantage in advertisement targeting, the equilibrium will be symmetric, in that it either involves both firms adopting user financing, or both opting for advertising financing.

We then compare the welfare under this scenario with that under our baseline framework, when firm G could use data to target advertisements.

Proposition 3.6. *Let Assumptions 1 and 2 hold. Consider a regulation that prevents firm G from accessing data possessed by its affiliates. We then have that*

(i) The market structure is advertising financed whenever $\beta > \beta_S$ (with the prices mim-

[15] However, in its decision, the commission argued that even if a merged entity can combine the data sets there will be extensive data available outside the reach of the merged entity which competitors can use.

icking those in Proposition 3.1(2)(a)), and is user financed otherwise (with the prices mimicking those in Proposition 3.1(2)(c)).

(ii) Turning to welfare implications:

(a) For $\beta_G(k) < \beta \leq 1$, both user and social welfare decreases relative to the baseline model.

(b) For $0 \leq \beta \leq \beta_G(k)$, both user and social welfare remains unchanged.

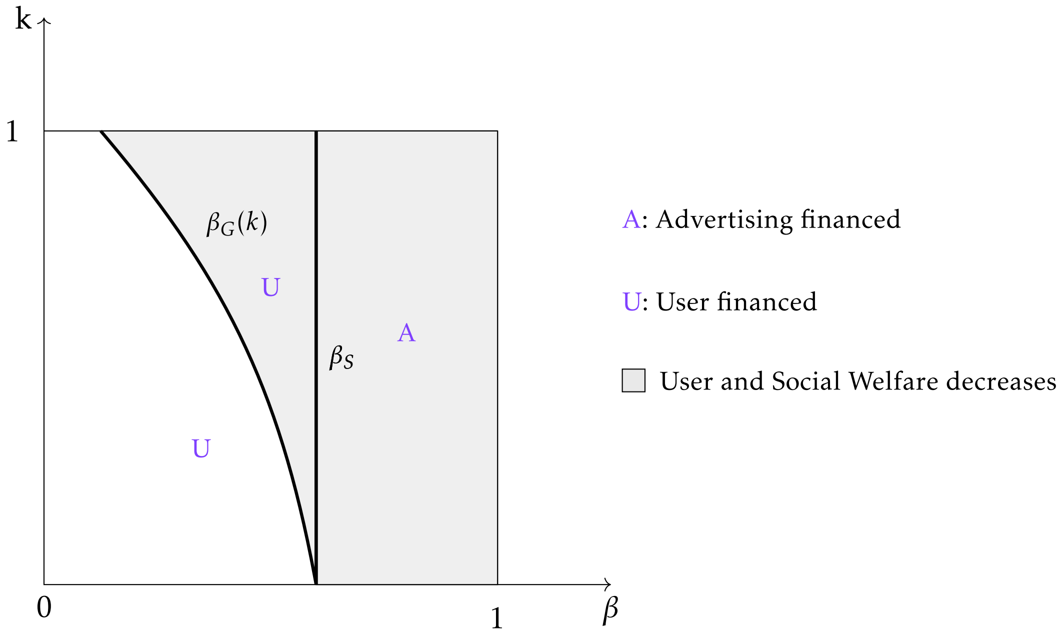


Figure 3.5: Restricting access to data and welfare ($t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.6$)

Thus, in this framework, privacy regulations that aim at curtailing the firm's ability to combine data sets for advertisement targeting is counter-productive. First, consider the effect on consumers' surplus. With no incentive to collect data, firm G has no reason to lower prices in market 1, which reduces consumers' surplus. As to the effect on aggregate surplus, with breaking up of firm's data sets, advertising surplus falls. It is this first order effect that dominates all other effects, including lower privacy costs. This is, of course, not to suggest that there is no role for such regulations. In particular, such regulations can lead to less collection of private data in the first place, thereby improving welfare.

3.4.3 Other Regulatory Changes

Finally, note that the GDPR was recently enacted in the European Union, and may become the world standard for data protection regimes in the future. Given its importance, we finally use our framework to briefly examine some other provisions in the GDPR. In article 20, the GDPR has introduced the right to data portability to reduce user side switching costs and increase the level of competition among online firms. In our framework, this can be formalized as reducing the transportation cost parameter τ. Our analysis, relegated to the appendix, shows that user and social welfare may decrease as a result. We find that such a measure may benefit users, but social welfare can decrease as firm G compensates for lower information collection through higher prices. Interestingly enough, the GDPR does not focus on privacy-enhancing technologies, e.g., those that reduce online theft and misuse. In our framework, this can be formalized as reducing the level of γ. Our analysis shows that, as expected, this raises both user and social welfare in the model. We summarize the main results.

Proposition 3.7. *Let assumptions 1 and 2 hold. Then*

(i) there exist thresholds $\beta_\tau(k)$, $\beta_\tau(k)^{'}$ such that i) $0 \leq \beta_\tau(k), \beta_\tau(k)^{'} \leq 1$, and ii) $\beta_\tau(k), \beta_\tau(k)^{'}$ are decreasing in k. An increase in competition intensity (i.e. a fall in τ) a) increases user welfare under all business models, b) decreases social welfare under advertising financed equilibirum for $\beta_\tau(k) \leq \beta \leq 1$, c) decreases social welfare under strategic differentiation for $\beta_\tau(k)^{'} \leq \beta \leq \beta_S$, and d) increases social welfare under user financed equilibrium.

(ii) a rise in marginal privacy cost (i.e. a rise in γ) reduces user and social welfare under all business models.

3.5 Discussion

3.5.1 Uncovered Market

We briefly discuss the effects of relaxing the full market coverage assumption !. If we allow for partial market coverage, then firms will operate as local monopolies in market 2 rather than compete, with some users staying out of the market. Our analysis shows that under partial market coverage, the results are qualitatively similar, in that we have advertisement financing when β is large, user financing when β is small, and strategic differentiation when β takes an intermediate value. Interestingly, however, the parameter space over which strategic differentiation can arise in equilibrium will contract. This is because data collection in market 1 gives firm G a competitive advantage in market 2. However, with the firms being local monopolies in market 2, firm G no longer has an incentive to collect a lot of data, which makes advertising financing less attractive.

Proposition 3.8. *Strategic differentiation is less likely under partial market coverage.*

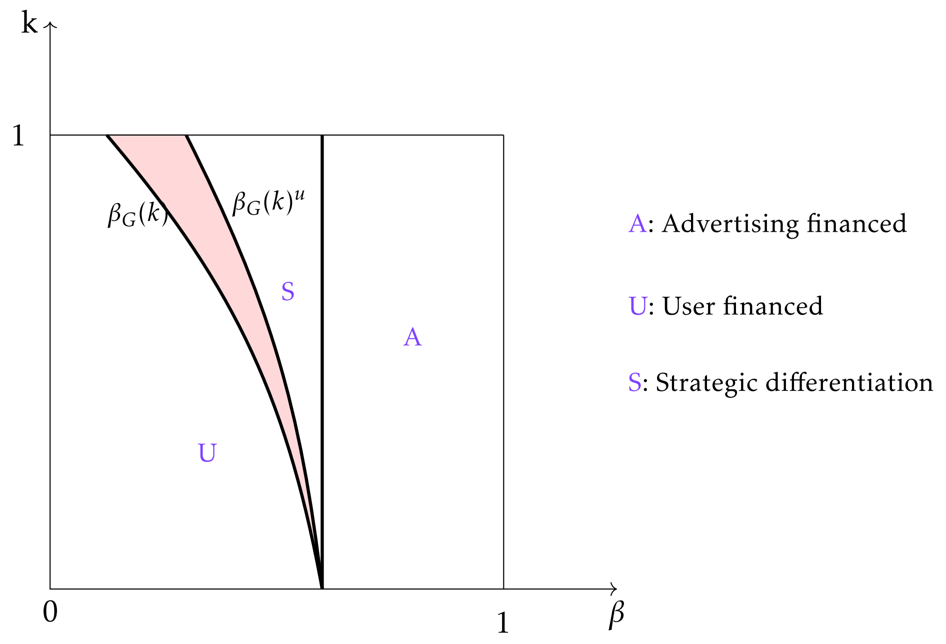


Figure 3.6: Uncovered Market and Choice of Business Model (X = 0.4**,** γ = 0.15**,** t = 0.6**,** τ = 0.5 **and** θ = 0.6)

Figure 3.6 illustrates the preceding result. $\beta_G(k)$ and $\beta_G(k)^{'}$ are the thresholds that comprise points along which firm G is indifferent between switching between advertisement financing and user financing under full and partial market coverage. The shaded region represents the area in which strategic differentiation emerges as an equilibrium under the covered market and the user financed emerges as an equilibrium under the uncovered market. So, partial market coverage works against the role of data as a competitive advantage.

3.5.2 When Market 1 is Not Attractive

In this subsection, we consider the case when θ is sufficiently small, i.e. $\theta \leq k\left[\frac{\tau}{2t} + \gamma\right]$.. The following lemma characterizes the choice of user price in market 1 by firm G and shows how it is affected by stand alone utility in market 1, i.e. θ.

Lemma .1. *If $\theta \leq k\left[\frac{\tau}{2t} + \gamma\right]$, then there exists a threshold $\widetilde{\beta}(k)$ such that $0 \leq \widetilde{\beta}(k) \leq 1$, $\widetilde{\beta}(k)$ is increasing in k, and firm G opts for $q^*_{G1} > 0$ if and only if*

$$0 \leq \widetilde{\beta}(k) < \beta \leq 1. \tag{3.17}$$

The above lemma shows that for firm G to set $q^*_{G1} = 0$ it must be that $0 < \beta \leq \widetilde{\beta}(k) < 1$ in k - β space. Having derived the threshold $\widetilde{\beta}(k)$, next we consider the competitive equilibirum and welfare. The business model equilibrium and their nature remains qualitatively the same as under proposition 3.1. Figure 3.7 describes this scenario. In this case, when firm G adopts advertising financing then it might charge zero price in market 1 for $\beta_G(k)^{'} \leq \beta \leq \widetilde{\beta}(k)$.

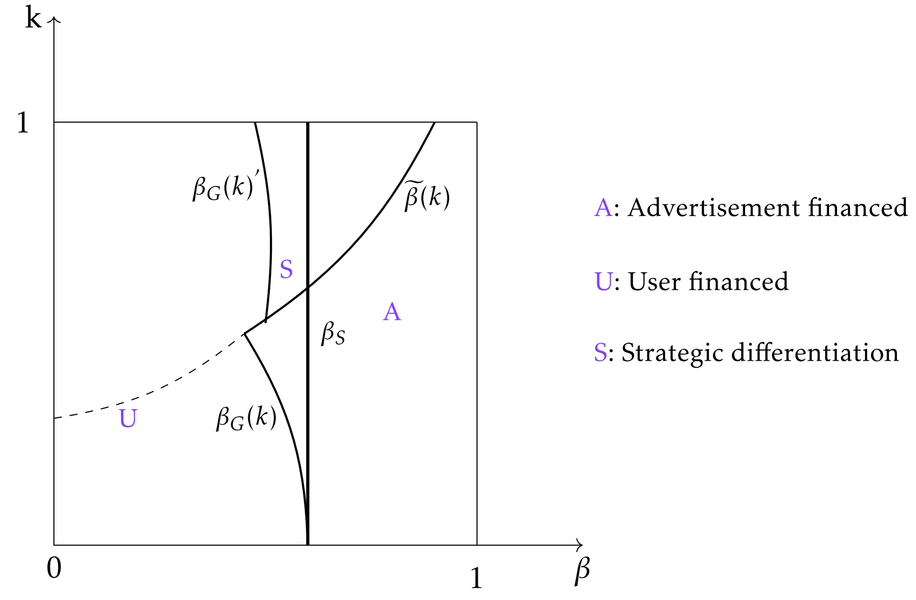


Figure 3.7: Choice of business model when market 1 is not attractive ($\gamma = .15$, $t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.2$)

The welfare comparison of alternate business models qualitatively remains the same as in the basenline model. The only difference that emerges is that comparison of social welfare under strategic differentiation (when $q^*_{G1} = 0$) and user financed equilibrium gives a slightly different result. It is summarized in the following proposition.

Proposition 3.9. *There exists a threshold* $\beta_E(k)'$*, and* $0 \leq \beta_E(k)' \leq \beta_S \leq 1$ *such that when* $q^*_{G1} = 0$*, then strategic differentiation maximizes social welfare for* $\beta_E(k)' \leq \beta \leq \beta_S$*. (refer figure* 3.8*)*

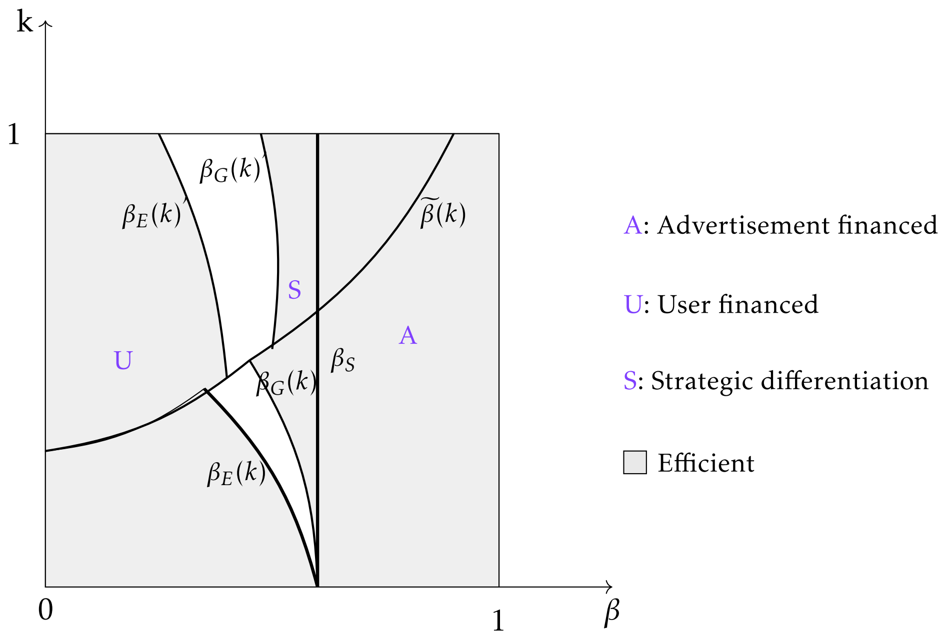


Figure 3.8: Social welfare comparison when market 1 is not attractive ($\gamma = .15$, $t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.2$)

Next, we consider how user and social welfare changes with a restriction on individual level data collection k under alternate business models. The following proposition characterizes the change in welfare as k decreases.

Proposition 3.10. *Let Assumption 1 holds and consider a market* (β, k)*. If* $\theta \leq k\left[\frac{\tau}{2t} + \gamma\right]$*, then there exist thresholds* $\widetilde{\beta}(k)$*,* $\hat{\beta}(k)$ *and* $\bar{\beta}$ *such that (refer figure 3.9)*

(i) For Max $\{\beta_G(k), \widetilde{\beta}(k)\} < \beta \leq \bar{\beta}$*, user and social welfare decreases with user control of data (Region I).*

(ii) For i) Max $\{\bar{\beta}, \widetilde{\beta}(k)\} < \beta \leq 1$*, ii)* $Max\{\beta_G(k)', \hat{\beta}(k)\} < \beta \leq \widetilde{\beta}(k)$*, iii)* $0 \leq \beta \leq \beta_G(k)$*, and iv)* $0 \leq \beta \leq \beta_G(k)'$*, user and social welfare increases with user control of data (Region II).*

(iii) For $\beta_G(k)' < \beta \leq \hat{\beta}(k)$ *user welfare increases and social welfare decreases with user control of data (Region III).*

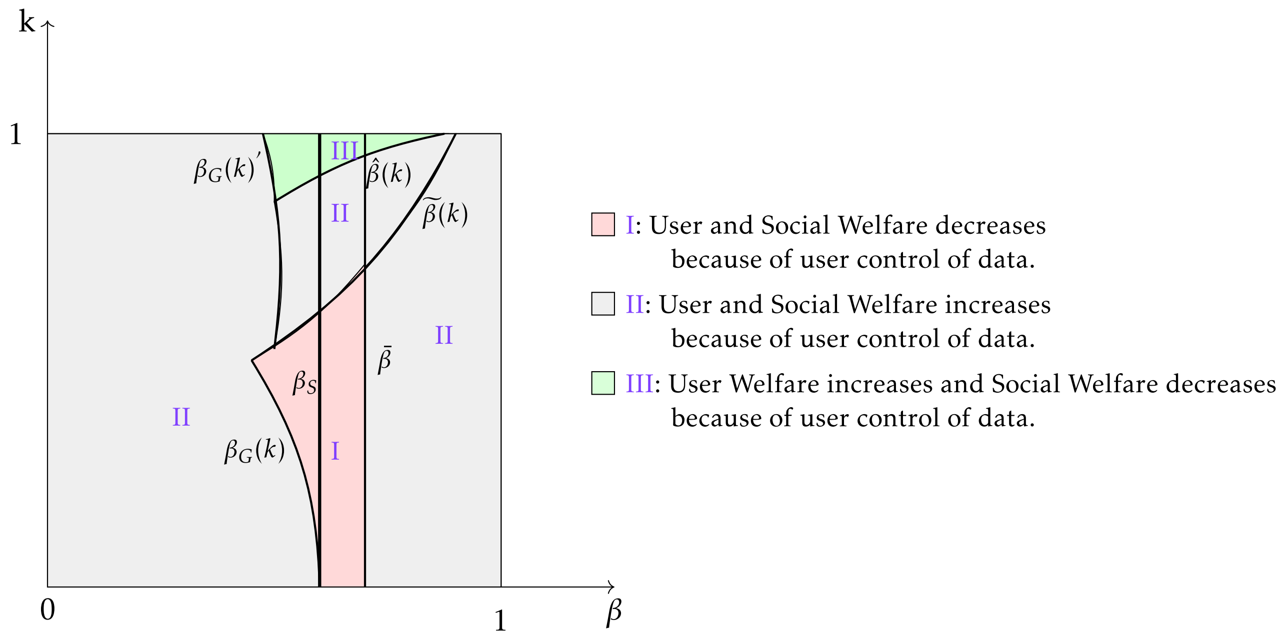


Figure 3.9: User control of data and welfare when market 1 is not atractive ($\gamma = .15$, $t = 0.6$, $\tau = 0.5$ **and** $\theta = 0.2$)

Figure 3.9 illustrates the preceding proposition. It clearly differentiates the three regions: I, II and III. The intuition for region I and II is the same as discussed for the case when $\theta > k\left[\frac{\tau}{2t} + \gamma\right]$. The new region is III, in which the advertising financed firm G optimally chooses q^*_{G1} equal to 0. In this region, the price discount effect of increasing k vanishes. So, the only impact on user welfare is increase in privacy cost as k increases. Hence, when user control data collection, then user welfare will increase as privacy costs are reduced. However, the advertisers gain a large surplus due to greater access to data over large number of users. This tend to offset privacy cost of increasing k. So, as k crosses the threshold value $\hat{\beta}(k)$, social welfare increases. Hence, user control of data collection in region III tend to reduce social welfare.

3.5.3 Endogenous Location

Empirically, we observe that search engines differ in the nature of their functioning. For instance, we see co-existence of a general search engine like Google along with a specialist search engine like Amazon. We can use the framework developed in this chapter to understand the locational choice of these platforms. An interesting question can be how the presence of data advantage from another market for platform G affects its location decision? We can introduce a new stage in the begining of the game at which platforms compete in locations in market 2. A complete analysis is not possible here. However, we can provide an insight into the possible equilibrium outcome. As firms locate close to each other on the Hotelling line, price competition or advertising competition intensifies, reducing their profits. However, in our model, data advantage from market 1 can help platform G to sustain greater losses from intesified competition. Thus, we expect firm G to locate closer to the center of the hotelling line. Whereas, firm S would choose a location near to point 1. In otherwords, firm G would choose a generalist profile, whereas firm S would compete through specialzing in a single domain. This depicts the empirical observation that we described above.

3.6 Conclusion

The present study provides a theory of endogenous adoption of business models in platform markets, finding that such choice is closely related to data collection by firms, as well as the efficacy of advertising in these markets. Moreover, we find that the theoretical predications are consistent with reality. We then use this framework to analyse two policy prescriptions that deal with privacy protection. Inspired by the GDPR, one possible policy is to empower users to control data collection. We find that such a policy can both strengthen privacy, as well as enhance welfare whenever advertiser targeting rates are either large, or small. For targeting rates are at an intermediate level however, such a policy is welfare reducing. Further, even when aggregate welfare increases, advertisers may lose out from such regulations, so that the interests of different sides will be misaligned. This suggests that while promising, careful analysis is required before implementing such a policy in any given market. We also examine the implications of restricting the access of merged firms to data owned by subsidiaries. Our analysis finds that this either reduces welfare, or leaves it unchanged.

3.7 Appendices

Appendix I: Proofs of the baseline model

Proof of Proposition 3.1

First Order Conditions

Since profit functions are continously differentiable, any optimal pair of prices and advertising quantities must satisfy the first order necessary conditions of firms' optimization problem. They are

$$\frac{\partial \pi_G}{\partial q_{G1}} = 1 - \frac{2q_{G1}}{\theta} - \frac{\gamma k}{\theta} - \left[\frac{k(1-\beta)}{\theta}\right]\left[\frac{\tau + tm_{S2} - tm_{G2} + q_{S2} - q_{G2}}{2\tau}\right]m_{G2} \leq 0, \quad (3.18)$$

$$\frac{\partial \pi_G}{\partial q_{G2}} = \frac{\tau + tm_{S2} - tm_{G2} + q_{S2} - 2q_{G2}}{2\tau} - \frac{[\beta + k(1-\beta)N_{G1}]m_{G2}}{2\tau} \leq 0, \quad (3.19)$$

$$\frac{\partial \pi_S}{\partial q_{S2}} = \frac{\tau + tm_{G2} - tm_{S2} + q_{G2} - 2q_{S2}}{2\tau} - \frac{\beta m_{S2}}{2\tau} \leq 0, \quad (3.20)$$

$$\frac{\partial \pi_G}{\partial m_{G2}} = -\frac{tq_{G2}}{2\tau} + [\beta + k(1-\beta)N_{G1}]\left[\frac{\tau + tm_{S2} - 2tm_{G2} + q_{S2} - q_{G2}}{2\tau}\right] \leq 0, \quad (3.21)$$

$$\frac{\partial \pi_S}{\partial m_{S2}} = -\frac{tq_{S2}}{2\tau} + \beta\left[\frac{\tau + tm_{G2} - 2tm_{S2} + q_{G2} - q_{S2}}{2\tau}\right] \leq 0. \quad (3.22)$$

The value of N_{G1} is as defined in equation (4.10). The strict inequality holds if the corresponding variable takes value zero. We characterize the solutions to the F.O.Cs through a series of claims.

Claim .1. *There does not exist any solution with* $q_{G2} > 0, q_{S2} > 0, m_{G2} > 0, m_{S2} > 0$.

Proof: Suppose to the contrary a solution with $q_{G2} > 0, q_{S2} > 0, m_{G2} > 0$, and $m_{S2} > 0$ exists. Then equations (3.19) to (3.22) hold with equality. Then, from equations (3.19) and (3.20) we can solve for q_{G2} and q_{S2}. Next substituting these into (3.21) and (3.22), we have that

$$3\tau + [\beta - t + k(1-\beta)N_{G1}]m_{G2} - (\beta - t)m_{S2} = 0,$$

$$3\tau - [\beta - t + k(1-\beta)N_{G1}]m_{G2} + (\beta - t)m_{S2} = 0.$$

Note that these two equations are inconsistent, and hence no solution exists. A contradiction. □

Claim .2. *There does not exist any solution in which firm i, where $i = G, S$, charges a positive price to both sides of the market, i.e., $q_{i2} > 0$ and $m_{i2} > 0$.*

Proof. Suppose to the contrary such a solution exists. Then, the possible candidates for a solution to the system of equations are:

(i) $q_{G1} \geq 0, q_{G2} > 0, m_{G2} > 0, q_{S2} \geq 0, m_{S2} \geq 0$.

(ii) $q_{G1} \geq 0, q_{G2} \geq 0, m_{G2} \geq 0, q_{S2} > 0, m_{S2} > 0$.

For case 1, for a given value of q_{G1}, it can be shown that the solution to the system of equations (3.18) - (3.22) are:

$$q_{G2} = \frac{3\tau[\beta + k(1-\beta)N_{G1}]}{\beta + k(1-\beta)N_{G1} - t}, \text{ and } m_{G2} = \frac{-3\tau}{\beta + k(1-\beta)N_{G1} - t}. \tag{3.24}$$

Thus the sign of q_{G2} is the negative of the sign of m_{G2}.

Whereas, for case 2, for a given value of q_{G1}, it can be shown that the solution to the system of equations (3.18) - (3.22) are:

$$q_{S2} = \frac{3\tau\beta}{\beta - t}, \text{ and } m_{S2} = \frac{-3\tau}{\beta - t}. \tag{3.25}$$

Thus the sign of q_{S2} is the negative of the sign of m_{S2}.

Thus in both these cases we arrive at a contradiction since the relevant variables must be non-negative. □

So, the only possible candidates for a solution to the system of F.O.Cs given in equations (3.18) - (3.22) are as follows:

(i) $q_{G1} \geq 0, q_{G2} = 0, q_{S2} = 0, m_{G2} > 0, m_{S2} > 0$.

(ii) $q_{G1} \geq 0, q_{G2} > 0, q_{S2} > 0, m_{G2} = 0, m_{S2} = 0$.

(iii) $q_{G1} \geq 0, q_{G2} > 0, q_{S2} = 0, m_{G2} = 0, m_{S2} > 0$.

(iv) $q_{G1} \geq 0, q_{G2} = 0, q_{S2} > 0, m_{G2} > 0, m_{S2} = 0$.

We next take each candidate equilibrium in turn and examine the necessary conditions for each of them. We find that there exists parameter values under which equilibria 1, 2 and 4 can be sustained, whereas candidate 3 can never arise in equilibrium. Further, the parameter values sustaining equilibria 1, 2 and 4 generate parameter restrictions that are exhaustive and mutually exclusive.

We next identify necessary conditions that follow from each of these candidate equilibria. To begin with note that under assumption 2, the value of q^*_{G1} is greater than 0.

(i) Consider candidate equilibrium 1, i.e., $q_{G1} > 0, q_{G2} = 0, q_{S2} = 0, m_{G2} > 0, m_{S2} > 0$: Setting equations (3.18),(3.21) and (3.22) equal to 0 gives $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)\tau}{4t}$, $m^*_{G2} = m^*_{S2} = \frac{\tau}{t}$.

Since in this equilibrium $q_{G2} = q_{S2} = 0$, we need (3.19) and (3.20) to be less than zero. Setting them equal to gives thresholds β_S and $\beta^{'}(k)$ such that:

1. $\beta_S = t$,

(3.26)

$$2.\ \beta^{'}(k) = -\left\{1 - \frac{k}{2} - \frac{k^2(\tau - \gamma t)}{2t\theta}\right\} * \left\{\frac{k^2\tau}{2t\theta}\right\}^{-1}$$

$$+ \left\{\sqrt{\left[1 - \frac{k}{2} - \frac{k^2(\tau - \gamma t)}{2t\theta}\right]^2 + \frac{k^2\tau}{t\theta}\left[t - \frac{k}{2} + \frac{k^2(2t\gamma - \tau)}{4t\theta}\right]}\right\} * \left\{\frac{k^2\tau}{2t\theta}\right\}^{-1}. \quad (3.27)$$

Now, for candidate 1 to be a solution to the system of F.O.Cs, $\beta > \beta_S$ and $\beta > \beta^{'}(k)$. If this holds, then F.O.Cs are satisfied for it, i.e. (3.18),(3.21) and (3.22) equal to 0, and (3.19) and (3.20) are less than zero. Wheras, if these parametric restrictions doesn't hold, i.e. $\beta < \beta_S$ and/or $\beta < \beta^{'}_k$, then F.O.Cs doesn't hold for candidate 1, i.e. (3.19) and/or (3.20) is not less than zero, and hence it cannot be a solution for this parameter region. Also, straightforward calculations will show that $\beta^{'}(k)$ is negative for $\beta > t$. So, the only binding constraint is $\beta > \beta_S$. If this holds, then parametric restrictions are satisfied and candidate 1 can be the solution to the system of F.O.Cs.

