Second-degree Price Discrimination by a Two-sided Monopoly Platform*

Doh-Shin Jeon† Byung-Cheol Kim‡ Domenico Menicucci§

September 22, 2017

Abstract

We study second-degree price discrimination by a two-sided monopoly platform. We find that the platform may optimally forgo price discrimination and offer a single contract on the side that generates strong externalities due to non-responsiveness (Guesnerie and Laffont 1984). However, under certain conditions, the platform may mitigate or remove this non-responsiveness by properly designing price discrimination on the other side. Our research also delivers a welfare analysis of price discrimination in two-sided markets. Then we provide two different applications of our theory: the net neutrality debate and an optimal mechanism design for an advertising platform mediating consumers and advertisers.

JEL codes: D4, D82, L5, M3

Key words: (second-degree) price discrimination, two-sided markets, non-responsiveness, type reversal, advertising platform, net neutrality.

---

* We appreciate helpful comments from Simon Anderson, Alessandro Bonatti, Eric Bond, Jay Pil Choi, Philippe Choné, Andrew Daughety, Lapo Filistrucchi, Renato Gomes, Bruno Jullien, Jinwoo Kim, Yassine Lefouili, Alejandro Molnar, Andras Niedermayer, Greg Taylor, Matthew Oliver, Jennifer Reinganum, Roland Strausz, and seminar attendants at the MaCCI Annual Conference 2017, the 9th IDEI-TSE-IAST Conference on the Economics of Intellectual Property, Software and the Internet, Media Economics Workshop at U of Zurich, ESMT Economics of Platform Workshop, Toulouse School of Economics, Vanderbilt U, and Seoul National U. We acknowledge the financial support from the Net Institute (www.NETinst.org). Jeon acknowledges the financial support from the Jean-Jacques Laffont Digital Chair.

† Toulouse School of Economics, University of Toulouse Capitole (IDEI). Manufacture de Tabacs, 21 Allée de Brienne - 31000 Toulouse, France. E-mail: dohshin.jeon@gmail.com.

‡ Department of Economics, Finance, & Legal Studies, Culverhouse College of Commerce, University of Alabama. Tuscaloosa, AL, 35487, USA. E-mail: byung-cheol.kim@ua.edu.

§ Dipartimento di Scienze per l'Economia e l'impresa, Università degli Studi di Firenze. Via delle Pandette 9, I-50127 Firenze (FI), Italy. E-mail: domenico.menicucci@unifi.it.

1 Introduction

Many two-sided platforms mediating interactions between two different groups practice price discrimination ('PD' for shorthand) against one or both groups of agents. However, little economic analysis has yet been put forward regarding second-degree price discrimination by a two-sided platform, despite the fact that second-degree price discrimination by a monopolist is one of the best-known applications of the principal-agent theory.[1]

For example, the world's largest on-demand streaming service, YouTube, launched its ad-free premium version 'YouTube Red' for a subscription fee in October 2015 while maintaining the free version which contains advertisements.[2] Since YouTube advertisers pay different average-per-view costs depending on ad formats, ad amount and targeting, it suggests that YouTube now adopts PD towards both advertisers and users. Network neutrality regulation is another important example. The debate has primarily focused on whether a tiered-Internet should be allowed for Internet service providers (ISPs) vis-à-vis content providers while ISPs' menu pricing against residential consumers with different quality-price pairs remains uncontroversial. Thus, we can conceptualize the ongoing network neutrality debate as whether society would benefit from introducing PD on the side of content providers in the presence of PD on the side of residential broadband Internet-service subscribers.[3]

In this paper we adapt a canonical model of monopolistic screening à la Mussa and Rosen (1978) to a two-sided monopoly platform (Armstrong 2006; Rochet and Tirole 2006) and study the profit-maximizing mechanism of second-degree price discrimination. In particular, we address the following questions: When is it optimal for the monopoly platform to offer multiple options over a uniform choice to a given side? When does allowing a monopoly platform to use PD increase or reduce welfare? When does PD on one side complements or substitutes for PD on the other side?

A central concept in our paper is *non-responsiveness* (Guesnerie and Laffont 1984,

[1]The seminal papers include Maskin and Riley (1984) and Mussa and Rosen (1978), and there is a vast literature on non-linear pricing. See Armstrong (2015) and Wilson (1993) for in-depth reviews.

[2]There are other examples of menu including an ad-free premium service. Youku Tudou, China's biggest video site, allows subscribers to skip all ads at 20 RMB per month and Amazon Kindle users can choose "Special Offers" to avoid ads for a price ($15-$20). Menu pricing on the consumer side is also used by payment cards, newspapers, Spotify etc.

[3]There has been a fast development of the literature on net neutrality regulation embracing the two-sided market framework. See Greenstein, Peitz, and Valletti (2016) for a recent review on the economic literature about the regulation debate.

Laffont and Martimort 2002), which refers to a clash between the allocation that the principal desires to achieve and incentive compatible (or implementable) allocations; as a result, the principal finds it optimal to offer a pooling contract. In a standard principal-agent model, this conflict may arise when the agent's type directly affects the principal's utility. For instance, suppose that the principal is a benevolent regulator who cares not only about economic efficiency in production cost of a regulated firm but also about an externality measured by the amount of pollution the firm emits. Incentive compatibility requires that a low-cost firm should produce more output than a high-cost firm. However, if the higher cost results from greater efforts to reduce pollution, then the principal may want to induce the high-cost firm to produce more than the low-cost firm. However, such a "non-monotonic" quantity schedule clashes with the implementability condition (i.e., only increasing quality schedules can be implemented in a incentive-compatible way), which makes the principal adopt a pooling contract that is offered to both types of firms.

In this paper, we point out that this non-responsiveness situation can frequently appears to two-sided platforms which offer intermediation services involving cross-group interactions. Furthermore, we show that a two-sided platform may mitigate or remove non-responsiveness at one side by properly designing PD on the other side. Consider, for instance, a media platform that offers content to consumers who are exposed to advertisements delivered together. Suppose that there are rich (H type) and poor (L type) consumers. Without PD on the advertising side, the rich may suffer the higher average nuisance from advertisements so that they have a greater willingness to pay to avoid ads than the poor. However, from an advertiser's perspective, the rich consumers are more valuable than the poor ones. Therefore, the platform may prefer that H type consumers have a greater exposure to ads than L types, which is impossible to implement because of the incentive compatibility constraints. Can a two-sided platform show more ads to the rich consumers while still inducing self-selection? The answer is positive, provided that there are high-end advertisements that disturb the rich less than the poor. Then, by designing a PD on the advertising side that assigns more weight to such high-end ads than to low-end ones, the platform can make viewing ads on average less displeasing to H types than to L types.

One assumption implicitly made in the above example is that, although the rich consumers will pay more to avoid the ads on average, there must be some high-end ads that they find less offensive. We refer to this property as *type reversal.* Note also that above the platform's screening instrument (called "quality" in our model following Mussa and Rosen 1978) on one side is complementary to the one on the other side from the consumers perspective.[4]

In this paper we provide a canonical monopolistic screening model of a two-sided platform that allows for all possible combinations of type reversal or no reversal and complementarity or substitutability between the qualities on each side of a two-sided market.[5] In our model, a monopolistic two-sided platform offers a menu of price-quality pairs to a continuum of agents whose mass is normalized to one on each side. The utility that an agent i of side k $(= A, B)$ obtains from interacting with an agent j of the other side $l \neq k$ (and $l = A, B$) depends on each agent's type, denoted θ_i^k and θ_j^l, and the quality each agent receives, denoted q_i^k and q_j^l. We consider a two-type model: agent i of side k has either H or L type where H type is defined to have the greater *expected benefit* than L type from any given increase in q_i^k when interacting with all agents on the other side l.[6] The two qualities q_i^k and q_j^l can be complements or substitutes on a given side, which we refer to as "non-separable" case. By contrast, the "separable" case refers to the situation in which the qualities affect agent i's utility in a separate way.

We have four sets of novel results. First, we consider the separable case and characterize the first-best and the second-best allocations. In the separable case, we find that the first-best quality schedule on a given side k is *decreasing* (i.e., the quality chosen for H type is lower than the one for L type) if the L type of side k generates sufficiently larger positive externalities to the other side than the H type does. In the presence of asymmetric information, such decreasing schedule clashes with the implementability condition, which makes pooling optimal. More generally, asymmetric information creates not only the well-known own-side distortion but also new distortions due to the two-sidedness of the market as the information rent yielded to the H type of a given side can be affected by the quality schedule offered to the other side. Consequently, the standard result of "no distortion at top and a downward distortion at bottom" holds no longer. In addition, because of this new distortion, a non-responsiveness

[4]Suppose that the platform offers consumers the service to opt out of advertising at a certain fee, which clearly means a higher quality to consumers. The benefit from avoiding the ads should increase with the amount of advertisements which measures "quality" to advertisers. So, the two quality measures go to the same direction and form complements.

[5]The qualities are complements (substitutes) on side k if the cross-derivative of $u^k(q_i^k, q_j^l)$ is positive (negative) for $k \neq l$ and $k, l = A, B$. This should not be confused with the complementarity (substitution) between PD of both sides.

[6]One caveat is that the H type does not necessarily have the greater benefit than the L type for interacting with *each* given type of agents on the other side (see Section 2 for details).

can occur even if the first-best quality schedule is increasing—this is not possible in a one-sided market. In other words, two-sided interactions generate another source for non-responsiveness, different from the one identified by Guesnerie and Laffont (1984).

Second, we provide a welfare analysis of price discrimination in the separable case and apply it to the net neutrality regulation. Suppose that the platform introduces PD on side B. We isolate its effect on the welfare of the same side, which is the standard welfare effect in one-sided market, from its effect on that of the other side which is unique to a two-sided market. We find that introducing PD on side B raises (reduces) welfare of side A if H type agents of side B generate more (less) positive externalities to the agents of side A than L type agents of side B do. When applied to the net neutrality regulation, our results suggest that prohibiting price discrimination on the content side can be socially desirable if L-type content generates more positive externalities to consumers than H-type content. Here, H type content is expected to be more sensitive to traffic delay than L type content: e.g., H type content includes video and music streaming service, video conferencing, etc. If H type content has more market power and hence tends to use more micro-payments instead of advertising-based business model than L type content does, then the former would extract much more surplus from consumers than the latter does. In this case, H-type content could generate less positive externalities to consumers than L-type content and the net neutrality regulation could be socially desirable. Our welfare analysis captures that of Choi, Jeon and Kim (2015) as a special case.

Third, we characterize implementable allocations on one side given an arbitrary quality schedule on the other side. We find that the implementable allocations on side k are equal to the set of increasing quality schedules if any of the following conditions holds: (i) q_i^k and q_j^l are separable on side k, (ii) there is no type reversal on side k, (iii) there is no PD on side l with $l \neq k$. Then, we show that if the qualities are non-separable and there is a type reversal on side A, a decreasing schedule can be implemented on side A when some appropriate PD is introduced onto side B. The intuition for this result is as follows. Basically, the implementability condition on side A means that, given a quality schedule on side B, an L type's gain from choosing q_L^A instead of q_H^A must be greater than that of an H type. Consider a decreasing schedule on both sides, i.e., $q_L^k > q_H^k$ for $k = A, B$. Consider a particular kind of type reversal (later to be defined as 'type reversal with a positive sorting') such that the L type agent on side A gets the larger benefit than the H type when interacting with an L type on side B. If the two qualities are complements on side A, then the L type agent on side A can experience a much greater utility increase from choosing q_L^A instead of

q_H^A than the H type, which makes $q_L^A > q_H^A$ implementable.

Last, we illustrate our insights by providing an application to an advertising platform as we described earlier in the introduction. This application also allows us to answer the question of when introducing PD on one side complements or substitutes for PD on the other side.[7] We study how the optimal profit-maximizing mechanism varies with the intensity of the type reversal on the consumer side. We find that, for a low intensity of type reversal, the profit maximization requires that L type advertisers to advertise more than H type advertisers, which clashes with the implementability condition on the advertising side. Thus, a pooling contract becomes optimal on the advertising side. This implies that a strict PD on the advertising side will reduce the platform's profit and therefore PD on the advertising side substitutes for PD on the consumer side. By contrast, for a high intensity, profit maximization requires H type advertisers to advertise more than L type advertisers and can even require implementing a decreasing quality schedule on the consumer side (i.e., showing ads only to H type consumers). Then, PD on the advertising side is complementary to the PD on the consumer side as it not only allows implementation of a desirable discrimination on the advertising side but also a decreasing quality schedule on the consumer side.

■ **Related literature**

This article is related to several strands of literature. First, our paper is closely related to the second-degree PD in the principal-agent theory (e.g. Maskin and Riley 1984; Mussa and Rosen 1978) and to the concept of non-responsiveness. The non-responsiveness was developed by Guesnerie and Laffont (1984) and then was explored by Caillaud and Tirole (2004) in the context of financing an essential facility and by Jeon and Menicucci (2008) in allocating talented scientists between public and private sectors. To our knowledge, however, non-responsiveness has never been explored from the perspective of two-sided markets; our contribution is to identify a novel source for non-responsiveness that has to do with two-sidedness of various markets.

By now, there is a large literature on two-sided markets. Our paper is more closely related to the papers studying a monopoly two-sided platform (e.g., Caillaud and Jullien 2001, Rochet and Tirole 2003, Armstrong 2006, Rochet and Tirole 2006, Hagiu 2009, Jeon and Rochet 2010, Weyl 2010). They mostly study the optimal pricing structure by focusing on the number of users joining the platform on each side; for instance, Weyl (2010) considers a rich type space and identifies the Spence (1975)

[7] We say that PD on one side complements PD on the other side if the optimal mechanism does not involve any pooling.

distortion in that when deciding the level of participation on one side, the platform internalizes cross-side externalities to marginal rather than average users of the other side. To our knowledge, we are the first to study second-degree price discrimination of a monopoly platform: we introduce a screening instrument called 'quality' on each side and allow the platform to offer a menu on each side in a setup where the benefit that an agent obtains from interacting with another agent depends on the type of each agent and the quality that each agent chooses. We maintain the standard assumption in the literature that all agents on any given side interact with all (or a random subset of) agents on the other side (Armstrong 2006, Rochet and Tirole 2006, Weyl 2010). This framework allows us to analyze several important issues that have not been previously addressed such as non-responsiveness and complementarity or substitution between PD on one side and PD on the other side. Gomes and Pavan (2016) is slightly close to our paper in that they consider heterogeneous agents on both sides. But they study a centralized many-to-many matching and an optimal matching rule[8], which does not satisfy our assumption that all agents on one side interact with all agents on the other side. Their screening instrument is the matching rule while we consider qualities as screening instruments as in the standard monopolistic screening literature (Mussa and Rosen 1978). Choi, Jeon and Kim (2015) study second-degree PD of a two-sided monopoly platform in the context of the network neutrality regulations. However, they consider heterogeneous agents only on the content-provider side and assume homogeneous agents on the consumer side, while we consider heterogeneous agents on both sides. Moreover, they study neither non-responsiveness nor type reversal and their welfare analysis is a special case of ours.[9]

Finally, our application to an advertising platform is related to the two-sided market literature on advertising/media platforms (Gabszewicz, Laussel and Sonnac 2004; Anderson and Coate 2005; Peitz and Valletti 2008; Crampes, Haritchabalet and Jullien, 2009; Ambrus, Calvano and Reisinger 2016; Angelucci and Cage 2016). For instance, Angelucci and Cage (2016) study PD of a monopoly newspaper by selling subscription and individual issue to consumers but do not consider PD on the advertising side. Our

[8]They provide conditions on the primitives under which the optimal matching rule has a threshold structure such that each agent on one side is matched with all agents on the other side above a threshold type.

