Data Sharing in Platform Markets

Gaurav Jakhu

Economics and Planning Unit

Indian Statistical Institute

2019

Abstract

This paper analyses the strategic and welfare impact of voluntary data sharing through technology adoption in platform markets. In the game theoretic model formulated, it can affect user market through better service provision and advertising market through higher targeting rates. It is shown that, under technology 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, exclusive technology sharing regime paradoxically improves the welfare of all users. When technology sharing is endogenous, the upstream firm can have incentives to over invest in data exploitation to enforce technology adoption.

Keywords: Platforms, Vertical Markets, Data Sharing

JEL classification: D21, D42, L12, L13, L42, L51

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, online intermediaries like Google, Facebook are offering technology products that help other firms to improve their access to users. In return, they also get access to user profiles and information which improve their targeting abilities. 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 mobile web. It also provides Google access to user data through these websites. Another example is of Facebook offering instant articles technology under which content publishers get priority placement in Facbeook news feed and Facebook retains detailed user data collected from these websites. Finally, social logins that are being offered by online intermediaries 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 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, the more recent voluntary data sharing among unaffiliated online firms and the strategic and welfare implications of such data sharing has not been studied in the literature. In this paper, I develop a theoretical framework to understand how technology adoption affects the equilibrium level of data collection on 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 also helps us understand how a better and safe environment for user privacy can be attained.

I develop a game theoretic model in which there is a single upstream firm that users must join in order to access downstream firms. All firms derive revenue through two sources - advertisements and sale of data to third parties. The firms can choose the level of investment in data collection. This data takes the form of browsing histories, location information, personal information and so on. However, the collection of data imposes privacy cost on users which depends on the strength of privacy preferences and the firm's investment level. On the advertising side, advertisers 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 offer access to technology to the downstream firms. This technology serves as a mechanism for data sharing among these online 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 targeting rate of both upstream and downstream firm improves with technology sharing. The technology, if offered, is always adopted by the downstream firms. However, in an extension, I explicitly consider a technology adoption game analysing offer and adoption decision. I consider the business model of no advertisement targeting in which a firm uses user data for direct sale to third parties only.[1] So, the improvement in targeting rate of a firm is possible only through technology sharing.

Using this model, I examine the impact of technology sharing on the level of data exploitation by the upstream firm. The market structure is affected on both the user and the advertiser side. Whether data collection on the upstream firm increases or decreases depends, among other things, on the extent of improvement in targeting ability in both the markets. For “small" improvements in advertising targeting rate in the upstream market, data exploitation by the upstream firm under no sharing regime is lower than under technology sharing whereas for “large" improvements, it is the opposite. The intuition for this result stems from how advertising competition and hence advertising prices are affected under technology sharing. For small improvements in upstream market targeting rate, advertising competition intensifies and advertising prices fall under technology sharing which in turn increases data exploitation to reap profits. Whereas, for large improvement in the upstream market targeting rate, advertising competition softens and advertising prices rise which reduces data exploitation under technology sharing. Turning to the welfare analysis, it is shown that there are two main opposing effects that work on social welfare. First, technology sharing improves the targeting ability of the upstream firm which dominates any other effect and the advertiser revenue goes up. At the same time, the revenue from sale of data to third parties might fall due to technology sharing. However, it is shown that social welfare rises with technology sharing due to better probability of match over the user set.

Then, I extend the baseline model in several directions. First, I analyze the equilibrium relations under mixed business model - advertisement targeting with sale of data to third parties. Under it, conclusions from the baseline model still holds. The only difference is that data exploitation under

[1] In an extension, I also look at mixed business model that can exist with data collection. In this model firms use data for advertisement targeting and selling it to third parties.

technology sharing is always higher. Then, a comparison with the baseline model of no targeting shows that data exploitation is always higher under mixed business model compared to the baseline model.

In another extension, I consider the possibility of exclusive sharing along with non exclusive technology sharing in the baseline model. It is shown that for very small improvement in targeting rates, data exploitation is the highest under non exclusive sharing whereas for large improvement in targeting rates, data exploitation is the lowest under it. Turning to the welfare implications, it is shown that user welfare is highest under an exclusive regime whereas social welfare is highest under a non exclusive regime. The divergence in welfare analysis stems from the changes in transportation cost, 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 technology 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.

Finally, I consider an extension that endogenizes 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 cost for the users. The net effect determines whether over investment takes place or not. In markets with intermediate to large improvement in targeting rate of the upstream firm, strategic over investment can be profitable to enforce technology sharing.

2 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)). This paper differs from the preceding literature in that it focuses on examining the interaction between privacy and technological change. Here I study how adoption of technology by websites affects the privacy choice of an upstream firm. This issue is important as introduction of new technology in the market can affect the market power of upstream firm through unfair data collection from the users. This paper differs from 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 contractual arrangement between firms. In two sided market literature, a 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. The literature on the issue of technology adoption as a data sharing mechanism 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. My paper differs from this study. First, the model set up is different from theirs. I 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 exploitation of downstream firms. Whereas, I consider the privacy impact of these technology adoption decisions and how it affects the market outcomes.

The rest of the paper is organized as follows. Section 3 discusses the nature of technology sharing in greater detail. In Section 4, the baseline model is set up where platforms finance themselves through advertising and sale of data to third parties. Section 5 studies the equilibrium relations and welfare effects. Section 6 studies some extensions to the basic model and draw new insights from that. Section 7 concludes. Appendix 1 presents the proofs of the basic model, Appendix 2 the ones of the alternate business model and Appendix 3 the technology adoption game.

3 Nature of the technology

Technology sharing between platforms as a means to share user data has been growing in popularity. Such technology sharing can help overcome many issues that come up in digital markets. It can be used to manage passwords and login credentials through adoption of social login by websites, to improve the loading speed of content through AMP and instant article technology. 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 technology sharing between unaffiliated platforms can improve the targeting rates of both platforms. This is possible due to 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. Additionally, these websites can access other details such as users' likes, political and religious preferences, relationship status, location, photos, and even personal messages. Similarly, apps and websites using Google+ sign-in can access information of users' public profiles as well as their friend lists in order 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 be used to improve the targeting rate of advertisements on these platforms. Another instance is 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 the nature of 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 quite recent and has affected the 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 website 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 inevitably trigger concerns among users. Another source of concern can be how this technology sharing affects the data collection practices on the platform itself. As discussed earlier, platforms collect data about user activities and use it for personalized advertisements. The rate of data collection can be affected by these exogenous technology changes. The social desirability of these technology sharing regimes remains unclear. This paper examines the effect of technology sharing on user privacy and welfare.

4 The Model

Internet User

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, 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 is a decreasing function of the distance between firm i's position and the user's location $x\epsilon[0,1]$. 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

$$U_i(x) = \begin{cases} V + I*\theta - \gamma_m(m_0 + m_1) - \gamma_p(q_0 + q_1) - tx, & \text{if } i = 1 \\ V + I*\theta - \gamma_m(m_0 + m_2) - \gamma_p(q_0 + q_2) - t(1-x), & \text{if } i = 2 \end{cases} \quad (1)$$

where V measures utility from accessing the content; θ measures improvement in user utility from data sharing; I is an indicator function which takes value one if 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 nuisance cost of advertisements ; q_i is the level of investment in data exploitation by firm i = 0, 1 and 2; $\gamma_p > 0$ measures user sensitivity to privacy and $t > 0$ measures disutility from discrepancy between the user's location x and firm i's location.

Users dislike advertisements that are bundled with content on a firm and suffers a nuisance cost of $\gamma_m m_i$. This has been empirically validated in some media studies which found that advertising reduces users' utility (Wilbur 2008; Depken and Wilson 2004). Theoretical work has also characterised advertising as a nuisance to users (e.g., Anderson and Coate 2005). In addition to that, a user is concerned about how much data the firm collects about her and sells to third parties, modelled via a disutility of $\gamma_p q_i$, when q_i amount of data is collected from her. In this setup, the data is collected from the user when she joins firm $i = 0, 1$ and 2. For instance, it might be the personal information which the user provides to register on a platform or behavioural data like search history. For the rest of analysis, it is assumed that $q_1 = q_2 = q$ and q_0 may be chosen endogenously in the model. The reservation utility of the users is taken to be zero.

Advertiser

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 [Anderson and Coate (2005) and Crampes, Haritchabalet, and Jullien (2009)].

The probability that advertisers inform a singlehoming user on firm i is given by $\rho(m_i)$, where $\rho(m_i)$ depends on both m_i as well as whether firm i and 0 share data or not. For ease of exposition, I assume that

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

where $\alpha > 0$ measures 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. 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 increase in number of advertisements on the other firm $j \neq i$. This assumption is in line with earlier research work on platform markets [Coriniere and Taylor (2014); Hahn and Singer (2008) etc].