(ii) Consider candidate equilibrium 2, i.e. $q_{G1} \geq 0, q_{G2} > 0, q_{S2} > 0, m_{G2} = 0, m_{S2} = 0$: Then, setting equation (3.18), (3.19) and (3.20) equal to zero gives $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2}$, $q^*_{G2} = \tau$ and $q^*_{S2} = \tau$.

Next, since m_{G2} and m_{S2} equal zero, we need (3.21) and (3.22) to be less than zero. Setting them equal to zero gives thresholds $\beta^{''}(k)$ and β_S such that:

1. $\beta_S = t,$ (3.28)

2. $\beta^{''}(k) = \dfrac{t - k(\theta/2 - \gamma k/2)}{1 - k(\theta/2 - \gamma k/2)}.$ (3.29)

For candidate 2 to be a solution to the system of F.O.Cs, $\beta < \beta_S$ and $\beta \leq \beta^{''}(k)$. If this holds, then F.O.Cs are satisfied for it, i.e. (3.18), (3.19) and (3.20) equal to zero, and (3.21) and (3.22) are less than zero, and hence candidate 2 is a solution. Whereas, if $\beta > \beta_S$ and/or $\beta > \beta^{''}(k)$, then F.O.Cs are not satisfied, i,e. (3.21) and(3.22) will not be less than zero, and candidtate 2 cannot be a solution for this parameter region.

(iii) Consider candidate equilibrium 3, i.e., $q_{G1} \geq 0, q_{G2} > 0, q_{S2} = 0, m_{G2} = 0, m_{S2} > 0$: For it to be a soultion to the system of F.O.Cs, we need (3.18), (3.19), and (3.22) equal to zero, and (3.21) and (3.20) to be less than zero. These conditions can hold when $\beta > \beta_S$ and $\beta \leq \beta^{''}(k)$, where β_S and $\beta^{''}(k)$ are as defined in equations (3.26) and (3.29). However, both these constraints cannot be satisfied simultaneously since $\beta^{''}(k) \leq \beta_S$. Hence, either (3.21) or (3.20) is not less than zero, and

F.O.Cs are not satisfied. Therefore, it cannot be a solution.

(iv) Consider candidate equilibrium 4, i.e., $q_{G1} \geq 0, q_{G2} = 0, q_{S2} > 0, m_{G2} > 0, m_{S2} = 0$. For it to be a solution to the system of F.O.Cs, we need (3.18), (3.21), and (3.20) equal to zero, and (3.19) and (3.22) to be less than zero. This requires $\beta \leq \beta_S$ and $\beta > \beta^{'}(k)$, where β_S and $\beta^{'}(k)$ are as defined in (3.26) and (3.34). Since $\beta^{'}(k) \leq \beta_S$, there exists a parameter range for which candidate 4 can be the solution. Whereas, if $\beta > \beta_S$ and/or $\beta \leq \beta^{'}(k)$, then (3.19) and/or (3.22) will not be less than zero, and F.O.Cs are not satisfied.

The preceding discussion can be summarized into a claim as:

Claim .3. *There exist thresholds β_S, $\beta^{'}(k)$, and $\beta^{''}(k)$ such that $\beta_S \geq \beta^{''}(k) \geq \beta^{'}(k)$ and when $0 \leq \beta^{'}(k), \beta^{''}(k) \leq 1$, then*

(i) For $\beta_S \leq \beta \leq 1$, candidate 1 is the solution to the systen of F.O.Cs.

(ii) For $\beta^{'}(k) \leq \beta \leq \beta_S$, candidate 4 is the solution to the system of F.O.Cs.

(iii) For $0 \leq \beta \leq \beta^{''}(k)$, candidate 2 is the solution to the systen of F.O.Cs.

Thus, from the above the claim, it can be seen that, both candidate 2 and 4 are solution to the system of equations (3.18) - (3.22) for the range $\beta^{'}(k) < \beta \leq \beta^{''}(k)$.

Second Order Conditions

Next, we derive the sufficient conditions such that the equilibrium candidates derived from the system of F.O.Cs and outlined in claim 3 characterize a local maximum. Consider candidate 1. Let H_S denote the bordered Hessian for firm S, and H_G for firm G. Evaluating these matrices at this solution gives

$$H_S = \begin{bmatrix} 0 & -1 & 0 \\ -1 & \frac{-1}{\tau} & \frac{-(t+\beta)}{2\tau} \\ 0 & \frac{-(t+\beta)}{2}\tau & \frac{-\beta t}{\tau} \end{bmatrix}$$

$$H_G = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & \frac{-2}{\theta} & \frac{k(1-\beta)}{2\theta t} & 0 \\ -1 & \frac{k(1-\beta)}{2\theta t} & \frac{-1}{\tau} & \frac{-(t+\beta(q^*_{G1}))}{2\tau} \\ 0 & 0 & \frac{-(t+\beta(q^*_{G1}))}{2\tau} & \frac{-t\beta(q^*_{G1})}{\tau} \end{bmatrix}$$

Straightforward calculations will show that i) the determinant of H_S, i..e. $det(H_S) > 0$, and ii) the determinant of principal minor of order 3 of H_G is greater than zero, and of order 4 of H_G, i.e. $det(H_G)$ is less than zero. Therefore, bordered hessian matrices are negative definite and second order conditions are satisfied. Hence, candidate 1 is a local maxmimum.

Consider candidate 2. The bordeed Hessian matrices evaluated at this solution are

$$H_S = \begin{bmatrix} 0 & 0 & -1 \\ 0 & \frac{-1}{\tau} & \frac{-(t+\beta)}{2\tau} \\ -1 & \frac{-(t+\beta)}{2}\tau & \frac{-\beta t}{\tau} \end{bmatrix}$$

$$H_G = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & \frac{-2}{\theta} & 0 & \frac{-k(1-\beta)}{2\theta t} \\ 0 & 0 & \frac{-1}{\tau} & \frac{-(t+\beta(q^*_{G1}))}{2\tau} \\ -1 & \frac{-k(1-\beta)}{2\theta t} & \frac{-(t+\beta(q^*_{G1}))}{2\tau} & \frac{-t\beta(q^*_{G1})}{\tau} \end{bmatrix}$$

Using these, like in the previous case, straightforward calculations will show that the bordered hessian matrices are negative definite, and second order conditions are satisfied. Thus, it is a local maximum. Similarly, for candidate 4, in a similar way, the bordered matrices will be negative definite, and it is a local maximum.

From claim 3, it can be seen that, for the range $\beta^{'}(k) < \beta \leq \beta^{''}(k)$, we have two local maxima, i.e. candiate 2 and candiate 4. In these two soultions, firm S has the same business model. However, firm G adopts different business models. Comparing the two profit of firm G at these two local maxima gives a threshold $\beta_G(k)$ defined as

$$\beta_G(k) = -\left\{1 - \frac{k}{2} - \frac{k^2(\tau - 2\gamma t)}{4t\theta}\right\} * \left\{\frac{k^2\tau}{4t\theta}\right\}^{-1} + \left\{\sqrt{\left[1 - \frac{k}{2} - \frac{k^2(\tau - 2\gamma t)}{4t\theta}\right]^2 + \frac{k^2\tau}{2t\theta}\left[t - \frac{k}{2} + \frac{k^2(4t\gamma - \tau)}{8t\theta}\right]}\right\} * \left\{\frac{k^2\tau}{4t\theta}\right\}^{-1}, \quad (3.34)$$

such that for $\beta_G(k) < \beta \leq \beta_S$, firm G's profit is higher under candidate 4, and for $0 \leq \beta \leq \beta_G(k)$, firm G's profit is higher under candidate 2. Also, using parameter values, straightforward calculations will show that $\beta^{''}(k) \geq \beta_G(k) \geq \beta^{'}(k)$. Thus, for $\beta_G(k) < \beta \leq \beta_S$, candidate 4 is the equilibrium, and for $0 \leq \beta \leq \beta_G(k)$, candidate 2 is the equilibrium. Hence proved. □

Proof of Proposition 3.2

Let UW^i denote the user welfare under equilibirum outcome $i = A, S, U$. From propositon 3.1, it is written as

$$UW^A = X - \frac{5\tau}{4} + \frac{\theta}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right]^2 : \text{ under advertiser financed,} \quad (3.35)$$

$$UW^S = X - \frac{5\tau}{4} + \frac{\theta}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right]^2 : \text{ under strategic differentiation,} \quad (3.36)$$

$$UW^U = X - \frac{5\tau}{4} + \frac{\theta}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta}\right]^2 : \text{ under user financed.} \quad (3.37)$$

It can be seen that $UW^A = UW^S > UW^U$. Hence proved. □

Proof of Proposition 3.3

Let SW^i denote social welfare under equilibirum outcome $i = A, S, U$. It is written as

$$SW^A = X - \frac{5\tau}{4} + \frac{3\theta}{8} + \frac{\beta\tau}{t} + \frac{3k(1-\beta)\tau}{8t} + \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2} - \frac{3k\gamma}{4} - \frac{3k^2(1-\beta)\tau\gamma}{8t\theta} + \frac{3k^2\gamma^2}{8\theta}, \quad (3.38)$$

$$SW^S = X - \frac{3\tau}{4} + \frac{3\theta}{8} + \frac{\beta\tau}{2t} + \frac{3k(1-\beta)\tau}{8t} + \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2} - \frac{3k\gamma}{4} - \frac{3k^2(1-\beta)\tau\gamma}{8t\theta} + \frac{3k^2\gamma^2}{8\theta}, \quad (3.39)$$

$$SW^U = X - \frac{\tau}{4} + \frac{3\theta}{8} + \frac{3k^2\gamma^2}{8\theta} - \frac{3k\gamma}{4}. \quad (3.40)$$

Now, using above values, $SW^A - SW^S = \frac{\tau(\beta-t)}{2t}$. Therefore, for $\beta > t = \beta_S$, $SW^A > SW^S$ and for $\beta \leq t = \beta_S$, $SW^A \leq SW^S$. Using this result, we can compare SW^U with SW^A and SW^S separately .

i) For $\beta \leq \beta_S$, putting $SW^S - SW^U$ equal to 0 gives a threshold $\beta_E(k)$ such that

$$\beta_E(k) = -\left\{1 - \frac{3k}{4} - \frac{3k^2}{8t\theta}(\tau - 2t\gamma)\right\}\left\{\frac{3k^2\tau}{8t\theta}\right\}^{-1} + \left\{\sqrt{\left[1 - \frac{3k}{4} - \frac{3k^2}{8t\theta}(\tau - 2t\gamma)\right]^2 + \frac{3k^2\tau}{4t\theta}\left[t - \frac{3k}{4} + \frac{3k^2}{16t\theta}(4t\gamma - \tau)\right]}\right\}\left\{\frac{3k^2\tau}{8t\theta}\right\}. \quad (3.41)$$

Straightforward calculations will show that $\beta_E(k) < \beta_G(k)$. Therefore, for $\beta_E(k) < \beta \leq \beta_G(k)$, $SW^S > SW^U$ and for $0 < \beta \leq \beta_E(k)$, $SW^S \leq SW^U$.

ii) For $\beta \geq \beta_S$, putting $SW^A - SW^U = 0$ gives a threshold $\beta_{EA}(k)$ such that

$$\beta_{EA}(k) = -\left\{1 - \frac{3k}{8} - \frac{3k^2}{16t\theta}(\tau - 2t\gamma)\right\}\left\{\frac{3k^2\tau}{16t\theta}\right\}^{-1} + \left\{\sqrt{\left[1 - \frac{3k}{8} - \frac{3k^2}{16t\theta}(\tau - 2t\gamma)\right]^2 + \frac{3k^2\tau}{8t\theta}\left[t - \frac{3k}{8} + \frac{3k^2}{32t\theta}(4t\gamma - \tau)\right]}\right\}\left\{\frac{3k^2\tau}{16t\theta}\right\}. \quad (3.42)$$

It can be shown that $\beta_{EA}(k) \geq 0$ if and only if $\beta \leq \beta_S$. So, $SW^A > SW^U$ for all $\beta \in (\beta_S, 1]$. □

Appendix II: Privacy Regulations

Proof of Proposition 3.4

Suppose to the contrary, there exists an equilibirum, where $k_{v,c} > 0$ for a user (v,c). Since each user is inifinitesimally small, individual user's decision about the value of k doesn't affect the choice of advertising and user pricing. From the user utility given in (4.34), it can be seen that, a lower k improves the utility. So, the (v,c) user optimally deviates and sets $k_{v,c} = 0$. A contradiction. Hence proved □

Proof of Proposition 3.5

Using (3.14) and (3.15) the change in user welfare and social welfare can be written as

$$\frac{\partial UW}{\partial k} = \left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)}{2t\theta}\right]\left[\frac{(1-\beta)\tau}{4t} - \frac{\gamma}{2}\right], \tag{3.43}$$

$$\frac{\partial SW}{\partial k} = \frac{3}{4}\left[\frac{(1-\beta)\tau}{2t} - \gamma\right] + \frac{3k}{4}\left[\frac{(1-\beta)^2\tau^2}{4t^2\theta} - \frac{(1-\beta)\tau\gamma}{t\theta} + \frac{\gamma^2}{\theta}\right]. \tag{3.44}$$

Putting (3.43) equal to zero gives a threshold $\bar{\beta}$ such that

$$\bar{\beta} = 1 - \frac{2t\gamma}{\tau}. \tag{3.45}$$

Similarly, putting (3.44) equal to zero will give us a threshold value which can be approximated by the value $\bar{\beta}$. So, for $0 \leq \beta \leq \bar{\beta}$, user welfare and social welfare rises with k and falls otherwise. Hence proved. □

Proof of Proposition 3.6

The proof will follow the same line of argument as under proof of proposition 3.1. When firm G is not allowed to comibine data sets, then the profit of the firm will be

$$\pi_G = q_{G1}N_{G1} + q_{G2}N_{G2} + \beta m_{G2}N_{G2} : \text{ Firm G's profit, and} \tag{3.46}$$

$$\pi_{S2} = q_{S2}N_{S2} + \beta m_{S2}N_{S2} : \text{ Firm S's profit,} \tag{3.47}$$

where we have substitiuted the value for advertising prices from equation (3.16). Now, using equation (3.46) and (3.47), the F.O.Cs can be derived. The first order necessary conditions of firms' optimization problem are

$$1.\ \frac{\partial \pi_G}{\partial q_{G1}} = 1 - \frac{2q_{G1}}{\theta} - \frac{\gamma k}{\theta} \leq 0. \tag{3.48}$$

$$2.\ \frac{\partial \pi_G}{\partial q_{G2}} = \frac{\tau + tm_{S2} - tm_{G2} + q_{S2} - 2q_{G2}}{2\tau} - \frac{\beta m_{G2}}{2\tau} \leq 0. \tag{3.49}$$

$$3.\ \frac{\partial \pi_S}{\partial q_{S2}} = \frac{\tau + tm_{G2} - tm_{S2} + q_{G2} - 2q_{S2}}{2\tau} - \frac{\beta m_{S2}}{2\tau} \leq 0. \tag{3.50}$$

$$4.\ \frac{\partial \pi_G}{\partial m_{G2}} = -\frac{tq_{G2}}{2\tau} + \beta\left[\frac{\tau + tm_{S2} - 2tm_{G2} + q_{S2} - q_{G2}}{2\tau}\right] \leq 0. \tag{3.51}$$

$$5.\ \frac{\partial \pi_S}{\partial m_{S2}} = -\frac{tq_{S2}}{2\tau} + \beta\left[\frac{\tau + tm_{G2} - 2tm_{S2} + q_{G2} - q_{S2}}{2\tau}\right] \leq 0. \tag{3.52}$$

Similar to the argument given in proposition 3.1, there exists a threshold β_S as defined in (3.26) such that the solutions to the system of F.O.Cs (3.48) - (3.52) and the parametric restrictions for no deviation condition are

(i) For $\beta > \beta_S$, equilibirum values are $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2}$, $q^*_{G2} = 0$, $q^*_{S2} = 0$, $m^*_{G2} = \frac{\tau}{t}$, and $m^*_{S2} = \frac{\tau}{t}$.

(ii) For $\beta \leq \beta_S$, equilibrium values are $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2}$, $q^*_{G2} = \tau$, $q^*_{S2} = \tau$, $m^*_{G2} = 0$, and $m^*_{S2} = 0$.

The no deviation condition is satisfied for the threshold value β_S.

Next, we need to compare the user and social welfare with the regime when there is no regulation. Since markets are not connected, the value of user welfare and social welfare remains constant at

$$UW^* = X - \frac{5\tau}{4} + \frac{\theta}{8}\left[1 - \frac{\gamma k}{\theta}\right]^2, \tag{3.53}$$

$$SW^* = X - \frac{\tau}{4} + \frac{3\theta}{8}\left[1 - \frac{\gamma k}{\theta}\right]^2. \tag{3.54}$$

Now, comparing the value for UW^* and SW^* under no privacy regulation with those under privacy regulation will give us the required result. Hence proved. □

Proof of Proposition 3.7

Using (3.35), (3.36) and (3.37) the change in user w.r.t. τ is

$$\frac{\partial UW^A}{\partial \tau} = -\frac{5}{4} + \frac{k(1-\beta)}{t}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right], \tag{3.55}$$

$$\frac{\partial UW^S}{\partial \tau} = -\frac{3}{4} + \frac{k(1-\beta)}{t}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right], \tag{3.56}$$

$$\frac{\partial UW^U}{\partial \tau} = -\frac{5}{4}. \tag{3.57}$$

Straightforward calculations will show that for $0 \le k, \beta \le 1$ user welfare always falls with τ. Using (3.38), (3.39) and (3.40), the change in social welfare w.r.t τ is

$$\frac{\partial SW^A}{\partial \tau} = -\frac{5}{4} + \frac{\beta}{t} + \frac{3k(1-\beta)}{8t} + \frac{3k^2(1-\beta)^2\tau}{16\theta t} - \frac{3k^2(1-\beta)\gamma}{8t\theta}, \tag{3.58}$$

$$\frac{\partial SW^S}{\partial \tau} = -\frac{3}{4} + \frac{\beta}{t} + \frac{3k(1-\beta)}{8t} + \frac{3k^2(1-\beta)^2\tau}{16\theta t} - \frac{3k^2(1-\beta)\gamma}{8t\theta}, \tag{3.59}$$

$$\frac{\partial SW^U}{\partial \tau} = -\frac{1}{4}. \tag{3.60}$$

Using (3.58) and setting it equal to zero gives a threshold $\beta_\tau(k)$ such that

$$\beta_\tau(k) = -\left[\frac{1}{t} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right) - \frac{3k^2\tau^2}{8\theta t^2}\right] * \left[\frac{3k^2\tau^2}{8\theta t^2}\right]^{-1} + \left\{\sqrt{\left[\frac{1}{t} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right) - \frac{3k^2\tau^2}{8\theta t^2}\right]^2 + \frac{3k^2\tau^2}{4\theta t}\left[\frac{5}{4} - \frac{3k^2\tau^2}{16\theta t^2} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right)\right]}\right\} * \left[\frac{3k^2\tau^2}{8\theta t^2}\right]^{-1}. \tag{3.61}$$

Straightforward calculations will show that when $\beta_S \leq \beta_\tau(k) \leq 1$, then $\frac{\partial SW^A}{\partial \tau} > 0$ for $\beta_z(k) < \beta \leq 1$. Similarly, using (3.59) and setting it equal to 0 gives a threshold $\beta_\tau(k)^{'}$ such that

$$\beta_\tau(k)^{'} = -\left[\frac{1}{t} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right) - \frac{3k^2\tau^2}{8\theta t^2}\right] * \left[\frac{3k^2\tau^2}{8\theta t^2}\right]^{-1} + \left\{\sqrt{\left[\frac{1}{t} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right) - \frac{3k^2\tau^2}{8\theta t^2}\right]^2 + \frac{3k^2\tau^2}{4\theta t}\left[\frac{5}{4} - \frac{3k^2\tau^2}{16\theta t^2} - \frac{3k}{8t}\left(1 - \frac{k\gamma}{\theta}\right)\right]}\right\} * \left[\frac{3k^2\tau^2}{8\theta t^2}\right]^{-1}. \tag{3.62}$$

When $0 \leq \beta_\tau(k)^{'} \leq \beta_S$, then $\frac{\partial SW^S}{\partial \tau} > 0$ for $\beta_\tau(k)^{'} < \beta \leq \beta_S$. From (3.60) social welfare always decreaes with an increase in τ.

The change in user welfare w.r.t. γ is

$$\frac{\partial UW^A}{\partial \gamma} = -\frac{k}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right], \tag{3.63}$$

$$\frac{\partial UW^S}{\partial \gamma} = -\frac{k}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}\right], \tag{3.64}$$

$$\frac{\partial UW^A}{\partial \gamma} = -\frac{k}{2}\left[\frac{1}{2} - \frac{\gamma k}{2\theta}\right]. \tag{3.65}$$

Under assumption 2, straightforward calculations will show that user welfare falls with an increase in γ. Similarly the change in social welfare w.r.t γ is

$$\frac{\partial SW^A}{\partial \gamma} = -\frac{3k}{4} - \frac{3k^2(1-\beta)\tau}{8t\theta} + \frac{3k^2\gamma}{4\theta}, \tag{3.66}$$

$$\frac{\partial SW^S}{\partial \gamma} = -\frac{3k}{4} - \frac{3k^2(1-\beta)\tau}{8t\theta} + \frac{3k^2\gamma}{4\theta}, \tag{3.67}$$

$$\frac{\partial SW^U}{\partial \gamma} = \frac{3k^2\gamma}{4\theta} - \frac{3k}{4}. \tag{3.68}$$

Under assumption 2, straightforward calculations will show that social welfare decreases with an increase in γ. Hence proved. □

Appendix III: Uncovered Market

Proof of proposition 3.8

We need to derive the demand functions. The number of users joining firm G or not in market 1 is given by the equation (3.7). In market 2, now each firm $i = G, S$ is a local monopoly. A user indifferent between joining firm i or not is given by $U_{i2} = 0$. This gives location of the indifferent user c_i where

$$c_G = 1 - \left[\frac{X - tm_{S2} - q_{S2}}{\tau}\right] \text{ and } c_S = \frac{X - tm_{G2} - q_{S2}}{\tau}. \tag{3.69}$$

Using equation (3.69), the number of users joining each firm in market 2 is

$$N_{G2} = \frac{X - tm_{G2} - q_{S2}}{\tau}, \tag{3.70}$$

$$N_{S2} = \frac{X - tm_{S2} - q_{S2}}{\tau}. \tag{3.71}$$

It can be seen from (3.70) and 3.71 that each firm's demand is independent of the choice of strategic variables by the other firm. So, each firm's optimization problem can be solved individually. The following lemma summarizes the equilibrium outcome under uncovered market.

Lemma .2. *Let Assumption 2 holds.*

(i) There does not exist an equilibrium where either firm charges both sides of the market.

(ii) There exist thresholds β_S and $\beta_G(k)^u$, where $0 \leq \beta_G(k)^u \leq \beta_S \leq 1$, $\beta_G(k)^u$ is decreasing in k, and $\beta_S = \beta_G(0)^u$, such that given a market characterized by (β, k), we have that:

(a) For $\beta_S < \beta \leq 1$, equilibrium values are

$$q_{G1}^* = \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)X^2}{8t\tau}, q_{G2}^* = 0, q_{S2}^* = 0, m_{G2}^* = \frac{X}{2t} \text{ and } m_{S2}^* = \frac{X}{2t}, \quad (3.72)$$

(b) For $\beta \leq \beta_S$ and $\beta_G(k)^u < \beta \leq \beta_S$, equilibrium values are

$$q_{G1}^* = \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)X^2}{8t\tau}, q_{G2}^* = 0, q_{S2}^* = \frac{X}{2}, m_{G2}^* = \frac{X}{2t} \text{ and } m_{S2}^* = 0, \quad (3.73)$$

(c) For $\beta \leq \beta_S$ and $0 \leq k \leq \beta_G(k)^u$, equilibrium values are

$$q_{G1}^* = \frac{\theta}{2} - \frac{\gamma k}{2}, q_{G2}^* = \frac{X}{2}, q_{S2}^* = \frac{X}{2}, m_{G2}^* = 0 \text{ and } m_{S2}^* = 0. \quad (3.74)$$

Proof: The proof follows the same of argument as for proposition 3.1. The first order necessary conditions of firms' optimization problem are:

$$\frac{\partial \pi_G}{\partial q_{G1}} = 1 - \frac{2q_{G1}}{\theta} - \frac{k(1-\beta)}{\theta} m_{G2} N_{G2} \quad (3.75)$$

$$\frac{\partial \pi_G}{\partial q_{G2}} = \frac{X - tm_{G2} - 2q_{G2}}{\tau} + [\beta + k(1-\beta)N_{G1}] m_{G2} \left(-\frac{1}{\tau}\right) \quad (3.76)$$

$$\frac{\partial \pi_G}{\partial m_{G2}} = -\frac{tq_{G2}}{\tau} + [\beta + k(1-\beta)N_{G1}] \left(\frac{X - 2tm_{G2} - q_{G2}}{\tau}\right) \quad (3.77)$$

$$\frac{\partial \pi_S}{\partial q_{S2}} = \frac{X - tm_{S2} - 2q_{S2}}{\tau} + \beta m_{S2} \left(-\frac{1}{\tau}\right) \quad (3.78)$$

$$\frac{\partial \pi_S}{\partial m_{S2}} = -\frac{tq_{S2}}{\tau} + \beta \left(\frac{X - 2tm_{S2} - q_{S2}}{\tau}\right) \quad (3.79)$$

We characterize the solutions to the F.O.Cs through a series of claim.

Claim .4. *There does not exist any solution in which firm i, where $i = G, S$, charges positive price to both sides of the market, i.e. $q_{i2} > 0$ and $m_{i2} > 0$.*

Proof: Suppose firm G charges both sides in market 2. Then putting equations (3.76) and (3.77) equal to zero gives

$$q_{G2} = \frac{[\beta + k(1-\beta)N_{G1}]X}{\beta + k(1-\beta)N_{G1} - t} \text{ and } m_{G2} = \frac{-X}{\beta + k(1-\beta)N_{G1} - t}, \tag{3.80}$$

whereas if firm S charges both sides in market 2, then putting equations (3.78) and (3.79) equal to zero gives

$$q_{S2} = \frac{\beta X}{\beta - t} \text{ and } m_{G2} = \frac{-X}{\beta - t}, \tag{3.81}$$

From the above equations (3.80) and (3.81), it can be seen that if either firm $i = G, S$ opts for both strategic variables, then one of the strategic variable will be negative and we will get a contradiction.