[9]Böhme (2012) analyzes second-degree PD in a monopolistic screening model with network effects. Since he considers two types of agents only in one side, who are heterogeneous regarding their intrinsic utility from joining the platform, most remarks we made to Choi, Jeon and Kim (2015) apply to Böhme (2012).

contribution consists in studying the second-degree price discrimination of the platform by focusing on non-responsiveness and complementarity or substitution between PD on one side and PD on the other side.

The rest of the article is organized as follows. We set up the canonical model in Section 2. In Sections 3-5, we consider the separable case. We characterize the first-best allocation in Section 3 and the second-best allocation in Section 4. In Section 5, we perform the welfare analysis of price discrimination and provide the application to net neutrality debate. In Sections 6-7, we consider the non-separable case. In Section 6, we study the implementable allocations on side A for a given quality schedule on side B and in Section 7, we apply the insight from Section 6 to an advertising platform. We conclude in Section 8. All mathematical proofs not covered in the text are relegated to the Appendix.

2 A canonical principal-agent model in two-sided markets

We consider a canonical principal-agent model (Mussa and Rosen 1978, Laffont and Martimort 2002) and adapt it to a two-sided market where a monopoly platform as the principal designs a mechanism to mediate interactions between agents from two sides, $k = \mathrm{A}, \mathrm{B}$. On each side there is a mass one of agents. Let θ_i^k represent the type of agent i on side k. For simplicity we consider a two-type model: an agent has one of the two types, H or L, on each side, i.e., $\theta_i^k \in \{H, L\}$. Let $\nu_H^k \in (0,1)$ represent the fraction of H-types on side k. Let $\nu_L^k \equiv 1 - \nu_H^k$. Let q_i^k be the "quality" chosen for agent i of side k: although we call it quality, q_i^k be interpreted as quantity depending on applications. When an agent i of side k interacts with an agent j of side l with $k \neq l$ and $k, l = \mathrm{A}, \mathrm{B}$, the gross utility the agent i obtains can be represented as follows:

$$U_i^k(\theta_i^k, \theta_j^l, q_i^k, q_j^l) = \theta_{ij}^k u^k(q_i^k, q_j^l),$$

where the types interact in a multiplicative way with qualities as in Mussa and Rosen (1978) and θ_{ij}^k represents the consumption intensity of the agent i of side k as a function of both agents' types. Compared to the price discrimination in one-sided market of Mussa and Rosen (1978), there are two additional interactions from the two-sidedness of the market: both the type of and the quality of the agent j on the other side l matter.

We maintain the standard assumption in the two-sided market literature (Armstrong 2006, Rochet and Tirole 2006, Weyl 2010) that all agents on any given side interact with all (or a random subset of) agents on the other side. We focus on the situations in which any given agent receives a unique quality that does not depend on the types of the agents interacting with. Therefore, on each side k $(= \mathrm{A}, \mathrm{B})$, the platform offers the following menu of quality-price pairs $\{(q_H^k, p_H^k), (q_L^k, p_L^k)\}$ where $q_H^k \in \mathbb{R}^+$ denotes the quality for a H-type agent of side k and $p_H^k \in \mathbb{R}$ denotes a monetary payment from a H-type agent of side k to the platform.[10] Let $\mathbf{q} \equiv (q_H^A, q_L^A, q_H^B, q_L^B) \in \mathbb{R}_+^4$ denote the vector of quality specifications. If all agents accept the offer of the platform and self-select, agent i of side k obtains the following utility depending on his type

$$\begin{array}{lllll} \nu_H^l \theta_{HH}^k u^k(q_H^k, q_H^l) & + & \nu_L^l \theta_{HL}^k u^k(q_H^k, q_L^l) & - p_H^k, & \text{if } \theta_i^k = H; \\ \nu_H^l \theta_{LH}^k u^k(q_L^k, q_H^l) & + & \nu_L^l \theta_{LL}^k u^k(q_L^k, q_L^l) & - p_L^k, & \text{if } \theta_i^k = L. \end{array}$$

To provide more tangible interpretation for the parameters and variables in our model, let us consider three applications.

- Net neutrality regulation: In Section 5, we apply our model to the net neutrality debate. In the debate, a monopoly ISP mediates a group of network subscribers ($k = A$) with a group of content providers ($k = B$). The parameter θ_{ij}^A in this setting measures consumer i's preference intensity when she consumes content provided by content provider j, which is then multiplied by her utility that depends on the consumer i's choice of her residential Internet quality, q_i^A, and the sending content provider j's quality, q_j^B. Similarly, θ_{ji}^B measures content provider j's preference intensity relating to the revenue which is also affected by the types of i and j.

- Advertising platform: In Section 7, we apply the model to an advertising platform. In the application, $q_i^A \in \{0, 1\}$ and $q_i^A = 1$ means no exposure to advertising like YouTube Red and $q_i^A = 0$ means exposure to advertising like basic YouTube. θ_{ij}^A captures consumer i's nuisance from advertiser j's advertisement. On side B, q_j^B represents advertising amount of advertiser j and θ_{ji}^B measures j's advertising revenue which is jointly affected by consumer i's type such as income and advertiser j's type such as the advertised product's characteristics.

- Privacy protection and targeted advertising: Consider consumer privacy protec-

[10] If the quality chosen for an agent of side k can vary depending on the type of the agent of side l he is interacting with, we can consider the following mechanism $\{(q_{HH}^k, q_{HL}^k, p_H^k), (q_{LH}^k, q_{LL}^k, p_L^k)\}$ in which q_{HL}^k represent the quality for an H-type agent of side k when he interacts with an L-type agent of side l. Such mechanism can be relevant for targeted advertising in which the amount of advertising for a given type of advertiser varies depending on the type of consumer.

tion design by an online-advertising platform who uses the information released from consumers to increase efficiency in targeted advertising. In this environment, q_i^A captures the level of privacy designed for consumer i and q_j^B the advertising amount by advertiser j, while the intensity measures typify the match-based preferences by consumers and advertisers.[11]

We assume that the utility function $u^k : \mathbb{R}_+^2 \to \mathbb{R}$ is strictly increasing in q^k, and concave in (q^k, q^l) with $k \neq l$ and $k, l = \mathrm{A}, \mathrm{B}$. Note that u^k may increase or decrease with q^l. For example, it is increasing in q^l in net neutrality application, but decreasing in q^l in the application to an advertising platform. Let u_m^k denote the partial derivative of u^k with respect to its m-th variable, for $m = 1, 2$. Moreover, we define u_{12}^k as follows:

$$u_{12}^k(q_i^k, q_j^l) \equiv \frac{\partial^2 u^k(q_i^k, q_j^l)}{\partial q_i^k \partial q_j^l}.$$

We assume that u_{12}^k has the same sign for each q_i^k and q_j^l. For a given side k the qualities are said to be independent if $u_{12}^k = 0$, complements if $u_{12}^k > 0$, and substitutes if $u_{12}^k < 0$. The costs of producing q_i^A and q_j^B are respectively denoted by $C^A(q_i^A)$ and $C^B(q_j^B)$. We assume that both cost functions are strictly increasing and convex.

Depending on the match of types, we have the following four parameters of consumption intensity on side k:

$k \backslash l$	H	L
H	θ_{HH}^k	θ_{HL}^k
L	θ_{LH}^k	θ_{LL}^k

To give a standard meaning to the H and L types, we introduce the following notation and assumption:

Assumption 1. $\theta_H^k \equiv \nu_H^l \theta_{HH}^k + \nu_L^l \theta_{HL}^k > \theta_L^k \equiv \nu_H^l \theta_{LH}^k + \nu_L^l \theta_{LL}^k$ *with* $k \neq l$ *and* $k, l = A, B$.

Assumption 1 means that the H type on side k enjoys a higher increase in benefit from a marginal increase in u^k than the L type on side k when interacting with the agents on the other side l. Because of Assumption 1, we say that the quality schedule of side k is increasing if $q_L^k \leq q_H^k$ and decreasing if $q_L^k \geq q_H^k$. Basically, $q^k(\theta)$ is increasing in θ if a higher θ leads to a higher q^k.

[11]According to a recent settlement between the Federal Communications Commission (FCC) and Verizon Wireless in March 2016, the wireless company needs opt-in from users in order to employ its tracking system so-called "supercookies" for targeted advertising. In the model, q_i^A can have two binary values for opt-in and opt-out (default) choices.

Under Assumption 1, we can further identify three sub-cases depending on the signs of $\theta^k_{HH} - \theta^k_{LH}$ and of $\theta^k_{HL} - \theta^k_{LL}$.[12]

Definition (type reversal) We say that on side k, there is

$$\begin{cases} \text{no type reversal} & \text{if } \theta^k_{HH} - \theta^k_{LH} > 0 \text{ and } \theta^k_{HL} - \theta^k_{LL} > 0; \\ \text{type reversal with a positive sorting} & \text{if } \theta^k_{HH} - \theta^k_{LH} > 0 > \theta^k_{HL} - \theta^k_{LL}; \\ \text{type reversal with a negative sorting} & \text{if } \theta^k_{HL} - \theta^k_{LL} > 0 > \theta^k_{HH} - \theta^k_{LH}. \end{cases}$$

Type reversal arises on side k if an L type obtains the greater marginal benefit compared to an H type when interacting with a *particular* type of side $l(\neq k)$, despite that by definition the H type will always enjoys the greater average benefit than an L type when interacting with *all* the agents of side l. If this particular type of side l is L (H), then we have type reversal with a positive (negative) sorting.

For simplicity, we assume that it is optimal for the platform to induce full participation of all agents on both sides. Therefore, no PD on side k means that the platform offers a sole option $(q^k_H, p^k_H) = (q^k_L, p^k_L)$ on side k which satisfies the participation constraints of both types.[13]

Even with a two-type model of the multiplicative specification, our model is characterized by quite a few parameters of $\Theta^k \equiv \{\theta^k_{HH}, \theta^k_{HL}, \theta^k_{LH},\ \theta^k_{LL}\},\ \nu^k$, and the utility function u^k is defined for each $k =$ A, B. For this reason, when necessary, we consider a simpler case by further specifying the model. In the case of $u^A_{12} = u^B_{12} = 0$, u^A and u^B become separable in the sense that there exist four single variable functions u^A_A, u^A_B, u^B_B, and u^B_A satisfying

$$u^A(q^A, q^B) = u^A_A(q^A) + u^A_B(q^B) \quad \text{and} \quad u^B(q^B, q^A) = u^B_B(q^B) + u^B_A(q^A)$$

where the superscripts refer to the side of the agent whose utility is computed whereas the subscripts refer to the side of which the quality affects the utility of the agent. In Sections 3-5, we focus on this separable case whereas in Sections 6-7 we consider the non-separable case. Note that in the separable case, under certain conditions, the optimal quality schedule on one side does not depend on the quality schedule on the

[12]We are implicitly assuming that $\theta^k_{HH} - \theta^k_{LH} \neq 0$ and $\theta^k_{HL} - \theta^k_{LL} \neq 0$ but this is immaterial and only for expositional brevity.

[13]This assumption is relaxed in the application to an advertising platform in Section 7 where the exclusion of a certain type of advertisers may arise in an optimal mechanism.

other side. This helps us to perform the positive and normative analysis of introducing PD on a given side independently of the quality schedule applied to the other side. However, we consider the non-separable case when we study complementarity and substitution between PD on one side and PD on the other side.

3 First-best in the separable case

In this section we characterize the first-best quality schedule that maximizes the total surplus given as follows:

$$\begin{aligned}\Pi^{FB}(\mathbf{q}) = &\nu_H^A \nu_H^B \left[\theta_{HH}^A u^A(q_H^A, q_H^B) + \theta_{HH}^B u^B(q_H^B, q_H^A)\right] \\ &+ \nu_H^A \nu_L^B \left[\theta_{HL}^A u^A(q_H^A, q_L^B) + \theta_{LH}^B u^B(q_L^B, q_H^A)\right] \\ &+ \nu_L^A \nu_H^B \left[\theta_{LH}^A u^A(q_L^A, q_H^B) + \theta_{HL}^B u^B(q_H^B, q_L^A)\right] \\ &+ \nu_L^A \nu_L^B \left[\theta_{LL}^A u^A(q_L^A, q_L^B) + \theta_{LL}^B u^B(q_L^B, q_L^A)\right] \\ &- \nu_H^A C^A(q_H^A) - \nu_L^A C^A(q_L^A) - \nu_H^B C^B(q_H^B) - \nu_L^B C^B(q_L^B),\end{aligned}$$

where each of the first four lines represents the total gross surplus from each matching pattern (H, H), (H, L), (L, H) and (L, L) while the last line measures the total costs.[14]

Given our assumptions, Π^{FB} is concave and therefore the FOCs characterize the first-best quality schedule. In the separable case, the first-best quality schedule on side A, denoted by $(q_H^{A,FB}, q_L^{A,FB})$, is determined by a system of the following two FOCs:

$$\theta_H^A u_A^{A\prime}(q_H^A) + (\nu_H^B \theta_{HH}^B + \nu_L^B \theta_{LH}^B) u_A^{B\prime}(q_H^A) = C^{A\prime}(q_H^A); \tag{1}$$

$$\theta_L^A u_A^{A\prime}(q_L^A) + (\nu_H^B \theta_{HL}^B + \nu_L^B \theta_{LL}^B) u_A^{B\prime}(q_L^A) = C^{A\prime}(q_L^A). \tag{2}$$

Note that neither of (1) and (2) depends on the quality schedule on side B.