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

$$\Pi(m_0, m_i) = [1 - (1 - \rho(m_i))(1 - \rho(m_0))];\ i = 1, 2. \tag{2}$$

This is the probability that the user is informed on atleast one of the firms. An important assumption here is that the probability of informing a multi-homing user on a single firm is equal to the probability of informing a single-homing user on the same firm. As can be seen from equation (2), $\frac{\partial \Pi^2}{\partial m_i \partial m_j} < 0$. This essentially captures the important assumption that the advertisement quantities on two different firms are imperfect substitutes. 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), the probability of reaching a user who has joined firm 1 is $[1 - (1 - \rho(m_0))(1 - \rho(m_1))]$ and the probability of reaching a user who has joined firm 2 is $[1 - (1 - \rho(m_0))(1 - \rho(m_2))]$.[2]

Now, using these probability functions we can find the revenue that the advertisers receives from purchasing advertisement quantity m_0, m_1 and m_2. The expected per user revenue on either

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

firm i = 1,2 is $\rho(m_0)N_i + [1-\rho(m_0)]\rho(m_i)N_i$. Thus, it equals the expected revenue from a user from joining firm 0, $\rho(m_0)N_i$, plus the additional revenue from the same user on the downstream firm i, i.e. $[1-\rho(m_0)]\rho(m_i)N_i$. The expected per user revenue can be written as

$$AW = \rho(m_0) + [1-\rho(m_0)][\rho(m_1)N_1] + [1-\rho(m_0)][\rho(m_2)N_2]. \tag{3}$$

In the preceding equation, $\rho(m_0)$ represents the revenue from reaching users on firm 0 and [1 - $\rho(m_0)][\rho(m_i)N_i]$ for i =1,2 represents additional revenue obtained from reaching users that are not informed on firm 0. Let P_i denote the advertising price paid for a unit of advertisement on firm $i \in \{0,1,2\}$. Thus, the expected profit of advertisers is

$$\pi_a = AW - P_0m_0 - P_1m_1 - P_2m_2, \tag{4}$$

where 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 marginal benefit from a unit of an advertisement i.e. $\frac{\partial R(.)}{\partial m_i}$ is equal to its marginal cost P_i. This implies that the prices are determined so as 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

All three firms monetize user data through selling it to third parties. Firm i has two strategic tools in hand to maximize profit - m_i and q_i. The profit of the firm when it sells data directly to third parties is written as

$$\pi_i(m_i, q_i) = P_im_i + Rq_i - \frac{1}{2}q_i^2, \tag{5}$$

where the first term in the preceding equation is revenue from advertisements on firm i; the second term captures revenue generation from sale of data to third parties; last term is the cost of investment in data exploitation. In the next section, we will examine the effect of technology sharing on user privacy and welfare under an alternate business model. To begin with, the business model is one with “no" advertisement targeting with direct sale of data and then extend the analysis to targeted advertisements in the next section.

Timing of the game

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

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

Stage 2: Firm 0 chooses quantity of advertisements m_0.[3]

Stage 3: Firm $i = 1, 2$ chooses quantity of ad 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.

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

I look for sub game perfect nash equilibrium of the game. When solving the game, the two different regimes are considered - a) no technology sharing, when firm 0 does not share the technology with the downstream firms; b) technology sharing, when firm 0 shares the technology and it is adopted by both downstream firms.

5 Equilibrium Analysis under “sale of data" model

The outcome of the game depends on the technology regime i.e. whether there is technology sharing or not. I make the following assupmtion for the rest of the analysis.

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

[3] 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 I stick to the timing in which firm 0 chooses q_0 first and then advertising decision is made.

The threhsold 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.

5.1 No Technology Sharing

Consider no technology sharing regime, indicated by a superscript “nt".

Efficiency Benchmark

First, the “efficient" level of data collection and advertising levels are obtained that sets as our benchmark. Assuming that full market coverage is socially optimal, a user will purchase from the nearest downstream firm on the Hotelling line. The participation decision 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 (1)) and $\hat{x}$ equals

$$\hat{x} = \frac{1}{2} + \frac{\gamma_m(m_2 - m_1)}{2t}. \tag{6}$$

From this, the demand for firm 1 is $\hat{x}$ and for firm 2 is $1 - \hat{x}$. Social welfare under no technology sharing is

$$SW^{nt} = \int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx + \pi_a + \pi_0 + \pi_1 + \pi_2, \tag{7}$$

where note that the focus is on the “second best" outcome where the social planner set q_0, m_0, m_1 and m_2 to maximize social welfare taking the users' participation decision as well as the advertisers' decision as given. Since advertising prices are just transfers, social welfare will be

$$SW^{nt} = \underbrace{\int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx}_{\text{User Surplus}} + \underbrace{AW^{nt}}_{\text{Advertiser Revenue}} + \underbrace{Rq_0 - \frac{1}{2}q_0^2 + 2Rq - q^2}_{\text{Revenue from Sale of Data.}}. \tag{8}$$

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 sum of revenue from sale of data to third parties.

Proposition 1. *The efficient solution is characterized by* $m_1 = m_2 = m_0 = \tilde{m}$*, and* $q_0 = \tilde{q_0}$*, where*

$$\tilde{m} = \frac{\alpha - \gamma_m}{\alpha^2}, \text{and } \tilde{q_0} = R - \gamma_p. \tag{9}$$

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 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. $\gamma_m + \alpha^2 \tilde{m}$ equals marginal social benefit α. It can be seen from the preceding equation that $\tilde{m} > 0$ if $\alpha > \gamma_m$ and $\tilde{q_0} > 0$ if $R > \gamma_p$ i.e only if marginal social benefit is sufficiently large. Otherwise the optimal value is 0.

Equilibrium

I now solve for the equilibrium when there is no technology sharing.

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

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

Stage 4: The prices of advertisements will be determined. As discussed earlier, advertisers are price taker in all markets and will place advertisements such that $P_i = MR_i$. The advertiser profit is given by equation (4). Substituting (2) into (4), we have that

$$\pi_a = \alpha m_0 + [1 - \alpha m_0][\alpha m_1 N_1 + \alpha m_2 N_2] - P_0 m_0 - P_1 m_1 - P_2 m_2. \tag{10}$$

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

$$P_0^{nt} = \alpha[1 - \alpha m_1 N_1 - \alpha m_2 N_2], \tag{11}$$

$$P_i^{nt} = [1 - \alpha m_0]\alpha N_i, \; i = 1, 2. \tag{12}$$

Note from the preceding equations (11) and (12) 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 functions and advertising price given in equation (12). The maximization problem of firm i is

$$\underset{m_0}{Max}\, \pi_i \text{ subject to } U_i(\hat{x}(m_i, m_j); m_0) \geq 0. \tag{13}$$

This gives us the optimal choice of m_i^{nt} as

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

Stage 2: Firm 0 operate 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 obtains zero utility in equilibrium. Under assumption 1, this yields the solution

$$m_0^{nt} = \frac{V - t - \gamma_p(q_0 + q)}{\gamma_m}. \tag{14}$$

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

$$m_1^{nt} = m_2^{nt} = \frac{t}{2\gamma_m} \tag{15}$$

This choice of m_0^{nt} is intuitive. If firm 0 inceases its advertising to $m_0^{nt} + \epsilon$, then market becomes uncovered and under assumption 1, it is not profitable for firm 0 to do that. Whereas, if it reduces the advertising level then the market still remains covered and it suffers a loss.

Stage 1: It can be seen that the optimal m_0^{nt} falls with 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 sale of data to third parties. Now, substituting (14) in (5) gives firm 0's profit as

$$\pi_0^{nt} = \alpha m_0^{nt}\left[1 - \frac{\alpha t}{2\gamma_m}\right] + Rq_0^{nt} - \frac{1}{2}(q_0^{nt})^2. \tag{16}$$

Maximizing (16) with respect to q_0 gives the the equilibrium level of data exploitation

$$q_0^{nt} = R - \frac{\alpha\gamma_p}{\gamma_m}\left[1 - \frac{\alpha t}{2\gamma_m}\right]. \tag{17}$$

5.2 Technology Sharing

Consider the case when firm 0 shares technology with both the downstream firms. 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

Assuming that social planner keeps the market covered, the indifferent user's location is still at $\hat{x}$ such that

$$\hat{x} = \frac{1}{2} + \frac{\gamma_m(m_2 - m_1)}{2t}. \tag{18}$$

Using this, social welfare under technology sharing is

$$SW^t = \underbrace{\int_0^{\hat{x}} U_1(x)dx + \int_{\hat{x}}^1 U_2(x)dx}_{\text{User Surplus}} + \underbrace{AW^t}_{\text{Advertiser Revenue}} + \underbrace{Rq_0 - \frac{1}{2}q_0^2 + 2Rq - q^2}_{\text{Revenue from Sale of Data}}. \tag{19}$$

Using equation (53), and substituting the values for $U_1(x), U_2(x)$ and AW^t in it, the optimal values can be found.

Proposition 2. *The efficient solution is characterized by* $m_1 = m_2 = m_0 = \tilde{m}$ *and* $q_0 = \tilde{q_0}$, *where characterized by*

$$\tilde{m} = \frac{\alpha(1+\beta_0) - \gamma_m}{\alpha^2(1+\beta_0)(1+\beta)}, \text{ and } \tilde{q_0} = R - \gamma_p. \tag{20}$$

The efficient level of data collection remains unaffected by technology sharing. This is due to the way data is monetised here. The only source of monetisation is sale of data to third parties. However, efficient advertising level depends on targeting rates of both markets. A higher β_0 increases $\tilde{m}$ and a higher β reduces $\tilde{m}$.