Next, we can solve for solutions to the system of F.O.Cs for each firm separately. For firm S, there exists a threshold β_S as defined in (3.26) such that for i) for $0 \le \beta \le \beta_S$, it chooses $q^*_{S2} = \frac{X}{2}$, and $m^*_{S2} = 0$, and ii) for $\beta_S \le \beta \le 1$, it chooses $q_{S2} = 0$, and $m^*_{S2} = \frac{X}{2t}$. For firm G, there exists a threshold $\beta_G(k)^u$ given as

$$\beta_G(k)^u = -\left\{\left[1 - \frac{k}{2} - \frac{k^2X^2}{8t\tau\theta} + \frac{k^2\gamma}{2\theta}\right]\right\}\left\{\frac{k^2X^2}{8t\tau\theta}\right\}^{-1} + \left\{\sqrt{\left[1 - \frac{k}{2} - \frac{k^2X^2}{8t\tau\theta} + \frac{k^2\gamma}{2\theta}\right]^2 + \frac{k^2X^2}{4t\tau\theta}\left[t - \frac{k}{2} + \frac{k^2\gamma}{2\theta}\right]}\right\}\left\{\frac{k^2X^2}{8t\tau\theta}\right\}^{-1}, \tag{3.82}$$

such that for i) $0 \le \beta \le \beta_G(k)^u$, firm G chooses $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2}, q^*_{G2} = \frac{X}{2}, m^*_{G2} = 0$, and ii) for $\beta_G(k)^u \le \beta \le 1$, firm G chooses $q^*_{G1} = \frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)X^2}{8t\tau}, q^*_{G2} = 0, m^*_{G2} = \frac{X}{2t}$. Further, straightforward calculations will show that, for sufficiently small X, $\beta_G(k)^u > \beta_G(k)$ (as defined in (3.34)). Also, second order conditions for local maximum hold at these equilibrium values. Hence proved. □

Appendix IV: When market 1 is not attractive, i.e. case when $\theta \leq k\left[\frac{\tau}{2t}+\gamma\right]$.

Proof of Lemma .1

In order to derive $\widetilde{\beta}(k)$, set $q^*_{G1}=0$ when firm G chooses advertising financing. This gives

$$\widetilde{\beta}(k) = 1-[\theta-\gamma k]*\left[\frac{k\tau}{2t}\right]^{-1} \tag{3.83}$$

Straightforward calculations will show that $0 \leq \widetilde{\beta}(k) \leq 1$ if $\theta \leq k\left[\frac{\tau}{2t}+\gamma\right]$. Hence proved. □

Proof of proposition 3.9

If $q^*_{G1}=0$, then user welfare is

.

$$UW^A = X-\frac{5\tau}{4}+\frac{\theta}{2}\left[1-\frac{\gamma k}{\theta}\right]^2 : \text{ under advertiser financed,}$$

$$UW^S = X-\frac{5\tau}{4}+\frac{\theta}{2}\left[1-\frac{\gamma k}{\theta}\right]^2 : \text{ under strategic differentiation,}$$

$$UW^U = X-\frac{5\tau}{4}+\frac{\theta}{2}\left[\frac{1}{2}-\frac{\gamma k}{2\theta}\right]^2 : \text{ under user financed.}$$

It can be seen that $UW^A = UW^S > UW^U$. Similarly, if $q^*_{G1}=0$ then equation (4.45) gives social welfare as

$$SW^A = \frac{\theta}{2}\left[1-\frac{\gamma k}{\theta}\right]^2 + X - \frac{5\tau}{4} + k\left[1-\frac{\gamma k}{\theta}\right]\frac{\tau}{2t} + \frac{\beta\tau}{2t}\left[2-k\left(1-\frac{\gamma k}{\theta}\right)\right],$$

$$SW^S = \frac{\theta}{2}\left[1-\frac{\gamma k}{\theta}\right]^2 + X - \frac{3\tau}{4} + k\left[1-\frac{\gamma k}{\theta}\right]\frac{\tau}{2t} + \frac{\beta\tau}{2t}\left[1-k\left(1-\frac{\gamma k}{\theta}\right)\right],$$

$$SW^U = \frac{3\theta}{8}\left[1-\frac{\gamma k}{\theta}\right]^2 + X - \frac{\tau}{4}.$$

Now, using above values, similar to the case $q^*_{G1} > 0$, $SW^A - SW^S = \frac{\tau(\beta-t)}{2t}$. Therefore, for $\beta > t = \beta_S$, $SW^A > SW^S$, and for $\beta \le t = \beta_S$, $SW^A \le SW^S$. Using this result, we can compare SW^U with SW^A and SW^S separately .

i) For $\beta \le \beta_S$, putting $SW^S - SW^U$ equal to 0 gives a threshold $\beta_E(k)^{'}$ such that

$$\beta_E(k)^{'} = \left\{\frac{\tau}{2} - \frac{\theta}{8}\left[1-\frac{\gamma k}{\theta}\right]^2 - k\left[1-\frac{\gamma k}{\theta}\right]\frac{\tau}{2t}\right\} * \left\{\frac{\tau}{2t}\left[1-k\left(1-\frac{\gamma k}{\theta}\right)\right]\right\}^{-1}. \tag{3.84}$$

Straightforward calculations will show that $\beta_E(k)^{'} < \beta_G(k)^{'}$. Therefore, as shown under case 1, for $\beta_E(k)^{'} < \beta < \beta_G(k)^{'}$, $SW^S > SW^U$ and for $0 < \beta \le \beta_E(k)^{'}$, $SW^S \le SW^U$.

ii) For $\beta > \beta_S$, putting $SW^A - SW^U$ equal to 0 gives a threshold $\beta_{EA}(k)^{'}$ such that

$$\beta_{EA}(k)^{'} = \left\{\tau - \frac{\theta}{8}\left[1-\frac{\gamma k}{\theta}\right]^2 - k\left[1-\frac{\gamma k}{\theta}\right]\frac{\tau}{2t}\right\}\left\{\frac{\tau}{2t}\left[2-k\left(1-\frac{\gamma k}{\theta}\right)\right]\right\}^{-1}. \tag{3.85}$$

Straightforward calculations will show that $\beta_{EA}(k)^{'} \ge 0$ if and only if $\beta \le \beta_S$. So, $SW^A > SW^U$ for all $\beta \in (\beta_S, 1]$. Hence proved. □

Proof of proposition 3.10

When $q^*_{G1} = 0$, then the change in user welfare and social welfare are

$$\frac{\partial UW}{\partial k} = -\gamma\left[1-\frac{\gamma k}{\theta}\right], \tag{3.86}$$

$$\frac{\partial SW}{\partial k} = -\gamma\left[1-\frac{\gamma k}{\theta}\right] + \left[1-\frac{2\gamma k}{\theta}\right]\frac{\tau}{2t} - \frac{\beta\tau}{2t}\left[1-\frac{2\gamma k}{\theta}\right]. \tag{3.87}$$

From (3.86), it can be seen that user welfare always falls with rise in k. Putting (3.87) equal to 0 gives threshold $\hat{\beta}(k)$ such that

$$\hat{\beta}(k) = 1 - \frac{2t\gamma(1-\gamma k/\theta)}{\tau(1-2\gamma k/\theta)}. \tag{3.88}$$

So, when $0 \leq \hat{\beta}(k) \leq \widetilde{\beta}(k)$, then i) for $\hat{\beta}(k) < \beta \leq \widetilde{\beta}(k)$, social welfare falls as k increases and ii) for $\beta(k)^{'} < \beta \leq \hat{\beta}(k)$, social welfare rises as k increases. Hence proved. □

Chapter 4

Bundling in Platform Markets in the Presence of Data Advantage

4.1 Introduction

Recent antitrust cases have focussed on tying practices employed by dominant firms.[1] There are many high profile tying cases under scrutiny across jurisdictions. Google has been under investigation for its alleged anti-competitive practices in search and mobile operating system. The EU launched a formal investigation against Google for analysing the search bias claim that in search results it favors its comparison shopping website vis-à-vis other comparison shopping websites in the European Economic Area (EEA).[2] If this is true, then Google's practices can artificially divert traffic from other websites to its website hindering the growth of other platforms. Another sphere of Google's dominance is mobile operating systems, where it leads the market with over 80 percent market share. Other leading Google applications on mobile devices are Google Maps, Google Search, YouTube, etc. The antitrust complaint against Google is based on the fact that Google requires mobile device manufacturers to sign "Mobile

[1] In this paper, a firm in market 2 would mean a two-sided platform connecting advertisers and users.
[2] Statement of Objections of the EU Commission published on April 14, 2015.

Application Development Agreement (MADA)" among other agreements.[3] This paper aims to understand the bundling strategy employed by a firm to extend its dominance from one market to another.

A key feature highlighted in this paper is the role of user information as a strategic asset and its effect on equilibrium outcome and welfare. A firm can collect information about users that relate to personal information (user IP address, location), demographic information and behavioural information (online browsing, interests, etc.). An online firm can use this information to target advertisements catering to consumer needs and interests. This can improve the probability that a user would buy the advertised product. For example, a search engine can target advertisements based on search queries entered by the users. Facebook shows advertisements that can be targeted based on a user's characteristics. In addition to generating benefits for advertisers, a user can also benefit from big data through improved services. A search engine can change the answers to user queries based on the user information it collects. A social networking site can highlight the news feed that a user would be most interested in. How can this big data be collected? A platform can understand user behaviour from the user history on its interface. Alternatively, they could make a predictive analysis of a user through the information that a user would leave on other platforms. They can use the acquired data advantage from such moves to entrench its position in the core sector.

Combining these two empirical facts about internet firms, in this paper, the focus is on the use of data advantage to bundle products and services across markets. A firm with dominance in market 1 sells its good in market 2 with valuable data sets created in market 1. This is the mechanism underlying targeted advertising where advertising slots are accompanied with valuable information about consumers to increase advertising effectiveness. Improvement in user services because of better information avail-

[3] See Edelman and Geradin (2016). It is observed that, first, manufacturers must "pre-install all Google applications" that Google specifies. Second, Google requires that these pre-installed apps be placed prominently on mobile devices. Third, Google requires that Google search "must be set as the default search engine for all Web search access points," ruling out the possibility of any other search engine being the default.

able about them will not be considered. Only improvement in the predictive power of user attitudes and its impact on advertising technology will be part of our analysis. We examine how cross usage of data across two markets on the advertising side affects the private and social incentive to bundle services on the user side. Finally, implications for competition policy and regulation are drawn.

This paper uses the model developed in the previous chapter. There are two markets - market 1 and market 2. Firm G is a monopolist in market 1 and a duopolist in market 2 with firm S as its rival. In market 2, users single-home and advertisers multi-home. Users dislike advertisements and have identical intrinsic value for two firms. However, firm G has a data advantage from its presence in market 1. This provides it with user information relevant to functioning in core platform market 2. On the advertising side, firms compete in advertising quantities and sell the advertising space to advertisers. The crucial feature of advertising technology is that there is a difference in returns from placing advertisements in the two firms. Firm G has data regarding users which lead to a more efficient and better advertising technology on its platform. Thus, an advertisement in firm G has a higher probability of reaching a user.

The contribution of this paper to the literature on two-sided markets is to show how the presence of user data advantage affects bundling in platform markets and analyse its welfare implications. The underlying mechanism that affects the profitability of bundling in this model is the following. On the one hand, bundling increases user participation in market 1 and thus, aggregate user information available to firm G. This improves advertisement targeting, which increases advertising revenue on firm G in the tied market, i.e., market 2, making bundling profitable. On the other hand, its profit decline in the tying market, i.e., market 1 due to bundled discount. The balance between the gains and loss depends on the parameter values. Bundling is profitable when, everything else equal, i) investment in data collection is sufficiently large, and/or ii) advertising targeting rate is large.

Next, a normative analysis is done to draw policy implications. Various effects -

positive or negative, work on social welfare. When the firms have adopted an advertising financed model, then bundling increases the nuisance cost of advertisements and transportation cost and reduces advertisers' revenue on the rival firm S. Also, it increases advertising revenue on firm G, and users gain through bundled discount. The net effect depends on the interplay of these forces. A careful examination of the parameter regions shows that there are regions where private and social incentives diverge and converge. In markets with a small investment in data collection, a profitable bundling may lead to a fall in social welfare. Whereas, in markets with a large investment in data collection, bundling is profitable and improves social welfare. However, if the advertising targeting rate is sufficiently large, then bundling is always profitable. So, including the role of data in bundling decisions affects the policy implications that can be drawn. Lastly, it is shown that there can be a divergence between the user and social welfare, i.e., bundling may reduce social welfare but increases user welfare.

4.1.1 Related literature

This paper is related to many strands of literature. First, this paper is related to the literature on two-sided markets (e.g., Armstrong (2006), Caillaud and Jullien (2003), Rochet and Tirole (2003)), and, more generally, to the effect of network effects on the products and regulatory design in online markets (e.g., Bloch and Demange (2018), Lu and Poddar (2018)). Second, this study contributes to the understanding of bundling incentives in two sided markets. The leverage theory of bundling has a well established intellectual history and many papers have studied bundling as an entry deterrence device (e.g., Whinston (1990), Choi and Stefanadis (2001), Carlton and Waldman (2002), Nalebuff (2004)). In addition, a few papers have focussed on bundling in platform markets. Amelio and Jullien (2012) and Choi and Jeon (2016) considered models with platforms that are unable to charge negative prices. They examined the incentives of a monopolist to tie its monopolized product with product facing competition in two-sided markets and derive its welfare implications. The novel mechanism that makes bundling profitable in these papers is the ability to overcome non-negative price constraints. Since the rival is constrained to set non-negative prices, it limits aggressive response by the rival and additional profits are generated. Choi (2010) studied tying in two-sided markets when each platform has some exclusive content to offer to consumers. It shows that tying can improve social welfare if multi-homing is allowed on the content provider side. Corniere and Taylor (2017) set up a slightly different model in which platforms can set negative prices. There are application developers and users on two sides interacting through a platform - smartphone manufacturers. Applications derive benefits for their developers, and developers can offer payments to the device manufacturers in exchange for being installed. They show that bundling reduces rival application developers' willingness to pay manufacturers for inclusion on their devices and allows a multi-application developer to capture a larger share of industry profit.

The policy stand on bundling in two sided markets is divided. Consider Google's bundling practices in search and mobile operating. Defenders[4] argue that bundling is not anti-competitive for two main reasons - i) there are no restrictions on multi-homing on the user side, and ii) bundling helps the firm to innovate, and it's a product improvement. Whereas, opponents[5] of bundling argue that i) users are exposed to a large quantity of advertisements, ii) bundling imposes restrictions on the advertiser side, and iii) bundling allows the platform to gather a huge amount of data from complementary markets and reinforce its dominant position in the core market.

This paper formalizes the argument substantiated in Newman (2014) on control of user data in platform markets. From a theoretical standpoint, despite the importance of user level data in affecting market outcomes, none of the studies mentioned above consider the role of data in strategic decision making. To fill the gap, this paper explicitly considers the role of data advantage in bundling decisions and analyse its welfare implications. This would help in exploring the market conditions under which a platform, present in multiple markets, can use user level data for leveraging market power. Bundling of a monopolized product with one side of the platform mar-

[4]Refer Bork and Sidak (2012), and Manne and D. Wright (2010).
[5]Refer Newman (2014), and Edelman and Geradin (2016).

ket can expand the set of users to whom the tying firm can sell on the other side of the market. This increases the user data set and advertisement targeting is possible over a larger user base. Thus, additional advertising revenue can be captured through bundling.

The rest of the paper is organized as follows. Section 4.2 briefly discusses the baseline model. Section 4.3 discusses the market outcomes and bundling incentive. Section 4.4 conducts a welfare analysis and discusses policy implications. Section 4.5 discusses a few extensions of the baseline model. Section 4.6 concludes. All proofs are in the appendix.

4.2 The Model

The model is based on the framework developed in the previous chapter. Here, we summarize the agents and their payoffs. There are three sets of agents, two firms, G and S, a unit mass of users, and a unit mass of advertisers. Firm G is a monopolist in market 1, whereas firm G and S compete à la Hotelling in market 2, with firm G located at point 0 and firm S located at point 1 of the unit Hotelling interval. Let the two products sold by firm G to the users be denoted as $G1$ in market 1 and $G2$ in market 2. Similarly, let $S2$ denote the product consumed by users when they join firm S. Here, the focus is only on *advertising financed platforms*. So, the users are charged a zero price for the goods consumed by them in market 2. However, they are charged a positive price in market 1. On the advertising side, firms set advertising quantities, m_{G2} and m_{S2} respectively. The profit functions are

$$\pi_G = q_{G1}N_{G1} + p_{G2}m_{G2} : \text{ Firm G's profit,} \tag{4.1}$$

$$\text{and, } \pi_S = p_{S2}m_{S2} : \text{ Firm S's profit.} \tag{4.2}$$

As defined in the previous chapter, there is a unit mass of users with each user defined by a pair (v, c), where v is her valuation for good $G1$ in market 1, and c defines her location in market 2. Thus, her net utility in market 1 is

$$U_{G1}(v) = v\theta - q_{G1}, \tag{4.3}$$

where q_{G1} is the price charged by firm G in this market. Here, it is assumed that users are *privacy insensitive*. So, there are no privacy costs of data collection in market 1.[6] In market 2, she gets a gross utility of X irrespective of which firm she joins and faces a transportation cost of τc if she joins firm G, and $\tau(1-c)$ if she joins firm S. Also, she is exposed to an advertising level of m_{G2}, with a total disutility of tm_{G2} if she joins firm G, and an advertising level of m_{S2}, with a total disutility of tm_{S2} if she joins firm S. Thus, her net utility in market 2, U_{i2}, $i = \{G, S\}$, equals

$$\begin{cases} X - tm_{G2} - q_{G2} - \tau c, & \text{if she joins firm } G, \\ X - tm_{S2} - q_{S2} - \tau(1-c), & \text{if she joins firm } S. \end{cases} \tag{4.4}$$

As highlighted in the previous chapter, nuisance and transportation cost establishes a strategic interdependence between consumer demand for the two firms in market 2. On the advertising side, there is a unit mass of identical advertisers all of whom want to generate attention for its product or service through placing advertisements in firms G and S in market 2. As discussed in the previous chapter, the probability of informing a single user depends on whether the concerned firm has data regarding the user or not. In case the firm has no such data, then the probability that one advertisement informs a single user is β, $\beta > 0$, whereas if the firm has such data, then this probability is higher at $\beta + k(1-\beta)$, where $k > 0$, is some exogenous level of investment in data collection by the firm. The probability of informing a user in the model also captures the network effects on the advertiser side. It depicts the marginal increase in value from placing advertisements on a platform, when an additional user joins that

[6] In an extension, we consider the case when users are privacy sensitive. The main results will remain the same with some new results.

platform. The larger the user base, the higher the value from placing advertisements on the platform. Therefore, the gross advertising revenue from placing m_{S2} advertisements in firm S is

$$\beta m_{S2} N_{S2}, \tag{4.5}$$

since firm S has no presence in market 1, and thus no data advantage. Firm G has data regarding $N_{G1}N_{G2}$ users who join it in market 2 and no data regarding $(1-N_{G1})N_{G2}$ users who join it in market 2. Following the same line of argument as in the previous chapter, the gross advertising revenue from placing m_{G2} advertisements in firm G is

$$[(\beta + k(1-\beta))N_{G1} + \beta(1-N_{G1})]m_{G2}N_{G2} = [\beta + k(1-\beta)N_{G1}]m_{G2}N_{G2}. \tag{4.6}$$

As highlighted in the previous chapter, advertising revenue of firm G is dependent upon the aggreagte data collected in market 1, which establishes an interdependence between the two markets. The total gross revenue from placing m_{G2} and m_{S2} advertisements in both the firms is

$$R(m_{G2}, m_{S2}) = [\beta + k(1-\beta)N_{G1}]m_{G2}N_{G2} + \beta m_{S2}N_{S2}. \tag{4.7}$$

The advertisers are charged a price p_{G2} by firm G and p_{S2} by firm S. Therefore, the total advertising profit from placing advertisements in the two firms is

$$\pi_a = R(m_{G2}, m_{S2}) - p_{G2}m_{G2} - p_{S2}m_{S2}, \tag{4.8}$$

The timing of the game is as follows:

Stage 1: Firms G and S simultaneously choose user prices and the advertising levels, i.e., firm G chooses q_{G1}, q_{G2} and m_{G2}, and firm S chooses q_{S2} and m_{S2}.

Stage 2: Observing firms' choices, users decide i) whether or not to join firm G in market 1, and ii) which firm to join in market 2. Whereas, prices p_{G2} and p_{S2} adjust so that the advertising market clears. A strategy for the advertiser is to decide whether to join firm G or S or both. Since advertisers are price takers in the model, they will place advertisements in firm $i = G, S$, as long as these prices equal their marginal benefit from a unit of an advertisement in firm i.

The solution concept used is subgame perfect Nash equilibrium (henceforth equilibrium).

4.3 Equilibrium Analysis

4.3.1 Independent Pricing

In this section, we characterize equilibrium under independent pricing. A type (v, c) user can make her purchase decision independently in the two markets. Therefore. her choice set consists of four options:

(i) $G1G2$: Buy good 1 in market 1 and join firm G in market 2.

(ii) $G1S2$: Buy good 1 in market 1 and join firm S in market 2.

(iii) $G2$: Do not buy good 1 and join firm G in market 2.

(iv) $S2$: Do not buy good 1 and join firm S in market 2.

At stage 2, users make participation decisions. In market 1, an indifferent user is defined by $\hat{v}$ such that

$$\hat{v}\theta - q_{G1} = 0. \tag{4.9}$$

This gives demand for good 1 in market 1 N_{G1} as

$$N_{G1} = 1 - \frac{q_{G1}}{\theta}. \tag{4.10}$$

In market 2, the user indifferent between participating on firm G and firm S is defined by the location $\hat{c} \in [0,1]$ such that $U_{G2}(\hat{c}) = U_{S2}(\hat{c})$. This gives

$$\hat{c} = \frac{1}{2} + + \frac{tm_{S2} - tm_{G2}}{2\tau}. \tag{4.11}$$

Using this, the demand for firm G, defined by N_{G2} and for firm S, defined by N_{S2} are

$$N_{G2} = \hat{c} = \frac{1}{2} + + \frac{tm_{S2} - tm_{G2}}{2\tau}, \tag{4.12}$$

$$N_{S2} = 1 - \hat{c} = \frac{1}{2} + \frac{tm_{G2} - tm_{S2}}{2\tau}. \tag{4.13}$$

On the advertising side, the advertisers would participate in the advertising market as long as the marginal benefit of an advertisement is equal to its marginal cost p_i. Using equation (4.8), the inverse advertising demand functions can be written as

$$p_{G2} = [\beta + k(1-\beta)N_{G1}]N_{G2}, \tag{4.14}$$

$$p_{S2} = \beta N_{S2}. \tag{4.15}$$

At stage 1, using the inverse advertising demand functions defined in equations (4.14) and (4.15) and the user demand functions defined in equations (4.10), (4.12) and (4.13), the two firms would maximize their profits w.r.t. user prices and advertising quantities. The optimal strategies of the two firms are characterized in the following proposition.

Proposition 4.1. *Under independent pricing, there exists a unique equilibrium such that the optimal price and advertising quantities are characterized by $q_{G1} = q^*_{G1}$, $m_{G2} = m^*_{G2}$, and $m_{S2} = m^*_{S2}$, where*

$$q^*_{G1} = \frac{\theta}{2} - \frac{k(1-\beta)\tau}{4t},\ m^*_{G2} = \frac{\tau}{t},\ and\ m^*_{S2} = \frac{\tau}{t}. \tag{4.16}$$

4.3.2 Bundling

In this section, we examine the bundling decision of firm G. At the outset, we will consider pure bundling decision, in that firm G sells both its product as a pure bundle to the users. This means that any user joining market 1 also joins firm G in market 2. The users' choice set is now reduced to

(i) Bundled product of firm G: $G1G2$

(ii) Firm S's product: $S2$

Let q_G be the price of bundled product $G1G2$ sold by firm G to the users and N_G be the demand for the bundled good. The rest of the notations are the same as under the no bundling case. Now, at stage 2, the decision is to choose one of the above items in the choice set. If a (v,c) type user purchases the bundled product at a price q_G, her utility is

$$U_G(v,c) = v\theta + X - tm_{G2} - q_G - \tau c, \tag{4.17}$$

whereas if a (v,c) type user joins firm S, her utility is

$$U_{S2}(c) = X - tm_{S2} - \tau(1-c). \tag{4.18}$$

An indifferent user is defined by a pair (v,c) satisfying the following equality

$$v\theta + X - tm_{G2} - q_G - \tau c = X - tm_{S2} - \tau(1-c). \tag{4.19}$$

Using the preceding equation, the indifferent user is described by the line $f(v)$ such that

$$f(v) = \frac{v\theta + \tau}{2\tau} + \frac{t(m_{S2} - m_{G2})}{2\tau} - \frac{q_G}{2\tau}. \tag{4.20}$$

We can draw $f(v)$ in $v - c$ space to describe the demand for firm G and firm S.

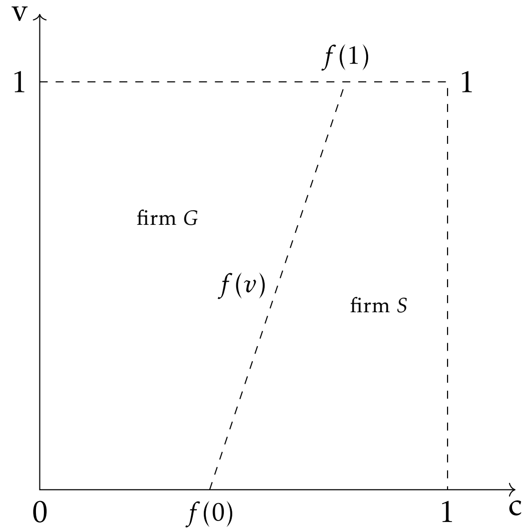


Figure 4.1: Demand under Bundling Case

For tractability, it is assumed that $0 < f(0), f(1) < 1$. Using figure 4.1, the demand for firm G, i.e., N_G and for firm S, i.e., N_{S2} can be written as

$$N_G = \int_0^1 f(v)dv = \frac{\theta + 2\tau}{4\tau} + \frac{tm_{S2} - tm_{G2}}{2\tau} - \frac{q_G}{2\tau}, \tag{4.21}$$

$$N_{S2} = \int_0^1 (1 - f(v))dv = \frac{2\tau - \theta}{4\tau} + \frac{tm_{G2} - tm_{S2}}{2\tau} + \frac{q_G}{2\tau}. \tag{4.22}$$

The crucial difference under the bundling case is the change in firm G's profit from the advertisers. Earlier, in market 2, the two kinds of users joined firm G - those who participated in market 1 and those who didn't. Now, with bundling, the only set of users on firm G in market 2 are those who have participated in market 1, i.e., N_G. Thus, firm G has access to information regarding all users who join it in market 2. In other words, the probability of informing a user who has joined firm G in market 2 is $\beta + k(1 - \beta)$. Bundling affects the total revenue of an advertiser on firm G. The gross advertising revenue from placing m_{G2} advertisements in firm G is

$$[\beta + k(1-\beta)]m_{G2}N_G, \tag{4.23}$$

whereas the gross advertising revenue on firm S remains the same at $\beta m_{S2}N_{S2}$. Thus, the total revenue from placing m_{G2} and m_{S2} advertisements is

$$R(m_{G2}, m_{S2}) = [\beta + k(1-\beta)]m_{G2}N_G + \beta m_{S2}N_{S2}. \tag{4.24}$$

Using the preceding equation, the inverse advertising demand functions are

$$p_{G2} = [\beta + k(1-\beta)]N_G, \tag{4.25}$$

$$p_{S2} = \beta N_{S2}. \tag{4.26}$$

At stage 1, using the user demand functions given in equations (4.21) and (4.22), and the advertising prices given in equations (4.25) and (4.26), and putting the values for them in the profit functions (4.1) and (4.2), firm G chooses the user price q_G, and firms G and S choose the advertising quantities, m_{G2} and m_{S2} to maximize their profits. The main insights are summarized in the following proposition.

Proposition 4.2. *When firm G could bundle its two products, then there exists a unique equilibrium such that the optimal price and advertising quantities are characterised by* $q_G = \widetilde{q}_G$, $m_{G2} = \widetilde{m}_{G2}$, *and* $m_{S2} = \widetilde{m}_{S2}$, *where*

$$\widetilde{q}_G = 0,\ \widetilde{m}_{G2} = \frac{6\tau + \theta}{6t},\ \textit{and}\ \widetilde{m}_{S2} = \frac{6\tau - \theta}{6t}. \tag{4.27}$$

4.3.3 Private Incentive to Bundle

Now, firm G's incentive to bundle the two products $G1$ and $G2$ is analysed. The result is explained through the following proposition and illustrated in figure 4.2.