We find that the first-best quality schedule is decreasing on side A (i.e., $q_L^{A,FB} > q_H^{A,FB}$) if and only if

$$(\theta_H^A - \theta_L^A) u_A^{A\prime}(q_H^{A,FB}) + \left[\nu_H^B \left(\theta_{HH}^B - \theta_{HL}^B\right) + \nu_L^B \left(\theta_{LH}^B - \theta_{LL}^B\right)\right] u_A^{B\prime}(q_H^{A,FB}) < 0. \tag{3}$$

For a clearer interpretation of (3), for now let us consider the special case in which

[14] An alternative cost function is

$$C^A(\nu_H^A q_H^A + \nu_L^A q_L^A) + C^B(\nu_H^B q_H^B + \nu_L^B q_L^B).$$

Qualitative results in this paper remain robust regardless of which cost function is chosen.

there exists a $\beta_A^B > 0$ such that

$$u_A^B(q) = \beta_A^B u_A^A(q) \qquad \text{for each } q > 0, \tag{4}$$

where β_A^B captures the intensity of the externalities from quality of side A to side B. Then, the condition specified at (3) becomes

$$\underbrace{\left(\theta_H^A - \theta_L^A\right)}_{(\dagger)>0} + \underbrace{\beta_A^B \left[\nu_H^B \left(\theta_{HH}^B - \theta_{HL}^B\right) + \nu_L^B \left(\theta_{LH}^B - \theta_{LL}^B\right)\right]}_{(\ddagger)\gtrless 0} < 0.$$

The first bracketed term denoted by (†) represents the change in the private benefit that a side A agent would experience when her type changes from L to H. This term is positive under Assumption 1. By contrast, the second term (‡) represents the change in the externality onto the agents on side B from the same type change. If $\nu_H^B \left(\theta_{HH}^B - \theta_{HL}^B\right) + \nu_L^B \left(\theta_{LH}^B - \theta_{LL}^B\right) > 0$, it captures the situation that the H type of side A generates a greater positive or a smaller negative externality to side B than the L type of side A does. Both the own benefit to side A and the externality onto side B take positive values, which means the first-best quality schedule is increasing such that $q_L^{A,FB} < q_H^{A,FB}$. However, if $\nu_H^B \left(\theta_{HH}^B - \theta_{HL}^B\right) + \nu_L^B \left(\theta_{LH}^B - \theta_{LL}^B\right) < 0$ and β_A^B is sufficiently large, the externality term (‡) is negatively large enough that Condition (3) is satisfied. Then, the first-best schedule is decreasing, i.e., $q_L^{A,FB} > q_H^{A,FB}$. Such decreasing quality schedule can arise only if the L type agent of side A generates a sufficiently greater externality to side B than the H type agent of side A does.

Proposition 1. *(First-best) Suppose that Assumption 1 holds and both u^A and u^B are separable (i.e., $u_{12}^A = u_{12}^B = 0$).*

(i) The first-best quality schedule on side A, $(q_H^{A,FB}, q_L^{A,FB})$ is determined by (1) and (2) independently of that on side B, $(q_H^{B,FB}, q_L^{B,FB})$.

(ii) We have a decreasing first-best quality schedule $q_L^{A,FB} > q_H^{A,FB}$ if and only if inequality (3) holds. This is when an H type's gain in terms of private benefit relative to that of an L type on the same side A is smaller than an L type's contribution in terms of externality to the other side B relative to that of an H type.

(iii) Parallel statements can be made regarding $q_H^{B,FB}$ and $q_L^{B,FB}$.

Decreasing quality schedule matters only for the side that generates large positive externalities. For instance, if A side alone generates positive externalities such that the platform makes most revenue from side B, the first-best quality schedule on side B is always increasing.

The analysis of the first-best in the separable case can be extended to some non-separable cases. For instance, we can prove that for specific u_A and u_B in which the effect of complements/substitutes is represented by the product of qualities, a small degree of complementarity (substitution) on both sides increases (decreases) each first-best quality compared to the separable case.[15]

4 Second-best in the separable case

In this section, we study the second-best mechanism in the separable case and identify the distortions generated by asymmetric information. By doing so, we can clearly identify two different sources for non-responsiveness: one is known from Guesnerie and Laffont (1984), but the other is newly emerging due to the two-sidedness of the market.

The platform's optimization problem is given by:

$$\max_{\{(q_H^k,p_H^k),(q_L^k,p_L^k)\}} \nu_H^A\left[p_H^A - C^A(q_H^A)\right] + \nu_L^A\left[p_L^A - C^A(q_L^A)\right] + \nu_H^B\left[p_H^B - C^B(q_H^B)\right] + \nu_L^B\left[p_L^B - C^B(q_L^B)\right]$$

subject to

$$\begin{aligned}
&(\mathrm{IR}_H^k) \qquad \nu_H^l \theta_{HH}^k u^k(q_H^k, q_H^l) + \nu_L^l \theta_{HL}^k u^k(q_H^k, q_L^l) - p_H^k \geq 0;\\
&(\mathrm{IR}_L^k) \qquad \nu_H^l \theta_{LH}^k u^k(q_L^k, q_H^l) + \nu_L^l \theta_{LL}^k u^k(q_L^k, q_L^l) - p_L^k \geq 0;
\end{aligned}$$

$$(\mathrm{IC}_H^k)\ \nu_H^l \theta_{HH}^k u^k(q_H^k, q_H^l) + \nu_L^l \theta_{HL}^k u^k(q_H^k, q_L^l) - p_H^k \geq \nu_H^l \theta_{HH}^k u^k(q_L^k, q_H^l) + \nu_L^l \theta_{HL}^k u^k(q_L^k, q_L^l) - p_L^k;$$

$$(\mathrm{IC}_L^k)\ \nu_H^l \theta_{LH}^k u^k(q_L^k, q_H^l) + \nu_L^l \theta_{LL}^k u^k(q_L^k, q_L^l) - p_L^k \geq \nu_H^l \theta_{LH}^k u^k(q_H^k, q_H^l) + \nu_L^l \theta_{LL}^k u^k(q_H^k, q_L^l) - p_H^k;$$

where $k, l = A, B$ and $k \neq l$.

When u^A and u^B are separable, the above constraints get simplified as follows:

$$\begin{aligned}
&(\mathrm{IR}_H^k) \qquad \theta_H^k u_k^k(q_H^k) + \nu_H^l \theta_{HH}^k u_l^k(q_H^l) + \nu_L^l \theta_{HL}^k u_l^k(q_L^l) - p_H^k \geq 0;\\
&(\mathrm{IR}_L^k) \qquad \theta_L^k u_k^k(q_L^k) + \nu_H^l \theta_{LH}^k u_l^k(q_H^l) + \nu_L^l \theta_{LL}^k u_l^k(q_L^l) - p_L^k \geq 0;\\
&(\mathrm{IC}_H^k) \qquad \theta_H^k u_k^k(q_H^k) - p_H^k \geq \theta_H^k u_k^k(q_L^k) - p_L^k;\\
&(\mathrm{IC}_L^k) \qquad \theta_L^k u_k^k(q_L^k) - p_L^k \geq \theta_L^k u_k^k(q_H^k) - p_H^k.
\end{aligned}$$

Notice that the two IC constraints (IC_H^k) and (IC_L^k) are independent of $\left(q_H^l, q_L^l\right)$ (be-

[15] The analysis can be obtained from the authors upon request.

cause of $u^k_{12} = 0$), but $\left(q^l_H, q^l_L\right)$ affects both IR constraints (IR^k_H) and (IR^k_L). By adding up the two incentive constraints on side k, we obtain the implementability condition on side k, which is equivalent to the monotonicity constraint $q^k_H \geq q^k_L$.

To solve the platform's optimization problem, we impose the following property on each single variable utility function:

Assumption 2. *$u^k_l(0) = 0$, $u^{k\prime}_l(q) > 0$, $u^{k\prime\prime}_l(q) \leq 0$ for each q, for each $k, l = A, B$.*

We focus on the standard case in which (IR^k_L) and (IC^k_H) are binding while (IR^k_H) is redundant.

Lemma 1. *Suppose that u^A and u^B are separable and there is no type reversal or type reversal with a positive sorting for a given side k. Then, under Assumptions 1 and 2, (IR^k_H) is redundant, (IR^k_L) and (IC^k_H) bind in the optimal mechanism, and (IC^k_L) is equivalent to $q^k_H \geq q^k_L$.*

Remark 1. *We can prove Lemma 1 under weaker assumptions such that it covers non-separable u^A and u^B. For the details, see the proof of Lemma 1 in the Appendix.*

Lemma 1 allows to use (IR^k_L) and (IC^k_H) to pin down the agents' payments on side k, and then we can write the expression for H-type's information rent as follows:

$$\Omega^k_H = (\theta^k_H - \theta^k_L)u^k_k(q^k_L) + \nu^l_H(\theta^k_{HH} - \theta^k_{LH})u^k_l(q^l_H) + \nu^l_L(\theta^k_{HL} - \theta^k_{LL})u^k_l(q^l_L).$$

Therefore, the platform's original problem is equivalent to maximizing the following objective subject to the monotonicity constraints $q^A_L \leq q^A_H$ and $q^B_L \leq q^B_H$:

$$\hat{\Pi}(\mathbf{q}) \equiv \Pi^{FB}(\mathbf{q}) - \nu^A_H\Omega^A_H - \nu^B_H\Omega^B_H. \tag{5}$$

Let $\hat{\mathbf{q}}$ denote the maximizer of $\hat{\Pi}$ when the monotonicity constraints are neglected. When we focus on side A, from the first-order conditions, we have:

$$\theta^A_H u^{A\prime}_A(\hat{q}^A_H) + \theta^B_{LH} u^{B\prime}_A(\hat{q}^A_H) = C^{A\prime}(\hat{q}^A_H); \tag{6}$$

$$\theta^{Av}_L u^{A\prime}_A(\hat{q}^A_L) + \theta^B_{LL} u^{B\prime}_A(\hat{q}^A_L) = C^{A\prime}(\hat{q}^A_L), \tag{7}$$

where $\theta^{Av}_L \equiv \theta^A_L - \left(\theta^A_H - \theta^A_L\right) v^A_H/\nu^A_L$ is the virtual valuation of an L-type of side A, which is smaller than θ^A_L under Assumption 1. Assume momentarily $\hat{q}^A_H \geq \hat{q}^A_L$, which implies $q^{A,SB}_H = \hat{q}^A_H$ and $q^{A,SB}_L = \hat{q}^A_L$.

To identify the distortions generated by asymmetric information, we compare the FOCs of the first-best (1) and (2) with these second-best FOCs (6) and (7). First, regarding the utility that the quality of side A generates to the same side, θ_L^A is replaced by the virtual valuation θ_L^{Av}, which is well-known from the price discrimination in one-sided market. Second, regarding the utility that the quality of side A generates to the other side, $\nu_H^B\theta_{HH}^B + \nu_L^B\theta_{LH}^B$ is replaced with θ_{LH}^B, and $\nu_H^B\theta_{HL}^B + \nu_L^B\theta_{LL}^B$ with θ_{LL}^B. This distortion is unique in a two-sided market and is interpreted as the Spence (1975) distortion in a two-sided market by Weyl (2010): the monopoly platform evaluates the externality to side B agents with the valuation of the marginal type (i.e., the L type) instead of using the average valuation of all agents. This occurs because the payment of type H on side B is determined not by (IR_H^B) but by (IC_H^B) (see Lemma 1). For this reason, even a distortion at the top arises whenever $\theta_{HH}^B \neq \theta_{LH}^B$; precisely, we have $q_H^{A,FB} \gtreqless q_H^{A,SB}$ if and only if $\theta_{HH}^B \gtreqless \theta_{LH}^B$. For instance, if an H type of side B obtains greater benefit than an L type of the same side B from interacting with an H type agent of side A, then it is optimal for the platform to introduce a downward distortion in q_H^A in order to extract more rent from the H type agents on side B.

To make a concrete comparison of $q_L^{A,FB}$ with $q_L^{A,SB}$, let us consider the specific case in which (4) holds. Then, we find $q_L^{A,FB} < q_L^{A,SB}$ if and only if

$$\theta_L^A - \theta_L^{Av} = \frac{\nu_H^A}{\nu_L^A}(\theta_H^A - \theta_L^A) < \beta_A^B \nu_H^B (\theta_{LL}^B - \theta_{HL}^B). \tag{8}$$

This means that there will be an upward distortion at the bottom on side A when there is type reversal with a strong positive sorting on side B (i.e., $\theta_{LL}^B - \theta_{HL}^B > 0$) such that inequality (8) holds.

In summary, we report that (i) no distortion at the top is not valid any more: there can be either an upward or a downward distortion at the top, (ii) the same is true for the distortion at the bottom though a downward distortion will be the case except for type reversal with a sufficiently large positive sorting.

Now we turn to the monotonicity constraint. If $\widehat{q}_H^A < \widehat{q}_L^A$, then $(\widehat{q}_H^A, \widehat{q}_L^A)$ cannot be the optimal schedule (that is, $(q_H^{A,SB}, q_L^{A,SB}) \neq (\widehat{q}_H^A, \widehat{q}_L^A)$) and the monotonicity constraint binds, which means a pooling contract ($q_H^{A,SB} = q_L^{A,SB}$) is required in the optimal mechanism. This case occurs if and only if the following condition holds:

$$(\theta_H^A - \theta_L^{Av}) u_A^{A\prime}(\widehat{q}_H^A) + (\theta_{LH}^B - \theta_{LL}^B) u_A^{B\prime}(\widehat{q}_H^A) < 0. \tag{9}$$

Again considering the special case in which (4) holds, this condition is equivalent to

$$\theta_H^A + \beta_A^B \theta_{LH}^B < \theta_L^{Av} + \beta_A^B \theta_{LL}^B, \tag{10}$$

which can be decomposed as follows:

$$\begin{aligned}
&\underbrace{\theta_H^A - \theta_L^A + \beta_A^B(\nu_H^B\theta_{HH}^B + \nu_L^B\theta_{LH}^B - \nu_H^B\theta_{HL}^B - \nu_L^B\theta_{LL}^B)}_{\text{First best term}} \\
< \quad & \underbrace{-\frac{\nu_H^A}{\nu_L^A}(\theta_H^A - \theta_L^A)}_{\text{distortion in one-side market}} \quad + \underbrace{\beta_A^B\nu_H^B\left[\left(\theta_{HH}^B - \theta_{LH}^B\right) - \left(\theta_{HL}^B - \theta_{LL}^B\right)\right]}_{\text{distortion due to two-sidedness}}. \qquad (11)
\end{aligned}$$

Suppose that there is no distortion due to the two-sidedness of the market. Then, a necessary condition for (11) to be satisfied is that the left hand side in (11) is negative, which means that the first-best schedule must be at least strictly decreasing. This is exactly what happens in one-sided market as in Guesnerie and Laffont (1984). We notice here, however, that the condition (11) can be satisfied even if the first-best schedule is rather increasing as long as the distortion from the two-sidedness is strong enough.[16] This result has one important implication: the distortion associated with two-sidedness expands the possibility of non-responsiveness.