Equilibrium

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

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

Turning to *stage 4*, the advertiser now has a higher probability of match on all the firms. Substituting (2) in (4), its profit function is given as

$$\pi_a = \alpha(1+\beta_0)m_0 + [1 - \alpha(1+\beta_0)m_0][\alpha(1+\beta)m_1N_1 + \alpha(1+\beta)m_2N_2] - P_0m_0 - P_1m_1 - P_2m_2. \tag{21}$$

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

$$P_0^t = \alpha(1+\beta_0) - \alpha(1+\beta_0)[\alpha(1+\beta)m_1N_1 + \alpha(1+\beta)m_2N_2], \tag{22}$$

$$P_i^t = [1-\alpha(1+\beta_0)m_0]\alpha(1+\beta)m_iN_i, i = 1,2. \tag{23}$$

At *stage 3*, each firm i =1,2 maximizes its profit function π_i using the inverse demand functions and advertising price given in equation (23). This gives us the optimal choice of m_i^t as

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

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

$$m_0^t = \frac{V-t+\theta-\gamma_p(q_0+q)}{\gamma_m}. \tag{24}$$

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

$$\pi_0^t = \alpha(1+\beta_0)m_0^t\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right] + Rq_0^t - \frac{1}{2}(q_0^t)^2. \tag{25}$$

Maximizing (25) gives the equilibrium level of q_0 as

$$q_0^t = R - \alpha(1+\beta_0)\frac{\gamma_p}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]. \tag{26}$$

6 Technology Sharing versus No Technology Sharing

6.1 Advertising and Data Exploitation Comparison

Advertising Levels

We start by comparing no technology and technology sharing for the advertising levels. Under no technology sharing, advertising quantity at firm 0 is given by (14) and under technology sharing, it is given by (24). Thus, it can be seen that *advertising level of firm 0 is higher under technology sharing relative to no technology sharing regime*.

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 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 (24) and (24), it can be seen that equilibrium advertising on 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 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 downstream market is lower and on upstream market is higher than the social optimum.

Another source of difference that can come from is the improvement in targeting rates under technology sharing regime. On the downstream market, parameter β_0 increases advertising level and parameter β reduces advertising level under social optimum. However, m_i^* is independent of targeting rates. So, for sufficiently high β_0 or low β, there can be under provision of advertisements in market equilibrium by the downstream firms.

Data Exploitation

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

Proposition 3. *For sufficiently small value of β, there exist $\bar{\beta_0}$ such that* *i) For* $\beta_0 \leq \bar{\beta}_0$, $q_0^t \geq q_0^{nt}$,

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

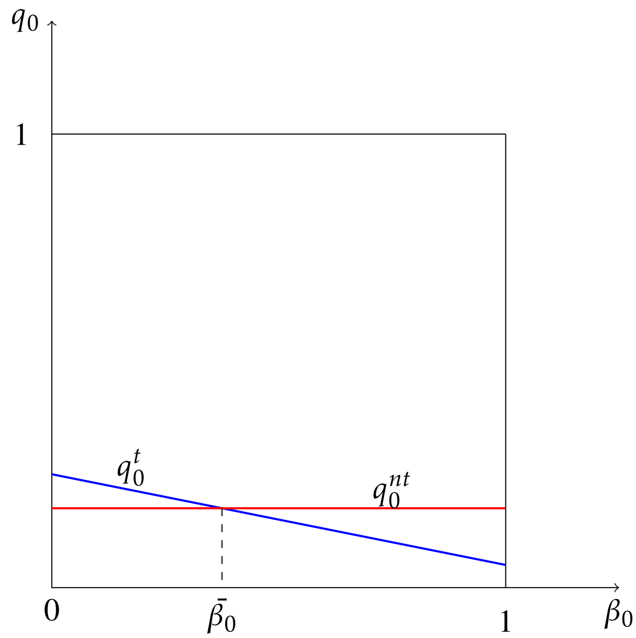

Figure 1: Data exploitation under different regimes

q_0^t is the unconstrained level of investment in data exploitation under technology sharing and q_0^{nt} is the unconstrained level under no technology 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 (1) graphically depicts the optimal choice of data exploitation under a specific range of parameter values. The above proposition suggests that data sharing through technology adoption can affect data exploitation on the dominant platform either way. In markets where targeting rate on upstream firm can rise significantly (large β_0), technology sharing can reduce the rate of data exploitation. In order to interpret the result in preceding proposition, we need to look at how advertising price is affected by technology sharing. Advertising price charged to the advertiser and data collected from the users are substitutes in the model. So, a rise in advertising price reduces data collection. This is the mechanism that works under technology sharing. However, for low values of β_0, the reduction is not too much and data collection under technology sharing still remains higher than under no technology sharing regime. However, for higher values of β_0, advertising price rises sufficiently reducing data collection under technology sharing below the level under no technology sharing regime.

Next, comparing the level of data collection under market outcome with efficient level $\tilde{q_0}$, 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.

6.2 Welfare Analysis

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

$$SW^i = UW^i + AW^i + Rq_0^i + 2Rq - \frac{1}{2}(q_0^i)^2 - q^2, \tag{27}$$

where UW^i is user welfare under regime i; AW^i is advertisers' revenue under regime i; Rq_0 is the revenue of firm 0 from sale of data to third parties; $2Rq$ is the revenue of downstream firms from sale of data to third parties; and the last two terms comprise the total cost of data exploitation.

User Welfare

First, I investigate how user surplus changes under different scenarios. It is defined as the integral over all purchasing users of their utility.

$$\begin{aligned} UW^i = &\int_0^{\hat{x}} [V + I * \theta - \gamma_p(q_0^i + q) - \gamma_m(m_0^i + m_1^i) - tx]dx \\ &+ \int_{\hat{x}}^1 [V + I * \theta - \gamma_p(q_0^i + q) - \gamma(m_0^i + m_2^i) - t(1-x)]dx, \end{aligned} \tag{28}$$

where $i = t$ or nt.

Proposition 4. *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. Technology sharing improves the quality of services for the firms. So, users who join firm i receives a higher utility. Also, it raises the advertisements by firm i which raises the nuisance cost of these users.

The users are exposed to total ads $m_0^t + t/\gamma_m$ under technology sharing whereas they are exposed to $m_0^{nt} + t/\gamma_m$ under no technology sharing. Since, advertising level on firm 0 is such that the market remains covered in the model, any surplus utility from technology sharing cancels out. Also, the practice does not distort the distribution of users across two downstream firms. Taken together, the overall loss from higher nuisance cost and gain from improvement in service balance out each other and the user welfare remains the same.

Advertiser Revenue

Here, I look at how advertiser revenue changes with technology sharing. Substituting the values for $\rho(m_i)$ in equation (3) under the two regimes and taking the difference of two, the change in advertiser revenue $\Delta AW = AW^t - AW^{nt}$ can be written as

$$\Delta AW = \alpha(1+\beta_0)m_0^t\left[1 - \frac{\alpha(1+\beta)t}{\gamma_m}\right] + \frac{\alpha\beta t}{\gamma_m} - \alpha m_0^{nt}\left[1 - \frac{\alpha t}{\gamma_m}\right]. \tag{29}$$

Following can be stated about how advertiser revenue changes with technology sharing.

Proposition 5. *For sufficiently large θ*
i) Advertiser revenue rises with technology sharing.
ii) The difference in advertiser revenue expands with an increase in the value of parameter β_0, β and contracts with an increase in γ_p.

Sale of Data

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

$$\Delta R = Rq_0^t - Rq_0^{nt} + \frac{1}{2}(q_0^{nt})^2 - \frac{1}{2}(q_0^t)^2. \tag{30}$$

Substituting the values for q_0^{nt} and q_0^t 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 sale of data is summarized as follows

Proposition 6. *i) There exist β_0^r such that a) for $\beta_0 < \beta_0^r$, revenue from sale of data rises with technology sharing, and b) for $\beta \geq \beta_0^r$ revenue from sale of data falls with technology sharing.*
ii) The difference in revenue from sale of data can contract with an increase in β_0, expand with an increase in β, and contract with an increase in γ_p.

Taken together the components of social welfare can move in different directions. The major advantage of technology 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 technology sharing. Another concern, from firm point of view, is that technology sharing reduces revenue from sale of data. However, data sharing improves probability of match and this effect can potentially outweigh the other concerns, namely loss of privacy, increased nuisance cost of advertisements and fall in revenue from sale of data to third parties.

Proposition 7. *Social welfare is higher under technology sharing i.e. $SW^t > SW^{nt}$.*

7 Extensions

7.1 Alternate Business Model

In this subsection,we examine a mixed business model where firms use user data both for advertisement targeting and selling it directly to third parties. This requires targeting at firm i to be a function of q_i, $\alpha + \delta q_i$ 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. The advertiser profit under no technology sharing is

$$(\alpha + \delta q_0)m_0 + [1 - (\alpha + \delta q_0)m_0][(\alpha + \delta q)(m_1 N_1 + m_2 N_2))] - P_0 m_0 - P_1 m_1 - P_2 m r_2, \quad (31)$$

whereas under technology sharing it is

$$(\alpha + \delta q_0)(1 + \beta_0)m_0 + [1 - (\alpha + \delta q_0)(1 + \beta_0)m_0][(\alpha + \delta q)(1 + \beta)(m_1 N_1 + m_2 N_2)] - P_0 m_0 - P_1 m_1 - P_2 m_2. \quad (32)$$

As earlier, the profit function of firm i is $P_i(q_i)m_i + Rq_i$. Next, the model is solved using backward induction. For sufficiently small δ, data exploitation is higher under technology sharing vis a vis no sharing. An interesting point of difference is the level of data exploitation and welfare effects relative to the baseline model.