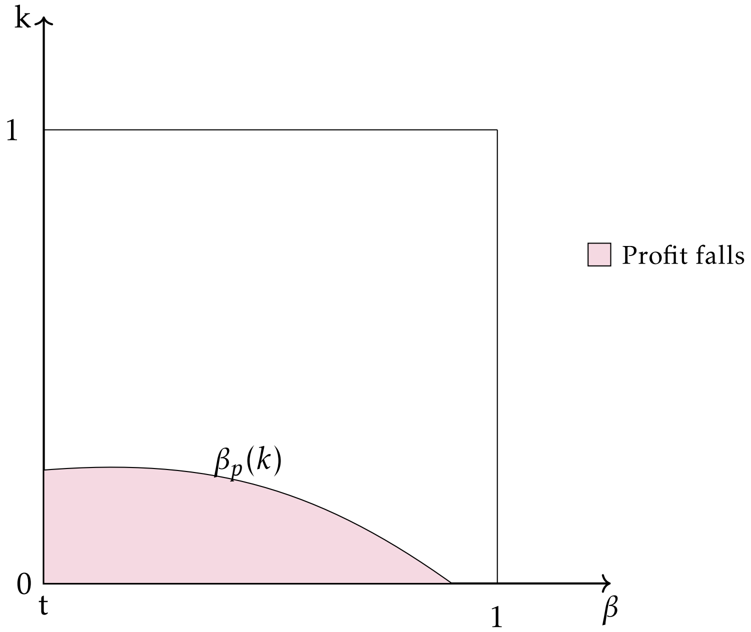
Proposition 4.3. *When firm G could adopt bundling, then there exists a threshold $\beta_p(k)$, where $0 \leq \beta_p(k) \leq 1$, and $\beta_p(k)$ is decreasing in k, such that given a market characterized by $(\beta.k)$, we have that:*

(i) Bundling reduces firm S's profit.

(ii) Bundling is profitable for firm G iff $\beta_p(k) \leq \beta \leq 1$.

The parameter of interest is k which measures the level of investment in data collection. Bundling has a trade-off for firm G - gain in advertising revenue in market 2 and loss of subscription revenue in market 1. The former results from the collection of data regarding users and thus the higher advertising price that firm G can get from the advertisers. Whereas, the latter is due to overall bundling discounts and foregone user revenue in market 1 to increase the market share. When the investment in data collection (large k) is high, then, through bundling, firm G can earn higher advertising revenues. This is sufficient to offset any loss from market 1. When k is small, the increase in advertising revenue because of bundling may not offset the revenue loss in market 1 and firm $G's$ profit may decline. However, even for a small k, bundling can be profitable if β is sufficiently large. This makes possible sufficient advertising revenue to offset foregone revenue.

To provide more insights about the conditions under which bundling is profitable, we investigate how the critical threshold $\beta_p(k)$ changes, ceteris paribus, with respect to a change in nuisance cost of advertisements, t. The nuisance cost parameter t affects firm G's profit in two ways. First, under independent pricing, the price charged to users in market 1 goes up, i.e., $\frac{\partial q^*_{G1}}{\partial t} > 0$. This implies that bundled discounts will be higher for users and thus greater loss to firm G from bundling. In market 2, t increases the nuisance cost of advertisements to the users. Thus, equilibrium advertising levels go down, which reduces gain in the advertising revenue through bundling. So, a rise in t reduces firm G's profit from bundling.



The figure is drawn for parameter values $t = 0.6$, $\beta = 0.7$, and $\theta = 0.6$. On the horizontal axis, the parameter β ranges from t to 1. Recall from the previous chapter that for $\beta \geq t$, advertising financing is the equilibrium. This would also hold under the bundling regime. So, we restrict the relevant parameter space for β such that $\beta \in [t, 1]$. The threshold $\beta_p(k)$ represents the loci of points along which firm G's profit under independent pricing and bundling are the same.

Figure 4.2: Private Incentive to Bundle

4.4 Welfare Analysis

4.4.1 Social Welfare

Based on the equilibrium analysis done in the last section, now the change in social welfare as a result of bundling is examined. It is defined as the sum of users' surplus, advertisers' profit and firms' profit. Since prices are just transfers in the model, it equals the sum of users' surplus and advertisers' revenue. Social welfare under independent pricing is

$$SW^* = \int_{\frac{q^*_{G1}}{\theta}}^{1} v\theta dv + \int_0^{\hat{c}} \left[X - tm^*_{G2} - \tau c\right] dc + \int_{\hat{c}}^{1} \left[X - tm^*_{S2} - \tau(1-c)\right] dc$$
$$+ [\beta + k(1-\beta)N^*_{G1}]m^*_{G2}N^*_{G2} + \beta m^*_{S2}N^*_{S2}, \quad (4.28)$$

whereas social welfare under bundling is

$$\widetilde{SW} = \int_0^1 \int_0^{f(v)} [v\theta + X - t\widetilde{m}_{G2} - \tau c] dc\, dv + \int_0^1 \int_{f(v)}^1 [X - t\widetilde{m}_{S2} - \tau(1-c)] dc\, dv$$
$$+ [\beta + k(1-\beta)]\widetilde{m}_{G2}\widetilde{N}_{G2} + \beta \widetilde{m}_{S2}\widetilde{N}_{S2}, \quad (4.29)$$

where $\int_0^1 f(v)$ is as defined in equation (4.20) and the rest of the notations are the same as defined in earlier sections. The change in social welfare ΔSW is defined as

$$\Delta SW = \underbrace{\int_0^1 v\theta f(v) dv - \int_{\frac{q^*_{G1}}{\theta}}^1 v\theta dv}_{\text{change in users' surplus in market 1}} \underbrace{- [t\widetilde{m}_{G2}\widetilde{N}_G + t\widetilde{m}_{S2}\widetilde{N}_{S2}] + [tm^*_{G2}N^*_{G2} + tm^*_{S2}N^*_{S2}]}_{\text{change in nuisance cost}} +$$
$$\underbrace{\tau \int_0^1 [f(v) - f(v)^2] dv - \tau[N^*_{G2} - (N^*_{G2})^2]}_{\text{change in transportation cost}} + \underbrace{\beta \widetilde{m}_{S2}\widetilde{N}_{S2} - \beta m^*_{S2}N^*_{S2}}_{\text{change in advertisers' revenue on firm S}}$$
$$+ \underbrace{[\beta + k(1-\beta)]\widetilde{m}_{G2}\widetilde{N}_{G2} - [\beta + k(1-\beta)N^*_{G1}]m^*_{G2}N^*_{G2}}_{\text{change in advertisers' revenue on firm G}}. \quad (4.30)$$

So, the trade-offs from bundling and its impact on social welfare would depend on the following effects:

(i) Change in users' surplus in market 1: It may increase or decrease depending on the parameter values.

(ii) Change in nuisance cost of advertisements: increases.

(iii) Change in transportation cost: increases.

(iv) Change in advertisers' revenue: It may increase or decrease depending on the parameter values.

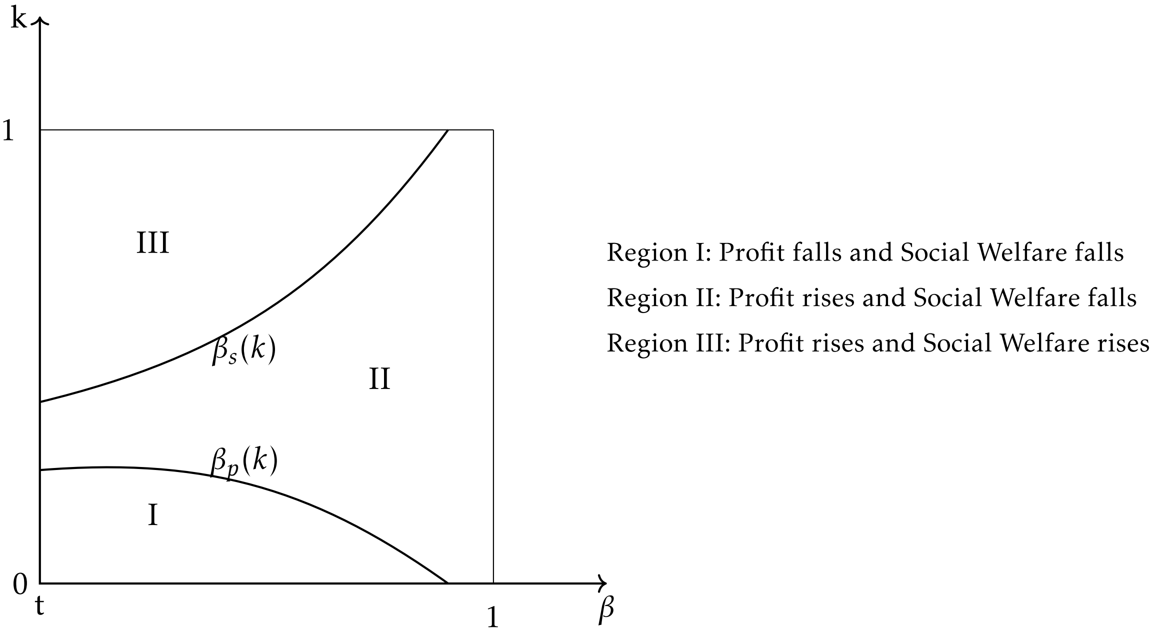
The following proposition summarizes the main results on welfare analysis and is illustrated in figure 4.3 below.

Proposition 4.4. *When firm G adopts bundling, then there exists a threshold $\beta_s(k)$, where*

$\beta_s(k) \geq 0$, and $\beta_s(k)$ is increasing in k, such that in a market characterized by (β, k), we have that bundling:

(i) increases social welfare for $0 \leq \beta \leq \beta_s(k)$, and

(ii) reduces social welfare for $\beta_s(k) < \beta \leq 1$.



The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$, and $\theta = 0.6$. The $\beta_p(k)$ is as defined previously. The threshold $\beta_s(k)$ represents the loci of points along which social welfare under independent pricing and bundling are the same. In region II, there is a divergence between social and private incentive to bundle.

Figure 4.3: Change in Social Welfare

The net effect on welfare would depend on the strength of parameters k and β, among other parameters. For large k, bundling would improve advertisers' revenue from placing advertisements in firm G which might be sufficient to overcome distortions arising from a rise in nuisance cost, transportation cost, bundled discounts, and fall in advertising revenue on firm S. Whereas when k is small, the data advantage is not very significant and thus, the gains from bundling would be lower, which would not be sufficient to improve overall welfare. However, for β approaching to 1, this effect vanishes and bundling always improves social welfare.

When nuisance cost parameter t increases, then bundling is more likely to decrease social welfare. Intuitively, this happens because equilibrium advertising levels will be low and in market 1, user surplus in the absence of bundling would be higher.

4.4.2 User Welfare

Competition authorities dealing with bundling cases in two sided markets are interested not just in how regulation affects social welfare, but also how user welfare changes with intervention. So, in this section, user welfare effects are analysed. User welfare under independent pricing is

$$UW^* = \int_{\frac{q^*_{G1}}{\theta}}^{1} (v\theta - q^*_{G1})dv + \int_0^{\hat{c}} \left[X - tm^*_{G2} - \tau c\right]dc + \int_{\hat{c}}^{1} \left[X - tm^*_{S2} - \tau(1-c)\right]dc, \quad (4.31)$$

whereas user welfare under bundling is

$$\widetilde{UW} = \int_0^1 \int_0^{f(v)} [v\theta + X - t\widetilde{m}_{G2} - \widetilde{q}_{G2} - \tau c]dc\,dv + \int_0^1 \int_{f(v)}^{1} [X - t\widetilde{m}_{S2} - \tau(1-c)]dc\,dv, \quad (4.32)$$

where $\int_0^1 f(v)$ is as defined in equation (4.20) and the rest of the notations are same as defined in earlier sections. Using equations (4.32) and (4.31), the change in user welfare (ΔUW) can be written as

$$\Delta UW = \underbrace{\int_0^1 v\theta f(v)dv - \int_{\frac{q^*_{G1}}{\theta}}^{1} \left[v\theta - q^*_{G1}\right]dv}_{\text{change in users' surplus in market 1}} - \underbrace{[t\widetilde{m}_{G2}\widetilde{N}_G + t\widetilde{m}_{S2}\widetilde{N}_{S2}] + [tm^*_{G2}N^*_{G2} + tm^*_{S2}N^*_{S2}]}_{\text{change in nuisance cost}}$$

$$+ \underbrace{\tau \int_0^1 [f(v) - f(v)^2]dv - \tau[N^*_{G2} - (N^*_{G2})^2]}_{\text{change in transportation cost}}.$$

So, the net change in user welfare would depend on three effects:

(i) Change in net user surplus in market 1: increases because of bundled discounts.

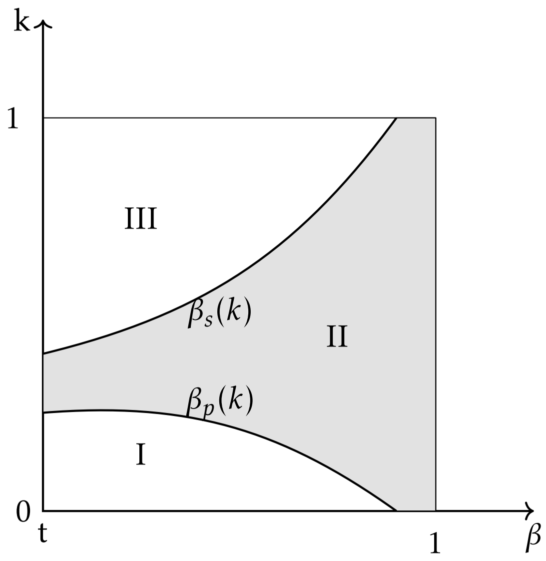
(ii) Change in aggregate nuisance cost of advertisements: increases.

(iii) Change in aggregate transportation cost: increases.

The net effect on aggregate user surplus would depend on whether bundled discounts to users from market 1 are sufficient to offset distortions arising in market 2 because of a rise in aggregate nuisance and trasnportation cost. The main insight is as follows:

Proposition 4.5. *When firm G adopts bundling, then user welfare increases.*

Next, a comparison with social welfare shows that there can be parameter regions, where bundling improves social welfare but at the cost of reduced user welfare. Figure 4.4 below highlights this result.



Region	**I**	**II**	**III**
Firm G's Profit	Falls	Rises	Rises
Social Welfare	Falls	Falls	Rises
User Welfare	Rises	Rises	Rises

The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$ and $\theta = 0.6$. The thresholds $\beta_p(k)$ and $\beta_s(k)$ are as defined previously. In shaded region II, there can be a divergence between regulatory intervention depending on the welfare standard.

Figure 4.4: Change in Firm *G*'s Profit, Social Welfare, and User Welfare

There are three subregions differentiated based on how bundling affects firm *G*'s profit, user and social welfare. In region I, bundling does not generate sufficient advertising revenue. So, both firms and society lose from bundling. In Region II, firm *G*'s profit rises but social welfare falls and user welfare rises with bundling. Thus, there can be a divergence between the two welfare standards. User welfare increases because of bundled discounts that are sufficient to cover a rise in aggregate nuisance and transportation cost in market 2. However, k is not sufficiently high. So, the gain in advertising revenue does not offset distortions in market 2, i.e., higher aggregate nuisance and transportation cost. In region III, investment in data collection and exploitation k is very high. In this region, the rise in advertisers' revenue and bundled discounts are sufficient to offset the cost of distortion in the markets. This increases firm G's profit, user and social welfare.

From figure 4.4, it is clear that the parameter region where the firm and society would disagree depends on the welfare standard. In this case, if competition authorities follow a total welfare standard, bundling would be prohibited in region II, but users would lose. The following proposition summarizes the analysis.

Proposition 4.6. *When firm G can bundle its two products G1 and G2, then there exist thresholds $\beta_p(k)$ and $\beta_s(k)$ such that in a market characterized by (β, k), we have that bundling:*

(i) is not profitable, increases user welfare, but decreases social welfare for $0 \leq \beta < \beta_p(k)$;

(ii) is profitable, increases user welfare, but decreases social welfare for Max $\{\beta_p(k), \beta_s(k)\} \leq \beta \leq 1$;

(iii) is profitable and increases both user and social welfare for $0 \leq \beta \leq \beta_s(k)$.

4.4.3 Policy Implication

From the above analysis, it is clear that how an antitrust intervention affects the market would depend, among other things, on i) the level of investment in data collection, and ii) advertising targeting rate. Here, we consider mandatory unbundling. This approach has been followed by competition authorities in previous tying cases in two sided markets.[7] The impact of mandatory unbundling is summarized in the following corollary.

Corollary 4.1. *The implication for antitrust is (Refer Figure 4.4):*

(i) In markets with large k, the incentive to bundle based on data advantage exists and both user and social welfare increase. Policy intervention becomes less likely.

[7] The investigation of Microsoft by the European Commission (EC) was one of the biggest antitrust tying cases in the European Union. Briefly, Microsoft was accused of abusing its dominant position in the PC Operating System market in two ways. The first related to compatibility issues in the work group server market, and the second to Microsoft's practice of tying its Windows Media Player (WMP) to the Windows OS. On 24th May 2004, European Commission, in its decision on Microsoft case, required Microsoft to offer a version of its Windows client PC OS without WMP to PC manufacturers. Similarly, mandatory unbundling has been decided upon in the Google Android case where the European Commission in its decision has required, among other things, Google to stop forcing manufacturers to pre-install Chrome and Google search in order to offer the Google Play Store on their handsets.

(ii) *In markets with sufficiently small k, the incentive to bundle may or may not exist. If it exists, then social welfare can fall. Since user welfare always increases with bundling, there can be a dilemma for the regulation based on the welfare standard.*

(iii) *When β is large, then policy intervention is complex as it depends on which welfare standard regulator follows.*

4.5 Discussion and Extensions

4.5.1 Endogenous Prices

In this section, we extend the baseline framework in which firms could endogenously set the user prices in market 2. The rest of the specifications are the same as in the baseline model. Under independent pricing, let q_{i2} be the price charged to the users by firm i in market 2. Under bundling, let q_G be the price charged to the users for the bundled good and q_{S2} be the price charged by firm S under both regimes. The main results regarding competition between firms[8] are:

Proposition 4.7. *When firm G could bundle its two goods and prices are endogenous, then there exist β_S and $\widetilde{\beta}_G(k)$, where $0 \leq \widetilde{\beta}_G(k) \leq 1$, such that in a market characterized by (β, k), we have that:*

(i) *for $\beta_S \leq \beta \leq 1$, advertising financed is the equilibrium;*

(ii) *For $\widetilde{\beta}_G(k) < \beta \leq \beta_S$, strategic differentiation is the equilibrium;*

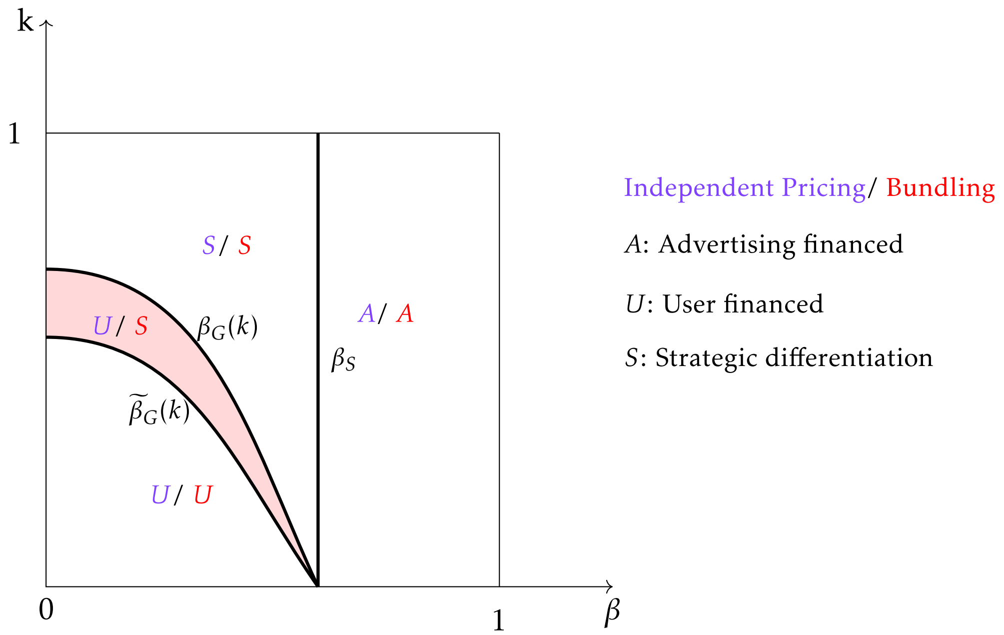
(iii) *For $0 < \beta \leq \widetilde{\beta}_G(k)$, user financed is the equilibrium; and*

(iv) *the parameter space over which strategic differentiation is the equilibrium increases under bundling.*

Figure 4.5 illustrates the results. It shows the parameter space over which different business models would emerge under the two regimes. For the range $\widetilde{\beta}_G(k) < \beta < \beta_G(k)$, a user financed model will emerge as an equilibrium under independent pricing, whereas under bundling, strategic differentiation would emerge as an equilibrium.

[8]For the analysis when firms set prices and advertising quantities independently, refer to chapter 2.

The intuition behind this result is that firm G's best response would change when it could bundle the two goods. Now, it can resort to advertisements as the source of revenue for a larger range of β. A user's utility is interlinked under bundling case, i.e., the addition of utilities from the two goods. So, some users with high v who were earlier not joining firm G due to high τ can now join it. The total demand for firm G and user base over which it can collect data increases. Thus, firm G can rely on advertising revenue for a larger parameter space (the shaded region in figure 4.5).



The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$, and $\theta = 0.6$. The threshold β_S represents the loci of points along which firm S is indifferent between switching business model under independent pricing and bundling. Whereas, the thresholds $\beta_G(k)$ and $\widetilde{\beta}_G(k)$ represent the loci of points along which firm G is indifferent between switching business model under independent pricing and bundling.

Figure 4.5: Endogenous Pricing: Choice of Business Model under Independent Pricing and Bundling

Next, we consider the change in firm G's profit, user welfare and social welfare under independent pricing and bundling. The main results are as follows:

Proposition 4.8. *When firm G can bundle its two products, then there exist thresholds β_S, $\beta_G(k)$, $\widetilde{\beta}_G(k)$, $\beta_p(k)$, $\beta_s(k)$, $\beta_{su}(k)$, $\beta_{ss}(k)$, and $\beta_u(k)$, where a) $\beta_s(k)$ and $\beta_u(k)$ are increasing in k, and b) $\beta_G(k)$, $\widetilde{\beta}_G(k)$, $\beta_p(k)$, $\beta_{su}(k)$, and $\beta_{ss}(k)$ are decreasing in k, such that in a market* *characterized by* (β, k)*, we that bundling:*

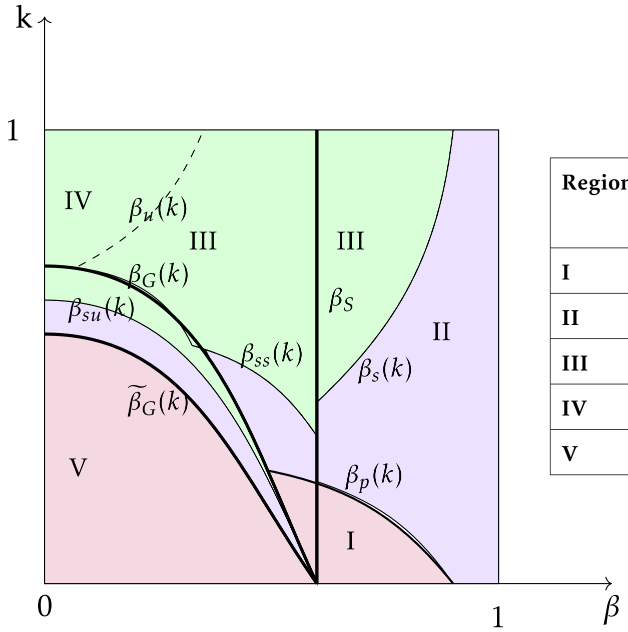
(i) is not profitable and decreases social welfare but increases user welfare for $\beta_G(k) < \beta \leq \beta_p(k)$*;*

(ii) is profitable, increases user welfare but decreases social welfare for a) Max $\{\beta_s(k), \beta_p(k)\} < \beta \leq 1$*, b)* $\widetilde{\beta}_G(k) < \beta \leq \beta_{su}(k)$*, and c)* $\beta_G(k) < \beta \leq \beta_{ss}(k)$*;*

(iii) is profitable and increases both social and user welfare for a) Max $\{\beta_u(k), \beta_G(k), \beta_{ss}(k)\} < \beta \leq \beta_s(k)$*, and b)* $\beta_{su}(k) < \beta \leq \beta_G(k)$*;*

(iv) is profitable and increases social welfare but decreases user welfare for $0 \leq \beta \leq \beta_u(k)$*;*

(v) is not profitable and decreases both user and social welfare for $0 \leq \beta \leq \widetilde{\beta}_G(k)$*.*



Region	Firm G's Profit	User Welfare	Social Welfare
I	Falls	Rises	Falls
II	Rises	Rises	Falls
III	Rises	Rises	Rises
IV	Rises	Falls	Rises
V	Falls	Falls	Falls

The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$, and $\theta = 0.6$. The thresholds β_S, $\beta_G(k)$, $\beta_p(k)$ and $\beta_s(k)$ are as defined previously. The thresholds $\beta_G(k)$ and $\widetilde{\beta}_G(k)$ represent the loci of points along which firm G is indifferent between switching business model under independent pricing and bundling, respectively. The thresholds $\beta_s(k)$, $\beta_{su}(k)$ and $\beta_{ss}(k)$ represent the loci of points along which social welfare is the same under both regimes. The threshold $\beta_u(k)$ represents the loci of points along which user welfare is the same under both regimes. In shaded regions II and IV, there can be a divergence between regulatory intervention depending on the welfare standard.

Figure 4.6: Endogenous Pricing: Change in Firm *G*'s Profit, Social Welfare and User Welfare

Figure 4.6 illustrates the above result. Properties of regions I, II and III are the same as described previously. The new regions are IV and V. In region IV, the equilibrium is strategic differentiation under both regimes. The crucial difference is that user welfare can fall in this region. The bundled discounts do not offset the rise in aggregate transportation and nuisance cost. Social welfare rises because k is very high which means a significant rise in advertising revenue. So, the important regions that highlight the dilemma for policymakers are regions II and IV. In these two regions, social and user welfare would move in opposite directions.

4.5.2 Uncovered Market

In this subsection, we consider the effect of relaxing the covered market assumption. Under partial market coverage in market 2, firms will no longer compete for users and will work as local monopolies. This will have a direct effect on the advertising revenue that they can get in equilibrium. So, although the market expands with bundling, the gain in advertising revenue will be lower now. The net effect depends on whether the gains are sufficient to cover losses in market 1 due to bundled discounts. The main result is as follows:

Proposition 4.9. *Under partial market coverage, bundling is not profitable for firm G.*

4.5.3 Restricting Access to Data Owned by Subsidiaries

In this section, we consider how the main results of this paper are affected when there is no data sharing between affiliated firms. This can happen when a regulator imposes restrictions on access to data owned by a firm's subsidiary, as considered in the previous chapter. Suppose firm G, due to regulatory restrictions, cannot use data regarding users to target advertisements in market 2. The market shares of the two firms will remain the same as in equations (4.10), (4.12) and (4.13) under independent pricing, and as in equations (4.21) and (4.22) under bundling. On the advertising side, the gross advertising revenue under the two regimes will be defined as

$$R(m_{G2}, m_{S2}) = \beta m_{G2} N_{G2} + \beta m_{S2} N_{S2}. \tag{4.33}$$

Relegating the firms' optimization problem to the appendix, here, we summarize the main results on firms' profit and welfare.