Lastly, we briefly mention the case of pooling for the completeness of our analysis. For the pooling, we have $q_H^{A,SB} = q_L^{A,SB} \equiv q^{A,SB}$ and from (6) and (7) the optimal pooling schedule is characterized by

$$\theta_L^A u_A^{A\prime}(q) + \theta_L^B u_A^{B\prime}(q) = C^{A\prime}(q). \qquad (12)$$

We summarize thus far results as follows:

Proposition 2. *Suppose that u^A and u^B are separable and that for a given side A there is no type reversal or type reversal with a positive sorting. Then, under Assumptions 1 and 2, we find:*

(i) (Quality distortions) The second-best optimal mechanism is such that the quality schedule on side A is increasing ($q_H^{A,SB} \geq q_L^{A,SB}$) if and only if (9) is satisfied with the reverse inequality. For an increasing second-best schedule,

(a) Both an upward and a downward distortions are feasible at the top: $q_H^{A,FB} \gtreqless q_H^{A,SB}$ if and only if $\theta_{HH}^B \gtreqless \theta_{LH}^B$.

(b) The distortion at the bottom can be even upward: for instance, when $u_A^B(q) =$

[16] $\beta_A^B\nu_H^B\left(\theta_{HH}^B - \theta_{LH}^B\right)$ captures the Spence distortion in $q_H^{A,SB}$ and $\beta_A^B\nu_H^B(\theta_{HL}^B - \theta_{LL}^B)$ the Spence distortion in $q_L^{A,SB}$.

$\beta_A^B u_A^A(q)$ holds, $q_L^{A,SB} > q_L^{A,FB}$ if and only if

$$\beta_A^B \nu_H^B(\theta_{LL}^B - \theta_{HL}^B) > \theta_L^A - \theta_L^{Av} = \frac{\nu_H^A}{\nu_L^A}(\theta_H^A - \theta_L^A).$$

(ii) (Non-responsiveness) A pooling scheme on side A becomes optimal if (9) holds. Interestingly, the pooling can occur in the optimum even when the first-best quality schedule is increasing.

Note that pooling can occur only on the side which generates strong externalities because (11) will be never satisfied under Assumption 1 when β_A^B is close to zero.

The analysis of the second-best can be extended to some non-separable cases. In the standard case in which the monotonicity constraint is not binding, the analysis of the second-best is the same as that of the first best but for the fact that the valuations of the L-type replaced by the virtual ones:

$$\theta_{LH}^{jv} = \theta_{LH}^j - \frac{\nu_H^j}{\nu_L^j}(\theta_{HH}^j - \theta_{LH}^j); \quad \theta_{LL}^{jv} = \theta_{LL}^j - \frac{\nu_H^j}{\nu_L^j}(\theta_{HL}^j - \theta_{LL}^j).$$

Hence, for the specific non-separable u_A, u_B mentioned at the end of Section 3, a small degree of complementarity (substitution) on both sides increases (decreases) each second-best quality compared to the separable case.

5 Welfare analysis

In this section, we study the welfare effect of introducing PD on side B in the separable case and then apply our result to the net neutrality debate.

5.1 Welfare analysis of price discrimination in the separable case

We consider the setup in which Lemma 1 holds. Recall that the second-best quality schedule $(q_L^{A,SB}, q_H^{A,SB})$ is determined independently of PD on side B. This implies that, for the welfare analysis, we are allowed to consider only the part of the welfare that is affected by $\left(q_H^B, q_L^B\right)$, which is given by

$$\begin{aligned} W(q_H^B, q_L^B) \equiv\ & v_H^B \left\{\theta_H^B u_B^B(q_H^B) + \left(v_H^A\theta_{HH}^A + v_L^A\theta_{LH}^A\right) u_B^A(q_H^B) - C^B(q_H^B)\right\} \\ & + v_L^B \left\{\theta_L^B u_B^B(q_L^B) + \left(v_H^A\theta_{HL}^A + v_L^A\theta_{LL}^A\right) u_B^A(q_L^B) - C^B(q_L^B)\right\}. \end{aligned}$$

We assume that when PD is allowed, the platform chooses a strictly increasing schedule on side B, $q_H^{B,SB} > q_L^{B,SB}$. When PD is prohibited on side B, then the platform chooses the pooling schedule $q_H^B = q_L^B = q^{B,SB}$, which is determined by

$$\theta_L^B u_B^{B\prime}(q) + \theta_L^A u_B^{A\prime}(q) = C^{B\prime}(q).$$

Hence, we have the following relationship:

$$q_H^{B,SB} > q^{B,SB} > q_L^{B,SB}.$$

For simplicity we focus on the case of linear relationship $u_B^A(q) = \beta_B^A u_B^B(q)$ with $\beta_B^A > 0$ and let $u_B^B(q) = u(q)$ and $\beta_B^A = \beta$.

In order to isolate the effects of the PD on the same side from the effects on the other side, we conduct the welfare analysis in two steps. First, we analyze the effect on the total welfare in the hypothetical and extreme case of $\theta_{HH}^A = \theta_{HL}^A = \theta_{LH}^A = \theta_{LL}^A$. Second, we relax this restriction and analyze the effects of PD on the welfare of side A. The first step allows us to rediscover the welfare effect of PD in one-sided market; the second grants us to isolate the welfare effect on the other side, which is distinctive in a two-sided market.

5.1.1 Welfare effect on the own side

Suppose $\theta_{HH}^A = \theta_{HL}^A = \theta_{LH}^A = \theta_{LL}^A = \theta^A$. Then, we have:

$$W(q_H^B, q_L^B) = v_H^B \left\{ (\theta_H^B + \beta\theta^A) u(q_H^B) - C^B(q_H^B) \right\} + v_L^B \left\{ (\theta_L^B + \beta\theta^A) u(q_L^B) - C^B(q_L^B) \right\}$$

In this case, the welfare effect of PD is the same as the effect in a one-sided market in which an H type's marginal valuation is given by $\left(\theta_H^B + \beta\theta^A\right) u'$ and an L type's marginal valuation is given by $\left(\theta_L^B + \beta\theta^A\right) u'$. Under PD, we have no distortion at the top, $q_H^{B,SB} = q_H^{B,FB}$, and a downward distortion at the bottom, $q_L^{B,SB} < q_L^{B,FB}$. Specifically, when PD is introduced, the welfare for the H type increases because PD will remove the distortion at the top under the pooling scheme (in math, $q^{B,SB} \neq q_H^{B,FB}$ $\Rightarrow q_H^{B,SB} = q_H^{B,FB}$), but the welfare for the L type drops because PD will lead to a distortion at the bottom (in math, $q^{B,SB} = q_L^{B,FB} \Rightarrow q_L^{B,SB} < q_L^{B,FB}$).

Hence, the net effect of PD depends on whether the increase in welfare of the H types is smaller or larger than the reduction in welfare of the L types. PD would increase welfare if v_L^B is close to one because then the virtual valuation $\theta_L^{Bv} \equiv \theta_L^B - \left(\theta_H^B - \theta_L^B\right) v_H^B / \nu_L^B$ is close to θ_L^B and hence the welfare under PD is close to the first-best welfare.

To further pin down a condition under which PD reduces welfare, suppose that v_L^B

is such that $\theta_L^{Bv} + \beta\theta^A = 0$, which implies $q_L^{B,SB} = 0$ and this restriction is equivalent to $v_L^B = \frac{\theta_H^B - \theta_L^B}{\theta_L^B + \beta\theta^A} v_H^B$. Then, PD will lower welfare if the following inequality holds

$$\begin{aligned}
& \underbrace{v_L^B \left[(\theta_L^B + \beta\theta^A) u(q^{B,SB}) - C^B(q^{B,SB}) \right]}_{\text{welfare decrease from the L types}} \\
= \;& \frac{\theta_H^B - \theta_L^B}{\theta_L^B + \beta\theta^A} v_H^B \left[(\theta_L^B + \beta\theta^A) u(q^{B,SB}) - C^B(q^{B,SB}) \right] \qquad (13) \\
> \;& \underbrace{v_H^B \Big[(\theta_H^B + \beta\theta^A) u(q_H^{B,SB}) - C^B(q_H^{B,SB}) - \big((\theta_H^B + \beta\theta^A) u(q^{B,SB}) - C^B(q^{B,SB}) \big) \Big]}_{\text{welfare increase from the H types}}.
\end{aligned}$$

This inequality holds if θ_H^B is slightly larger than θ_L^B.[17] Then because the welfare enhancement from the H types is negligible relative to the welfare reduction from the L types who are assigned with $q_L^{B,SB} = 0$ under PD, the total welfare is lower under PD.

5.1.2 Welfare effect on the other side

Let us now study how the PD on side B affects the welfare of side A when the restriction of $\theta_{HH}^A = \theta_{HL}^A = \theta_{LH}^A = \theta_{LL}^A$ is relaxed. First, let us define $W^A(q_H^B, q_L^B)$ and $\Delta W^A(q_H^B, q_L^B)$ as follows:

$$\begin{aligned}
W^A(q_H^B, q_L^B) &\equiv v_H^B \left(v_H^A \theta_{HH}^A + v_L^A \theta_{LH}^A \right) u_B^A(q_H^B) + v_L^B \left(v_H^A \theta_{HL}^A + v_L^A \theta_{LL}^A \right) u_B^A(q_L^B); \\
\Delta W^A &\equiv W^A(q_H^{B,SB}, q_L^{B,SB}) - W^A(q^{B,SB}, q^{B,SB}).
\end{aligned}$$

$W^A(q_H^B, q_L^B)$ measures the welfare for the agents on side A that is associated with the quality schedule on side B; ΔW^A denotes the difference in this measure as PD is introduced onto side B. Hence,

$$\frac{\Delta W^A}{\beta} = v_H^B \left(v_H^A \theta_{HH}^A + v_L^A \theta_{LH}^A \right) \Big[u(q_H^{B,SB}) - u(q^{B,SB}) \Big] - v_L^B \left(v_H^A \theta_{HL}^A + v_L^A \theta_{LL}^A \right) \Big[u(q^{B,SB}) - u(q_L^{B,SB}) \Big].$$

The parameters $(\theta_{HH}^A, \theta_{HL}^A, \theta_{LH}^A, \theta_{LL}^A)$ affect ΔW^A directly but also indirectly through the changes in the quality variables on side B. The first-order conditions for the quality variables on side B are analogous to those of side A, (6), (7), and (12), from which we can check that $\theta_{HH}^A, \theta_{HL}^A$ do not appear but $\theta_{LH}^A, \theta_{LL}^A$ do. This implies that only direct

[17] In details, the inequality relationship in (13) becomes an equality if $\theta_H^B = \theta_L^B$, and the derivative of the left hand side with respect to θ_H^B is greater than the derivative of the right hand side with respect to θ_H^B when $\theta_H^B = \theta_L^B$.

effects matter for θ^A_{HH} and θ^A_{HL}. Hence, we find that ΔW^A increases with θ^A_{HH} and decreases with θ^A_{HL} because of $u(q_H^{B,SB}) - u(q^{B,SB}) > 0$ and $u(q^{B,SB}) - u(q_L^{B,SB}) < 0$ respectively.

On the contrary, because of the indirect effects, it is not straightforward to see how ΔW^A changes with θ^A_{LH} and with θ^A_{LL}. As we show in the proof of Proposition 3 in the Appendix, we find it useful to define a function $f(q)$ such that

$$f(q) = \frac{u'(q)}{\frac{d}{dq}\left(\frac{C'(q)}{u'(q)}\right)} = \frac{(u'(q))^2}{C''(q) - \frac{C'(q)}{u'(q)}u''(q)}.$$

Then, we can show that ΔW^A increases with θ^A_{LH} if $v^A_H\theta^A_{HH} + v^A_L\theta^A_{LH} \geq v^A_H\theta^A_{HL} + v^A_L\theta^A_{LL}$ and $f(q)$ is weakly increasing.[18] Similarly, the effect of θ^A_{LL} on ΔW^A is identified.

We summarize thus far findings as below.

Proposition 3. *Suppose that the platform introduces PD on side B. Consider* $u^A_B(q) = \beta u^B_B(q)$ *with* $\beta > 0$ *and let* $u^B_B(q) = u(q)$*. Recall* ΔW^A *measure the change in welfare of side A from PD on side B. Then, we find*

- *(i)* ΔW^A *strictly increases with* θ^A_{HH} *and strictly decreases with* θ^A_{HL}.
- *(ii) If* $v^A_H\theta^A_{HH} + v^A_L\theta^A_{LH} \geq v^A_H\theta^A_{HL} + v^A_L\theta^A_{LL}$ *and* $f(q)$ *is weakly increasing,* ΔW^A *strictly increases with* θ^A_{LH} *and strictly decreases with* θ^A_{LL}.

When the conditions of Proposition 3(ii) hold, ΔW^A increases with $\theta^A_{HH} - \theta^A_{HL}$ and $\theta^A_{LH} - \theta^A_{LL}$. The insight is clear. For instance, $\theta^A_{HH} - \theta^A_{HL}$ measures the difference in the externalities onto the H and L type agents of side A from an H type agent of side B. Therefore, introducing PD on side B increases (decreases) welfare of side A if H type agents of side B generate more (less) positive externalities to the agents of side A than L type agents of side B do.

As the conditions in Proposition 3(ii) are largely sufficient, we consider a CARA utility $u(q) = 1 - e^{-\alpha q}$ and a quadratic cost $C(q) = \frac{1}{2}q^2$ in order to provide a tighter tangible condition. Then, we have $f(q) = \alpha^2 \frac{e^{-2q\alpha}}{1+q\alpha}$, which is decreasing. However, we find that if $v^A_H\theta^A_{HH} + v^A_L\theta^A_{LH} \geq v^A_H\theta^A_{HL} + v^A_L\theta^A_{LL}$, ΔW^A strictly increases with θ^A_{LH} (and strictly decreases with θ^A_{LL}) if and only if the risk aversion is small enough, which is precisely given by $\frac{v^A_L}{\beta\left[v^A_H\theta^A_{HL} + v^A_L\theta^A_{LL}\right]} > 3\alpha^2$.

[18] For instance, $f(q)$ is increasing if u is linear and C' is concave.

5.2 Application to the net neutrality debate

Now let us apply our results to the net neutrality debate. In the debate, the monopoly platform is an Internet service provider (ISP) which mediates interactions between consumers on side A and content providers on side B. q_i^A represents residential Internet quality sold to consumer i, which depends on the download speed and capacity. ISPs practice price discrimination on the consumer side and this has never made controversy in the debate. The key issue is whether the regulator should allow ISPs to practice price discrimination on the content side.