Proposition 8. *Let $q_0^{k,i}$ be the level of data exploitation under business model $k = s, m$ and regime $i = nt, t$. Then there exist $R_1 > 0$ and $R_2 > 0$ with $R_1 < R_2$ such that*
i) For $R \leq R_1$, $q_0^{s,i} < q_0^{m,i}$.
ii) For $R_1 < R \leq R_2$, $q_0^{s,nt} > q_0^{m,nt}$ and $q_0^{s,t} < q_0^{m,t}$.
iii) For $R > R_2$, $q_0^{s,i} > q_0^{m,i}$.

The above proposition shows that with an increase in revenue from sale of data (a higher R), data exploitation under sale of data model increases faster than under mixed business model. To understand this, we need to look at the trade offs that affect the choice of data exploitation under mixed business model. An increase in q_0 increases both revenue from sale of data and targeting rate. The latter effect was not present under sale of data model. The marginal cost of q_0 is also higher under mixed model. So, at margin, the effect of parameter R on optimal value is weighed down by a larger marginal cost. A higher R has a more pronounced effect on choice of q_0 under sale of data model.

7.2 Exclusive Technology Sharing

In the baseline model, the only way to share technology was to offer it to all downstream firms. Suppose, now in addition to that, there can also be a regime"et" in which firm 0 exclusively shares technology with 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

$$N_i = \begin{cases} [t+\theta+\gamma_m(m_j-m_i)]/2t, & \text{if } 2V+\theta-2\gamma_p(q_0+q)-\gamma_m(2m_0+m_i+m_j)-t \geq 0, \\ [V+\theta-\gamma_m(m_0+m_i)-\gamma_p(q_0+q)]/t, & \text{otherwise,} \end{cases} \tag{33}$$

whereas, the demand for firm j is given as

$$N_j = \begin{cases} [t-\theta+\gamma_m(m_i-m_j)]/2t, & \text{if } 2V+\theta-2\gamma_p(q_0+q)-\gamma_m(2m_0+m_i+m_j)-t \geq 0, \\ [V-\gamma_m(m_0+m_i)-\gamma_p(q_0+q)]/t, & \text{otherwise.} \end{cases} \tag{34}$$

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

$$\alpha(1+\beta_0)m_0N_i+\alpha m_0N_j+[1-\alpha(1+\beta_0)m_0]\alpha(1+\beta)m_iN_i+ \\ [1-\alpha m_0]\alpha m_jN_j-P_0m_0-P_im_i-P_jm_j. \tag{35}$$

Figure 2 below gives us a comparison of the level of data exploitation under different regimes. The equilibrium level of data exploitation under regime "t" and "et" decreases monotonically with 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 a vis other regimes. As a result, advertising prices increases faster under regime t. So, when β_0 increases advertising prices rises more under regime"t" than under regime"et", the optimal q_0^t falls faster. This shows that there will be 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, following can be summarised about the level of data exploitation under different regimes.

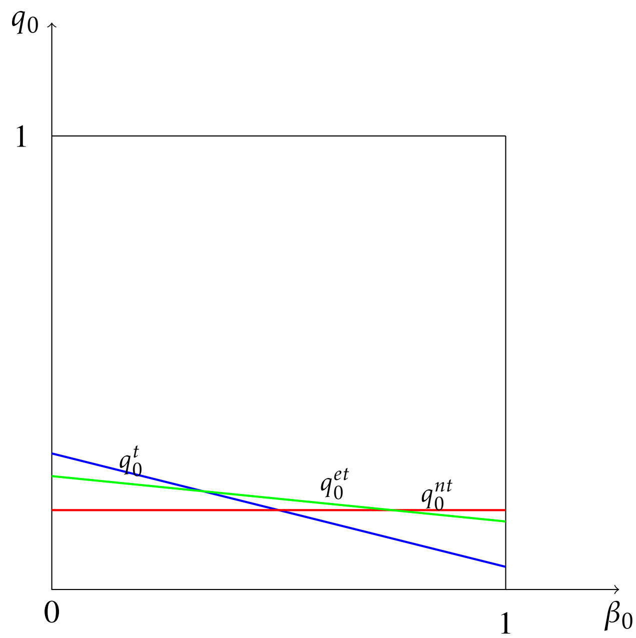

Figure 2: Data exploitation under different regimes

q_0^t (blue line) is the unconstrained level of investment in data exploitation under technology sharing; q_0^{nt} (red line) is the unconstrained level under no technology sharing regime; and q_0^{et} (green line) is the unconstrained level under exclusive technology 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 $\beta = 0.8$.

Proposition 9. *If improvement in targeting rate of downstream firms is sufficiently large then for small values of β_0, data exploitation is highest under technology sharing whereas for large values of β_0 it is highest under no technology sharing.*

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

Proposition 10. *User welfare is the 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 technology 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 receives 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 less advertisements by firm j. The practice also distorts the distribution of users across 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 a distortion. However, due to 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.

7.3 Technology Adoption Decision

In the baseline model, technology sharing was taken as exogenous. In this section, I extend the model and introduce a technology adoption stage. Since data exploitation is a longer term choice variable in the model, a new stage is introduced before 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 to share technology or not.

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

Stage 3: Firm 0 chooses quantity of advertisements m_0 and users decide whether to join firm 0 or not.

Stage 4: Firm 1 and 2 choose 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) whhether to join firm 0 or not, and ii) which downstream firm to join or do not join either.

The market outcome of technology offer and adoption game is described in figure 3 below. As can be seen from the figure, there are thresholds q^a and q^r that are sufficient to delineate the market outcomes. There are four sub regions which have different properties in terms of either nature of offer, adoption or profitability of downstream firms' profit. The technology is offered and adopted by both firms in 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,

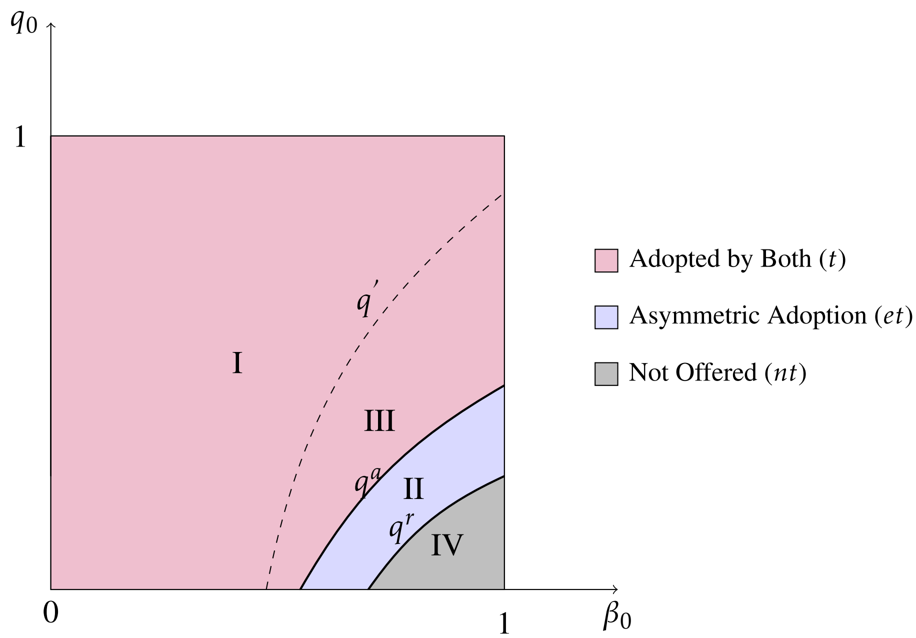

Figure 3: Technology Sharing

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 due to 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.

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 Kramer et al (2018). 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 change in parameter values. The two parameters of interest in this model are - user sensitivity to privacy measured by γ_p and improvement in 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 based are summarised.

It should be noted that the parameter γ_p relates to vertical competition in the market for users. It affects all thresholds and qualitatively 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 increases then, paradoxically, thresholds q^a and q^r decreases. This can be understood from the likely effect of β on advertising competition. As β increases, advertising market becomes more competitive under technology sharing. So, each firms' profit from asymmetirc adoption would be higher. Hence, in the equilibrium, prisoner dilemma outcome becomes more likely as each firm end 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 11. *When the technology offer and adoption decision is 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 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 unconstrained 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^t and q^{nt}) are shown in figure 4. Firm 0 has the option to choose an unconstrained level of data exploitation or strategically reduce or increase the investment to enforce a particular regime. From figure 4, it can be seen that for low values of β_0, firm 0 can choose unconstrained levels q^{nt} or q^t. For high values of β_0, firm 0 can again choose either q^{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