Proposition 4.10. *When firm G cannot use data from another market to target advertisements, then there exist $\bar{\beta} \in [0,1]$, such that in a market characterized by (β, k), we have that bundling:*

(i) increases firm G's profit for $\beta > \bar{\beta}$;

(ii) reduces firm S's profit; and

(iii) increases user welfare but decreases social welfare.

Thus, what we can conclude from the above proposition is that bundling is profitable for a smaller parameter space. Intuitively, since firm G cannot use data to target advertisements it can earn higher revenue from bundling only if β is sufficiently high. User welfare always rises because of bundled discounts that they get. However, the rise in transportation cost, nuisance cost and the fall in advertising revenue on firm S offset the gains and reduces social welfare.

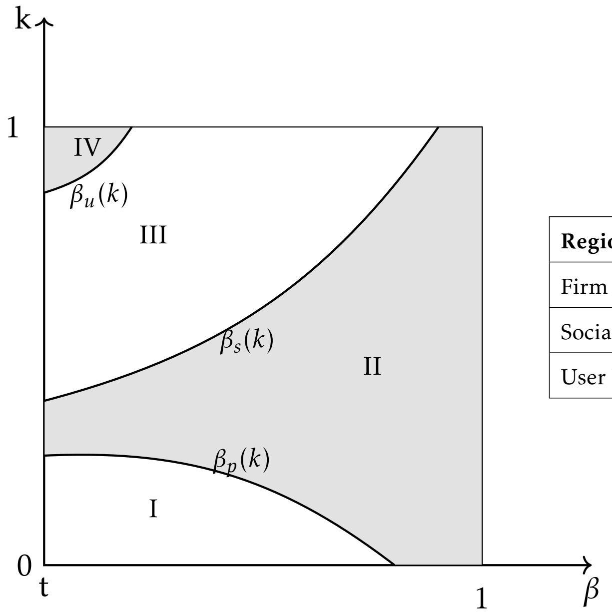
4.5.4 Privacy Costs

In this section, we examine the equilibrium when users are privacy sensitive, i.e., in market 1 there is a cost of data collection that is imposed upon the users. For simplicity, as in the previous chapter, the privacy cost is assumed to be proportional to the level of data collection and is given by γk, where $\gamma > 0$, is per unit cost of data collection for the users. The utility of a user in market 1 is

$$U_{G1}(v) = v\theta - q_{G1} - \gamma k, \tag{4.34}$$

The utility functions in market 2 are the same as defined in (4.4). Relegating the algebra to the appendix, here we present the main results. Two cases can be differentiated depending on the value of θ. There exists a threshold $\bar{\theta}(k)$ such that when

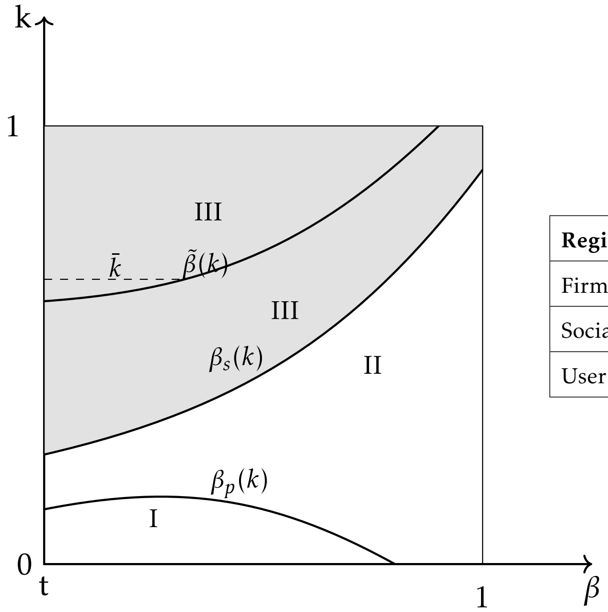
$\theta > \bar{\theta}(k)$, then $q^*_{G1} > 0$ and when $\theta < \bar{\theta}(k)$, then $q^*_{G1} = 0$. Figure 4.7 presents the welfare and equilibrium results for the case when $\theta > \bar{\theta}(k)$. The main difference from the case when users are privacy insensitive is that user welfare can fall with bundling, as shown in region IV. The policy implication when $k = 0$ remains the same as in the baseline model. However, when $k = 1$, policy implication becomes complex. In markets with small β, user welfare falls and social welfare rises; with intermediate β user welfare rises and social welfare rises; and in markets with large β user welfare rises and social welfare falls.



Region	**I**	**II**	**III**	**IV**
Firm G's Profit	Falls	Rises	Rises	Rises
Social Welfare	Falls	Falls	Rises	Rises
User Welfare	Rises	Rises	Rises	Falls

The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$. $\gamma_p = 0.15$ and $\theta = 0.6$.

Figure 4.7: Privacy Costs and Large θ: Change in Firm G's Profit, Social Welfare and User Welfare



Region	I	II	III
Firm G's Profit	Falls	Rises	Rises
Social Welfare	Falls	Falls	Rises
User Welfare	Falls	Falls	Falls

The figure is drawn for parameter values $t = 0.6$, $\tau = 0.5$, $\gamma_p = 0.15$ and $\theta = 0.2$.

Figure 4.8: Privacy Costs and Small θ: Change in Firm G's Profit, Social Welfare and User Welfare

Figure 4.8 represents the case when $\theta < \bar{\theta}$. In this case, as discussed in the previous chapter, when $\beta < \widetilde{\beta}(k)$, $q^*_{G1} = 0$. The policy implication is much clear when $k = 1$ as user welfare falls and social welfare rises for all $\beta \in [0, 1]$. Another interesting insight is that firm S can also profit from bundling. This can happen when $k > \bar{k}$. Intuitively, when k is large, then bundling has two opposite effects on firm S's market share. First, it reduces firm S's market share as firm G bundles the two goods. Second, now privacy costs pushes some users to join firm S who were earlier with firm G in market 2. The net effect of the two depends on the value of k. Hence, firm S can also profit from bundling.

4.5.5 Endogenous Location

The model can also be extended to analyze the endogenous location decisions in market 2. We briefly discuss the equilibrium outcome in this case. As discussed in the previous chapter, firm G would locate close to the centre of the hotelling line. This is because it can collect and monetize data collected from market 1, which can help it to sustain intensified competition in market 2. In the presence of bundling, we expect a similar outcome, but over a larger parameter space. Intuitively, bundling helps firm G to collect data over a larger user set, raising its advertising revenue. Thus. it can sustain losses from intensified competition because of choosing a location closer to the other firm for a larger parameter space.

4.6 Application and Conclusion

In this paper, a multi-product firm has to decide whether to sell two goods independently or as a bundle. A useful application of the results presented is Google's strategy to bundle the Android operating system with its other apps. It can be argued that through MADA requirements, Google can leverage its dominance in the mobile sector to maintain and strengthen its dominance in the search advertising sector. From the beginning, Google has offered Android to hardware manufacturers at no cost. It intends to make no profit from the sale of Android phones to users. Instead, it is used as an indirect tool to attract as much attention as possible from users on other platforms such as Google search, Maps, YouTube, etc. It can use this attention to gain advertising revenue. The mechanism can be explained as follows. Through Android phones, Google has access to critical location data. It can decipher the location where people were when they made the searches. This information, in complement to other information collected, gives Google a big data advantage over its rivals. It not only marginally increases users' search results but raises the willingness to pay by advertisers. This gives Google a premium per click on advertisements compared to any rival in the keyword-based advertising sector. It controls 85 percent of search ad revenues and over 90 percent of mobile advertising revenue.[9] This surge in advertising revenue is not only a result of rising user market share but, more importantly, from the premium price on each click these users make on an advertisement. Thus, in the presence of cross-market data advantage, bundling Android with other apps is a profitable

[9]See Newman (2014), How Mobile Supports Google's Online Adverting Dominance: Why the European Union Competition Authority Should Take Action, Data Justice Policy Brief.

strategy for Google.

Based on a simple model, we showed that the profitability of pure bundling depends on the interplay of investment in data collection and the advertising targeting rate. Bundling is not profitable when investment in data collection and targeting rate are small.

Next, how bundling affects social welfare was evaluated. The parameter regions where private and social incentives to bundle coincide or diverge were specified. The firm can have excess incentives to bundle vis-à-vis a social planner. Lastly, the user welfare component can behave differently with bundling. There exist parametric regions where user welfare and social welfare move in the opposite directions.

4.7 Appendices

Appendix I: Proofs of Baseline Model

Proof of Proposition 4.1

Under independent pricing, the first-order necessary conditions of firms' optimization problem are:

$$1. \frac{\partial \pi_G}{\partial q_{G1}} = 1 - \frac{2q_{G1}}{\theta} - \left[\frac{k(1-\beta)}{\theta}\right]\left[\frac{\tau + tm_{S2} - tm_{G2}}{2\tau}\right] m_{G2} \leq 0. \tag{4.35}$$

$$2. \frac{\partial \pi_G}{\partial m_{G2}} = \frac{\tau + tm_{S2} - 2tm_{G2}}{2\tau} \leq 0. \tag{4.36}$$

$$3. \frac{\partial \pi_S}{\partial m_{S2}} = \frac{\tau + tm_{G2} - 2tm_{S2}}{2\tau} \leq 0. \tag{4.37}$$

Solving the above system of F.O.Cs simultaneously gives us the required solution.

Proof of Proposition 4.2

Under bundling, the first-order necessary conditions of firms' optimization problem are:

$$1. \frac{\partial \pi_G}{\partial q_G} = 2\tau + \theta + 2(tm_{S2} - tm_{G2}) - 2q_G - 2[\beta + k(1-\beta)]m_{G2} \leq 0.$$

$$2. \frac{\partial \pi_G}{\partial m_{G2}} = -2tq_G + [\beta + k(1-\beta)][2\tau + \theta + 2(tm_{S2} - 2tm_{G2}) - 2q_G] \leq 0.$$

$$3. \frac{\partial \pi_S}{\partial m_{S2}} = \beta[2\tau - \theta + 2(tm_{G2} - 2tm_{S2}) + 2q_G] \leq 0..$$

It is a system of linear equations. It can be shown that there does not exist any solution with $q_G^* > 0$, $m_{G2}^* > 0$, and $m_{S2}^* > 0$. The only solution to the above system is with $q_G^* = 0$, $m_{G2}^* > 0$, and $m_{S2}^* > 0$.

Proof of Proposition 4.3

Firm G's profit under the two regimes are

$$\pi_G^* = q_{G1}^* N_{G1}^* + [\beta + k(1-\beta)N_{G1}]N_{G2}^* m_{G2}^* : \text{ Independent Pricing,} \tag{4.38}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\widetilde{N}_G \widetilde{m}_G : \text{ Bundling.} \tag{4.39}$$

Putting in the equilibrium values for strategic variables in equations (4.38) and (4.39) gives

$$\pi_G^* = \frac{[k(1-\beta)\tau]^2}{16t^2\theta} + \frac{k(1-\beta)\tau}{4t} + \frac{\beta\tau}{2t} + \frac{\theta}{4} : \text{ Independent Pricing,} \tag{4.40}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\left[\frac{(6\tau+\theta)^2}{72\tau t}\right] : \text{ Bundling.} \tag{4.41}$$

Now, $\widetilde{\pi}_G = \pi_G^*$ at $\beta_p(k)$, which equals

$$\beta_p(k) = \left\{\frac{(1-k)(12\tau\theta+\theta^2)}{72\tau t} - \frac{\tau k}{4t} + \frac{k^2\tau^2}{8t^2\theta}\right\} * \left\{\frac{k^2\tau^2}{8t^2\theta}\right\}^{-1}$$
$$-\left\{\sqrt{\left[\frac{(1-k)(12\tau\theta+\theta^2)}{72\tau t} - \frac{\tau k}{4t} + \frac{k^2\tau^2}{8t^2\theta}\right]^2 - \frac{k^2\tau^2}{4t^2\theta}\left[\frac{k^2\tau^2}{16t^2\theta} + \frac{\theta}{4} - \frac{k\tau}{4t}\right]}\right\} * \left\{\frac{k^2\tau^2}{8t^2\theta}\right\}^{-1}, \tag{4.42}$$

such that i) for $\beta_S < \beta \leq \beta_p(k) \Rightarrow \widetilde{\pi}_G \leq \pi_G^*$, and ii) for $\beta_p(k) < \beta \leq 1 \Rightarrow \widetilde{\pi}_G > \pi_G^*$. Firm S's profit under the two scenarios are

$$\pi_S^* = \beta m_S^* N_S^* = \frac{\beta}{2t} : \text{ Independent Pricing,} \tag{4.43}$$

$$\widetilde{\pi}_S = \beta \widetilde{m}_S \widetilde{N}_S = \frac{\beta(6\tau-\theta)^2}{72t\tau} : \text{ Bundling.} \tag{4.44}$$

It can be shown that, for all $k, \beta \in [0,1] \Rightarrow \widetilde{\pi}_S < \pi_S^*$. Hence proved.

Proof of Proposition 4.4

Using the equilibrium values for strategic variables, social welfare is

$$SW^* = X - \frac{5\tau}{4} + \frac{3\theta}{8} + \frac{\beta\tau}{t} + \frac{3k(1-\beta)\tau}{8t} + \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2} : \text{ Independent Pricing,} \quad (4.45)$$

$$\widetilde{SW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)^2}{36\tau} + [\beta + k(1-\beta)]\frac{(6\tau+\theta)^2}{72\tau t} + \frac{\beta(6\tau-\theta)^2}{72\tau t} : \text{ Bundling.} \quad (4.46)$$

The change in social welfare is

$$\Delta SW = \frac{-\theta}{8} + \frac{\theta^2(t+\beta)}{36\tau t} + \frac{k(1-\beta)}{72\tau t}\left[9\tau^2 + \theta^2 + 12\tau t\right] - \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2}. \quad (4.47)$$

Putting the preceding equation equal to 0 gives the threshold $\beta_s(k)$ where

$$\beta_s(k) = \left\{\frac{\theta^2}{36\tau t} + \frac{k^2\tau^2}{16t^2\theta} - \frac{k[9\tau^2+\theta^2+12\tau\theta]}{72\tau t}\right\}\left\{\frac{3k^2\tau^2}{16t^2\theta}\right\}^{-1} -$$

$$\left\{\sqrt{\left[\frac{\theta^2}{36\tau t} + \frac{k^2\tau^2}{16t^2\theta} - \frac{k[9\tau^2+\theta^2+12\tau\theta]}{72\tau t}\right]^2 - \frac{3k^2\tau^2}{8t^2\theta} * C_1}\right\}\left\{\frac{3k^2\tau^2}{16t^2\theta}\right\}^{-1},$$

$$\text{where } C_1 = \left[\frac{\theta}{8} + \frac{3k^2\tau^2}{32t^2\theta} - \frac{\theta^2}{36} - \frac{k[9\tau^2+\theta^2+12\tau\theta]}{72\tau t}\right],$$

such that i) $\beta_s(k) > \beta_p(k)$, and ii) $\beta_s(k)$ is decreasing in k. So, i) for $0 \le \beta \le \beta_s(k)$, $\Delta SW < 0$, and ii) for $\beta_s(k) < \beta \le 1$, $\Delta SW > 0$. Hence proved.

Proof of Proposition 4.5

Putting the equilibrium values in equations (4.31) and (4.32), user welfare under the two regimes is

1. $UW^* = X - \frac{5\tau}{4} + \frac{\theta}{8} + \frac{k(1-\beta)\tau}{8t} + \frac{k^2(1-\beta)^2\tau^2}{32\theta t^2}$: Independent Pricing, (4.48)

2. $\widetilde{UW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)}{36\tau}$: Bundling. (4.49)

Using equations (4.48) and (4.49), the change is user welfare is

$$\Delta UW = \frac{\theta^2}{36\tau} + \frac{\theta}{8} - \frac{k(1-\beta)\tau}{8t} - \frac{k^2(1-\beta)^2\tau^2}{32\theta t^2}. \tag{4.50}$$

Setting the preceding equation equal to 0 gives a threshold $\beta_u(k)$ such that

$$\beta_u(k) = \left\{\frac{k\tau}{8t} + \frac{k^2\tau^2}{16t^2\theta}\right\} * \left\{\frac{k^2\tau^2}{16t^2\theta}\right\}^{-1} - \left\{\sqrt{\left[\frac{k\tau}{8t} + \frac{k^2\tau^2}{16t^2\theta}\right]^2 - \frac{k^2\tau^2}{8t^2\theta}\left[\frac{k\tau}{8t} + \frac{k^2\tau^2}{32t^2\theta} - \frac{\theta}{8} - \frac{\theta^2}{36\tau}\right]}\right\} * \left\{\frac{k^2\tau^2}{16t^2\theta}\right\}^{-1}. \tag{4.51}$$

User welfare will increase for $\beta_u(k) \le \beta \le 1$. Now, consider the R.H.S in the preceding equation. It can be shown that $\beta_u(k) < \beta_S$. Thus, user welfare increases. Hence proved.

Appendix II: Endogenous Prices

Proof of Proposition 4.7

For independent pricing, solution to the firms' optimization problem are as specified in the previous chapter. As mentioned there, there exist thresholds β_S and $\beta_G(k)$ such that

(i) If $0 \le \beta_S < \beta \le 1$, the equilibrium is advertising financed, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{\left(\frac{\theta}{2} - \frac{k(1-\beta)\tau}{4t}, 0, \frac{\tau}{t}\right); \left(0, \frac{\tau}{t}\right)\right\}.$$

(ii) If $0 \le \beta \le \beta_G(k)$, the equilibrium is user financed, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{\left(\frac{\theta}{2}, \tau, 0\right); (\tau, 0)\right\}.$$

(iii) If $\beta_G(k) < \beta \le \beta_S$, the equilibrium is strategic differentiation, and

$$\{(q^*_{G1}, q^*_{G2}, m^*_{G2}); (q^*_{S2}, m^*_{S2})\} = \left\{\left(\frac{\theta}{2} - \frac{k(1-\beta)\tau}{4t}, 0, \frac{\tau}{t}\right); (\tau, 0)\right\},$$

where $\beta_G(k) = \left\{ -\left[1 - \frac{k}{2} - \frac{k^2\tau}{t}\right] + \sqrt{\left[1 - \frac{k}{2} - \frac{k^2\tau}{t}\right]^2 - \frac{2k^2\tau}{t}\left[\frac{k}{2} + \frac{k^2\tau}{2t} - t\right]} \right\} * \left\{\frac{k^2\tau}{t}\right\}^{-1}$.

Next, the equilibrium under bundling is specified when firms can endogenously set prices in market 2. A type (v, c)'s user utility is

$$U_{i2} = \begin{cases} v\theta + X - tm_{G2} - q_{G2} - \tau c, & \text{if she joins firm } G, \\ X - tm_{S2} - q_{S2} - \tau(1-c), & \text{if she joins firm } S. \end{cases}$$

Using this, the market share of firm i, i.e., N_i is

$$N_G = \frac{\theta + 2\tau}{4\tau} + \frac{tm_{S2} - tm_{G2}}{2\tau} + \frac{q_{S2} - q_G}{2\tau}, \text{ and} \tag{4.52}$$

$$N_{S2} = \frac{2\tau - \theta}{4\tau} + \frac{tm_{G2} - tm_{S2}}{2\tau} + \frac{q_G - q_{S2}}{2\tau}. \tag{4.53}$$

Each firm i will maximize its profit w.r.t. user price q_{i2} and advertising quantity m_{i2}. The first-order necessary conditions of firms' optimization problem under bundling are:

$$\frac{\partial \pi_G}{\partial q_G} = 2\tau + \theta + 2(tm_{S2} - tm_{G2}) + 2(q_{S2} - 2q_G) - 2[\beta + k(1-\beta)]m_{G2} \leq 0, \tag{4.54a}$$

$$\frac{\partial \pi_S}{\partial q_{S2}} = 2\tau - \theta + 2(tm_{G2} - tm_{S2}) + 2(q_G - 2q_{S2}) - 2\beta m_{S2} \leq 0, \tag{4.54b}$$

$$\frac{\partial \pi_G}{\partial m_{G2}} = -2tq_G + [\beta + k(1-\beta)][2\tau + \theta + 2(tm_{S2} - 2tm_{G2}) + 2(q_{S2} - q_G)] \leq 0, \tag{4.54c}$$

$$\frac{\partial \pi_S}{\partial m_{S2}} = -2tq_{S2} + \beta[2\tau - \theta + 2(tm_{G2} - 2tm_{S2}) + 2(q_G - q_{S2})] \leq 0. \tag{4.54d}$$

It is a system of linear equations. The solutions are obtained through the following claims:

Claim 1: There does not exist any solution with $\widetilde{q}_G > 0$, $\widetilde{q}_{S2} > 0$, $\widetilde{m}_{G2} > 0$, and $\widetilde{m}_{S2} > 0$.

Claim 2: There does not exist any solution in which firm i relies on both strategic variables.

Claim 3: Since each firm will find it optimal to charge one side, the possible candidates for the solution to the system of F.O.Cs (4.54a) - (4.54d) are:

(i) $\widetilde{q}_G = 0, \widetilde{q}_{S2} = 0, \widetilde{m}_{G2} > 0$, and $\widetilde{m}_{S2} > 0$.

(ii) $\widetilde{q}_G > 0, \widetilde{q}_{S2} > 0, \widetilde{m}_{G2} = 0$, and $\widetilde{m}_{S2} = 0$.

(iii) $\widetilde{q}_G > 0, \widetilde{q}_{S2} = 0, \widetilde{m}_{G2} = 0$, and $\widetilde{m}_{S2} > 0$.

(iv) $\widetilde{q}_G = 0, \widetilde{q}_{S2} > 0, \widetilde{m}_{G2} > 0$, and $\widetilde{m}_{S2} = 0$.

The next step is to check the parametric conditions under which F.O.Cs hold for each candidate solution.

1. Consider that candidate 1 is the solution. Then, putting equations (4.54c) and (4.54d) equal to zero gives $\widetilde{m}_{G2} = \frac{6\tau+\theta}{6t}$, and $\widetilde{m}_{S2} = \frac{6\tau-\theta}{6t}$. Since, $\widetilde{q}_G = \widetilde{q}_{S2} = 0$, equations (4.54a) and (4.54b) must be negative. This gives thresholds β_S and $\widetilde{\beta}_G(k)$ such that

$$\beta_S = t, \text{ and } \widetilde{\beta}_G(k) = \frac{t-\beta}{1-\beta}. \tag{4.55}$$

For candidate 1 to be the equilibrium, $\beta > \beta_S$. Since, $\widetilde{\beta}_G(k) \leq \beta_S$ for all $k \in [0,1]$, the only binding constraint is $\beta > \beta_S$.

2. Consider that candidate 2 is the equilibrium. Then putting equations (4.54a), and (4.54b) equal to 0 gives $\widetilde{q}_G = \frac{6\tau+\theta}{6}$ and $\widetilde{q}_{S2} = \frac{6\tau-\theta}{6}$. Since $\widetilde{m}_{G2} = \widetilde{m}_{S2} = 0$, equations (4.54c) and (4.54d) must be less than 0. This gives thresholds β_S and $\widetilde{\beta}_G(k)$ as defined in equation (4.55). For candidate 2 to be the equilibrium, it must be that $0 \leq \beta \leq \widetilde{\beta}_G(k)$.

3. Consider that candidate 3 is an equilibrium. Then, in a similar way, we get thresholds β_S and $\widetilde{\beta}_G(k)$. For it to be the equilibrium, $\beta > \beta_S$ and $\widetilde{\beta}_G(k) < \beta < \beta_S$. Since both constraints cannot be satisfied at the same time, we get a contradiction.

4. Consider that candidate 4 is the equilibrium. Then, putting equations (4.54c) and (4.54b) equal to 0 gives $\widetilde{m}_{G2} = \frac{6\tau+\theta}{6t}$ and $\widetilde{q}_{S2} = \frac{6\tau-\theta}{6t}$. Also, $\widetilde{q}_G = \widetilde{m}_{S2} = 0$ requires equations (4.54a) and (4.54d) to be less than 0. This gives thresholds β_S and $\widetilde{\beta}_G(k)$ as defined in (4.55). For it be the equilibrium, it must be that $\widetilde{\beta}_G(k) < \beta \leq \beta_S$.

Hence, the equilibrium solutions are candidate 1, 2 and 4. There does not exist multiple equilibria and candidate 3 does not satisfy F.O.Cs for any parameter range. Now, since $\widetilde{\beta}_G(k) \leq \beta_G(k)$, the parameter range over which strategic differentiation is the solution expands under bundling.

Proof of Proposition 4.8

In order to prove this proposition, we compare the firms' profit, social welfare and user welfare under the two regimes. There can be four sub-cases depending on the equilibrium business model under each regime. For $0 < \beta_S < \beta \leq 1$, we are in the case when advertising financed model is the equilibrium business model under independent pricing and bundling. The analysis for this case has already been done in appendix I. So, we restrict our analysis to the case when $\beta \leq \beta_S$.

Case i: when $\beta_G(k) < \beta \leq \beta_S$, i.e., strategic differentiation is the equilibrium under both regimes.

Firm G's profit under the two scenarios are the same as in equations (4.40) and (4.41). Equality of the two profit equations gives the threshold $\beta_p(k)$ as defined in equation (4.42). So, this implies that i) for $\beta_G(k) < \beta \leq \beta_p(k)$, $\widetilde{\pi}_G \leq \pi_G^*$, and ii) for $\text{Max}\{\beta_G(k), \beta_p(k)\} < \beta \leq \beta_S$, $\widetilde{\pi}_G > \pi_G^*$.

Next, firm S's profit under the two regimes are

$$\pi_S^* = \frac{\tau}{2} : \text{ Independent Pricing,} \tag{4.56}$$

$$\widetilde{\pi}_S = \frac{(6\tau - \theta)^2}{72\tau} : \text{ Bundling.} \tag{4.57}$$

It can be seen that, for all $k, \beta \in [0, 1]$, $\widetilde{\pi}_S < \pi_S^*$.

Social welfare under the two regimes is

$$SW^* = X - \frac{3\tau}{4} + \frac{3\theta}{8} + \frac{\beta\tau}{2t} + \frac{3k(1-\beta)\tau}{8t} + \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2} : \text{ Independent Pricing,} \quad (4.58)$$

$$\widetilde{SW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)^2}{36\tau} + [\beta + k(1-\beta)]\frac{(6\tau+\theta)^2}{72\tau t} + \frac{(6\tau-\theta)^2}{72\tau} : \text{ Bundling.} \quad (4.59)$$

Using equations (4.58) and (4.59), the change in social welfare is

$$\Delta SW = \frac{-7\theta}{24} + \frac{\beta\theta^2}{72\tau t} + \frac{\beta\theta}{6t} + \frac{k(1-\beta)}{72\tau t}\left[9\tau^2 + \theta^2 + 12\tau t\right] - \frac{3k^2(1-\beta)^2\tau^2}{32\theta t^2}. \quad (4.60)$$

Putting the preceding equation equal to 0 gives a threshold $\beta_{ss}(k)$ such that

$$\beta_{ss}(k) = \left\{\frac{\theta^2}{72\tau t} + \frac{\theta}{6t} - \frac{k(9\tau^2+\theta^2+12\tau t)}{72\tau t} + \frac{3k^2\tau^2}{16\theta t^2}\right\} * \left\{\frac{3k^2\tau^2}{16\theta t^2}\right\}^{-1} -$$

$$\left\{\sqrt{\left[\frac{\theta^2}{72\tau t} + \frac{\theta}{6t} - \frac{k(9\tau^2+\theta^2+12\tau t)}{72\tau t} + \frac{3k^2\tau^2}{16\theta t^2}\right]^2 - \frac{3k^2\tau^2}{8\theta t^2}C_2}\right\} * \left\{\frac{3k^2\tau^2}{16\theta t^2}\right\}^{-1},$$

$$\text{where } C_2 = \left[\frac{3k^2\tau^2}{32\theta t^2} + \frac{7\theta}{24} - \frac{k(9\tau^2+\theta+12\tau t)}{72\tau t}\right].$$

When $\beta_{ss}(k) > \beta_G(k)$, then i) for $\beta_G(k) < \beta \leq \beta_{ss}(k)$, $\Delta SW < 0$, and ii) for Max $\{\beta_G(k), \beta_{ss}(k)\} < \beta \leq \beta_S$, $\Delta SW > 0$.