So, let us focus on the welfare consequences of introducing PD on the content side. In this application, we consider a specific case of $u_B^B(q) = q$, $u_B^A(q) = \beta_B^A q$ with $\beta_B^A > 0$ and $C^B(q) = \frac{1}{2}q^2$. Then, we obtain

$$\begin{aligned} q_H^{B,SB} &= \theta_H^B + \beta_B^A \theta_{LH}^A \\ q_L^{B,SB} &= \theta_L^{Bv} + \beta_B^A \theta_{LL}^A \\ q^{B,SB} &= v_H^B q_H^{B,SB} + v_L^B q_L^{B,SB} = \theta_L^B + \beta_B^A \left(v_H^B \theta_{LH}^A + v_L^B \theta_{LL}^A \right). \end{aligned}$$

Furthermore, we can readily find that

$$q_H^{B,FB} - q^{B,FB} = \theta_H^B - \theta_L^B + \beta_B^A v_H^A \left(\theta_{HH}^A - \theta_{HL}^A \right) + \beta_B^A v_L^A \left(\theta_{LH}^A - \theta_{LL}^A \right) > 0; \quad (14)$$

$$q_H^{B,SB} - q^{B,SB} = \theta_H^B - \theta_L^B + \beta_B^A v_L^B \left(\theta_{LH}^A - \theta_{LL}^A \right) > 0. \quad (15)$$

We can reasonably assume that H type content such as real-time video streaming service is more sensitive to traffic delay than L type content such as text-based websites. Therefore, it is natural to assume that in the first-best case a benevolent social planner prefers providing a higher quality to H type content than to L type content so that (14) is satisfied. We also hear that most ISPs prefer to practice PD on the content side, which implies that (15) should be satisfied as well. In this circumstance, one should be interested in studying the conditions under which PD on the content side reduces overall welfare while (14) and (15) are satisfied.

The change in welfare of side A as we move from no PD to allowing PD on the content side is derived as

$$\Delta W^A = \beta_B^A v_H^B \left[\theta_H^B - \theta_L^B + \beta_B^A v_L^B \left(\theta_{LH}^A - \theta_{LL}^A \right) \right] \left[v_H^A \left(\theta_{HH}^A - \theta_{HL}^A \right) + v_L^A \left(\theta_{LH}^A - \theta_{LL}^A \right) \right].$$

From (15), we know $q_H^{B,SB} - q^{B,SB} = \theta_H^B - \theta_L^B + \beta_B^A v_L^B \left(\theta_{LH}^A - \theta_{LL}^A \right) > 0$. Hence, the sign of ΔW^A has the same sign of the term $\left[v_H^A \left(\theta_{HH}^A - \theta_{HL}^A \right) + v_L^A \left(\theta_{LH}^A - \theta_{LL}^A \right) \right]$, which is the difference between the externalities on consumers generated by an H type content and those generated by an L type content. This is altogether consistent with Proposition 3.

The welfare of side B that depends on $\left(q_H^B, q_L^B\right)$ is

$$W^B(q_H^B, q_L^B) = v_H^B \left[\theta_H^B u_B^B(q_H^B) - C^B(q_H^B)\right] + v_L^B \left[\theta_L^B u_B^B(q_L^B) - C^B(q_L^B)\right].$$

Hence, the change in the welfare of side B is

$$\begin{aligned}\Delta W^B &\equiv W_B^B(q_H^{B,SB}, q_L^{B,SB}) - W_B^B(q^{B,SB}, q^{B,SB}) \\ &= \frac{v_H^B}{2v_L^B}\left[\theta_H^B - \theta_L^B + \beta_B^A v_L^B \left(\theta_{LH}^A - \theta_{LL}^A\right)\right]\left[\left(2v_L^B - 1\right)\left(\theta_H^B - \theta_L^B\right) - v_L^B \left(\theta_{LH}^A - \theta_{LL}^A\right)\right].\end{aligned}$$

As before, ΔW^B has the same sign as $\left[\left(2v_L^B - 1\right)\left(\theta_H^B - \theta_L^B\right) - v_L^B(\theta_{LH}^A - \theta_{LL}^A)\right]$, which incorporate two opposite effects. On the one hand, an increase in q_H^B by $q_H^{B,SB} - q^{B,SB}$ and the corresponding decrease in q_L^B by $v_H^B(q_H^{B,SB} - q^{B,SB})/v_L^B$ raises overall surplus by $v_H^B(\theta_H^B - \theta_L^B)(q_H^{B,SB} - q^{B,SB})$. On the other hand, the same change increases overall cost by $v_H^B(q_H^{B,SB} - q^{B,SB})^2/2v_L^B$. To understand why this term decreases with v_L^B, suppose $\theta_{LH}^A = \theta_{LL}^A$. Then, as v_L^B changes, $q_H^{B,SB}$ and $q^{B,SB}$ are constant but $q_L^{B,SB}$ increases with v_L^B because the downward distortion in $q_L^{B,SB}$ lessens as v_L^B goes up. As a result, the quality spread diminishes as v_L^B enlarges, which in turn reduces the overall cost under PD because the cost function is convex. For instance, when $\theta_{LH}^A = \theta_{LL}^A$, $\Delta W^B > 0$ if and only if $v_L^B > 1/2$, which is consistent with the result in Section 5.1.1.

Taken together, the PD on the content side reduces total welfare from both sides if the following inequality holds:

$$\beta_B^A \left[v_H^A \left(\theta_{HH}^A - \theta_{HL}^A\right) + v_L^A \left(\theta_{LH}^A - \theta_{LL}^A\right)\right] + \frac{\left(2v_L^B - 1\right)\left(\theta_H^B - \theta_L^B\right) - v_L^B(\theta_{LH}^A - \theta_{LL}^A)}{2v_L^B} < 0.$$

To isolate different effects, suppose first $\theta_{HH}^A - \theta_{HL}^A = \theta_{LH}^A - \theta_{LL}^A = 0$. Therefore, $\Delta W^A = 0$. Then, (14) and (15) are trivially satisfied and the PD on the content side reduces welfare if $1/2 > v_L^B$. This is because when $1/2 > v_L^B$, PD introduces too much downward distortion in $q_L^{B,SB}$, which is a standard logic against PD in one-sided market.

Now in order to shut down the above effect, consider the case of $1/2 = v_L^B$. In addition, assume $\theta_{HL}^A - \theta_{HH}^A = \theta_{LL}^A - \theta_{LH}^A = \delta$. Then, the PD on the content side decreases total welfare if and only if $\left(\beta_B^A - 1/2\right)\delta > 0$. For instance, if $\beta_B^A > 1/2$, then the condition is satisfied whenever $\delta > 0$. In this case, L type content generates more positive externalities to consumers than H type content does and hence allowing PD on the content side decreases welfare on the consumer side. Then, (14) and (15) are satisfied if

$$\theta_H^B - \theta_L^B > \beta_B^A \delta.$$

Hence, the gap in the externalities δ should be mild enough such that it is optimal to assign a higher quality to H type content in the first best.

In summary, when applied to the net neutrality debate, our results generate a new rationale for maintaining net neutrality regulation: when L type content generates more positive externalities to consumers than H type content, it is socially desirable to maintain net neutrality. In real world, such situation can arise if H type content has more market power and hence tends to use more micro-payments instead of ad-based business model than L type content does such that the former extracts much more surplus from consumers than the latter.

Lastly, we connect our general model of two-sided PD with Choi-Jeon-Kim (2015) which show how business models on the content side would affect the overall welfare with and without PD on the content side. One notable contribution of this article is to find the results of Choi-Jeon-Kim as a special case of our more general consideration.

Remark 2. *Choi-Jeon-Kim (2015) consider a two-sided market with heterogeneous content providers but homogeneous consumers. In this setup, they introduce a parameter $\alpha \in [0,1]$ that represents division of surplus between content side and consumer side. Specifically, their model can be captured as a special case of our model as follows:*

$$\theta_H^B = \alpha S_H, \quad \theta_L^B = \alpha S_L, \quad \theta_{HH}^A = \theta_{LH}^A = (1-\alpha) S_H, \quad \theta_{HL}^A = \theta_{LL}^A = (1-\alpha) S_L, \quad \beta_B^A = 1$$

where $S_H = \theta_{HH}^A + \theta_H^B > S_L = \theta_{HL}^A + \theta_L^B > 0$ holds. Then, (14) and (15) are trivially satisfied. They find that even if welfare is higher with PD for $\alpha = 0$ and $\alpha = 1$,[19] PD can reduce welfare for an intermediate range of α. However, in their model, $\Delta W^A > 0$ for $\alpha < 1$ as $\theta_{HH}^A - \theta_{HL}^A = \theta_{LH}^A - \theta_{LL}^A = (1-\alpha)(S_H - S_L) > 0$. Therefore, introducing PD on the content side always increases welfare on the consumer side in their model, which implies that our general model captures some effects that theirs can not.

[19] They assume that without PD, the monopoly platform excludes L type content for $\alpha = 1$ and this is why welfare is higher with PD for $\alpha = 1$.

6 Implementable allocations in the non-separable case

In this section we consider the non-separable case and characterize the implementable allocations on side A with type reversal given an arbitrary quality schedule on side B. In the next section we provide an application of our result to an advertising platform.

By summing the incentive constraints in the second-best problem, (IC_H^k) and (IC_L^k), and considering $k = A$, we find the implementability condition on side A as follows:

$$\begin{aligned}\Phi^A := \ & \nu_H^B \left[\theta_{HH}^A - \theta_{LH}^A\right] \left[u^A(q_H^A, q_H^B) - u^A(q_L^A, q_H^B)\right] \\ & + \nu_L^B \left[\theta_{HL}^A - \theta_{LL}^A\right] \left[u^A(q_H^A, q_L^B) - u^A(q_L^A, q_L^B)\right] \geq 0.\end{aligned} \tag{16}$$

Let us take (q_H^B, q_L^B) as given and assume a type reversal of positive sorting on side A, i.e., $(\theta_{HH}^A - \theta_{LH}^A) > 0 > (\theta_{HL}^A - \theta_{LL}^A)$. Later, we briefly describe how our results extend to the case of negative sorting. Let F denote the set of (q_H^A, q_L^A) satisfying the implementability condition on side A (16) for a given pair of (q_H^B, q_L^B).[20] In order to describe F, we let

$$\begin{cases} M \text{ (from ``monotonic'')} & \text{denote the set of } (q_H^A, q_L^A) \text{ such that } q_H^A > q_L^A \geq 0; \\ N \text{ (from ``non-monotonic'')} & \text{denote the set of } (q_H^A, q_L^A) \text{ such that } 0 \leq q_H^A < q_L^A; \\ D \text{ (from ``diagonal'')} & \text{denote the set of } (q_H^A, q_L^A) \text{ such that } 0 \leq q_H^A = q_L^A. \end{cases}$$

Since $\Phi^A = 0$ at each point satisfying $q_H^A = q_L^A$, it is obvious that $D \subseteq F$. Moreover it is immediate to identify F if Φ^A is strictly monotonic with respect to q_H^A. Precisely, if Φ^A is strictly increasing in q_H^A then $F = M \cup D$; if Φ^A is strictly decreasing with respect to q_H^A, then $F = N \cup D$.

If u^A is separable, then

$$\Phi^A = \left(\theta_H^A - \theta_L^A\right)\left(u_A^A(q_H^A) - u_A^A(q_L^A)\right),$$

which is strictly increasing in q_H^A by Assumption 1, and thus $F = M \cup D$. This result does not depend on whether the type reversal occurs with a positive sorting or a negative sorting.

Considering a positive sorting and complementarity between the qualities, we have

$$\frac{\partial \Phi^A}{\partial q_H^A} = \nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) u_1^A(q_H^A, q_H^B) + \nu_L^B \left(\theta_{HL}^A - \theta_{LL}^A\right) u_1^A(q_H^A, q_L^B).$$

[20]Hence, F depends on (q_H^B, q_L^B) though our notation does not make it explicit.

Therefore, Φ^A is strictly increasing in q_H^A if $q_H^B \geq q_L^B$ (since in this case $u_1^A(q_H^A, q_H^B) \geq u_1^A(q_H^A, q_L^B)$ and Assumption 1 holds) or if $q_H^B < q_L^B$ and $|\nu_L^B \left(\theta_{HL}^A - \theta_{LL}^A\right)|$ is close to zero and/or the effect of complementarity is small. Conversely, if $q_H^B < q_L^B$, $|\nu_L^B \left(\theta_{HL}^A - \theta_{LL}^A\right)|$ is close to $\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right)$ (i.e., $\theta_H^A - \theta_L^A$ is close to zero) and the effect of complementarity is strong, then Φ^A is strictly decreasing with respect to q_H^A.

In the case of substitutes, we obtain opposite results: Φ^A is strictly increasing in q_H^A if $q_H^B \leq q_L^B$ (again $u_1^A(q_H^A, q_H^B) \geq u_1^A(q_H^A, q_L^B)$ and Assumption 1 holds), or if $q_H^B > q_L^B$ and $|\nu_L^B \left(\theta_{HL}^A - \theta_{LL}^A\right)|$ is close to zero and/or the effect of substitution is small. Conversely, if $q_H^B > q_L^B$, $|\nu_L^B \left(\theta_{HL}^A - \theta_{LL}^A\right)|$ is close to $\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right)$ and the effect of substitution is strong, then Φ^A is strictly decreasing with respect to q_H^A.

The case of a negative sorting (i.e., $\theta_{HL}^A - \theta_{LL}^A > 0 > \theta_{HH}^A - \theta_{LH}^A$) is symmetric to that of positive sorting. The following proposition summarizes our results.

Proposition 4. *Suppose that Assumption 1 holds, there is type reversal on side A, and quality schedule on side B $\left(q_H^B, q_L^B\right)$ is given.*

(i) If u^A is separable, the implementable set of (q_H^A, q_L^A) equals the set of increasing quality schedules (i.e., $F = M \cup D$) regardless of type reversal.

(ii) Suppose that on side A, qualities are complements (resp. substitutes) and type reversal occurs with a positive (resp. negative) sorting.

(a) The implementable set of (q_H^A, q_L^A) is equal to the set of increasing quality schedules (i.e., $F = M \cup D$) if $q_H^B \geq q_L^B$.

(b) The implementable set of (q_H^A, q_L^A) is equal to the set of decreasing (or constant) quality schedules $F = N \cup D$ if $q_H^B < q_L^B$, the complementarity (resp. the substitution) is sufficiently strong and $\theta_H^A - \theta_L^A$ is close to zero.