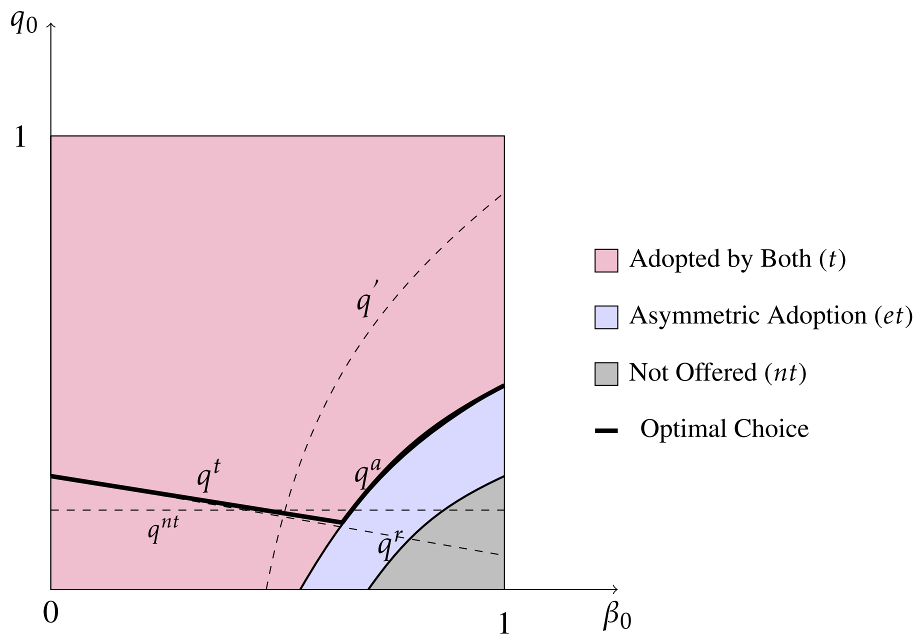

Figure 4: Strategic choice of data exploitation

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 $\beta = 0.8$. q_t and q^{nt} are the unconstrained level of investment in data exploitation udner no technology sharing and technology sharing regime.

line in figure 4. For low to intermediate value of β_0, firm 0 optimally chooses unconstrained q^t.[4] For high value of β_0, firm 0 strategically increases q_0 to q^a to enforce technology adoption by both firms.[5] This important insight is summarized in the following proposition.

Proposition 12. *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.*

The intuition for this result can be seen as follows. The parameter β_0 also affects the advertising competition. So, when β_0 is large then firm 0 can increase q_0 to a level to induce technology sharing and gain from higher probability of matches despite increased competition in the advertising market.[6]

[4] If we take a value of β sufficiently high. It might be the case that firm 0 chooses q^{nt} instead of q^t to avoid adverse ad competition.

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

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

8 Policy Implications and Directions for Future Research

The analysis done in the paper has many interesting policy implications. First, from the results in the paper it is clear that how privacy is affected by technology sharing depends on the business model. Social welfare is always higher when there is advertisement targeting and technology sharing. So, the focus of the intervention should not be on how data is monetized but whether firms share data or not. The analysis suggest that data sharing can be beneficial to the society.

In an extension, it is shown that exclusive sharing can raise user welfare. An exclusive offer benefits the users of the winning firm. It also benefits the excluded users because of lower nuisance cost of advertisements. So, a ban on exclusive offer may reduce the welfare of all users in the model. Also, regulation prohibiting discrimination in technology offer can lead to less offers of the technology. In figure 3, the area where downstream firms themselves adopted technology asymmetrically, a non discrimination rule can reduce the technology adoption.

The baseline model can also be extended in different directions. One of the limitations of the baseline model is that there is a single upstream firm. If there were multiple upstream firms offering contracts to downstream firms then an important question is whether a single firm or multiple firms gain access to data from the downstream firms through technology adoption. Second, vertical integration is an important form of market organisation in digital markets. Firms like Google, Facebook have presence in multiple markets. So, it will be interesting to understand how vertical integration affects the technology adoption and welfare in the model. Third, another form of regulation that can be studied is 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.

References

Armstrong, M., & Wright, J. (2007). Two-sided markets, competitive bottlenecks and exclusive contracts. *Economic Theory*, *32*(2), 353–380.

Bloch, F., & Demange, G. (2018). Taxation and privacy protection on internet platforms. *Journal of Public Economic Theory*, *20*(1), 52-66.

Bourreau, M., Caillaud, B., & De Nijs, R. (2018). Taxation of a digital monopoly platform. *Journal of Public Economic Theory*, *20*(1), 40-51.

Casadesus-Masanell, R., & Hervas-Drane, A. (2015). Competing with privacy. *Management Science*, *61*(1), 229-246.

Chou, C.-f., & Shy, O. (1990). Network effects without network externalities. *International Journal of Industrial Organization*, *8*(2), 259 - 270.

Chou, C.-f., & Shy, O. (1993). Partial compatibility and supporting services. *Economics Letters*, *41*(2), 193 - 197.

Chou, C.-f., & 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., & Gandal, N. (1992). Network effects, software provision, and standardization. *The Journal of Industrial Economics*, *40*(1), 85–103.

Church, J., & Gandal, N. (1993). Complementary network externalities and technological adoption. *International Journal of Industrial Organization*, *11*(2), 239 - 260.

Church, J., & Gandal, N. (2000). Systems competition, vertical merger, and foreclosure. *Journal of Economics & Management Strategy*, *9*(1), 25-51.

Hagiu, A., & Lee, R. S. (2011). Exclusivity and control. *Journal of Economics & Management Strategy*, *20*(3), 679-708.

Hogerdon, C., & Ka Yat Yuen, S. (2009). Platform competition with 'must-have' components*. *The Journal of Industrial Economics*, *57*(2), 294-318.

Krämer, J., Schnurr, D., & Wohlfarth, M. (2019). Winners, losers, and facebook: The role of social logins in the online advertising ecosystem. *Management Science*, *65*(4), 1678-1699.

Stennek, J. (2014). Exclusive quality – Why exclusive distribution may benefit the TV-viewers. *Information Economics and Policy*, *26*(C), 42-57.

Weeds, H. (2016). Tv wars: Exclusive content and platform competition in pay tv. *The Economic Journal*, *126*(594), 1600-1633.

Appendix I: Derivation of Covered Market Condition

In this appendix, I derive the covered market condition that is stated in assumption 1. There are two different regimes - no technology sharing and technology sharing, that we have to consider.

No Technology 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

$$N_i = \frac{V - \gamma_p(q + q_0) - \gamma_m(m_0 + m_1)}{t} \tag{36}$$

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 advertising level m_0. This gives

$$m_i^u = \frac{V - \gamma_p(q + q_0) - \gamma_m m_0^u}{2\gamma_m}, \tag{37}$$

$$m_0^u = \frac{-2(\gamma_m - \alpha\tilde{V}) + \sqrt{4(\alpha\tilde{V} - \gamma_m)^2 + 3\alpha(2\gamma_m\tilde{V} - \alpha\tilde{V}^2)}}{3\alpha\gamma_m}, \tag{38}$$

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

$$\bar{V} = \frac{t(4\gamma_m - 3\alpha t)}{2(\gamma_m - \alpha t)} + \gamma_p(q_0 + q), \tag{39}$$

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

Technology Sharing

Consider technology 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

$$N_i = \frac{V + \theta - \gamma_p(q + q_0) - \gamma_m(m_0 + m_1)}{t} \tag{40}$$

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 advertising level m_0. This gives

$$m_i^u = \frac{\tilde{V} + \theta - \gamma_m m_0^u}{2\gamma_m}, \tag{41}$$

$$m_0^u = \frac{-2(\gamma_m - \alpha(1+\beta)(\tilde{V} + \theta) + \sqrt{4[\alpha(1+\beta)(\tilde{V}+\theta) - \gamma_m]^2 + 3\alpha(1+\beta)[2\gamma_m(\tilde{V}+\theta) - \alpha(1+\beta)(\tilde{V}+\theta)^2]}}{3\alpha(1+\beta)\gamma_m}, \tag{42}$$

where $\tilde{V} = V - \gamma_p(q_0 + q)$. Using the value of m_0^u from the preceding equation, it can be seen that market is covered if $\tilde{V} + \theta - \gamma_m m_0^u > t$. This gives a threshold $V^{'}$ which equals

$$V^{'} = \frac{t[4\gamma_m - 3\alpha(1+\beta)t]}{2[\gamma_m - \alpha(1+\beta)t]} - \theta + \gamma_p(q_0 + q), \tag{43}$$

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 (39). So, $V^{'}$ is binding. Hence, if market is covered under technology sharing then it remains covered under no technology sharing as well.

Appendix II: Existence of Nash Equilibrium

No we need to show that the equilibirum advertising levels and data exploitation under the two regimes are a sub game perfect nash equilibrium.

No Technology Sharing

We need to show that the equilibirum advertising levels and data exploitation specified in equations (14), (15) and (17) constitute a sub game perfect nash equilibrium. The proof proceeds in two parts.

First, given q_0^{nt}, m_0^{nt} and m_j, a downstream firm i will have no incentive to deviate. Second, given firm 1 and 2's best repsonse functions. firm 0 cannot do any better by deviating from m_0^{nt} and q_0^{nt}.

1. No downsteam 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 (5). This gives 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, market becomes uncovered. Reducing its price is not profitable as profit function is concave. So, when deviating, firm i cannot do an unconstrained optimization with covered market. It will make market uncovered and user demand for it will be as given in equation (40). This will give best deviation advertising level as $m_i = \frac{t}{2\gamma_m}$. So, firm i cannot set a higher advertising level by deviating. Similarly it can be show that firm j cannot do any better by deviating.