The user welfare under both regimes is the same as given in equations (4.48) and (4.49) under advertising financed equilibrium. Therefore, following the same line of argument, i) for $0 \leq \beta \leq \beta_u(k)$, user welfare falls, and ii) for $\beta_u(k) < \beta \leq 1$, user welfare rises.

Case ii: when $\widetilde{\beta}_G(k) < \beta \leq \beta_G(k)$, i.e., user financed is the equilibrium under independent pricing and strategic differentiation is the equilibrium under bundling.

Consider the case when $\widetilde{\beta}_G(k) < \beta \leq \beta_G(k)$. Firm G's profit under the two regimes are

$$\pi_G^* = q_{G1}^* N_{G1}^* + q_{G2}^* N_{G2}^* : \text{ Independent Pricing,} \tag{4.61}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\widetilde{m}_G \widetilde{N}_G : \text{ Bundling.} \tag{4.62}$$

Putting in the values for strategic variables under the two cases gives

$$\pi_G^* = \frac{\theta}{4} + \frac{\tau}{2} : \text{ Independent Pricing,} \tag{4.63}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\left[\frac{(6\tau+\theta)^2}{72\tau t}\right] : \text{ Bundling.} \tag{4.64}$$

Equating $\widetilde{\pi}_G = \pi_G^*$ gives a threshold $\beta_{psu}(k)$ such that

$$\beta_{psu} = \frac{1}{1-k}\left\{\left[\frac{\theta}{4} + \frac{\tau}{2}\right]\left[\frac{(6\tau+\theta)^2}{72\tau t}\right]^{-1} - k\right\}. \tag{4.65}$$

It can shown that when $\widetilde{\beta}_G(k) < \beta_{psu}(k) \leq \beta_G(k)$, then i) for $\widetilde{\beta}_G(k) \leq \beta \leq \beta_{psu}(k)$, $\widetilde{\pi}_G \leq \pi_G^*$, and ii) for $\beta_{psu}(k) < \beta \leq \beta_G(k)$, $\widetilde{\pi}_G > \pi_G^*$. Whereas, for firm S, profit equations under the two regimes are as in equations (4.56) and (4.57). Thus, its profit falls under bundling.

Social welfare under the two regimes is

$$SW^* = X + \frac{\theta}{2} - \frac{1}{8} - \frac{\tau}{4} + \frac{1}{32\theta} : \text{ Independent Pricing,} \tag{4.66}$$

$$S\tilde{W} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)^2}{36\tau} + [\beta + k(1-\beta)]\frac{(6\tau+\theta)^2}{72\tau t} + \frac{(6\tau-\theta)^2}{72\tau} : \text{ Bundling.} \tag{4.67}$$

Using equations (4.66) and (4.67), the change is social welfare is

$$\Delta SW = \frac{k(1-\beta)(6\tau+\theta)^2}{72\tau t} + \frac{\beta\theta^2}{72\tau t} + \frac{\beta\theta}{6t} + \frac{\beta\tau}{2t} + \frac{1}{32\theta} + \frac{1}{8} - \frac{\tau}{2} - \frac{5\theta}{12}. \tag{4.68}$$

Putting the preceding equation equal to 0 gives a threshold $\beta_{su}(k)$ such that

$$\beta_{su}(k) = \left\{\frac{\tau}{2} + \frac{5\theta}{12} + \frac{1}{32\theta} - \frac{1}{8} - \frac{k(6\tau+\theta)^2}{72\tau t}\right\} * \left\{\frac{\theta^2}{72\tau t} + \frac{\theta}{6t} + \frac{\tau}{2t} - \frac{k(6\tau+\theta)^2}{72\tau t}\right\}^{-1}. \quad (4.69)$$

It can be shown that when $\widetilde{\beta}_G(k) < \beta_{su}(k) \leq \beta_G(k)$, then i) for $\widetilde{\beta}_G(k) < \beta \leq \beta_{su}(k)$, $\Delta SW < 0$, and ii) for $\beta_{su}(k) < \beta \leq \beta_G(k)$, $\Delta SW > 0$.

User welfare under the two regimes is

$$UW^* = X - \frac{5\tau}{4} + \frac{\theta}{8} : \text{ Independent Pricing,} \quad (4.70)$$

$$\widetilde{UW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)}{36\tau} : \text{ Bundling.} \quad (4.71)$$

It can be shown that $\widetilde{UW} > UW^*$.

Case iii: when $0 \leq \beta \leq \widetilde{\beta}_G(k)$, *i.e., user financed is the equilibrium under both regimes.*

Firm G's profit under independent pricing is as given in equation (4.63). Whereas, under bundling, it is

$$\widetilde{\pi}_G = \widetilde{q}_G \widetilde{N}_G = \frac{(6\tau+\theta)^2}{72\tau}. \quad (4.72)$$

Simple calculations will show that $\widetilde{\pi}_G < \pi^*_G$ for all $k, \beta \in [0,1]$. Firm S's profit will fall too.

Social welfare under the two regimes is

$$SW^* = X - \frac{\tau}{4} + \frac{\theta}{2} - \frac{1}{8} + \frac{1}{32\theta} : \text{ Independent Pricing,} \quad (4.73)$$

$$\widetilde{SW} = X - \frac{\tau}{4} + \frac{\theta}{4} + \frac{\theta^2}{24\tau} : \text{ Bundling.} \quad (4.74)$$

Using equations (4.73) and (4.74), the change is social welfare is

$$\Delta SW = -\frac{\theta}{4} + \frac{1}{8} - \frac{1}{32\theta} + \frac{\theta^2}{24\tau}. \quad (4.75)$$

From the preceding equation, it can be seen that, for sufficiently small θ, $\Delta SW < 0$.

User welfare under the two regimes is

$$UW^{*} = X - \frac{5\tau}{4} + \frac{\theta}{8} : \text{ Independent Pricing,} \tag{4.76}$$

$$\widetilde{UW} = X - \frac{3\tau}{2} + \frac{\theta}{4} + \frac{(3\tau - \theta)^2}{36\tau} : \text{ Bundling.} \tag{4.77}$$

Using equations (4.76) and (4.77), it can be shown that user welfare falls with bundling.

Appendix III: Uncovered Market

Proof of Proposition 4.10

When there is partial market coverage, then a user indifferent between joining firm $i = G, S$, or not is given by $U_i \geq 0$, i.e., she will join firm i if she gets a non-negative utility. There is no direct competition between firms. So, each firm will operate as a monopoly and maximize its profit without taking into consideration other firm's move.

The market shares of the two firms under independent pricing are

$$N_{G1} = 1 - \frac{q_{G1}}{\theta},\ N_{G2} = \frac{X - tm_{G2} - q_G}{\tau},\ \text{and } N_{S2} = \frac{X - tm_{S2}}{\tau}. \tag{4.78}$$

Solving each firms' optimization problem gives equilibrium values as

$$q^{*}_{G1} = \frac{\theta}{2} - \frac{k(1-\beta)X^2}{8t\tau},\ m^{*}_{G2} = \frac{X}{2t},\ \text{and } m^{*}_{S2} = \frac{X}{2t}. \tag{4.79}$$

Next, we consider the bundling regime. The market shares of the two firms under bundling are

$$N_G = \frac{v\theta + X - tm_{G2} - q_G}{\tau}, \text{ and } N_{S2} = \frac{X - tm_{S2}}{\tau}. \tag{4.80}$$

The solution to the first-order conditions of each firm's optimization problem gives equilibrium values as

$$\widetilde{q}_G = 0, \; \widetilde{m}_{G2} = \frac{\theta + 2X}{4t}, \text{ and } \widetilde{m}_{S2} = \frac{X}{2t}. \tag{4.81}$$

Firm G's profit under the two regimes is

$$\pi_G^* = \frac{\theta}{4} + \frac{\beta X^2}{4\tau t} + \frac{k^2(1-\beta)^2 X^4}{64\tau^2 t^2 \theta} + \frac{k(1-\beta)X^2}{8\tau t}, \tag{4.82}$$

$$\widetilde{\pi}_G = \frac{[\beta + k(1-\beta)](\theta + 2X)^2}{16\tau t}. \tag{4.83}$$

The change in profit due to bundling is

$$\Delta\pi_G = \frac{\beta(\theta^2 - 4X)}{16\tau t} + \frac{k(1-\beta)(\theta^2 + 2X^2 + 4X)}{16\tau t} - \frac{k^2(1-\beta)^2 X^4}{64\tau^2 t^2 \theta} - \frac{\theta}{4}. \tag{4.84}$$

Simple calculations will show that $\Delta\pi_G < 0$ for all $k, \beta \in [0,1]$.

Appendix IV: Restricting Access to Data Owned by Subsidiaries

When there is a restriction to access data from an affiliated firm, then firms' profit functions are

$$\pi_G = q_{G1} N_{G1} + \beta m_{G2} N_{G2}, \tag{4.85}$$

$$\pi_S = \beta m_{S2} N_{S2}. \tag{4.86}$$

Under bundling $N_{G1} = N_{G2} = N_G$. So, the targeting rate on both firms are equal to β. The equilibrium profit of the firms are

$$\pi_G^* = \frac{\theta}{4} + \frac{\beta\tau}{2t}, \text{ and } \pi_S^* = \frac{\beta\tau}{2t} : \text{ under Independent Pricing,} \tag{4.87}$$

$$\widetilde{\pi}_G = \frac{\beta(6\tau+\theta)^2}{72\tau t}, \text{ and } \widetilde{\pi}_S = \frac{\beta(6\tau-t)^2}{72\tau t} : \text{ under Bundling.} \tag{4.88}$$

It can be shown that there exists $\bar{\beta}$, $\bar{\beta} > \beta_S$ such that $\widetilde{\pi}_G > \pi_G^*$ for all $\beta \in [\bar{\beta}, 1]$, and $\widetilde{\pi}_G < \pi_G^*$ for all $\beta \in [\beta_S, \bar{\beta}]$, where

$$\bar{\beta} = \frac{\theta}{4}\left[\frac{(6\tau+\theta)^2}{72\tau t} - \frac{\tau}{2t}\right]^{-1}. \tag{4.89}$$

User welfare under the two regimes is

$$UW^* = X + \frac{\theta}{8} - \frac{5\tau}{4}, \tag{4.90}$$

$$\widetilde{UW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} + \frac{(3\tau-\theta)^2}{36\tau}. \tag{4.91}$$

For all $\beta \in [0,1]$, $\widetilde{UW} > UW^*$. Social welfare under the two regimes is

$$SW^* = X + \frac{3\theta}{8} - \frac{5\tau}{4} + \frac{\beta\tau}{t}, \tag{4.92}$$

$$\widetilde{SW} = X - \frac{5\tau}{4} + \frac{\theta}{4} + \frac{\theta^2}{36\tau} + \frac{\beta\tau}{t} + \frac{\beta\theta^2}{36\tau t}. \tag{4.93}$$

The change in social welfare is less than 0 for all $\beta \in [0,1]$. Hence proved.

Appendix V: Privacy Costs

The equilibrium advertising quantities and market shares under independent pricing were derived in the previous chapter. They are

$$\{q^*_{G1}, m^*_{G2}; m^*_{S2}\} = \left\{\frac{\theta}{2} - \frac{\gamma k}{2} - \frac{k(1-\beta)\tau}{4t}, \frac{\tau}{t}; \frac{\tau}{t}\right\},$$

$$\{N^*_{G1}, N^*_{G2}; N^*_{S2}\} = \left\{\frac{1}{2} - \frac{\gamma k}{2\theta} + \frac{k(1-\beta)\tau}{4t\theta}, \frac{1}{2}; \frac{1}{2}\right\}.$$

The equilibrium advertsiing quantities and market shares under bundling are

$$\{\widetilde{q}_G, \widetilde{m}_{G2}; \widetilde{m}_{S2}\} = \left\{0, \frac{6\tau + \theta - 2\gamma k}{6t}; \frac{6\tau - \theta + 2\gamma k}{6t}\right\},$$

$$\{\widetilde{N}_G, \widetilde{N}_S\} = \left\{\frac{6\tau + \theta - 2\gamma k}{12}, \frac{6\tau - \theta + 2\gamma k}{12}\right\}$$

There can be two cases depending on whether $q^*_{G1} > 0$ or $q^*_{G1} = 0$.

Case I: $q^*_{G1} > 0$

Recall from previous chapter that when θ is sufficiently large, then firm G always set $q^*_{G1} > 0$. The profit of firm G under the two regimes is

$$\pi^*_G = \frac{\theta}{4} + \frac{\beta\tau}{2t} + \frac{k(1-\beta)\tau}{4t} - \frac{k\gamma}{2} + \frac{k^2(1-\beta)^2\tau^2}{16t^2\theta} - \frac{k^2(1-\beta)\gamma\tau}{4t\theta} + \frac{\gamma^2 k^2}{4\theta} : \text{ Independent Pricing,}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\frac{[6\tau + \theta - 2\gamma k]^2}{72\tau t} : \text{ Bundling.}$$

Whereas, profit of firm S under the two regimes is

$$\pi^*_S = \frac{\beta\tau}{2t} \text{ Independent Pricing,}$$

$$\widetilde{\pi}_S = \frac{\beta[6\tau - \theta + 2\gamma k]}{6t} \text{ Bundling.}$$

Now, equating $\pi^*_G = \widetilde{\pi}_G$ gives a threshold $\beta_p(k)$ such that

$$\beta_p(k) = \left\{\frac{\tau}{4t}(2-k) - \frac{(1-k)(6\tau+\theta-2\gamma k)^2}{72\tau t} - \frac{k^2\tau^2}{8t^2\theta} + \frac{k^2\gamma\tau}{4t\theta}\right\} * \left\{\frac{k^2\tau^2}{8t^2\theta}\right\}^{-1} -$$

$$\left\{\sqrt{\left[\frac{\tau}{4t}(2-k) - \frac{(1-k)(6\tau+\theta-2\gamma k)^2}{72\tau t} - \frac{k^2\tau^2}{8t^2\theta} + \frac{k^2\gamma\tau}{4t\theta}\right]^2 + \frac{k^2\tau^2}{4t^2\theta}C_3}\right\} * \left\{\frac{k^2\tau^2}{8t^2\theta}\right\}^{-1},$$

$$\text{where } C_3 = \frac{k(6\tau+\theta-2\gamma k)^2}{72\tau t} - \frac{\theta}{4} - \frac{k\tau}{4t} - \frac{k^2\tau^2}{16t^2\theta} + \frac{k\gamma}{2} + \frac{k^2\gamma\tau}{24t\theta} - \frac{k^2\gamma^2}{4\theta}.$$

So, for $\beta_p(k) < \beta \leq 1$, $\widetilde{\pi}_G > \pi_G^*$ and it falls otherwise with bundling. Next, we compare user and social welfare under the two regimes. User welfare is

$$UW^* = X - \frac{5\tau}{4} + \frac{\theta}{8}\left(1-\frac{\gamma k}{\theta}\right)^2 + \frac{k^2\tau^2(1-\beta)^2}{32t^2\theta} + \left(1-\frac{\gamma k}{\theta}\right)\frac{k(1-\beta)\tau}{8t} : \text{ Independent Pricing,}$$

$$\widetilde{UW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} - \frac{\gamma k}{3} - \frac{\gamma k\theta}{12\tau} + \frac{(3\tau-\theta-\gamma k)^2}{36\tau} : \text{ Bundling.}$$

Equating $UW^* = \widetilde{UW}$ gives a threshold $\beta_u(k)$ such that

$$\beta_u(k) = \left\{\left[\frac{k^2\tau^2}{16t^2\theta} + \left(1-\frac{\gamma k}{\theta}\right)\frac{k\tau}{8t}\right] - \sqrt{\left[\frac{k^2\tau^2}{16t^2\theta} + \left(1-\frac{\gamma k}{\theta}\right)\frac{k\tau}{8t}\right]^2 - \frac{k^2\tau^2}{8t^2\theta}C_4}\right\} * \left\{\frac{k^2\tau^2}{16t^2\theta}\right\}^{-1},$$

$$\text{where } C_4 = \left(1-\frac{\gamma k}{\theta}\right)\frac{k\tau}{8t} + \frac{k^2\tau^2}{32t^2\theta} + \frac{\theta}{8}\left(1-\frac{\gamma k}{\theta}\right)^2 + \frac{\gamma k\theta}{12\tau} + \frac{\gamma k}{3} + \frac{\tau}{4} - \frac{5\theta}{12} - \frac{(3\tau-\theta-\gamma k)^2}{36\tau}.$$

User welfare will increase for $\beta_u(k) < \beta \leq 1$ and it falls otherwise. Social welfare under the two regimes is

$$SW^* = X - \frac{5\tau}{4} + \frac{\beta\tau}{t} + \frac{3\theta}{8} + \frac{3k(1-\beta)\tau}{8t} + \frac{3k^2(1-\beta)^2\tau^2}{32t^2\theta} - \frac{3k\gamma}{4} - \frac{3k^2(1-\beta)\tau\gamma}{8t\theta} + \frac{3k^2\gamma^2}{8\theta} : \text{ IP,}$$

$$\widetilde{SW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} - \frac{\gamma k}{3} - \frac{\gamma k\theta}{12\tau} + \frac{(3\tau-\theta-\gamma k)^2}{36\tau} + \frac{k(1-\beta)(6\tau+\theta-2\gamma k)^2}{72\tau t} + \beta\left[\frac{(6\tau+\theta-2\gamma k)^2}{72\tau t} + \frac{(6\tau-\theta+2\gamma k)^2}{72\tau t}\right]: \text{ Bundling.}$$

Equating $SW^* = \widetilde{SW}$ gives a threshold $\beta_s(k)$ such that

$$\beta_s(k) = \left\{\left[\frac{3k\tau}{8t} + \frac{3k^2\tau}{8t\theta}\left(\frac{\tau}{2t}-\gamma\right) - \frac{\tau}{t} + \frac{(6\tau+\theta-2\gamma k)^2(1-k)}{72\tau t} + \frac{(6\tau-\theta+2\gamma k)^2}{72\tau t}\right]\right\}\left\{\frac{3k^2\tau^2}{16t^2\theta}\right\}^{-1}$$

$$+\left\{\sqrt{\left[\frac{3k\tau}{8t} + \frac{3k^2\tau}{8t\theta}\left(\frac{\tau}{2t}-\gamma\right) - \frac{\tau}{t} + \frac{(6\tau+\theta-2\gamma k)^2(1-k)}{72\tau t} + \frac{(6\tau-\theta+2\gamma k)^2}{72\tau t}\right]^2 + \frac{3k^2\tau^2}{8t^2\theta}C_3}\right\}\left\{\frac{3k^2\tau^2}{16t^2\theta}\right\}^{-1},$$

$$\text{where } C_3 = \frac{3k^2}{8\theta}\left(\frac{\tau\gamma}{t} - \frac{\tau^2}{\gamma t} - \gamma^2\right) + \frac{(3\tau-\theta-\gamma k)^2}{36\tau} + \frac{(6\tau+\theta-2\gamma k)^2 k}{72\tau t} - \frac{3k\tau}{8t} - \frac{\tau}{4} - \frac{\gamma k\theta}{12\tau} + \frac{\theta}{24} + \frac{5k\gamma}{12}.$$

Using the preceding equation, it can be shown that for $0 \le \beta \le \beta_s(k)$, $\widetilde{SW} > SW^*$ and falls otherwise with bundling.

Case II: $q^*_{G1} = 0$

Recall from previous chapter that when θ is sufficiently small, then there exists a threshold $\widetilde{\beta}(k)$, such that when $\widetilde{\beta}(k) \in [0,1]$, then firm G can set $q^*_{G1} = 0$. The profit of firm G under the two regimes is

$$\pi^*_G = \left[\beta + k(1-\beta)\left(1 - \frac{\gamma k}{\theta}\right)\right]\frac{\tau}{2t}: \text{ Independent Pricing,}$$

$$\widetilde{\pi}_G = [\beta + k(1-\beta)]\frac{(6\tau+\theta-2\gamma k)^2}{72\tau t}: \text{ Bundling.}$$

Equating $\pi^*_G = \widetilde{\pi}_G$ gives a threshold $\beta_p(k)'$ such that

$$\beta_p(k)' = \left\{\frac{k(6\tau+\theta-2\gamma k)^2}{72\tau t} - \frac{k\tau}{2t}\left(1-\frac{\gamma k}{\theta}\right)\right\} * \left\{\frac{\tau}{2t} - \frac{k\tau}{2t}\left(1-\frac{\gamma k}{\theta}\right) - \frac{(1-k)(6\tau+\theta-2\gamma k)^2}{72\tau t}\right\}^{-1}.$$

It can be shown that $0 < \widetilde{\beta}(k) < \beta_p(k)^{'} \leq 1$. So, for all $\beta_S \leq \beta \leq \widetilde{\beta}(k)$, $\widetilde{\pi}_G > \pi_G^*$ when $q_{G1}^* = 0$. Profit of firm S under the two regimes is

$$\pi_S^* = \frac{\tau}{2t}, \text{ and } \widetilde{\pi}_S = \frac{(6\tau - \theta + 2\gamma k)^2}{72\tau t}.$$

Equating $\widetilde{\pi}_S = \pi_S^*$ gives a threshold $\bar{k} = \theta/2\gamma$ such that for $k > \bar{k}$, firm S also gains from bundling. Next, we compare user and social welfare. User welfare under the two regimes is

$$UW^* = \frac{\theta}{2}\left(1 - \frac{\gamma k}{\theta}\right)^2 + X - \frac{5\tau}{4}: \text{ Independent Pricing,}$$

$$\widetilde{UW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} - \frac{\gamma k}{3} - \frac{\gamma k\theta}{12\tau} + \frac{(3\tau - \theta - \gamma k)}{36\tau}: \text{ Bundling.}$$

It can be shown that $UW^* > \widetilde{UW}$ for all $\beta \in [0,1]$. Social welfare under the two regimes is

$$SW^* = X - \frac{5\tau}{4} + \frac{\theta}{2}\left(1 - \frac{\gamma k}{\theta}\right)^2 + \frac{k\tau}{2t}\left(1 - \frac{\gamma k}{\theta}\right) + \frac{\beta\tau}{2t}\left[2 - k\left(1 - \frac{\gamma k}{\theta}\right)\right]: \text{ Independent Pricing,}$$

$$\widetilde{SW} = X - \frac{3\tau}{2} + \frac{5\theta}{12} - \frac{\gamma k}{3} - \frac{\gamma k\theta}{12\tau} + \frac{(3\tau - \theta - \gamma k)^2}{36\tau} + \frac{k(6\tau + \theta - 2\gamma k)^2}{72\tau t} + \beta\left[\frac{(1-k)(6\tau + \theta - 2\gamma k)^2}{72\tau t} + \frac{(6\tau - \theta + 2\gamma k)^2}{72\tau t}\right]: \text{ Bundling.}$$

The change in social welfare is

$$\Delta SW = \beta\left\{\frac{(1-k)(6\tau + \theta - 2\gamma k)^2}{72\tau t} + \frac{(6\tau - \theta + 2\gamma k)}{72\tau t} - \frac{\tau}{2t}\left[2 - k\left(1 - \frac{\gamma k}{\theta}\right)\right]\right\} + \frac{5\theta}{12} - \frac{\tau}{4} - \frac{\gamma k}{3} - \frac{\gamma k\theta}{12\tau} + \frac{(3\tau - \theta - \gamma k)^2}{36\tau} - \frac{\theta}{2}\left(1 - \frac{\gamma k}{\theta}\right)^2 + \frac{k(6\tau + \theta - 2\gamma k)^2}{72\tau t} - \frac{k\tau}{2t}\left(1 - \frac{\gamma k}{\theta}\right).$$

It can be shown that $\Delta SW > 0$ for all $\beta \in [0,1]$. Hence proved.

Chapter 5

Dilemma in antitrust enforcement: How use of economics can guide enforcement rules in multi sided markets

Gaurav Jakhu* **and** **Payal Malik†**

ABSTRACT

The peculiarities of multi-sided markets set them apart from the traditional market paradigm entrenched in the antitrust policy framework. The working of platforms has posed antitrust issues and concerns, which cannot be dealt using the current rules. This does not imply that the new learnings on the economics of platforms cannot be incorporated into the adjudication of competition law or that Indian competition law is inflexible to accommodate the new insights. This article discusses how the conduct of platforms raises competitive ambiguity. Strategies that are considered harmful in standard economic literature may have a pro-competitive rationale. Such markets, generally, have a tendency to become concentrated with rapid innovation by firms. In this dynamic environment, the authorities should not pursue an ex ante agenda in trying to address a possibility of consumer harm and intervene only when anticompetitive harm is clearly identifiable. The enforcement should exclusively target objectionable activities that hurt consumers (not protect some competitors) leaving other pro-competitive conduct that benefit consumers unregulated.

KEYWORDS: multi-sided platforms, regulation, antitrust, innovation
JEL CLASSIFICATIONS: K21, L10, L40, L44, L51

This is a pre-copyedited, author-produced version of an article accepted for publication in "Journal of Antitrust Enforcement" following peer review. The version of record "Gaurav Jakhu & Payal Malik. Dilemma in antitrust enforcement: how use of economics can guide enforcement rules in multi-sided markets. Journal of Antitrust Enforcement (2017) 5 (2): 260-275" is available online at:

URL: http://www.oxfordjournals.org/licensing/index.html?doi=10.1093/jaenfo/jnx003

DOI: 10.1093/jaenfo/jnx003

*Gaurav Jakhu, Research Fellow, Economics and Planning Unit, Indian Statistical Institute, Delhi Centre, India. Email:gaurav.jakhu87@gmail.com.

†Payal Malik, Senior Advisor/Economist, Competition Law, Public Policy and Economic Regulation and Associate Professor of Economics PGDAV college, University of Delhi. and former Advisor, Economics and Head Economics Division, Competition Commission of India. Email: payal.malik@gmail.com.

Introduction

Multi sided markets permeate a host of our everyday activities. They are rapidly burgeoning today with the evolution of Information Communication Technologies (ICT) and new business strategies to organize complex markets. Their relevance has increased over time in academic writings[1] and also, to some extent, in antitrust enforcement across the globe. Noticeable examples of such markets with its multiple sides include credit cards (merchants and cardholders); newspapers (advertisers and readers); online auction portals (buyers and sellers); social media (users, advertisers, third-party game or content developers and affiliated third-party sites); shopping malls (retail stores and customers); computer operating systems (hardware manufacturers, software application developers and computer users); and search engines (advertisers, content owners and search engine users). Although multi sided platforms (MSPs)[2] emerge in a range of distinct industries, as illustrated above, they share certain common features. The defining characteristic of an MSP is that it resolves a transaction problem and creates value by bring together multiple groups that cannot deal with each other in the absence of it. The presence of cross group network effects sets it apart from the traditional one sided market like car market, the beverages market, etc. which are subject to textbook demand and supply functions. In simple terms, it means the value to an agent (from the transaction) on one side of the market depends on the participation on the other side(s).[3] Strong positive indirect network effects make such markets prone to tipping and concentration, with few firms emerging as giants with a large market share. This is evident, for instance, from the meteoric rise of online platforms with a continual burst of investment in technological development. The leaders in this innovative race have acquired substantial market power and this market leadership position of a few firms in this space has spurred a debate among academics and policy makers alike.