(iii) Suppose that qualities are substitutes (resp. complements) and type reversal occurs with a positive (resp. negative) sorting. Then, the same statements as above in (ii) can be made for (a) if $q_H^B \leq q_L^B$ (b) if $q_H^B > q_L^B$, the substitution (resp. the complementarity) is sufficiently strong and $\theta_H^A - \theta_L^A$ is close to zero.

Proposition 4 identifies when implementing a decreasing schedule on side A requires an increasing (or a decreasing) schedule on side B. Let us provide the intuition. The implementability condition (16) means that given the quality schedule on side B, when $q_H^A < q_L^A$, the L type's utility gain from receiving q_L^A instead of q_H^A must be larger than the H type's gain from doing the same. Consequently, absent PD on side B, a decreasing schedule (i.e., $q_H^A < q_L^A$) is not implementable by the definition of the H and L types. However, this is no longer the case if we introduce PD on side B under type reversal on side A. Consider type reversal with a positive sorting and suppose that the qualities are complements. If $q_H^B < q_L^B$, a decreasing schedule (i.e., $q_H^A < q_L^A$) becomes now implementable as an L type's utility gain can be larger than an H type's one when the quality increases from q_H^A to q_L^A. This is because an L type enjoys a high marginal utility from interacting with an L type when the qualities are complements and type reversal with a positive sorting arises ($\theta_{HL}^A < \theta_{LL}^A$). Symmetrically, if qualities are substitutes and there is type reversal with a positive sorting, implementing a decreasing schedule on side A requires $q_H^B > q_L^B$.

The discussions of the case with no type reversal and the above proposition give us sufficient conditions for F to equal the set of the increasing schedules:

Corollary 1. *The implementable set of (q_H^A, q_L^A) is equal to the set of increasing quality schedules if at least one of the following conditions is satisfied:*

(i) There is no type reversal on side A;

(ii) u^A is separable;

(iii) There is no PD on side B (i.e., $q_H^B = q_L^B$).

7 Application to an advertising platform

We here provide an application to demonstrate how our insight obtained in the previous section plays out in a more realistic two-sided market of a media platform that mediates content users and advertisers via content. This application also allows us to answer the question of when PD on one side complements or substitutes for PD on the other side. We consider private information on both sides.

7.1 The model

There is a mass one of consumers on side A and a mass one of advertisers on side B. Agents on each side have two different types, H and L. To reduce the number of parameters, we consider the equal population of each type on both sides, i.e., $\nu_H^A = \nu_H^B = 1/2$. On side A the platform offers a menu of quality-price pairs (q_H, p_H^A) and (q_L, p_L^A), with $(q_H, q_L) \in \{0,1\}^2$ where '1' means high quality or no nuisance from advertising and '0' means low quality or nuisance from advertising. On side B, we consider that each advertiser has multiple products to advertise and hence demands for multiple ad slots per consumer. Therefore, the platform offers a menu of advertising level (per consumer) and price: (a_H, p_H^B) and (a_L, p_L^B), with $\{a_H, a_L\} \in \mathbb{R}_+^2$.[21] Each consumer earns a constant utility $u_0 > 0$ from consuming the content offered by the platform if he does not receive any advertising. Consumer i suffers disutility from advertiser j's ads which is given by $\alpha_{ij}\psi(a_j)$ with $\alpha_{ij} > 0$ where we assume $\psi(\cdot)\,(\geq 0)$ is increasing and convex. Then, consumer i's gross utility when $q_i = 0$ is given as[22]

$$\begin{cases} u_0 - \frac{1}{2}\alpha_{HH}\psi(a_H) - \frac{1}{2}\alpha_{HL}\psi(a_L), & \text{if } \theta_i^A = H; \\ u_0 - \frac{1}{2}\alpha_{LH}\psi(a_H) - \frac{1}{2}\alpha_{LL}\psi(a_L), & \text{if } \theta_i^A = L. \end{cases}$$

In a similar manner let $\beta_{ji}R(a_j)$ with $\beta_{ji} > 0$ represent the revenue that advertiser j earns from consumer i when $q_i = 0$ where $R(\cdot)\,(\geq 0)$ is increasing and concave. Then, advertiser j's expected revenue from joining the platform is given by

$$\begin{cases} \frac{1}{2}\beta_{HH}R(a_H) + \frac{1}{2}\beta_{HL}R(a_H), & \text{if } \theta_j^B = H; \\ \frac{1}{2}\beta_{LH}R(a_L) + \frac{1}{2}\beta_{LL}R(a_L), & \text{if } \theta_j^B = L. \end{cases}$$

Then we impose the following assumptions on the parameters for two-sided interactions.

Assumption 3.

$$\begin{cases} (i)\ \alpha_{HH} + \alpha_{HL} > \alpha_{LH} + \alpha_{LL} \\ (ii)\ \alpha_{HH} < \alpha_{LH}, \quad \alpha_{HL} > \alpha_{LL} \\ (iii)\ \beta_{HH} > \beta_{LH}, \quad \beta_{HL} > \beta_{LL} \\ (iv)\ \beta_{HH} > \beta_{HL}, \quad \beta_{LH} > \beta_{LL}. \end{cases}$$

The first inequality (i) means that an H type consumer suffers more from nuisance than an L type in expected terms, which is equivalent to Assumption 1 applied to side A. The two inequalities in the second line (ii) introduce type reversal on side A. Conditional on receiving the ads from H type advertisers, an H type consumer's nuisance is smaller than an L type consumer's nuisance. Against L type advertisers, by contrast, the opposite holds. The inequalities in the third line (iii) means that an H type advertiser generates more revenue than an L type no matter what the consumer type they interact with: in other words, there is no type reversal on the advertiser side.

[21] Targeted advertising is not considered. In the case of targeted advertising, the mechanism on the advertising side would be (a_H^H, a_H^L, p_H^B) and (a_L^H, a_L^L, p_L^B) where for instance a_H^H $\left(a_H^L\right)$ refers to the amount of advertising per H (L) type consumer done by an H type advertisier.

[22] Alternatively, we can consider that for instance, the disutility of a H-type equals $\psi(\alpha_{HH}\frac{a_H}{2} + \alpha_{HL}\frac{a_L}{2})$ but the our results are not affected.

Hence, this assumption is stronger than Assumption 1 applied to side B. Lastly, (iv) means that an H type consumer is more valuable than an L type consumer in terms of advertising revenue for both types of advertisers.

In terms of the taxonomy introduced in the canonical model, the type reversal on side A is of a negative sorting because we have $\theta^A_{HH} - \theta^A_{LH} = \alpha_{HH} - \alpha_{LH} < 0$ and $\theta^A_{HL} - \theta^A_{LL} = \alpha_{HL} - \alpha_{LL} > 0$. In addition, the two screening instruments (q, a) are complements on side A: from $u^A(q,a) = -\psi(a) \cdot (1-q)$, we have $u^A_2(1,a) - u^A_2(0,a) = \psi'(a) > 0$.[23]

We assume that the platform is not viable without selling advertising, which means $(q_H, q_L) = (1,1)$ is never optimal. In what follows, we characterize the optimal contracts for the linear specification of $\psi(a) = R(a) = a$ for simplicity. We restrict our attention to non-negative consumer prices of $p^A_H \geq 0,\ p^A_L \geq 0$ because a negative price may induce consumers to take the money and run without consumption.

7.2 The optimal mechanism

Consider the general case in which the platform can propose a menu (including a pooling contract) on each side. On the advertising side B, we have $a_H \geq a_L$ from the implementability condition; in addition, the binding IR^B_L and IC^B_H imply

$$p^B_L = \frac{1}{2}\left(\beta_{LH}(1-q_H) + \beta_{LL}(1-q_L)\right) a_L, \tag{17}$$

$$p^B_H = \frac{1}{2}\left(\beta_{HH}(1-q_H) + \beta_{HL}(1-q_L)\right) a_H - \frac{1}{2}\left((\beta_{HH} - \beta_{LH})(1-q_H) + (\beta_{HL} - \beta_{LL})(1-q_L)\right) a_L. \tag{18}$$

On side A, we have the following four constraints:

$$\begin{aligned}
(\text{IC}^A_L) \qquad & u_0 - \frac{1}{2}\alpha_{LL}(1-q_L)a_L - \frac{1}{2}\alpha_{LH}(1-q_L)a_H - p^A_L \\
\geq\ & u_0 - \frac{1}{2}\alpha_{LL}(1-q_H)a_L - \frac{1}{2}\alpha_{LH}(1-q_H)a_H - p^A_H; \\
(\text{IC}^A_H) \qquad & u_0 - \frac{1}{2}\alpha_{HL}(1-q_H)a_L - \frac{1}{2}\alpha_{HH}(1-q_H)a_H - p^A_H \\
\geq\ & u_0 - \frac{1}{2}\alpha_{HL}(1-q_L)a_L - \frac{1}{2}\alpha_{HH}(1-q_L)a_H - p^A_L; \\
(\text{IR}^A_L) \qquad & u_0 - \frac{1}{2}\alpha_{LL}(1-q_L)a_L - \frac{1}{2}\alpha_{LH}(1-q_L)a_H - p^A_L \geq 0; \\
(\text{IR}^A_H) \qquad & u_0 - \frac{1}{2}\alpha_{HL}(1-q_H)a_L - \frac{1}{2}\alpha_{HH}(1-q_H)a_H - p^A_H \geq 0.
\end{aligned}$$

[23] $u^B(\cdot)$ can be written as follows: $u^B(q,a) = R(a) \cdot (1-q)$.

Adding IC_L^A to IC_H^A leads to the inequality

$$(q_H - q_L)(a_L - \rho a_H) \geq 0, \tag{19}$$

where $\rho \equiv \frac{\alpha_{LH}-\alpha_{HH}}{\alpha_{HL}-\alpha_{LL}} \in (0,1)$. Given a negative sorting and complementarity between qualities, according to Proposition 4, implementing a decreasing schedule on side A ($q_H \leq q_L$) requires an increasing schedule $a_H > a_L$ on side B (in fact, a sufficiently increasing schedule in this application, i.e., $\rho a_H > a_L$).

Suppose that IR_L^A binds at $q_L = 0$ which pins down p_L^A equal to $p_L^A = u_0 - \frac{1}{2}\alpha_{LL}a_L - \frac{1}{2}\alpha_{LH}a_H$. As $p_L^A \geq 0$ must hold, $\alpha_{LL}a_L + \alpha_{LH}a_H \leq 2u_0$ must be satisfied. It means that the upper limit of advertising levels consistent with the binding IR_L^A, $p_L^A \geq 0$ and the implementability condition $a_H \geq a_L$ is represented by the line EA in Figure 1. Similarly, the binding IR_H^A at $q_H = 0$ with $p_H^A \geq 0$ requires $\alpha_{HL}a_L + \alpha_{HH}a_H \leq 2u_0$; the corresponding upper limit of advertising levels is represented by the line CD in Figure 1. Type reversal with a negative sorting implies that EA crosses CD from above as a_H increases.

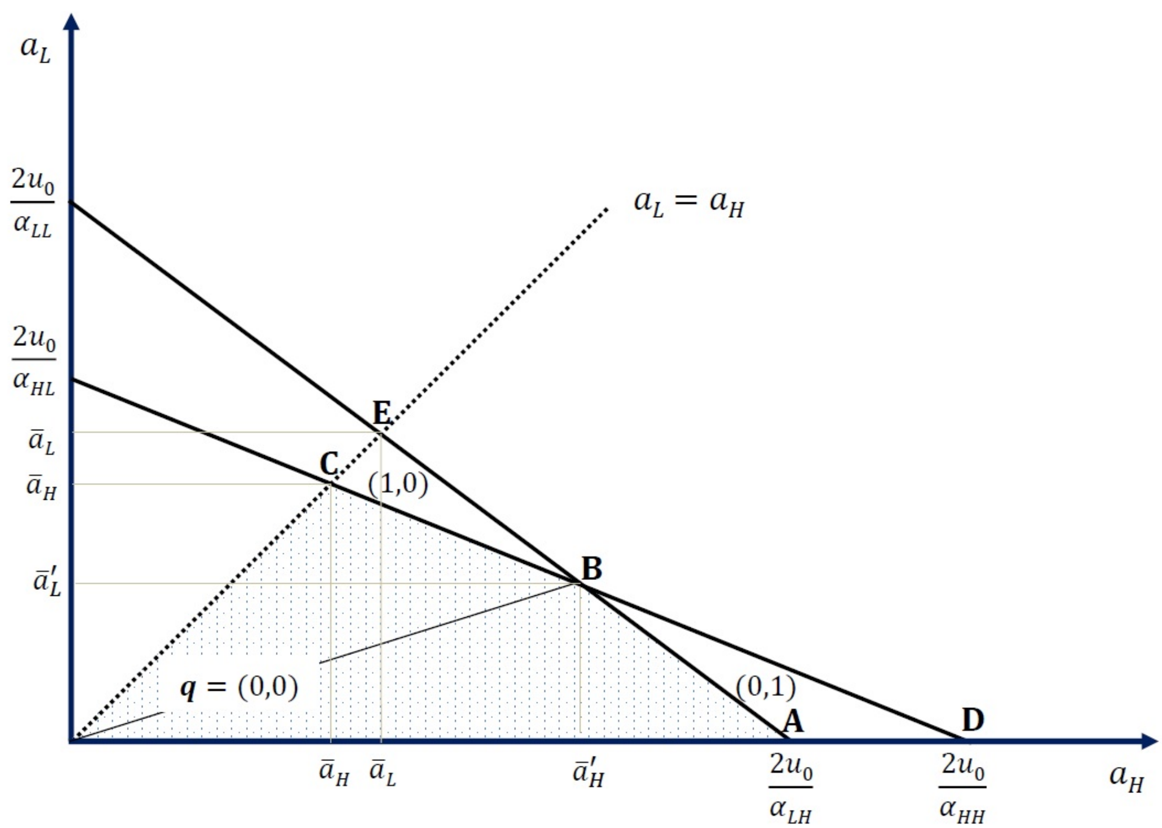

Figure 1: Optimal candidate contracts for the advertising platform

■ **Allocation** $(q_H, q_L) = (0,0)$

When $(q_H, q_L) = (0,0)$, IC_H^A and IC_L^A do not impose any restriction on (a_H, a_L) but it implies $p_H^A = p_L^A = p^A$ and the binding participation constraint on side A is determined by the sign of $a_L - \rho a_H$. Precisely, if $a_L < \rho a_H$ then IR_L^A binds, and p^A is equal to $u_0 - \frac{1}{2}\alpha_{LL}a_L - \frac{1}{2}\alpha_{LH}a_H$. Conversely, if $a_L \geq \rho a_H$ then IR_H^A binds, and p^A is equal to $u_0 - \frac{1}{2}\alpha_{HL}a_L - \frac{1}{2}\alpha_{HH}a_H$. Given (17) and (18), we find that the platform's profit is computed as

$$\begin{aligned}\pi(0,0,a_H,a_L) \quad &\equiv \quad u_0 + \frac{1}{4}(\beta_{HH} + \beta_{HL})a_H + \frac{1}{4}(2\beta_{LH} + 2\beta_{LL} - \beta_{HH} - \beta_{HL})a_L \\ &\quad -\frac{1}{2}\begin{cases} \alpha_{HH}a_H + \alpha_{HL}a_L & \text{if } a_H \geq a_L \geq \frac{\alpha_{LH}-\alpha_{HH}}{\alpha_{HL}-\alpha_{LL}}a_H \\ \alpha_{LH}a_H + \alpha_{LL}a_L & \text{if } a_L < \frac{\alpha_{LH}-\alpha_{HH}}{\alpha_{HL}-\alpha_{LL}}a_H \end{cases}\end{aligned}$$

and it must be evaluated over the polygonal set with vertices $(0,0)$, A, B, C. So, our attention can be limited to the points A, B and C.