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 profit function is concave, it is sufficient 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 $m_d^d = m_0^{nt} + \epsilon$,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 covereage. Since, at 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 keep market covered. Under partial market coverage, firm i sets its advertising level at $m_i^d = m_i^u$ given in equation (37) and it equals

$$m_i^d = \frac{t - \gamma_m \epsilon}{2\gamma_m}. \tag{44}$$

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

$$\pi_0(m_0^d) - \pi_0(m_0^{nt}) = \alpha m_0^{nt}\left[\alpha\epsilon - \frac{\alpha\epsilon^2\gamma_m}{2t} - \frac{\gamma_m\epsilon}{t}\right] + \alpha\epsilon\left[1 - \frac{\alpha t}{2\gamma_m} - \frac{\gamma_m\epsilon}{t} + \alpha\epsilon - \frac{\alpha\epsilon^2\gamma_m}{2t}\right] \quad (45)$$

Putting in the value for m_0^{nt} from equation (14), the expression in preceding equation is less than 0 if

$$V \geq t\left[1 - \frac{\alpha/2}{\gamma_m/t + \alpha\epsilon\gamma_m/2t - \alpha}\right] + \gamma_p(q_0^{nt} + q). \quad (46)$$

Given assumption 1, this will hold. So, firm 0 will not deviate to m_0^d. Since, profit function is concave in its arguuments, firm 0 can do no better by deviating from q_0^{nt}. Hence proved.

Technology Sharing

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

$$V \geq t\left[1 - \frac{\alpha(1+\beta)/2}{\gamma_m/t + \alpha(1+\beta)\epsilon\gamma_m/2t - \alpha(1+\beta)}\right] - \theta + \gamma_p(q_0^t + q). \quad (47)$$

This condition will be satisfied if $V \geq V^{'}$. Hence proved.

Appendix III: Proofs of Baseline Model

Proof of proposition 1

Social welfare under no technology sharing regime can be written as

$$\begin{aligned} SW^{nt} = V - \gamma_m m_0 - \gamma_p(q_0 + q) - \gamma_m m_1\hat{x} - \gamma_m m_2(1-\hat{x}) - \frac{t}{2} + t\hat{x} - t\hat{x}^2 \\ + \alpha m_0 + [1 - \alpha m_0][\alpha m_1\hat{x} + \alpha m_2(1-\hat{x})] + Rq_0 - \frac{1}{2}q_0^2 + 2Rq - q^2. \quad (48) \end{aligned}$$

Efficient solution is found by maximizing (48) w.rt q_0, m_0, m_1 and m_2. The first order conditions yield

$$1.\ \frac{\partial SW}{\partial q_0} = -\gamma_p + R - q_0 = 0, \tag{49}$$

$$2.\ \frac{\partial SW}{\partial m_0} = -\gamma_m + \alpha - \alpha[\alpha m_1 \hat{x} + \alpha m_2(1-\hat{x})] = 0, \tag{50}$$

$$3.\ \frac{\partial SW}{\partial m_1} = -\gamma_m\left[\frac{1}{2} + \frac{\gamma_m(m_2 - 2m_1)}{2t}\right] - \gamma_m m_2\left[\frac{\gamma_m}{2t}\right] - t\left[\frac{\gamma_m}{2t}\right] + 2t\hat{x}\left[\frac{\gamma_m}{2t}\right] \\ + [1-\alpha m_0]\left\{\alpha\left[\frac{1}{2} + \frac{\gamma_m(m_2 - 2m_1)}{2t}\right] + \alpha m_2\left[\frac{\gamma_m}{2t}\right]\right\} = 0, \tag{51}$$

$$4.\ \frac{\partial SW}{\partial m_2} = -\gamma_m\left[\frac{1}{2} + \frac{\gamma_m(m_1 - 2m_2)}{2t}\right] - \gamma_m m_1\left[\frac{\gamma_m}{2t}\right] + t\left[\frac{\gamma_m}{2t}\right] - 2t\hat{x}\left[\frac{\gamma_m}{2t}\right] \\ + [1-\alpha m_0]\left\{\alpha\left[\frac{1}{2} + \frac{\gamma_m(m_1 - 2m_2)}{2t}\right] + \alpha m_1\left[\frac{\gamma_m}{2t}\right]\right\} = 0. \tag{52}$$

Solving the F.O.Cs (49) - (52) simultaneously gives us the required solution.

Proof of proposition 2

Social welfare under technology sharing is

$$SW^t = V + \theta - \gamma_m m_0 - \gamma_p(q_0 + q) - \gamma_m m_1\hat{x} - \gamma_m m_2(1-\hat{x}) - \frac{t}{2} + t\hat{x} - t\hat{x}^2 + \alpha(1+\beta_0)m_0 \\ + [1-\alpha(1+\beta_0)m_0][\alpha(1+\beta)m_1\hat{x} + \alpha(1+\beta)m_2(1-\hat{x})] + Rq_0 - \frac{1}{2}q_0^2 + 2Rq - q^2. \tag{53}$$

The solution can be found by maximizing (53) w.r.t q_0, m_0, m_1 and m_2. Then using first order conditions we can get the required solution given in the main text.

Proof of proposition 3

The difference in optimal choice under the two regimes is

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

$$\text{Now, } q_0^t - q_0^{nt} > 0 \text{ if } \beta_0 < \frac{1-\alpha t/2\gamma_m}{1-\alpha(1+\beta)t/2\gamma_m} - 1. \tag{54}$$

The R.H.S in equation (54) is defined as the threshold $\bar{\beta_0}$. Since $\beta \in [0,1]$, it can be seen that $0 \le \bar{\beta_0} < 1$. Hence proved.

Proof of proposition 4

User welfare can be written as

$$UW^i = \int_0^{\hat{x}_i} [V + I*\theta - \gamma_p(q_0^i + q) - \gamma_m(m_0^i + m_1^i) - tx]dx \\ + \int_{\hat{x}_i}^1 [V + I*\theta - \gamma_p(q_0^i + q) - \gamma(m_0^i + m_2^i) - t(1-x)]dx, \tag{55}$$

where $i = t, nt$; $\hat{x}_i$ is the market share under regime i; and m_j^i is the advertising level on firm $j \in \{0,1,2\}$ under regime i. Equation (55) can be rewritten as

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

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 5

Change in the advertiser revenue due to technology sharing is

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

After some algebra, it can be written as

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

Now, putting in the values for m_0^t and m_0^{nt} as given in (24) and (14), we get

$$\Delta AW = \left\{\frac{V-t-\gamma_p(R+q)}{\gamma_m}\right\}\left\{\alpha(1+\beta_0)\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]-\alpha\left[1-\frac{\alpha t}{2\gamma_m}\right]\right\}+\frac{\alpha\beta t}{2\gamma_m}+$$
$$\frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]-\left\{\frac{\alpha\gamma_p}{\gamma_m}\left[1-\frac{\alpha t}{\gamma_m}\right]\right\}^2+\left\{\frac{\alpha(1+\beta_0)\gamma_p}{\gamma_m}\left[1-\frac{\alpha(1+\beta)t}{\gamma_m}\right]\right\}^2. \quad (56)$$

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. $$\frac{\partial AW}{\partial\beta_0} = \alpha\left\{\frac{\bar{V}+\theta-\gamma_p(q+q_0^t)}{\gamma_m}+\frac{\alpha\gamma_p^2(1+\beta_0)}{\gamma_m^2}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right\}\left[1-\frac{\alpha(1+\beta)t}{2\gamma_m}\right],$$

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

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

Since $\beta \in []0,1$, the derivative of ΔAW w.r.t. β_0, β and γ_p will be greater than 0. Hence proved.

Proof of proposition 6

The change in revenue from 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 (17) and (26), ΔR can be written as

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

Setting $\Delta R = 0$ gives a threshold $\bar{\beta_0}$ such that

$$\bar{\beta_0} = \frac{1-\alpha t/2\gamma_m}{1-\alpha(1+\beta)t/2\gamma_m} - 1. \tag{57}$$

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

It can be seen that $\frac{\partial \Delta R}{\partial \beta_0} > 0$ and $\frac{\partial \Delta}{\partial \beta} < 0$. There exist $\bar{\beta_0}$ as defined in (57) such that if $\beta_0 < \bar{\beta_0}$ then $\frac{\partial \Delta R}{\partial \gamma_p} > 0$ and for $\beta_0 \geq \bar{\beta_0}$, it is less tnan 0. Hence Proved.

Proof of proposition]7

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 (56) and (8). User welfare under regime "t" and "nt" are equal. So, $\Delta UW = 0$. This implies that

$$\Delta SW = \left\{\frac{V - t - \gamma_p(R+q)}{\gamma_m}\right\}\left\{\alpha(1+\beta_0)\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right] - \alpha\left[1 - \frac{\alpha t}{2\gamma_m}\right]\right\} + \frac{\alpha\beta t}{2\gamma_m} + \frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]. \quad (58)$$

For sufficiently high V as we have assumed, (58) is greater than 0. Hence proved.