Platforms raise heightened policy concern as the strategies that yield demand side scale economies could foreclose the market for potential new entrants. The platform that grows first can accelerate further with feedback effects between sides. So, it can have an irreversible effect on the market structure with near monopolization of the market in question. Such foreclosure effect is more plausible for MSPs, as achieving a minimum mass of participants is often essential to survive. However, this view is simplistic and overlooks the competitive ambiguity in platform strategies, giving rise to both pro-competitive and anti-competitive potential. The strategies that yield market power help achieve demand side efficiencies, increasing the value to consumers. Moreover, competition between several platforms may not be necessarily beneficial to the consumer when compared to monopolistic market structures. So these markets lead to totally opposite result with innovation and concentration being positively related and in direct conflict with the “traditionalist” view. The application of competition rules, in the presence of ambiguity, might generate false positive results.

In what follows, we will argue that competition authorities and regulators, in making a reliable assessment, should account for complexities of price and non-price decisions in platform markets

---

1 The seminal article by Rochet and Tirole (2003) on “Two sided markets” initiated the research agenda with hundreds of papers on this topic till date.

2 We would use the word multi-sided markets, multi sided platforms and platform markets interchangeably.

3 A newspaper with a larger reader base would be more attractive to an advertiser.

before coming down to conclusions on "observed behaviour" of these firms in network industries. Specifically, we analyze the conduct of stock exchanges, search engines, telecom operators (these are some markets that have come under the scrutiny of both the antitrust and telecommunication regulator in India) and a few other digital platforms based on the current economic literature on MSPs. The attempt is to link new insights that economic theories can provide with antitrust and regulatory approaches in India and draw conceivable principles for effective enforcement.

Network effects and cost based pricing principle

The pricing practices of platform markets and its associated welfare impact have been a long debated subject in antitrust scrutiny.[4] The prevalent thinking among the competition authorities is that too low prices on one side have a predatory intent[5] to exclude rivals and to reinforce its position in the market.[6] And if the price is greater than the incremental cost of providing services to customers, it is taken as evidence of exercising market power. This view is not meaningful in the presence of interdependent demands among various participants of a platform. Demand inter-linkages across the sides of a platform make the prices interlinked too. Theoretical models have robustly established that profit optimization by a platform will consider indirect network effects across groups and price responsiveness by agents on all sides.[7] Thus, the skewed price structure in a competitive equilibrium is common in such markets. This is in line with the objective of a platform to harness network effects and make itself a viable functional entity. It is not the exclusionary intent, but the penetration pricing to attract sufficient mass on one side, thus expanding the reach on other sides and building over a critical mass. Low pricing on one side, in the presence of demand side externalities, can be accompanied by higher prices on other sides.

This pricing principle and interdependent nature of groups can be illustrated through the working of a stock exchange industry. A stock exchange function as a platform market with two participating sides - liquidity demanders and liquidity suppliers. It acts as a supporting platform that facilitates interactions (or transactions) among the agents it serves, exhibiting significant network effects such that the members on one side are better off with greater members on the

4 On 16th November 2009, MCX Stock Exchange Ltd. filed a case under section 19(1)(a) of the Competition Act, 2002 against the National Stock Exchange India Ltd. (NSE), DotEx International Ltd. (DotEx) and Omnesys Technologies Pvt. Ltd. (Omnesys). The case concerned potential pricing and exclusionary conduct abuses by the NSE and other opposite parties with an aim of protecting its market share following the entry of a new competitor (MCX here) in the Indian securities market. Specifically, the source of antitrust concern was in the currency futures exchange services. Briefly, three conduct of NSE were alleged to be anticompetitive by MCX. First, since August 2006, NSE charged no transaction fee on all currency future trades executed on its platform. Second, NSE was charging no admission fee for membership in its CD segment as compared to charging of membership fee in the equity, F&O and debt segments. NSE also did not collect the annual subscription charges and an advance minimum transaction charges in respect of CD segment. Third, NSE did not charge any fee for providing the data feed in respect of its CD segment ever since the commencement of the segment.

5 Competition Act, 2002 defines predatory pricing as "the sale of goods or provision of services, at a price which is below the cost, as may be determined by regulations, of production of the goods or provision of services, with a view to reduce competition or eliminate the competitors."

6 Classical definition of predation is an act of deliberately sacrificing short term revenues, by charging a price below cost, for the purpose of driving rivals out of the market and then recouping the losses.

7 Mark Armstrong, 'Competition in Two-Sided Markets,'(2006) Rand Journal of Economics 37.
Jean-Charles Rochet & Jean Tirole, Platform Competition in Two-Sided Markets, (2003) J. Eur. Econ. Ass'n, 990.

other side. A liquid stock exchange increases the attractiveness of the exchange to the buyers and sellers on any segment.[8] Cross side network effects arise from the interdependent nature of demand among asset buyers and asset sellers, i.e., each side would prefer greater participation from the other side. Thus, an exchange altering its fee on one segment would affect participation in other segments as well.[9] With demand interdependencies, the potential sources of constraints can come from any or all segments operated by a stock exchange. The provision of stock exchange services as the relevant market[10] can serve as an alternative approach for competition assessment.[11] In this set up, the observed behavior of zero price/ below cost price on one of its segments cannot be considered equivalent to anti competitive conduct.[12] The crucial point to

8 The stock exchange performs two fundamental functions: reducing search costs and reducing shared costs. Search costs are incurred by the participants before a "best match" can be determined. The stock exchange, through aggregation of potential candidates, reduces the two sided asymmetric information problem and makes the transaction easier for both sides. The network effects flows in both directions: the more liquidity demanders attract more liquidity suppliers and vice versa. Shared costs are the costs which are common to all sides. Stock exchanges eliminate the shared or duplicate costs by avoiding the need for a barter kind of trade between the agents.

9 For instance, raising the transaction fee substantially on CD segment will reduce the participation there along with overall lower liquidity on the exchange which might reduce the appeal of other segments.

10 The traditional market definition and market power approach involves: defining "the relevant market" which constitutes a set of products constraining the pricing power of the product/products in questions; Calculating firms' shares in the defined market; and inferring significant market power mainly from whether the share is high. Consider the standard "SNNIP test or the critical loss approach to market definition": "Given a small increase in price, the "critical loss" is the amount of output reduction that the hypothetical monopolist needs for its reduced profits from lost sales to exceed its revenues from higher prices on retained sales. The "actual loss" is the output reduction that would result from the small increase in price. Under this approach, if the actual loss is greater than the critical loss, then the products sold by the hypothetical monopolist are a relevant market." In MCX Stock Exchange case, the view that market for stock exchange services forms the relevant market was rejected by the commission altogether. However, there was a contention about whether (over the counter (OTC) services be included in the relevant market. The commission took the view that the "Currency Derivative (CD) segment is the relevant market." The Commission states: "The CD market was futures derivative market where underlying securities are the currencies. OTC market, however, includes various products such as forwards, swaps and options for hedging currency risks. Functionally, the products may be considered as similar, they are different in terms of characteristics and participants. There was differentiation from the OTC market in terms of settlement on maturity, settlement period, counterparty risk, size of market lot and participation, among other things. The equity and equity derivative segments or WDM segment were essentially for investors or speculators who seek to gain from price movements of equities. However, the OTC segment was basically for importers and exporters having contractual exposures and who tried to hedge their risks emanating from fluctuations of exchange rates. The OTC products are not traded on exchanges and only specified entities can participate in this market. A product that is not being traded cannot be said to be a part of any market the two are operating in. Thus, CD market was fundamentally distinct from other segments of the capital market. The relevant market in this case should be the services offered in CD segment."

11 This view is asserted in the COMPAT order in MCX Stock Exchange case. However, COMPAT, in its order, nowhere defines the multi-sided nature of an exchange. It states that "SEBI allowed trading on stock exchanges of – (1) equity; (2) debts; (3) futures; (4) options; and (5) currency derivatives. All the stock exchanges provide trading services in respect of these products, though at the relevant time, the MCX-SX was providing the service only in CD segment. When the competition question comes, it would have to be understood, as to in what manner and what conditions these services were offered by various stock exchanges including NSE and MCX-SX. The very existence of the institution of the stock exchange is for providing services to the speculators, brokers and all those interested in those products. Therefore, what is important is a service not the segments in which the stock exchanges deal."

12 This fallacy is seen in the Commission's Order. The Commission in its final order took NSE's zero pricing in the CD segment as anticompetitive. It states: "NSE incurs various costs in the CD segment- advertisement and publicity, clearing and settlement charges and depreciation. Various waivers are, therefore, evidence of predatory

grasp is that the stock exchange functions as an MSP offering trading services in relation to various assets to the trading on a particular segment functioning as a two sided market. Adopting a price structure reflecting the underlying costs of various segments would only be efficient by chance. The objective of a stock exchange is to improve its liquidity, as traders prefer market with greater liquidity. It would jointly determine the price charged on all its segments to maximize the overall level of transactions. The fees charged for different segments are not done in isolation, but with an intent to promote the overall functioning of the stock exchange and enhance its liquidity position. This could imply "skewed pricing"[13] with one segment charged below cost, and other segments charged positive prices, above cost, to maximize overall profit. Acknowledging the presence of network effects, as mentioned above, in the NSE stock exchange case, a transaction on the CD segment can improve the overall depth and liquidity of the stock exchange and thus the cost based pricing seems to be not the appropriate strategy. An optimal price structure can be zero pricing on CD segment and positive fees for others. This form of pricing which seems predatory at first instance can actually be explained as part of its business strategy as an MSP to maximize the overall value of the platform.

Skewed price structure can also arise when a stock exchange is in its infancy and has to deal with "chicken and egg problem", i.e., designing strategies to get both sides on board and obtaining a critical mass to attract agents on all the sides.[14] [15]One strategy commonly employed is giving services for free on one side to get a critical mass of participation on that side which because of network effects, encourages participation from the other side. Interestingly, Facebook built up its network based on this principle with initial free services to Ivy leagues and later spread its user base outside them. Another strategy is to lower the costs of participation for consumers on one side. Microsoft, for example, invests in application developers by developing tools that make it easier for developers to write applications using Microsoft operating systems. Another component of the design problem is fixed costs of running a segment. A stock exchange has to bear the costs of maintaining and developing a particular segment. A higher liquid segment could generate economies of scale distributing costs over a large range of participants. Thus, fee waiver employed on one segment might have pro competitive potential. In fact, it can be viewed as an essential part of overcoming the potential trap that a platform can get entrenched in. It's a strategy to overcome the "chicken and egg problem" common in MSPs and getting sufficient participation from buyers and sellers on its trading platform. Additionally, a stock exchange has

---

pricing with an intent to lessen/eliminate competition. Transaction fee waiver is not linked to marker development. Reduction of admission fees appears to be a part of a larger scheme of an arrangement to keep all revenue streams in the CD segment to almost zero level, to drive out competitors. Data feed waiver was clearly aimed at consolidating the market share of NSE. As waivers have been continuing for over two years without having any reasonable relationship with costs and without any developmental basis, amounts to predatory pricing."

13 Skewed pricing is highly prevalent in network industries.

14 An important characteristic of two-sided markets is that the demand on each side vanishes if there is no demand on the other—regardless of what the price is. Men will not go to dating clubs that women do not attend because they cannot get a date. Merchants will not take a payment card if no customer carries it because no transaction will materialize. Computer users will not use an operating system that does not have applications they need to run. Sellers of corporate bonds will not use a trading mechanism that does not have any buyers. In all these cases, the businesses that participate in these industries have to figure out ways to get both sides on board. (Evans, 2003).

15 E.Glen Weyl, A Price Theory of Multi-Sided Platforms,(2010) American Economic Review 100 "Such strategies are known as insulating strategies."

both indirect network effects present on any one segment and direct network effects for the exchange as a whole (due to higher liquidity on the exchange). It is possible that, with strong direct network effects, the waiver would be welfare enhancing, as it could stimulate trading and liquidity on one segment which would enhance the liquidity standing of the stock exchange.

Dynamic competition and standard market power analysis

Many MSPs are characterized by continuous R&D efforts with drastic innovation. In contrast to the old economy, industries, platforms engage in dynamic competition in the market usually through research-and-development (R&D) competition to create a product, service, or feature that would replace the existing product and thus create market dominance. This evolutionary nature of MSP is clearly evident in the rise of online platforms such as social networking sites, search engines, e-commerce websites, etc. These internet platforms pose difficulties for policymakers, regulators and competition authorities and challenge the traditional paradigm of market power analysis. Antitrust intervention, based on inferences drawn from standard tools regarding online platforms with complex business structure,[16] can be risky harming consumer interest. The antitrust error can result from the mismatch between the traditional price-oriented approach and the more innovation-oriented approach required to deal with technologically dynamic platforms.

Static price/output competition on the margin in the market is less important. However, market leadership may be contestable because of the constant threat of drastic innovations by rivals. Narrowly focusing on prices, it might overlook the potential competitive constraints from the innovative or improved products. This is not to say that antitrust analysis does not consider innovation and non-price factors, but such concerns have remained secondary in enforcement. The primary focus is on the price effects of a firm’s behavior and static efficiency. Since, the competition is “for the market” in these industries, the potential rivals in the presence of a fierce and dominant competitor would raise their standards of requisite investment to challenge the established position. The platforms to engage in drastic innovation efforts, they must expect to

[16] In India, Google is under scanner and many cases are under investigation by CCI. The antitrust allegations against Google in India also encompass the above issues. In Case No. 07 & 30 of 2012, information under section 19(1)(a) of the Competition Act, 2002 was filed against M/s Google Inc. and M/s Google India Private Limited (collectively “Google”) by M/s Consim Info Private Limited and Consumer Unity & Trust Society (CUTS). Consim Info Pvt. Ltd**.** is engaged in the business of owning and running web portals, the business model of which is based on providing a platform for market users to interact and exchange information and take their relationship forward. The popular web portals of Consim include Bharatmatrimony.com, indiaproperty.com, & clickjobs.com. The informants illustrated the abuse of dominance by Google in the following markets: Algorithmic search markets in India and the search advertising market in India. Consim alleged that Google is allowing its competitors to advertise on words and phrases that Consim has trademarked. In addition, it alleged that Google attempts to leverage its market power to enhance the market position of vertical sites it owns (e.g. YouTube, Google Maps, or Google Books). In the algorithmic search, Google has been accused of lowering the ranking of competing vertical search engines. In search advertising, it lowered the “quality score” for sponsored links of competing vertical search providers. CUTS also alleged that Google was abusing its dominant position by practices like search bias, search manipulation, denial of access and creation of entry barriers for competing search engines etc. Further, Google has been accused of putting exclusionary obligations on its advertising and website partners. It is alleged that it has entered into exclusive vertical agreements with websites for embedded search to prevent competitors from achieving the necessary scale. It is also accused of blocking interoperability of advertising platforms by putting restrictions on data that can be inputted into and outputted from AdWords campaign, thus limiting multi homing by online advertisers.

earn an average rate of return which would be higher than the marginal costs. Thus, for investments to be made, they should produce temporary market power to charge supra-competitive price. In the present antitrust framework the enforcement that limits the ability of incumbent firms to gain market share would have an opposite effect of reducing the reward from a risky investment by the rivals, discouraging them from engaging in R&D in the first place and slowing down the overall innovation process. Thus, cautious and limited intervention in dynamic platform markets is a prudent choice.

An example of over-enforcement involves the Internet search market. With a tremendous flow of knowledge pouring in every day, a user faces the challenge of locating the appropriate data and facts in the vast ocean of know how. A search engine serves to aggregate information and enhance access of consumers to the aggregated content. A search engine is a digital platform through which multiple sides- advertisers, content developers and users- can interact with "unidirectional" network effects, i.e., advertisers value more users, but not vice versa. When a user types in a query in a search engine website, two distinct search results are produced- organic search, and a sponsored search. The organic search results are supplied to us at no cost and display the most relevant links to our query. It includes links to relevant websites or web pages and may also include images, videos and other information. A user can access the websites or other information through a click on the clickable links provided in the search results. The ranking of search results is achieved through the application of sophisticated search algorithms, which are frequently updated. The sponsored results are generated for the purpose of selling the positions in it to the advertisers; and content providers are subsidized by the search engines, as they can upload content for free.[17]

A search engine may become unshakable because of network effects on the demand side. A widely used search engine has a large user base. It creates confidence among the user regarding the reliability of search results, as others are relying upon it too. The presence of strong network effects favors gravitation to a single platform. The rapid technological evolution and economic structure of search platforms along with a tendency toward concentration raises important questions for antitrust policy. The main competition issues linked to dominance in the search engine market are the strategies targeted at reducing multi-homing, leveraging, and exploitative practices.[18] However, can a conventional antitrust analysis correctly determine the strength of a

[17] For instance, the amount Google charges for well positioned sponsored links is calculated according to a Vickrey second price keyword auction conducted through Google's AdWords platform. The auction is an automated process which happen each time a keyword is entered into Google's search engine. The other source of Google's income is from selling advertisements through its Adsense program. It allows placement of display advertisements on third-party websites.

[18] The French Competition Authority, in its Navx decision, dealt with the sudden closure of Navx's AdWords' account by Google for violation of its content policy. The FCA ruled that such closure without warning was discriminatory and nontransparent and asked Google to re-establish Navx account and to ensure the transparency of its content policy. Google committed to make the functioning of its AdWords service more transparent and predictable. It also noted that Google and its subsidiaries participated in the Ad Words service bidding, by purchasing keywords related to their activity, thus artificially raising the cost for competing vertical search engines or competitors of Google's ancillary services and increasing the traffic on its site (and consequently its advertising revenues). The European Commission investigated Google's practices over four concerns practices: Deceptive display, unauthorized content scrapping, exclusivity in advertising agreements, and portability of ad campaign data. EC, in its Statement of Objections, alleged that the company has abused its dominant position in the markets for general internet search services in the European Economic Area (EEA) by systematically

search platform's market power? It runs the risk of overlooking critical factors affecting competitive pressure. A search engine faces competitive constraints not only from current competitors, but also from radical innovations that would put out its own product. The same underlying positive network effects can allow a new entrant with a radical idea to outgrow the established search engine.

Another issue is restricting the relevant market narrowly to a general search and search based advertising. In the digital information world, users can shift to other digital intermediaries to access information. For instance, social networking sites such as Facebook; vertical search engines such as Amazon and eBay for product search and even mobile apps, with increasing penetration of smartphones, can be a substitute for a general search engine. On the other side, display ads and search based advertisements can work as substitutes in a firm's advertisement decision process. With rapidly evolving user online behavior, it would be difficult to define the relevant market in a narrow sense.

In internet search, interestingly, the new platforms can come not only from a startup with a revolutionary idea, but established firms from other businesses expanding it into a totally new field. For instance, studies suggest coming up of new search products by Apple or Facebook. The presence of network effects could spiral it off into a new profitable venture with more traffic. Thus, the static analysis of the dynamic search engine tells us little about the durability of its market power on the user side. On the advertising side also, market power is limited. Since a user can easily switch to other search engines, advertisers would follow the lead. For them, the value of advertising on a search engine is a function of the user traffic it attracts. In fact, studies have shown that advertisers have favored search sites such as Facebook and Amazon. Additionally, advertisers, multi-home, choose more than one search engine.[19]

Private Innovation and Ex ante Regulation

The argument outlined above suggests that antitrust enforcement should prioritize how a platform's strategy affects the innovation process. It might be true that the same strategy might harm competitors in the short term but have welfare gains of innovation in the long run. The regulator should be able to weigh the gains and losses from innovation strategies and intervene only when the overall social welfare would likely reduce. However, absent compelling evidence to substantiate on welfare gains, the competition authorities and regulators have advocated to regulate *ex ante* the behavior of online platforms. One form of government intervention has pressed upon the principle of neutrality at different levels of the internet ecosystem. It is based on the belief that the market itself cannot correct the anti-competitive effects, if any. We explain using examples of two crucial stand points of neutrality debate-network neutrality and search

favoring its own comparison shopping product in its general search results pages. Google had to make commitments to stop scraping and ad-word exclusivity and to give space for results from three competitors along with results from its own verticals. US FTC also outlined concerns over leveraging its own vertical search and exclusivity in AdWords. In its decision, the FTC rejected the allegation of search bias but, like in the EU, Google had to make commitments to stop scraping and ad-word. exclusivity.

19 Marina Lao , "Neutral" Search as a Basis for Antitrust Action? (2013) *Harvard Journal of Law and Technology* (Online Paper Series) 26. It is stated that "For an advertiser with paid links on both Google and Bing, an increase in the firm's advertising costs on Bing would be followed with a corresponding decrease in costs for the same advertising on Google, making possible multi-homing in search-based advertising."

neutrality- that this process of *ex ante* regulating the private innovation would do more harm than good.

Tim Wu coined the term net neutrality and stated: "Network neutrality is best defined as a network design principle. The idea is that a maximally useful public information network aspires to treat all content, sites, and platforms equally. This allows the network to carry every form of information and support every kind of application". Essentially, net neutrality means that all data packets carried over the internet are treated equally and no network operator can exercise control over it. The term "control" can mean differential pricing and discrimination among users, content sites and platforms. Can network neutrality ensure a level playing field between telcos and "over the top" (OTTs) services? It is true that an Internet Service Provider (ISP), working as a platform, can lead to a "competitive bottleneck"[20] situation. It arises when one side single-homes and the other side multi-homes. In equilibrium, there can be strong competition to sign up single-homing side and high prices can be charged to the multi-homing side. In the internet market, this situation can happen if content and application providers (CAPs) multi-home and internet users (IUs) single-home. Under this situation, hypothetically, ISPs could charge high prices to CAPs to allow access to their single-homing IUs resulting in a competitive bottleneck case. They can also apply traffic management, leading to prioritizing, degrading or blocking of content from CAPs which are a threat to a vertically integrated ISP with its own content or application.[21] It can also be used to limit traffic, which is a drain on its limited bandwidth capacity and use it for other profitable transmission of data. However, gradually, with new evidence, policymakers are realizing that there are countervailing factors limiting the ability of an ISP to abuse its market position. The possibility of an exclusive relation between an ISP and CAP in return for some compensation; bargaining power of large CAPs such as Google, Facebook, etc. over network providers; multi-homing even on IUs side (home and workplace internet access) constrain the market power that an ISP might have.

In the Indian context, ironically, however, a non-neutral internet regime is essential to pursue public policy objectives. India requires massive investment in network infrastructure- wire-line and wireless networks- to bridge the digital divide rapidly. Other than public resources, network owners must have sufficient incentives to invest in the network infrastructure to support the state's ability to raise funds. Many OTTs have gained market shares in India and put a strain on limited bandwidth capacity. This is a short term benefit to the consumers, but in the absence of earning profits through such opportunities, ISPs would not have an incentive to improve network infrastructure. The focus should be on efficient management of this infrastructure through CAPs pricing and generating infrastructure investment. In the long run, consumers also would prefer a network with better quality in terms of speed, array of content provided, etc. The most controversial of the non-neutral practices have been the zero rating tariff plans of the ISPs, where they provide access to some selected content for free.

---

20 A competitive bottleneck situation arises when a platform acts as a bottleneck monopoly between the sides and exploits one side to allow access to the other side. See Armstrong (2006).

21 Discriminatory treatment could emerge if an intermediary offers an OTT service such as VOIP telephone and instant messaging which competes with other OTT services. This is a standard vertical relationship issue in which a distributor has to sell its own and the competitor's product and many academic papers have dealt with it.

Another example of innovation changes producing mixed results, i.e., benefiting consumers and excluding rivals, is the algorithmic search innovation in internet search platforms. Concerned about the competitive implications, critics have highlighted "neutrality" as a governing principle to prevent a dominant search engine from abusing its dominant position in the general search. Search neutrality is a principle under which search engines should not discriminate between websites and provide results to search queries based on objective algorithmic standards. Is this hypothesis of leveraging market power in vertical search acceptable for search platforms? This issue becomes complicated in the presence of the underlying complex search algorithm. A search engine frequently updates its algorithm, in fact 500 times a year, to improve the end user experience.[22] The search has evolved from originally ten blue link format to an integrated information platform. This includes blending the organic results with search engines' own non web links such as maps, images, etc. or linking to a search engine's own vertical website at the top of the search result. This evolution of search has raised foreclosure concerns- foreclosing competition by displaying own vertical sites more prominently in search queries. This has also been called tying general search with a vertical search. It must be clarified that the way, tying is used in the internet is different from the standard industrial organization theory. When a user clicks on a vertical search link, there is no price discount he obtains. However, if it is located prominently on top in the search results, it is highly likely that the user would click on it for the search related queries. Thus, tying here means inducement to consume the tied good, i.e., vertical search.

What are the incentives of a non neutral search engine to degrade the user experience? A search engine can potentially degrade the user experience through distortion of search results. It could intentionally lower the ranking of potential advertisers to shift them into sponsored search links. It may also alter the ranking of organic results to place its own vertical links on top. But apart from the ability to degrade, it should also have the incentive to degrade the search results. The presence of such a bias is difficult to establish. Many papers have attempted to establish that a dominant search engine would indulge in search distortion. Some papers have shown that a search engine may provide suboptimal organic search results to increase profits from sponsored results.[23] Other papers have shown that it may also divert traffic to its own integrated verticals.[24]

Nevertheless, in order to build an antitrust theory of consumer harm, the relevant point to establish is that promoting own verticals is anti competitive. Experimental evidence on impact on consumer welfare is ambiguous. Luca and Wu (2015) show that if internal content is inferior to organic search results, then integration reduces consumer welfare.[25] But there can be many

22 Page rank is Google's complex proprietary search algorithm incorporating around 200 factors with different weights. It assesses the relative importance of web pages so that various links can be organized in an order and presented to the searcher.

23 Alexander White , Search Engines: Left Side Quality versus Right Side Profits (2014) IIJO.
Andrei Hagiu and Bruno Jullien, Why Do Intermediaries Divert Search? (2011) Rand Journal of Economics 42.
Xu, L., J. Chen, and A.Whinston, Oligopolistic Pricing with Online Search (2010) Journal of Management Information Systems 27(3).

24 EmanueleTarantino, A Simple Model of Vertical Search Engines Foreclosure. (2013) Telecommunications Policy 3.
B. Edelman and Z. Lai, Exclusive Preferential Placement as Search Diversion: Evidence from Flight Search. (2013) Harvard Business School NOM Unit Working Paper No. 13-087.

25 Michael Luca and Timothy Wu, Does Google Content Degrade Google Search? Experimental Evidence (2016) Harvard Business School working paper 16-035.

checks which can stop a search engine from degrading user experience beyond a certain threshold.[26] There are potential costs, which limits a search platform's behavior to excessively distort the user's search for securing greater advertising revenues. Any attempt by a search platform to tie the vertical search with the general search, which degrades the user's search experience, would have multiplier effects.

If the search results are too irrelevant to users, they can switch to other search platforms. Multi-homing is common is search queries market and a user can use more than one platform. They can switch to a rival search engine- Google, Bing, Yahoo for search results; directly access a vertical search engine for specialized search queries; or even use an entirely different platform like Facebook for finding content, people, etc. It is costless to compare the results with existing search engines.[27]

In the absence of consumer lock in, user demand would diminish over time, reducing advertising demand with a resulting decline in ad revenues. Therefore, a tie could be justified if the display of specialized search results along with general search results is a product improvement, which raises a user's value from a search. A search platform would tend to focus on enhancing the user's value through the redesign of its search results. The integration of maps, images, videos, etc., with organic search results would provide users a better and more accurate answer to its search query. This would increase consumer traffic on its platform and advertisers would place a higher value for ads on the platform, thus increasing its profits from sponsored results. If, on the other hand, users did not prefer specialized results with the general search results, the overall traffic would have reduced with substantial losses for the search engine.[28] It is likely that users prefer search non neutrality.