■ **Allocation** $(q_H, q_L) = (0,1)$

When $(q_H, q_L) = (0,1)$, the implementability condition (19) is simply given by $a_L \leq \rho a_H$ and IC_H^A and IR_L^A make IR_H^A redundant. Then we obtain $p_L^A = u_0$ and $p_H^A = u_0 - \frac{1}{2}\alpha_{HH}a_H - \frac{1}{2}\alpha_{HL}a_L$. As explained, $p_H^A \geq 0$ requires (a_H, a_L) to belong to the triangle which has vertices $(0,0), B, D$ in the graph; our attention can be limited to the points B and D, and the platform's profit is given by

$$\pi(0,1,a_H,a_L) \equiv u_0 + \frac{1}{4}\beta_{HH}a_H - \frac{1}{4}\alpha_{HH}a_H + (\frac{1}{2}\beta_{LH} - \frac{1}{4}\beta_{HH})a_L - \frac{1}{4}\alpha_{HL}a_L.$$

■ **Allocation** $(q_H, q_L) = (1,0)$

When $(q_H, q_L) = (1,0)$, (19) becomes $a_L \geq \rho a_H$ and IC_L^A and IR_H^A make IR_L^A redundant. Then we obtain $p_H^A = u_0$, $p_L^A = u_0 - \frac{1}{2}\alpha_{LH}a_H - \frac{1}{2}\alpha_{LL}a_L$ and the profit is

$$\pi(1,0,a_H,a_L) = u_0 + \frac{1}{4}\beta_{HL}a_H - \frac{1}{4}\alpha_{LH}a_H + (\frac{1}{2}\beta_{LL} - \frac{1}{4}\beta_{HL})a_L - \frac{1}{4}\alpha_{LL}a_L.$$

From $p_L^A \geq 0$, we need to restrict (a_H, a_L) to belong to the triangle which has vertices $(0,0), B, E$; our attention can be limited to the points B and E.

In summary, we have:

Lemma 2. *Consider the application with the linear specification and Assumption 3. The profit-maximizing mechanism is one among the following seven candidates:*

(a) Contracts A, B, or C with $(q_H, q_L) = (0,0)$;

(b) Contracts B or D with $(q_H, q_L) = (0, 1)$;

(c) Contracts B or E with $(q_H, q_L) = (1, 0)$.

7.3 No PD on side B

Now let us study the case in which advertisers face a single menu of $a_H = a_L = a$. Since this case is a special case of the more general case in the previous subsection and the solution can be easily understood from Figure 2 following the diagonal, we relegate the proof to the Appendix and provide only the result:

Lemma 3. *Consider the application with the linear specification and Assumption 3. Conditional on no price discrimination on the advertising side, the optimal mechanism is either Contract C with $(q_H, q_L) = (0, 0)$ or Contract E with $(q_H, q_L) = (1, 0)$.*

7.4 Comparison

The previous analysis has identified all possible candidates for the optimal mechanism. Because we have many parameters, for clear comparison among them, we reduce the number of parameters to one. By doing so, we can gain further insight about under which condition a particular contract becomes optimal and when PD on one side complements or substitutes for the one on the other side.

Let $u_0 = 1$ without loss of generality and consider the following set of values which satisfy all assumptions made in this section and use only one parameter δ:

$$\begin{cases} u_0 = 1, \quad \alpha_{HH} = 1 - \frac{1}{12}\delta, \qquad \alpha_{HL} = \frac{7}{9}, \qquad \alpha_{LH} = 1, \qquad \alpha_{LL} = \frac{1}{2} \\ \qquad \beta_{HH} = 1 + \frac{1}{6}\delta, \qquad \beta_{HL} = 1.01, \qquad \beta_{LH} = 1.01, \qquad \beta_{LL} = 1 \end{cases} \tag{20}$$

where $\delta \in (0.06, \frac{10}{3})$ to satisfy the assumptions of $\alpha_{HL} - \alpha_{LL} > \alpha_{LH} - \alpha_{HH}$ and $\beta_{HH} > \beta_{HL}$. Then the points A, B, C, D, E have coordinates $(2, 0)$, $(\frac{40}{3\delta+20}, \frac{12\delta}{3\delta+20})$, $(\frac{72}{64-3\delta}, \frac{72}{64-3\delta})$, $(\frac{24}{12-\delta}, 0)$, $(\frac{4}{3}, \frac{4}{3})$, respectively. Remarkably, δ captures the intensity of type reversal in that as it increases, the net surplus generated by a H type's watching a H type's advertisement becomes larger.

Let $\pi(q_H, q_L; A)$ represent the profit at point A given (q_H, q_L). Then, we find that $\pi(0, 0; B) > \pi(0, 1; B)$, $\pi(0, 0; B) > \pi(1, 0; B)$, and $\pi(0, 0; B) > \pi(0, 0; A)$. In other words, we can eliminate the three contracts $\pi(0, 1; B)$, $\pi(1, 0; B)$, and $\pi(0, 0; A)$ from consideration as they are strictly dominated by $\pi(0, 0; B)$. Comparing the surviving four candidates leads to:

Lemma 4. *Consider the application with the linear specification with parameters given by (20). Then, the optimal mechanism is*

(a) Contract E with $(q_H, q_L) = (1, 0)$ and $a_L = a_H$ if $\delta \in S_1 \equiv (0.06, 0.659)$

(b) Contract C with $(q_H, q_L) = (0, 0)$ and $a_L = a_H$ if $\delta \in S_2 \equiv (0.659, 0.993)$

(c) Contract B with $(q_H, q_L) = (0, 0)$ and $a_H > a_L > 0$ if $\delta \in S_3 \equiv (0.993, 1.887)$

(d) Contract D with $(q_H, q_L) = (0, 1)$ is $a_H > a_L = 0$ if $\delta \in S_4 \equiv (1.887, \frac{10}{3})$.

where the neighboring contracts are tied at each border value of δ.

For small enough $\delta \in S_1$, showing advertisements to H type consumers is not optimal as their nuisance cost is high relative to the advertising revenues generated from them. Conditional on advertising only to L type consumers, the platform ideally wants to implement $a_L > a_H$ on the advertising side as their nuisance from watching H type ads is much larger than the nuisance from L type ads. However, such a decreasing advertising schedule cannot be implemented on side B because of the implementability condition. Therefore, the platform chooses the uniform treatment of $a_L = a_H$, which leads to Contract E with $(q_H, q_L) = (1, 0)$.

As δ increases into S_2, it becomes optimal to show advertisements to both types of consumers. However, the platform still wants to choose $a_L > a_H$ and hence is constrained by the implementability condition on side B. This leads to pooling on both sides: $q_H = q_L = 0$ and $a_H = a_L = a_H$ which is Contract C.

As δ further increases and belongs to S_3, it is still optimal to show advertisements to both types of consumers but now δ is high enough that H type consumers generate much advertising revenue to H type advertisers while experiencing not much nuisance. Hence, the platform implements $a_H > a_L > 0$, which makes Contract B optimal.

Finally, for a high enough $\delta \in S_4$, H type consumers generates so much advertising revenue to H type advertisers while experiencing little nuisance that the platform wants to shutdown advertising to L type consumers and not to sell advertising service to L type advertisers, which leads to Contract D.

Now we turn to the original question: how the PD on both sides affects the profit of the platform compared to the PD on the consumer side only? For relatively small $\delta \in S_1 \cup S_2$, the optimal contract involves pooling on the advertising side: it can even involve pooling on both sides if Contract C is optimal. The platform wants to implement $a_L > a_H$ but it cannot due to the implementability condition, which makes $a_H = a_L$ second-best optimal. Therefore, forcing a strict PD on the advertising side would reduce the platform's overall profit. Hence, we can state that the PD on the advertising side substitutes for the PD on the consumer side in this case.

For relatively large $\delta \in S_3 \cup S_4$, the PD on the advertising side complements the PD on the consumer side for two reasons. First, given $(q_H, q_L) = (0, 0)$, since the platform wants to implement $a_H > a_L$, introducing the PD on the adverting side increases the platform's profit. This is a standard argument for a second-degree PD in one-sided market. Second, introducing PD on the adverting side can increase the profit by allowing the platform to implement a strictly decreasing schedule on the consumer side, which is unique in a two-sided market.

Proposition 5. *Consider the application to the advertising platform with the linear specification and parameters given by (20).*

(i) The PD on the advertising side substitutes for the PD on the consumer side if $\delta \in S_1 \cup S_2$

(ii) The PD on the advertising side complements the PD on the consumer side if $\delta \in S_3 \cup S_4$.

  (a) For $\delta \in S_3$*, it does so by implementing a strictly increasing advertising schedule without affecting the allocation on the consumer side.*

  (b) For $\delta \in S_4$*, it does so not only by implementing a strictly increasing advertising schedule but also by relaxing the implementability condition on the consumer side to implement a strictly decreasing quality schedule.*

The above result can provide some insight about actual business practices by many online media platforms. For instance, YouTube recently launched its long-discussed paid subscription service, YouTube Red. This kind of PD on the consumer side corresponds to the increasing schedule $(q_H, q_L) = (1, 0)$ in which H type consumers pay a certain fee to avoid the ads. Suppose for instance that YouTube Red means a change from Contract C with $(q_H, q_L) = (0, 0)$ (i.e., $\delta \in S_2$) to Contract E with $(q_H, q_L) = (1, 0)$ (i.e., $\delta \in S_1$). Then, this change involves an increase in advertising amount without changing its composition and an L type consumer gets worse off.

8 Concluding Remarks

A two-sided platform mediates interactions between two different groups. Each agent in a group plays a dual role of obtaining the private benefit from interactions and generating externalities to the agents of the other group. We show that this dual role may make pooling optimal on the side which generates strong externalities. In addition, we show that in the case of non-separable utility, properly designing price discrimination on one side may allow to mitigate or remove pooling on the other side because PD on one side may affect the set of implementable allocations on the other side. We also provide welfare analysis of price discrimination in a two-sided market. Our canonical model of two-sided PD can be applied to more specific environments as we demonstrated how our model can be adapted to net neutrality regulation and online media advertising platforms. It is to be hoped that our work will serve as a foundation from which studies of different specificity and greater depth may be undertaken for understanding price discrimination in two-sided markets.

References

[1] Ambrus, Attila, Emilio Calvano, and Markus Reisinger. (2016) "Either or Both Competition: A 'Two-Sided' Theory of Advertising with Overlapping Viewerships." *American Economic Journal: Microeconomics* 8(3): 189-222.

[2] Anderson, Simon and Stephen Coate (2005) "Market Provision of Broadcasting: A Welfare Analysis." *Review of Economic Studies*, 947-972.

[3] Angelucci, Charles, and Julia Cagé. (2016) "Newspapers in Times of Low Advertising Revenues." Columbia Business School Working Paper.

[4] Armstrong, Mark. (2006) "Competition in two-sided markets." *Rand Journal of Economics* 37: 668-691.

[5] Armstrong, Mark. (2015) "Non-linear Pricing." mimeo.

[6] Böhme, Enrico. (2012) "Second-Degree Price Discrimination on Two-Sided Markets." MPRA Paper 40951, University Library of Munich, Germany.

[7] Caillaud, Bernard and Bruno Jullien. (2001) "Competing Cybermediaries." *European Economic Review* 45: 797-808.

[8] Caillaud, Bernard and Jean Tirole. (2004) “Essential Facility Financing and Market Structure.” *Journal of Public Economics* 88: 667- 694.

[9] Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim. (2015) “Net Neutrality, Business Models, and Internet Interconnection.” *American Economic Journal: Microeconomics* 7(3): 104-141.

[10] Gabszewicz, Jean Jaskold, Didier Laussel, and Nathalie Sonnac. (2004) “Programming and Advertising Competition in the Broadcasting Industry.” *Journal of Economics and Management Strategy* 13 (4): 657â€“69.

[11] Gomes, Renato and Alessandro Pavan. (2016) “Many-to-many Matching and Price Discrimination.” *Theoretical Economics* 11: 1005-1052.

[12] Greenstein, Shane, Martin Peitz, and Tommaso Valletti. (2016) “Net Neutrality: A Fast Lane to Understanding the Trade-Offs.” *Journal of Economic Perspectives* 30(2): 127-50.

[13] Guesnerie, Roger and Jean-Jacques Laffont. (1984) “A Complete Solution to a Class of Principal–agent Problems with an Application to the Control of a Self-managed Firm.” *Journal of Public Economics* 25: 329-369.

[14] Hagiu, Andrei. (2009) “Two-sided Platforms: Product Variety and Pricing Stuctures.” *Journal of Economics and Management Strategy* 18(4): 1011–1043.

[15] Jeon, Doh-Shin and Domenico Menicucci. (2008) “Money, Fame and Allocation of Talent: Brain Drain and the Institution of Science.” *Journal of Economic Behavior and Organization* 66: 558-581.

[16] Jeon, Doh-Shin and Jean-Charles Rochet. (2010) “The Pricing of Academic Journals: A Two-sided Market Perspective.” *American Economic Journal: Microeconomics* 2(2): 222-255.

[17] Laffont, Jean-Jacques and David Martimort. (2002) The Theory of Incentives: The Principal-Agent Model. Princeton University Press.

[18] Maskin, Eric and John Riley. (1984) “Monopoly with Incomplete Information.” *Rand Journal of Economics* 15: 171-196.

[19] Mussa, Michael and Sherwin Rosen. (1978) “Monopoly and Product Quality.” *Journal of Economic Theory* 18: 301-317.

[20] Peitz, Martin and Tommaso Valletti. (2008) "Content and advertising in the media: Pay-tv versus free-to-air." *International Journal of Industrial Organization* 26 (4): 949â€“65.