Appendix IV: Extensions

Alternate Business Models

Proof of proposition 8

In order to prove this, we need to compare the level of data exploitation under alternate business models. Let $q_0^{i,k}$ be the level of data exploitation under regime i and business model k where $i = nt, t$ and $k = s$ (sale of data model), and m (mixed business model). The value of $q_0^{i,s}$ is given in the baseline model and for $q_0^{i,m}$ is solved below.[7]

The user demand function remains the same as under baseline model. The advertising prices under no technology sharing regime are as follows

$$P_0^{nt} = (\alpha + \delta q_0)m_0[1 - (\alpha + \delta q)(m_1 N_1 + m_2 N_2)], \quad (59)$$

$$P_i^{nt} = [1 - (\alpha + \delta q_0)m_0](\alpha + \delta q)m_i N_i, \; i = 1, 2. \quad (60)$$

Whereas, under technology sharing, advertising prices are

$$P_0^t = (\alpha + \delta q_0)(1 + \beta_0)[1 - (\alpha + \delta q)(1 + \beta)(m_1 N_1 + m_2 N_2)], \quad (61)$$

[7] The subscript m is omitted for ease of exposition.

$$P_i^t = [1-(\alpha+\delta q_0)(1+\beta_0)m_0](\alpha+\delta q)(1+\beta)m_i N_i,\ i=1,2 \tag{62}$$

Using the values of P_0, P_1 and P_2 we can solve for equilibrium advertising levels under the two regimes. This yields the same solution given in (14) and (24) in the baseline model. Next, we can find the equilibrium value of q_0. This yields

$$q_0^{nt} = \frac{\{1-(\alpha+\delta q)t/2\gamma_m\}\{[\delta(V-t-\gamma_p q)-\alpha\gamma_p]/\gamma_m\}+R}{1+2\gamma_p\delta[1-(\alpha+\delta q)t/2\gamma_m]/\gamma_m},$$
$$q_0^{t} = \frac{\{[1+\beta_0][1-(\alpha+\delta q)(1+\beta)t/2\gamma_m]\}\{[\delta(V-t+\theta-\gamma_p q)-\alpha\gamma_p)]/\gamma_m\}+R}{1+2\gamma_p\delta[1+\beta_0][1-(\alpha+\delta q)(1+\beta)t/2\gamma_m]/\gamma_m}$$

Now, we need to compare the equilibrium values of q_0 under the two business models. First, for no technology sharing regime, taking the difference $q_0^{s,nt} - q_0^{m,nt}$ and setting it equal to 0 gives a threshold R_1 such that

$$R_1 = \frac{\gamma_m}{2\delta\gamma_0}\left\{\frac{\delta(V-t-2\gamma_p q)-\alpha\gamma_p}{\gamma_m}\right\} + \frac{\alpha}{2\delta}\left\{\frac{1-\alpha t/2\gamma_m}{1-(\alpha+\delta q)t/2\gamma_m}\right\}\left\{1+\frac{2\delta\gamma_p}{\gamma_m}\left[1-\frac{(\alpha+\delta q)t}{2\gamma_m}\right]\right\}.$$

Similarly, for technology sharing regime, taking the difference $q_0^{m,t} - q_0^{s,t}$ and setting it equal to 0 gives a threshold R_2 such that

$$R_2 = \frac{\gamma_m}{2\delta\gamma_p}\left\{\frac{\delta(V-t+\theta-\gamma_p q)-\alpha\gamma_m}{\gamma_m}\right\} + \frac{\alpha}{2\delta}\left\{\frac{1-\alpha(1+\beta)t/2\gamma_m}{1-(\alpha+\delta q)(1+\beta)t/2\gamma_m}\right\}$$
$$\left\{1+\frac{2\delta\gamma_p(1+\beta_0)}{\gamma_m}\left[1-\frac{(\alpha+\delta q)(1+\beta)t}{2\gamma_m}\right]\right\}.$$

It can be shown that $R_2 > R_1$ for all $\beta_0 \in [0,1]$. Now, using the two thresholds, proposition can be proved.

Exclusive Technology Sharing

From equation (35), it is clear that targeting rate on firm 0 increases by β_0 and on firm i by β. Using this, the inverse advertising demand functions can be written as

$$P_0^{et} = \alpha(1+\beta_0)N_1 + \alpha N_2 - \alpha^2(1+\beta_0)(1+\beta)m_1 N_1 - \alpha^2 m_2 N_2,$$
$$P_i^{et} = [1-\alpha(1+\beta_0)m_0]\alpha(1+\beta)m_i N_i,$$
$$P_j^{et} = [1-\alpha m_0]\alpha m_j N_j.$$

The demand function for the downstream firms are given in equation (33)1 and (34). Using the advertising prices and demand functions, we can solve for the equilibrium advertising quantities. Downstream firms' best response are

$$m_1 = \begin{cases} (3t+\theta)/3\gamma_m, & \text{if } V - \gamma_p(q_0+q) - \gamma_m m_0 > 3t/2 - \theta/2, \\ [V+\theta-\gamma_p(q_0+q)-\gamma_m m_0]/2\gamma_m, & \text{if } 0 \le V - \gamma_p(q_0+q) - \gamma_m m_0 \le 3t/2 - \theta/2. \end{cases} \tag{63}$$

$$m_2 = \begin{cases} (3t-\theta)/3\gamma_m, & \text{if } V - \gamma_p(q_0+q) - \gamma_m m_0 > 3t/2 - \theta/2, \\ [V-\gamma_p(q_0+q)-\gamma_m m_0]/2\gamma_m, & \text{if } 0 \le V + \theta - \gamma_p(q_0+q) - \gamma_m m_0 \le 3t/2 - \theta/2. \end{cases} \tag{64}$$

At *stage 2*, using the best response functions given in (63) and (64), firm 0 set the optimal advertising level as[8]

$$m_0^{et} = \frac{V - t + \theta/2 - \gamma_p(q_0+q)}{\gamma_m}. \tag{65}$$

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

$$q_0^{et} = R - \alpha\frac{\gamma_p}{\gamma_m}\left\{1 + \frac{\beta_0(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right\}. \tag{66}$$

Proof of Proposition 9

The level of investment in data exploitation under different regimes are

[8] Under assumption 1, it will be optimal for firm 0 to set m_0 such that the market remains covered in equilibrium.

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

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

$$q_0^{et} - q_0^{nt} = \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} - \beta_0\left[\frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m}\right]. \quad (67)$$

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

$$\beta_0^{'} = \left\{\frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m}\right\}\left\{\frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m}\right\}^{-1}. \quad (68)$$

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

$$q_0^{et} - q_0^{t} = \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m} - \frac{\alpha\beta t}{2\gamma_m} + \beta_0\left[\frac{1}{2} - \frac{\theta}{4t} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(1+\beta)t}{2\gamma_m}\right] \quad (69)$$

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

$$\beta_0^{''} = \left\{\frac{\alpha\beta t}{2\gamma_m} - \frac{\alpha\theta^2}{8t\gamma_m} - \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m}\right\}\left\{\frac{1}{2} - \frac{\theta}{4t} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(1+\beta)t}{2\gamma_m}\right\}^{-1}. \quad (70)$$

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

Proof of proposition 10

User welfare is written as

$$UW^i = \int_0^{\hat{x}_i} [V + I * \theta - \gamma_p(q_0^i + q) - \gamma_m(m_0^i + m_1^i) - tx]dx$$
$$+ \int_{\hat{x}_i}^{1} [V + I * \theta - \gamma_p(q_0^i + q) - \gamma(m_0^i + m_2^i) - t(1-x)]dx, \quad (71)$$

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

$$UW^i = V + I * \theta - \gamma_m[m_0^i + m_1^i \hat{x}_i + m_2(1 - \hat{x}_i)] - \gamma_p(q_0^i + q)$$
$$- t\left(\frac{\hat{x}_i^2}{2}\right) - t(1 - \hat{x}_i) + t\left(\frac{1}{2} - \frac{\hat{x}_i^2}{2}\right). \quad (72)$$

The optimal value of m_1 and m_2 is $t/2\gamma_m$ under regime t and nt. Suppose, under regime et, firm 0 shared the technology exclusively with 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 demand for each firm i and j. Using the value of $\hat{x}$, the demand for firm i ($\hat{x}$) is $\frac{2t+\theta}{4t}$ and demand for firm j ($1 - \hat{x}$) is $\frac{2t-\theta}{4t}$. Using these market shares and advertising levels in the preceding equation, the user welfare under 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 technology sharing.