A tie can improve user value in another way, if users have an aversion for advertisements that are placed on the vertical websites and are indifferent between alternative content sites. De Corni`ere and Taylor (2014)[29] show that:

"Users are most likely to benefit from integration when the integrated content site and search engine are close substitutes in the advertising market and users have a strong aversion to advertisements, in which case any change in the search result quality is likely to be dominated by the equilibrium fall in ad volumes that follows integration. For homogeneous content sites if users are indifferent between publishers, then they are relatively unaffected by bias."

In the absence of any conclusive evidence on consumer welfare, is search neutrality a valid principle? First, implementing search neutrality would affect startup websites. Display of search results based on a purely objective criterion would imply that a search engine cannot manipulate search results to provide users information about new emerging start-ups that are growing fast by placing these websites among top search results and give satisfactory user search results. The

[26] Maurice E. Stucke and Ariel Ezrachi, When Competition fails to optimise quality. A look at search engines (2016) Yale J.L. & Tech. 70.
[27] Michael Katz Network effects and switching costs in online search (2011).
[28] Robert..H. Bork, and J.Gregory. Sidak What does the Chicago School Teach about Internet Search and Antitrust Treatment of Google? (2012) *Journal of Competition Law and Economics* 8.
[29] Alexandre de Corni`ere and Greg Taylor, Integration and search engine bias (2014) RAND Journal of Economics Vol. 45, No. 3.

dominant online platforms would remain at the top of search results, as these are the platforms with a strong installed user base. This affects the ability of new firms to enter the market and compete against the established online platforms. Furthermore, the application of search neutrality principle would inhibit product improvement and the evolution of a search engine. Under regulated search, a search platform's ability to compete would be limited. With less competition in the market, no search engine would have an incentive to invest and improve its search results.

The next issue to address is whether tying of general and vertical search, if profitable, would significantly foreclose the market for rival vertical search sites. The foreclosure effect could be justified only if it could be established that general search results are "essential" for users. The policymakers, in the absence of any benchmark to distinguish beneficial and harmful innovation, have relied upon the essential facility doctrine to pursue an ex ante regulatory framework. The concept of essential facility has been used in vertical foreclosure cases in which an upstream dominant firm refuses to supply firms in the downstream market with required services where it also operates. Therefore, we need to establish both dominance and indispensability of the upstream firm service. However, this concept runs into limitations when applied to a search platform. In order to build an anti competitive case based on the vertical structure of a chain, it is vital to establish that the search platform is dominant in the general search i.e. it must not face competitive constraints from rival engines. If general search is sufficiently competitive, then the degradation of search quality would threaten the existing high market share. Preliminary evidence has shown that rival search engines are just a click away. No websites is excluded from the general search results.

Even defining narrowly, the top results in general search cannot be considered an essential facility for the vertical search providers. Users have multiple alternative avenues to access information in the digital space. In Internet search, it is certainly plausible that some users consider general and vertical as complements, i.e., they rely on general search results to access the vertical search content. But others may directly turn to vertical sites such as Amazon and Trip Advisor, which are a good substitute for general search engines in certain product categories; they can access content through social networking sites and they can even use smartphone apps for online information. An important consideration is that brands are built over time and they cannot be adversely affected in a short time. Even in the presence of tying, users, in the short run, would not suddenly shift their behavior away from rival brands. Thus, in the absence of conclusive evidence, user preference for innovation, and competitive markets, ex ante regulation is unwarranted in the internet search. One last issue for an antitrust authority to consider is whether alternative search solutions to current tying exist to achieve the same efficiency objectives. For instance, sharing of top positions in search results; selected by users of their favorite provider, etc.[30] Implementing alternative formats would introduce technical engineering issues and burden for the dominant search engine to comply with. This would also have an impact on latency. Minimizing delay is significant for a search engine to attract user traffic.[31] Interpreting user queries by the third-party providers could take time, which would

30 A query could be transferred to a third party website who would then return a vertical search content based on relevance.

31 "A Google experiment reported in November 2006 revealed that a 0.5 second delay in generating the SERP caused a 20% drop in traffic, while a further experiment in 2009 found that slowing down the load time of the

delay the overall search results. The larger point being made here is that mandating the search engine to provide a level playing field for the sake of neutrality may be counterproductive for the consumer. If the end consumer the beneficiary of competition laws, then the whole objective of these laws get defeated if the search engine is mandated to be an essential facility.

Innovation changes in dynamic platform markets can help overcome incumbent's sunk cost advantage. If a new entrant can offer product features that are valuable to consumers, it could enter the market despite the incumbent's cost advantage. Examples of dominant platforms collapsing down when faced with a strong, innovative entrant are many. A compelling example is the evolution of social network platforms from the days of MySpace to Facebook. Myspace had very relaxed rules to govern the behaviour of users joining it. It didn't filter out the fake profiles from the genuine one. As a result of its lax policy, it initially had a spiraling effect with a large number of users gravitating to it and making profiles. As growth multiplied, the negative externalities set in and it gained the reputation of "bad and unsafe community" with objectionable content accessible freely on it without any regulation. This distracted advertisers from placing ads on it and over time advertising revenue declined, leading to the demise of MySpace. Facebook, on the other hand, entered with a strict standard to filter out the fake profiles and monitor bad behavior by users. On the advertiser side also, it took steps to regulate the placing of ads. With effective management of different sides, Facebook soon became the dominant social platform. Similar to Facebook, LinkedIn has adopted rules to prohibit members from developing fake profiles. This shows that entry barriers based on sunk costs is not a reliable indicator of market power in dynamic markets. It also emphasizes how self regulation by platforms could protect its investment. If dominant platforms are subject to public utility regulation, it is quite possible that radical innovation would never have emerged. Facebook and LinkedIn would never have entered in presence of *ex ante* regulation as it would have diminished profit opportunities.

Other successful self-regulatory practices adopted by some dominant platforms are: EBay requires that the bidder who won the item to buy it. Google established minimum compatibility standards for hardware devices so that its Android operating system provides a minimum consistent environment for application developers across devices. It also releases a Software Development Kit (SDK) to app developers so they can develop applications consistent meeting compatibility standards. In this way, it overcomes the negative externalities that could arise from fragmentation of android operating system across hardware devices that tend to differentiate the operating system to suit their needs.[32] This was possible as platforms have better information about the users and can mitigate the negative externalities raising its growth more quickly. Unlike public regulation, platforms will be subject to fewer constraints of administrative and legislative delays. Self regulation through innovation and product changes across platforms tend to promote positive externalities and limit the negative ones. Nevertheless, this does not mean that the ability to self regulate has no antitrust concern. But clearly, the antitrust policy should exercise caution in preventing platforms from engaging in self regulation and intervene based on a careful analysis.

SERP by 0.1 to 0.4 seconds over 4-6 weeks reduced the number of searches per user by on average 0.2% to 0.6%, which is of consequence given the level of Google searches." Source - England and Wales High Court (Chancery Division) Decisions, EU Ltd v Google Inc. & Ors [2016] EWHC 253.

32 David S. Evans, The Antitrust Analysis of Rules and Standards for Software Platforms (2014) at 32.

Antitrust enforcement

Economic theory of multi-sided markets is still at a nascent stage and there is a dearth of empirical evidence to validate predictions of theoretical models Antitrust authorities can still follow the general guidelines about market shares, market power, welfare impact, etc. developed so far in academic writings on platform markets for a more nuanced and sound policy analysis. An exception in antitrust judgements is the US court of appeals for the second circuit decision in *United States vs American Express Co.,* 838 F.3d 179 (2d Cir. 2016).[33] They incorporated economic principles of two sided markets in analyzing anti-competitive effects of a firms' conduct in the credit card market. Using this credit card case, Sidak and Willig (2016) have proposed a proper market definition and market power analysis for a two sided market taking all sides into account.[34] The Indian competition law, also, is robust enough to allow for nuanced policy analysis. However, there has been a debate as to whether the Indian competition law allows for an effects-based approach to determine an Abuse of Dominance allegations. We think that this debate is settled as is evident from the orders of the Commission, where in several cases it found entities to be dominant but the conduct was not found to be abusive. If the Commission did not use effects-based analysis, what was the basis for its conclusion that the alleged conduct was not abusive in these matters?[35]

An effects-based approach takes into consideration the fact that many business practices may have different effects in different circumstances: distorting competition in some cases and promoting efficiencies and innovation in others. A competition policy approach that directly confronts this duality will ensure that consumers are protected (through the prevention of behaviour that harms them) while promoting overall increased productivity and growth (since firms will not be discouraged in their search for efficiency).[36]

This duality is evident in the working of MSPs having complex business strategies , which leaves policymakers in a perplexed situation. Imposing a remedy when it is not required would have stifling effects on the evolution of platform markets. A few lessons can be drawn from the understanding of complexities of price and non-price decisions in the markets outlined in this paper. First, skewed pricing policies cannot be analyzed in the traditional welfare standard framework in which only the particular group getting price and output changes are examined. The interdependent nature of demand across groups would require considering feedback effects in the economic analysis of price effects. This analysis was missing in the NSE case and the impugned conduct was seen only in the context of the currency derivatives market ignoring the interdependence of demand across the various segments of a stock exchange. The economic models with independent demand assumptions are not suited to draw conclusions for platform business conduct. The welfare maximization in these models requires equating profit maximizing prices and costs. However, once interlinked demands are accounted for, the long run equilibrium

[33] http://law.justia.com/cases/federal/appellate-courts/ca2/15-1672/15-1672-2016-09-26.html.

[34] J.Gregory Sidak and Robert.D.Willig. Two Sided Market Definition and Competitive Effects for Credit Cards After *United States vs American Express. (*2016) The Criterion Journal on Innovation.

[35] Nicholas J. Franczyk, Jurisprudence development (Abuse of Dominance): Issues and Implications, Proceedings, Third International Conference on Competition Regulation and Competitiveness, February 5, 2016, New Delhi.

[36] Payal Malik, Competition Law in India: Developing Efficient Markets for Greater Good, VIKALPA The Journal for Decision Makers 41(2) 168–193.

prices can be above or below the costs. As a first step, antitrust authority can begin with considering the welfare effects across sides to avoid bias in enforcement.

The other source of type 1 error can occur from antitrust remedies in a technologically dynamic market. Regulating platforms purely based on effects on one side can lead to false positives and limiting the growth of network effects. One relevant change in an antitrust framework can be directly considering the competitive effects of innovation rather than indirectly relying on price and output effects. The presence of strong positive indirect network effects can tip the market in favor of one competitor. This would prevent a new entrant from achieving the required scale. However, from a societal point of view, tipping is not necessarily welfare reducing. Efficiencies arising from new product development might outweigh any market power derived from large market shares. The antitrust enforcer should not intervene every time a technological change takes place. Since such markets are characterized by sequential innovation, the antitrust enforcers and regulators should give more space to innovation changes and not rely on an interventionist approach. These in our view are Schumpeterian industries in which market changes correct itself over time with an overhaul of market structure. A recent example is of Apple iTunes controversy. Apple introduced digital rights management (DRM) technology “Fairplay” to protect its iTunes store from other rival stores. So, songs downloaded from rival online stores were not playable on Apple iPod. Competition complaints were filed against Apple, arguing access to Apple music player. The decision went in favor of Apple and it was not seen as an “essential facility”. Over time, with technological changes, the DRM controversy became futile as rival stores introduced technologies that allowed songs from their platform to be playable on I pod. The regulatory intervention was unnecessary as the market corrected itself over time. In the context of the search engines the welfare effects of the foreclosure of competing verticals may be at best ambiguous as improved search results for the final consumer cannot be compromised. In addition to that, the presence of dynamic competition, as discussed above, would discourage it from manipulating search results to place its own vertical links at the top.

The new approach that we are proposing for Indian jurisdiction is a cautious and limited interventionist approach. Even when a firm's conduct appears to be benevolent to customers, but harming competition, the regulator should refrain from intervening. Only when a product change or innovation has a clearly identified case for excessive harm to consumers with a further slowdown of innovation in the industry antitrust enforcement can be called for. That too, an ex post intervention when the conduct has taken place on a case-by-case basis. The famous Microsoft cases give examples of such conduct that clearly harmed consumers and made it difficult for them to use a downstream rival product- internet explorer and Workgroup server OS with the upstream Microsoft windows OS. In these cases, the intent was clearly to impede rival's innovation efforts and lock in consumers to its own system. Antitrust enforcement should be available when the product design stifles industry innovation. Even then, the antitrust authority should be able to demarcate between product design features that are genuine product improvements and features that are linked to anti-competitive effects. A pragmatic antitrust approach should be able to distinguish between objectives that are pro competitive and anti competitive. It should exclusively target objectionable activities that hurt consumers (not protect some competitors) leaving other pro-competitive conduct that benefit consumers unregulated.

This approach will be effective in preventing positive externalities from withering down and isolating anti competitive strategies. The implementation of it would require a collective effort among antitrust authorities, businesses and consumers. Only then the challenges arising from dominant platforms can be addressed effectively.

Chapter 6

Conclusion

This thesis aimed to understand and analyze the strategic and welfare implications of data sharing and network effects in platform markets. In doing so, using game-theoretic modelling, the first three chapters analyzed how data sharing works as a source of competitive advantage and its implications for privacy, market structure, and welfare. Whereas, the last chapter discussed how regulatory structure can be revamped to take into account the peculiarities of platform markets.

Chapter 2 examined how voluntary data sharing between unaffiliated firms affects the investment in data exploitation and welfare. It established that when firms are competing in the advertising market, then two opposite effects can affect the privacy choice of the upstream dominant firm. On one hand, data sharing improves the probability of a match over the users, raising advertising price that upstream firm can charge. On the other hand, it intensifies the advertising competition with the downstream firms, reducing advertising price in the upstream market. It was shown that the upstream firm can invest higher in data exploitation, especially in markets with a lower improvement in its advertising targeting rate. Nevertheless, despite privacy erosion, social welfare can rise. This follows because data sharing improves the service provision to users and the probability of a match over the users in the upstream firm. These results are robust to the alternative business model of advertisement targeting and sale of data. Thus, the focus of the intervention should not be on how data is monetized, but whether or not firms share data. We briefly discuss possible paths to pursue in this research. One, vertical integration is an important form of market organization in digital markets. Firms like Google, Facebook, etc., have a presence in multiple markets. So, it will be interesting to understand how vertical integration affects data sharing and welfare in the model. Second, another form of regulation that can be studied is the taxation of data revenues. This has been studied, for instance, by Bloch and Demange (2018), and Bourreau et al. (2018). However, no paper has analyzed the impact of taxation on technology adoption and data sharing. A tax will affect the incentive to offer technology exclusively or non-exclusively. It will also affect the strategic choice of data exploitation. The overall effect remains unclear.

Chapter 3 examined the link between business model choice and data collection in the platform markets. Moreover, it examined the competitive and welfare implications of alternative regulatory approaches to protect privacy, namely i) restricting access to data owned by subsidiaries, and ii) empowering users to control data collection. The net effect depends on the interplay of three forces - user discounts, privacy costs, and advertisement targeting. It is shown that, besides other market structures, market competition can lead to the co-existence of both advertising financed and user financed business model in the equilibrium. The welfare effects of regulation mediate through the market structure in existence. When the firm collecting data across multiple markets adopts advertising financed model in market 2, then there exists a negative relationship between aggregate data collection and user pricing in market 1. As a result, restricting access has been always welfare reducing because of the rise in user price which tends to offset the reduction in privacy costs. Whereas, empowering users can enhance the user and social welfare in markets with large advertising targeting rates. This follows because the reduction in aggregate privacy costs is sufficient to offset the rise in user price. Finally, a few possible areas for future research are worth considering. One simplifying feature of our model is that the advertising market is perfectly competitive. It would be of interest to examine where one or both sides have some market power. Another direction of future research can be to allow for advertisements in market 1 as well, i.e., a two-sided market structure. This can lead to new strategic effects on advertising prices.

Chapter 4 examined the impact of bundling on competition and welfare when the firm could collect data regarding users. The net effect depends on the interplay of bundled discount and increase in advertising revenue. Bundling is profitable when investment in data collection and/or advertising targeting rate are large. Moreover, user welfare and social welfare can move in the opposite direction depending on the level of investment in data collection. For small to intermediate investment in data collection, they can diverge. Whereas, for a large investment in data collection, they both can improve with bundling. Thus, the antitrust decision in bundling cases in such markets would vary depending on whether or not the role of data is considered. However, this model can be extended to consider some more observations. One interesting area is to consider mixed bundling by a firm. This would be analytically challenging but is an important area. In the presence of data advantage, mixed bundling might not be the dominant strategy for the firm. This is possible because under pure bundling the firm gains data over all the users, whereas under mixed bundling data is gained only over a subset of users. The net effect remains unclear. Another interesting modification can be to determine investment in data collection endogenously and how bundling affects it. This can bring new insights about the trade-off between non-price discrimination and privacy in platform markets and its impact on welfare.

Finally, chapter 5 linked current economic theory on platform markets with facts of the competition cases in Indian Jurisdiction. It discussed how network effects raise competitive ambiguity for the antitrust authorities about the working of these markets. In particular, it discussed two competition cases - the National Stock Exchange case and Google case. In doing so, it highlighted the shortcomings in CCI orders and approach and outlined a relevant approach to deal with cases in platform markets. It was argued that strategies that are considered harmful in standard economic literature may have a pro-competitive rationale. Moreover, such markets are innovation-driven and tend to be concentrated. In such dynamic markets, antitrust intervention should not be ex-ante and the authorities should intervene only when substantial consumer harm has been established.

Bibliography

Acquisti, A., John, L. K., and Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, 42(2):249–274.

Amelio, A. and Jullien, B. (2012). Tying and freebies in two-sided markets. *International Journal of Industrial Organization*, 30(5):436 – 446.

Anderson, S. P. and Coate, S. (2005). Market provision of broadcasting: A welfare analysis. *The Review of Economic Studies*, 72(4):947–972.

Anderson, S. P. and Jullien, B. (2016). The advertising-financed business model in two-sided media markets. TSE Working Papers 16-632, Toulouse School of Economics (TSE).

Armstrong, M. (2006). Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691.

Armstrong, M. and Weeds, H. (2007). Programme quality in subscription and advertising-funded television. *mimeo*.

Armstrong, M. and Wright, J. (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, 32(2):353–380.

Bansal, S. (2008). Choice and design of regulatory instruments in the presence of green consumers. *Resource and Energy Economics*, 30(3):345 – 368.

Bansal, S., Chakravarty, S., and Ramaswami, B. (2013). The informational and signaling impacts of labels: Experimental evidence from india on gm foods. *Environment and Development Economics*, 18.

Bergemann, D. and Bonatti, A. (2011). Targeting in advertising markets: implications for offline versus online media. *The RAND Journal of Economics*, 42(3):417–443.

Binns, R., Lyngs, U., Kleek, M. V., Zhao, J., Libert, T., and Shadbolt, N. (2018). Third party tracking in the mobile ecosystem. *WebSci'18: Proceedings of the 10th ACM Conference on Web Science*, pages 23–31.

Bloch, F. and Demange, G. (2018). Taxation and privacy protection on internet platforms. *Journal of Public Economic Theory*, 20(1):52–66.

Bork, R. H. and Sidak, J. G. (2012). What does the chicago school teach about internet search and the antitrust treatment of google? *Journal of Competition Law and Economics*, 8(4):663–700.

Bourreau, M., Caillaud, B., and De Nijs, R. (2018). Taxation of a digital monopoly platform. *Journal of Public Economic Theory*, 20(1):40–51.

Brill, J. (2011). The intersection of consumer protection and competition in the new world of privacy, ii. *Antitrust Chronicle*, 12.

Caillaud, B. and Jullien, B. (2003). Chicken egg: Competition among intermediation service providers. *The RAND Journal of Economics*, 34(2):309–328.

Calvano, E. and Polo, M. (2016). Strategic Differentiation by Business Models: Free-to-Air and Pay-TV. CSEF Working Papers 438, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1):47–73.

Carlton, D. W. and Waldman, M. (2002). The strategic use of tying to preserve and create market power in evolving industries. *The RAND Journal of Economics*, 33(2):194–220.

Casadesus-Masanell, R. and Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, 61(1):229–246.

Choi, J. P. (2010). Tying in two-sided markets with multihoming. *The Journal of Industrial Economics*, 58(3):607–626.

Choi, J. P. and Jeon, D. S. (2016). A leverage theory of tying in two-sided markets. CESifo Working Paper Series 6073, CESifo Group Munich.

Choi, J. P., Jeon, D. S., and Kim, B.-C. (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics*, 173:113 – 124.

Choi, J. P. and Stefanadis, C. (2001). Tying, investment, and the dynamic leverage theory. *The RAND Journal of Economics*, 32(1):52–71.

Chou, C. F. and Shy, O. (1990). Network effects without network externalities. *International Journal of Industrial Organization*, 8(2):259 – 270.

Chou, C. F. and Shy, O. (1993). Partial compatibility and supporting services. *Economics Letters*, 41(2):193 – 197.

Chou, C. F. and Shy, O. (1996). Do consumers gain or lose when more people buy the same brand. *European Journal of Political Economy*, 12(2):309–330.

Church, J. and Gandal, N. (1992). Network effects, software provision, and standardization. *The Journal of Industrial Economics*, 40(1):85–103.

Church, J. and Gandal, N. (1993). Complementary network externalities and technological adoption. *International Journal of Industrial Organization*, 11(2):239 – 260.

Church, J. and Gandal, N. (2000). Systems competition, vertical merger, and foreclosure. *Journal of Economics & Management Strategy*, 9(1):25–51.

Corniere, A. D. and Taylor, G. (2017). Application Bundling in System Markets. CEPR Discussion Papers 12129, C.E.P.R. Discussion Papers.

Cornière, A. D. and Taylor, G. (2014). Integration and search engine bias. *The RAND Journal of Economics*, 45(3):576–597.

Crampes, C., Haritchabalet, C., and Julien, B. (2009). Advertising, competition and entry in media industries. *The Journal of Industrial Economics*, 57(1):7–31.

de Cornière, A. and Sarvary, M. (2018). Social media and news : Attention capture via content bundling.

Depken, C. and Wilson, D. (2004). Is advertising a good or a bad? evidence from u.s. magazine subscriptions. *Journal of Business*, 77.

Dimakopoulos, D. and Sudaric, S. (2018). Privacy and platform competition. *International Journal of Industrial Organization*, 61:686 – 713.

Edelman, B. and Geradin, D. (2016). Android and competition law: exploring and assessing google's practices in mobile. *European Competition Journal*, 12(2-3):159–194.

Edelman, B. and Lai, Z. (2016). Design of search engine services: Channel interdependence in search engine results. *Journal of Marketing Research*, 53(6):881–900.

Gal-or, E. (1986). Information Transmission—Cournot and Bertrand Equilibria. *Review of Economic Studies*, 53(1):85–92.

Gal-Or, E. (1987). First mover disadvantages with private information. *The Review of Economic Studies*, 54(2):279–292.

Goldfarb, A. and Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, 57(1):3–43.

Goldschlag, N. and Tabarrok, A. (2018). Is regulation to blame for the decline in American entrepreneurship? *Economic Policy*, 33(93):5–44.

Hagiu, A. (2006). Pricing and commitment by two-sided platforms. *The RAND Journal of Economics*, 37(3):720–737.

Hagiu, A. and Lee, R. S. (2011). Exclusivity and control. *Journal of Economics & Management Strategy*, 20(3):679–708.

Hahn, H. W. and Singer, R. J. (2008). An antitrust analysis of google's proposed acquisition of doubleclick. *AEI-Brookings Joint Center for Regulatory Studies*, 24(7).

Hogerdon, C. and Ka Yat Yuen, S. (2009). Platform competition with 'must-have' components*. *The Journal of Industrial Economics*, 57(2):294–318.

Jaffe, A. B. and Palmer, K. (1997). Environmental regulation and innovation: A panel data study. *The Review of Economics and Statistics*, 79(4):610–619.

Jakhu, G. and Malik, P. (2017). Dilemma in antitrust enforcement: how use of economics can guide enforcement rules in multi-sided markets. *Journal of Antitrust Enforcement*, 5(2):260–275.

Kerber, W. (2016). Digital markets, data, and privacy: competition law, consumer law and data protection. *Journal of Intellectual Property Law Practice*, 11:jpw150.

Krämer, J., Schnurr, D., and Wohlfarth, M. (2019). Winners, losers, and facebook: The role of social logins in the online advertising ecosystem. *Management Science*, 65(4):1678–1699.

Lasserre, B. and Mundt, A. (2017). Competition law and big data: The enforcers' view. *Rivista Italiana di Antitrust / Italian Antitrust Review*, 4(1).

Lefouili, Y. and Toh, Y. (2017). Privacy and quality. TSE Working Papers 17-795, Toulouse School of Economics (TSE).

Lenard, T. M. and Rubin, P. H. (2010). In defense of data: Information and the costs of privacy. *Policy & Internet*, 2(1):149–183.

Lenard, T. M. and Rubin, P. H. (2015). Big data, privacy and the familiar solutions. *Journal of Law Economics & Policy*, 11(1).

Lerner, A. (2014). The role of big data in online platform competition. *SSRN Electronic Journal*.

Lu, Y. and Poddar, S. (2018). Strategic choice of network externality and its impact on digital piracy. *Review of Industrial Organization*, 52(1):139–160.

Manne, G. and D. Wright, J. (2010). Google and the limits of antitrust: The case against the antitrust case against google. *Harvard Journal of Law and Public Policy*, 34:171–244.

Mccan, D. and Hall, M. (2018). Blocking the data stalkers: Going beyond gdpr to tackle power in sharing economy. *New Economics Foundation, Report*.

Monopolkomission, G. (2015). Competition policy: the challenge of digital markets. *Special Report No 68*.

Nalebuff, B. (2004). Bundling as an entry barrier. *The Quarterly Journal of Economics*, 119(1):159–187.

Newman, N. (2013). Search, antitrust and the economics of the control of user data.

Newman, N. (2014). Search, antitrust, and the economics of the control of user data. *Yale Journal on Regulation*, 31(2):401–454.

Peitz, M. and Valletti, T. (2008). Content and advertising in the media: Pay-tv versus free-to-air. *International Journal of Industrial Organization*, 26(4):949 – 965.

Preibusch, S., Kübler, D., and Beresford, A. R. (2013). Price versus privacy: an experiment into the competitive advantage of collecting less personal information. *Electronic Commerce Research*, 13(4):423–455.

Prüfer, J. and Schottmuller, C. (2017). Competing with big data. Discussion Paper 006, Tilburg University, Tilburg Law and Economic Center.

Rochet, J. C. and Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029.

Sinha, U. B. (2013). On R & D information sharing and merger. *Economic Modelling*, 32:369 – 376.

Solove, D. (2004). The digital person: Technology and privacy in the information age. GWU legal Studies Research Paper 7, George Washington University (GWU) Law School.

Solove, D. (2012). Privacy self-management and the consent dilemma. *Harvard Law Review*, 126.

Stennek, J. (2014). Exclusive quality – Why exclusive distribution may benefit the TV-viewers. *Information Economics and Policy*, 26(C):42–57.

Strahilevitz, L. J. and Kugler, M. B. (2016). Is privacy policy language irrelevant to consumers? *The Journal of Legal Studies*, 45(S2):S69–S95.

Thöne, M., Rasch, A., and Wenzel, T. (2016). Business models in commercial media markets: Bargaining, advertising, and mixing. Annual conference 2016, augsburg: Demographic change.

Vives, X. (1990). Information and competitive advantage. *International Journal of Industrial Organization*, 8(1):17 – 35.

Weeds, H. (2016). Tv wars: Exclusive content and platform competition in pay tv. *The Economic Journal*, 126(594):1600–1633.

Whinston, M. D. (1990). Tying, foreclosure, and exclusion. *The American Economic Review*, 80(4):837–859.

Wilbur, K. C. (2008). A two-sided, empirical model of television advertising and viewing markets. *Marketing Science*, 27(3):356–378.

Z.Borgesius, F. and Poort, J. (2017). Online price discrimination and eu data privacy law. *Journal of Consumer Policy*, 40(3):347–366.