Social welfare under different regimes are

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

The difference in social welfare under regime t and et is

$$
\begin{aligned}
SW^t - SW^{et} = \frac{\alpha[V - t - \gamma_p(R+q)]}{\gamma_m}&\left\{(1+\beta_0)\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right.\\
-\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right] &+ \frac{\alpha(1+\beta_0)\theta}{\gamma_m}\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]\\
- \frac{\alpha\theta}{2\gamma_m}\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right]&\\
+ \frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m} - \frac{\theta^2}{16t}&\\
> \frac{\alpha[V+\theta - t - \gamma_p(R+q)]}{\gamma_m}\left\{(1+\beta_0)\left[1 - \frac{\alpha(1+\beta)t}{2\gamma_m}\right]\right.&\\
-\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right]&\\
+ \frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m} - \frac{\theta^2}{16t}. \qquad (73)&
\end{aligned}
$$

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

$$
\begin{aligned}
\left\{\left[\frac{\alpha(1+\beta)t}{2\gamma_m} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha(2t-\theta)^2}{16t\gamma_m}\right]\left[1 - \left(\frac{\alpha(V-\theta-t-\gamma_p(R+q))}{\gamma_m}\right)^{-1}\right] - \frac{\theta^2}{16t}\right\}&\\
\left\{\frac{1}{2} + \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m} - \frac{\theta}{4t} - \frac{\alpha(1+\beta)t}{2\gamma_m}\right\}^{-1}. \qquad (74)&
\end{aligned}
$$

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

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

$$
\begin{aligned}
SW^{nt} - SW^{et} = \frac{\alpha[V - t - \gamma_p(R+q)]}{\gamma_m}&\left\{\left[1 - \frac{\alpha t}{2\gamma_m}\right]\right.\\
-\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right]&\\
- \frac{\theta^2}{16t} - \frac{\alpha\theta}{2\gamma_m}\left[1+\beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right]&\\
- \frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} - \frac{\alpha\theta^2}{8t\gamma_m}, \qquad (75)&
\end{aligned}
$$

this can be re written as

$$SW^{nt} - SW^{et} = \left[\frac{\alpha\beta(2t+\theta)^2}{16t\gamma_m} + \frac{\alpha\theta^2}{8t\gamma_m}\right]\left[\frac{\alpha[V-t-\gamma_p(R+q)]}{\gamma_m} - 1\right] - (\beta_0)\left[\frac{1}{2} + \frac{\theta}{4t} - \frac{\alpha(1+\beta)(2t+\theta)^2}{16t\gamma_m}\right] - \frac{\theta^2}{16t} - \left[\frac{\alpha\theta}{2\gamma_m}\right]\left[1 + \beta_0\frac{(2t+\theta)}{4t} - \frac{\alpha}{16t\gamma_m}[(1+\beta_0)(1+\beta)(2t+\theta)^2 + (2t-\theta)^2]\right]. \quad (76)$$

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

Technology Adoption Game

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

$$\pi_i^{aa} = [1 - \alpha(1+\beta_0)m_0^{aa}]\alpha(1+\beta)m_i^{aa}N_i^{aa} : \text{ Both firms accept the technology,}$$
$$\pi_i^{ar} = [1 - \alpha(1+\beta_0)m_0^{ar}]\alpha(1+\beta)m_i^{ar}N_i^{ar} : \text{ Firm i accepts and firm j rejects the technology,}$$
$$\pi_i^{ra} = [1 - \alpha m_0^{ra}]\alpha m_i^{ra}N_i^{ra} : \text{ Firm i rejects and firm j accaepts the technology,}$$
$$\pi_i^{d} = [1 - \alpha m_0^{rr}]\alpha m_i^{rr}N_i^{rr} : \text{ Both firms reject the tehcnology.}$$

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 user and market share is equal to $N_i^{aa} = \frac{1}{2}$. If only 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 relevant user set. The values for m_i and m_0 under different scenarios are same as in previos sections. The profit functions given in (8) can be used to determine the thresholds q^a and q^r as outlined in section 7.

Proof of proposition 11

At stage 2b, firm 1 and 2 decide whether to accept or reject the technology offer if an offer is made. The firms move sequentially in the model and since the firms are symmetric, the order of moves doesn't affect the result. Each firm consider the impact of its decision on the anticipated profit that it can obtain. 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 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 $\pi_j^{aa} \geq \pi_j^{ar}$.[9] This inequality implies that, given firm i accepts, firm j's profit from accepting (π_j^{aa}) should be at least as large as firm j's profit from rejecting (π_j^{ar}) the offer. Equality of the two profits gives a threshold q^a such that

$$q^a = \frac{\gamma_m}{\gamma_p} * \left\{ \frac{V - t - \gamma_p q}{\gamma_m} - \frac{[1 + \beta - (2t - \theta)^2/4t^2]}{\alpha[(1 + \beta_0)(1 + \beta) - (2t - \theta)^2/4t^2]} + \frac{\theta[(1 + \beta_0)(1 + \beta) - (2t - \theta)^2/8t^2]}{\gamma_m[(1 + \beta_0)(1 + \beta) - (2t - \theta)^2/4t^2]} \right\}. \quad (77)$$

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

$$q^r = \frac{\gamma_m}{\gamma_p} * \left\{ \frac{V - t - \gamma_p q}{\gamma_m} - \frac{[(1 + \beta)(2t + \theta)^2/4t^2 - 1]}{\alpha[(1 + \beta_0)(1 + \beta)(2t + \theta)^2/4t^2 - 1]} + \frac{\theta[(1 + \beta_0)(1 + \beta)(2t + \theta)^2/8t^2]}{\gamma_m[(1 + \beta_0)(1 + \beta) + (2t + \theta)^2/4t^2 - 1]} \right\}. \quad (78)$$

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

[9] Since both downstream firms have symmetric profit functions the subscript i is dropped.

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 sub cases can be defined.

1. $0 \leq q_0 \leq q^r$: Firm j always reject the offer.
2. $q^r < q_0 < q^a$: Firm j rejects if firm i accepts and firm j accepts if firm i rejects.
3. $q^a \leq q_0 \leq 1$: Firm j always accept the offer.

Since downstream firms' profit functions are symmetric, firm i's choice will yield the same threshold as for firm j. In case 1, firm j always reject the offer. Firm i will accept the offer if its payoff from accepting the offer (π_i^{ar}) is at least as large as 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 2, 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

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

For sufficiently small α, R.H.S in the preceding equation is less than zero. So, firm i always accept the offer if firm j rejects it in the subsequent stage. In case 3, 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 1; firm i accepts and firm j rejects the offer under case 2; and both firms accept the offer under case 3.

Having derived the equilibrium adoption decision, 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 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

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

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

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

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

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 (78). 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 (77).

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. 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 exist $q^{'}$ such that for firm i

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

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

At *stage 1*, anticipating the effect of its choice on adoption decision and advertising levels in the subsequent stages, firm 0 decides whether to make an offer or not. The choice of firm 0 will depend on the level of data exploitation q_0 and hence, on three sub cases 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 or if it doesn't make an offer. A comparison of firm 0's profit under the two regimes is done to find out 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

$$q^o = V - t - \gamma_p q - \theta(1+\beta_0)[1 - \alpha(1+\beta)t/2\gamma_m]C^{-1}, \text{ where} \tag{79}$$

$$C = 1 - \alpha t/2\gamma_m - (1+\beta_0)(1 - \alpha(1+\beta)t/2\gamma_m). \tag{80}$$

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. Under case when both downstream firms reject the offer, firm 0 optimally chooses to make 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

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

where C is as defined in equation (80). It can be shown that, for sufficiently large $\bar{V}$, this partial derivatives with respect to γ_p is less than zero. Similarly, the partial derivative of offer threshold with respect to β is

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

Now, we need to evaluate the signs of the partial derivatives of adoption thresholds. To begin, the partial of q^q wr.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 exist a threshold β_0^{ao} such that

$$\beta_0^{ao} = \frac{(\theta)(2t-\theta)^2/8t^2 + \gamma_m[1+\beta-(2t-\theta)^2/4t^2]/\alpha - [(2t-\theta)^2/4t^2](V-t)}{(1+\beta)(V-t-\theta)} - 1, \tag{81}$$

and if $\beta_0 > \beta_0^{ao}$ then $\partial q^a/\partial\gamma_p < 0$. Also, setting the value of q^a given in (77) equal to 0, there exist a threshold β_0^a such that

$$\beta_0^a = \frac{(\theta)(2t-\theta)^2/8t^2 + \gamma_m[1+\beta-(2t-\theta)^2/4t^2]/\alpha + [(2t-\theta)^2/4t^2](V-t)}{(1+\beta)(V-t-\theta)} - 1, \tag{82}$$

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, $\partial q_0/\partial\gamma_p < 0$ 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

$$\beta_0^{ro} = \frac{\gamma_m[(1+\beta)(2t+\theta)^2/4t^2-1]/\alpha - (V-t)}{(1+\beta)[(V-t)(2t+\theta)^2/4t^2 + \theta(2t+\theta)^2/8t^2]} - 1, \tag{83}$$

and $\partial q^r/\partial\gamma_p < 0$ if $\beta_0 > \beta_0^{ro}$. Also, setting (78) equal to 0 gives a threshold β_0^r such that

$$\beta_0^r = \frac{\gamma_m[(1+\beta)(2t+\theta)^2/4t^2-1]/\alpha - (V-t-\gamma_p q)}{(1+\beta)[(V-t-\gamma_p q)(2t+\theta)^2/4t^2 + \theta(2t+\theta)^2/8t^2]} - 1, \tag{84}$$

and $q^r > 0$ if $\beta_0 > \beta_0^r$. It can be seen 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 13

In order to prove this proposition, we have to compare the profits that firm 0 can obtain from choosing either i) unconstrained investment levels or ii) strategically increasing or decreasing the data exploitation to enforce technology sharing or no sharing. From figure **??**, four sub cases can be defined:

Case a) 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 either choose unconstrained levels q_0^{nt} or q_0^t. If β is sufficiently large, then firm 0 might choose q_0^{nt}. Otherwise, for small values of β, firm 0 can choose q_0^t. This is the case that is shown in figure **??**.

Case b) When β_0 is intermediate and we are in the 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 in no incentive for it to deviate and choose $q_0 = 0$. So, it will choose unconstrained level q_0^t.

Case c) When β_0 is large and 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 the firms.

Case d) When β_0 is large and we are in the 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. If β is small then it will optimally choose q^a to enforce adoption by both firms. If β is large then it will optimally choose q^r to enforce asymmetric enforce adoption.

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