[21] Rochet, Jean-Charles and Jean Tirole. (2003) "Platform Competition in Two-sided Markets." *Journal of the European Economic Association* 1: 990-1029.

[22] Rochet, Jean-Charles and Jean Tirole. (2006) "Two-sided Markets: A Progress Report." *Rand Journal of Economics* 37: 645-667.

[23] Weyl, Glen. (2010) "A Price Theory of Two-sided Markets." *American Economic Review* 100(4): 1642-1672.

[24] Wilson, Robert. (1993) Non-linear pricing. Oxford University Press.

Appendix

A Mathematical Proofs

The proofs for other propositions and lemmas are discussed in the text. Thus, here we provide mathematical proofs for Lemma 1, Proposition 4, Lemma 3 and Lemma 4.

A.1 Proof of Lemma 1

Lemma 1 in the text provides sufficient conditions for the following claims to be true (notice that standard arguments imply that the second claim is a consequence of the first):

$$\begin{cases} (IR_H^k) \text{ is redundant} \\ (IR_L^k) \text{ and } (IC_H^k) \text{ bind in the optimal mechanism} \quad \text{for } k = A, B \\ (IC_L^k) \text{ is equivalent to } q_H^k \geq q_L^k \end{cases} \tag{A.1}$$

In fact, (A.1) can be proved under weaker assumptions than the ones described in Lemma 1. Precisely, consider the following conditions, which are both satisfied if u^k is separable and satisfies Assumption 2:

$$u^k(q^k, q^l) \geq 0 \quad \text{and} \quad u_1^k(q^k, q^l) > 0 \quad \text{for each} \quad q^k, q^l \tag{A.2}$$

$$u_2^k(q^k, q^l) > 0 \quad \text{for each} \quad q^k, q^l \tag{A.3}$$

Then we can state and prove a more general version of Lemma 1 than the one given in the text.

A general version of Lemma 1 The claims in (A.1) hold true if any of the following four sets of conditions is satified.

(i) On side k there is no type reversal and (A.2) is satisfied, for $k = A, B$.

(ii) On side A there is no type reversal and (A.2) is satisfied for $k = A$; on side B there is type reversal with positive sorting, (A.2)-(A.3) are satisfied for $k = B$, and $u_{12}^B \geq 0$, or $u_{12}^B < 0$ is close to zero and/or $v_L^A(\theta_{HL}^B - \theta_{LL}^B)$ is close to zero.

(iii) On side B there is no type reversal and (A.2) is satisfied for $k = B$; on side A there is type reversal with positive sorting, (A.2)-(A.3) are satisfied for $k = A$, and $u_{12}^A \geq 0$, or $u_{12}^A < 0$ is close to zero and/or $v_L^B(\theta_{HL}^A - \theta_{LL}^A)$ is close to zero.

(iv) On both sides there is type reversal with positive sorting, (A.2)-(A.3) are satisfied for $k = A, B$, for one side k we have u_{12}^k close to 0 and/or $v_L^l(\theta_{HL}^k - \theta_{LL}^k)$ close to zero, and for the other side l we have $u_{12}^l \geq 0$, or $u_{12}^l < 0$ is close to zero and/or $v_L^k(\theta_{HL}^l - \theta_{LL}^l)$ is close to zero.

Proof of the general version of Lemma 1

(i) Consider side A. We can combine (IR_L^A) and (IC_H^A) to find that (IR_H^A) holds if

$$v_H^B(\theta_{HH}^A - \theta_{LH}^A)u^A(q_L^A, q_H^B) + v_L^B(\theta_{HL}^A - \theta_{LL}^A)u^A(q_L^A, q_L^B) \geq 0 \qquad \text{(A.4)}$$

Since (A.2) is satisfied for $k = A$ and there is no type reversal, it follows that $v_H^B(\theta_{HH}^A - \theta_{LH}^A)u^A(q_L^A, q_H^B) \geq 0$ and $v_L^B(\theta_{HL}^A - \theta_{LL}^A)u^A(q_L^A, q_L^B) \geq 0$, hence (A.4) holds.

Adding up (IC_H^A), (IC_L^A) yields the inequality (16), and such inequality is equivalent to $q_H^A \geq q_L^A$ if

$$v_H^B(\theta_{HH}^A - \theta_{LH}^A)u_1^A(q_H^A, q_H^B) + v_L^B(\theta_{HL}^A - \theta_{LL}^A)u_1^A(q_H^A, q_L^B) > 0 \qquad \text{(A.5)}$$

Since (A.2) is satisfied for $k = A$ and there is no type reversal, it follows that $v_H^B(\theta_{HH}^A - \theta_{LH}^A)u_1^A(q_H^A, q_H^B) > 0$ and $v_L^B(\theta_{HL}^A - \theta_{LL}^A)u_1^A(q_H^A, q_L^B) > 0$, hence (A.5) holds. The proof for side B is analogous.

(ii) For side A we can argue exactly like in the proof of part (i) to prove that (A.4) and (A.5) is true.

For side B, we can combine (IR_L^B) and (IC_H^B) to find that (IR_H^B) holds if

$$v_H^A(\theta_{HH}^B - \theta_{LH}^B)u^B(q_L^B, q_H^A) + v_L^A(\theta_{HL}^B - \theta_{LL}^B)u^B(q_L^B, q_L^A) \geq 0 \tag{A.6}$$

Then we notice that $v_H^A(\theta_{HH}^B - \theta_{LH}^B)u^B(q_L^B, q_H^A) \geq v_H^A(\theta_{HH}^B - \theta_{LH}^B)u^B(q_L^B, q_L^A)$ given that $\theta_{HH}^B - \theta_{LH}^B > 0$, (A.3) holds for $k = B$, and $q_H^A \geq q_L^A$. Therefore the left hand side in (A.6) is at least as large as $v_H^A(\theta_{HH}^B - \theta_{LH}^B)u^B(q_L^B, q_L^A) + v_L^A(\theta_{HL}^B - \theta_{LL}^B)u^B(q_L^B, q_L^A) = \left(\theta_H^B - \theta_L^B\right) u^B(q_L^B, q_L^A)$, which is non negative because of Assumption 1 and (A.2) for $k = B$.

Adding up (IC_H^B), (IC_L^B) yields an inequality analogous to (16), and such inequality is equivalent to $q_H^B \geq q_L^B$ if

$$v_H^A(\theta_{HH}^B - \theta_{LH}^B)u_1^B(q_H^B, q_H^A) + v_L^A(\theta_{HL}^B - \theta_{LL}^B)u_1^B(q_H^B, q_L^A) > 0 \tag{A.7}$$

Given type reversal with positive sorting on side B, and given that (A.2) holds for $k = B$, the first term in (A.7) is positive and the second term is negative. In case that $u_{12}^B \geq 0$, we find that $v_H^A(\theta_{HH}^B - \theta_{LH}^B)u_1^B(q_H^B, q_H^A) \geq v_H^A(\theta_{HH}^B - \theta_{LH}^B)u_1^B(q_H^B, q_L^A)$, hence the left hand side of (A.7) is at least as large as $v_H^A(\theta_{HH}^B - \theta_{LH}^B)u_1^B(q_H^B, q_L^A) + v_L^A(\theta_{HL}^B - \theta_{LL}^B)u_1^B(q_H^B, q_L^A) = \left(\theta_H^B - \theta_L^B\right) u_1^B(q_H^B, q_L^A)$, which is positive by assumption 1 and (A.2) for $k = B$. In case that $u_{12}^B < 0$ is close to zero, then $u_1^B(q_H^B, q_H^A)$ is only sligthly smaller than $u_1^B(q_H^B, q_L^A)$, hence the above argument still applies. In case that $v_L^A(\theta_{HL}^B - \theta_{LL}^B)$ is close to zero, then the sign of the left hand side in (A.7) is determined by the term $v_H^A(\theta_{HH}^B - \theta_{LH}^B)u_1^B(q_H^B, q_H^A) > 0$.

(iii) The proof coincides with the proof given for part (ii), after switching the roles of A and B.

(iv) Suppose that u_{12}^B is close to 0. Then the left hand side in (A.7) is close to $\left(\theta_H^B - \theta_L^B\right) u_1^B(q_H^B, q_L^A)$, which is positive both if $q_H^A \geq q_L^A$ and if $q_H^A < q_L^A$. Likewise, if $v_L^A(\theta_{HL}^B - \theta_{LL}^B)$ is close to 0, then the left hand side in (A.7) is close to $v_H^A \left(\theta_{HH}^B - \theta_{LH}^B\right) u_1^B(q_H^B, q_H^A) > 0$, both if $q_H^A \geq q_L^A$ and if $q_H^A < q_L^A$. In either case we obtain $q_H^B \geq q_L^B$.

Then assume $u_{12}^A \geq 0$ and use $q_H^B \geq q_L^B$ to obtain that the left hand side in (A.5) is at least as large as $(\theta_H^A - \theta_L^A)u_1^A(q_H^A, q_L^B) > 0$. In alternative, if $u_{12}^A < 0$ is close to zero and/or $v_L^B(\theta_{HL}^A - \theta_{LL}^A)$ is close to zero, then it is still the case that (A.5) is satisfied. Therefore $q_H^A \geq q_L^A$.

Given $q_H^A \geq q_L^A$, we can prove that (A.6) holds by using $\theta_{HH}^B - \theta_{LH}^B > 0$ and (A.3) for $k = B$ as in the proof of part (ii). Likewise, given $q_H^B \geq q_L^B$, we can prove that (A.4) holds by using $\theta_{HH}^A - \theta_{LH}^A > 0$ and (A.3) for $k = A$. ■

A.2 Proof of Proposition 3

(i) The proof of (i) is omitted since it is straightforward.

(ii) Let $\tilde{\theta}_H^A \equiv v_H^A \theta_{HH}^A + v_L^A \theta_{LH}^A$ and $\tilde{\theta}_L^A \equiv v_H^A \theta_{HL}^A + v_L^A \theta_{LL}^A$.

- About the effect of θ_{LH}^A, we first notice

$$\frac{dq_H^B}{d\theta_{LH}^A} = \frac{\beta u'(q_H^B)}{C''(q_H^B) - (\theta_H^B + \beta\theta_{LH}^A)u''(q_H^B)}, \qquad \frac{dq^B}{d\theta_{LH}^A} = \frac{\beta v_H^B u'(q^B)}{C''(q^B) - (\theta_L^B + \beta\theta_L^A)u''(q^B)}.$$

Then we have

$$\begin{aligned}
\frac{d(\Delta W_B^A/\beta)}{d\theta_{LH}^A} &= v_H^B v_L^A \left(u(q_H^B) - u(q^B)\right) + v_H^B \tilde{\theta}_H^A u'(q_H^B)\frac{dq_H^B}{d\theta_{LH}^A} - (v_H^B \tilde{\theta}_H^A + v_L^B \tilde{\theta}_L^A)u'(q^B)\frac{dq^B}{d\theta_{LH}^A} \\
&= v_H^B v_L^A \left(u(q_H^B) - u(q^B)\right) + v_H^B \tilde{\theta}_H^A \beta \frac{(u'(q_H^B))^2}{C''(q_H^B) - (\theta_H^B + \beta\theta_{LH}^A)u''(q_H^B)} \\
&\quad - v_H^B (v_H^B \tilde{\theta}_H^A + v_L^B \tilde{\theta}_L^A)\beta \frac{(u'(q^B))^2}{C''(q^B) - (\theta_L^B + \beta\theta_L^A)u''(q^B)} \\
&= v_H^B \left\{ v_L^A \left(u(q_H^B) - u(q^B)\right) + \left[\tilde{\theta}_H^A - (v_H^B \tilde{\theta}_H^A + v_L^B \tilde{\theta}_L^A)\right] \beta f(q) \right\}
\end{aligned}$$

Hence if $f(q) = \frac{(u'(q))^2}{C''(q) - \frac{C'(q)}{u'(q)}u''(q)}$ is increasing, and $\tilde{\theta}_H^A \geq \tilde{\theta}_L^A$, then $\frac{d(\Delta W_B^A/\beta)}{d\theta_{LH}^A} > 0$.

- About the effect of θ_{LL}^A, we notice

$$\frac{dq_L^B}{d\theta_{LL}^A} = \frac{\beta u'(q_L^B)}{C''(q_L^B) - (\theta_L^{Bv} + \beta\theta_{LL}^A)u''(q_L^B)}, \qquad \frac{dq^B}{d\theta_{LL}^A} = \frac{\beta v_L^B u'(q^B)}{C''(q^B) - (\theta_L^B + \beta\theta_L^A)u''(q^B)}.$$

Then we have

$$\begin{aligned}
\frac{d(\Delta W_B^A/\beta)}{d\theta_{LL}^A} &= v_L^B \left\{ \left(-v_L^A \left[u(q^B) - u(q_L^B)\right] + \left[\tilde{\theta}_L^A - (v_H^B \tilde{\theta}_H^A + v_L^B \tilde{\theta}_L^A)\right] \beta f(q) \right) \right\} \\
&= -v_L^B v_L^A \left(u(q^B) - u(q_L^B)\right) + v_L^B \tilde{\theta}_L^A \beta \frac{(u'(q_L^B))^2}{C''(q_L^B) - (\theta_L^{Bv} + \beta\theta_{LL}^A)u''(q_L^B)} \\
&\quad - v_L^B (v_H^B \tilde{\theta}_H^A + v_L^B \tilde{\theta}_L^A)\beta \frac{(u'(q^B))^2}{C''(q^B) - (\theta_L^B + \beta\theta_L^A)u''(q^B)}
\end{aligned}$$

Then we find that $d(\Delta W_B^A/\beta) < 0$ if $f(q)$ is increasing and $\tilde{\theta}_H^A \geq \tilde{\theta}_L^A$. ■

A.3 Proof of Proposition 4

Here we prove a more detailed version of Proposition 4, and for that purpose we let

$$\phi(q_H^A) = \nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) u_1^A(q_H^A, q_H^B) + \nu_L^B(\theta_{HL} - \theta_{LL}) u_1^A(q_H^A, q_L^B)$$

denote the derivative of Φ^A with respect to q_H^A: notice that ϕ does not depend on q_L^A. As the case of negative sorting is symmetric to the case of positive sorting, we only provide the proof for the positive sorting of which the statement is refined as follows.

Refined version of Proposition 4(ii)

(i) Suppose that $u_{12} > 0$ and $u_{112} \geq 0$ (not needed for part (a)).

 (a) When $q_H^B \geq q_L^B$, we have $F = M \cup D$.

 (b) When $q_H^B < q_L^B$, we have that

  (b1) If $\phi(0) \leq 0$, then $F = N \cup D$ if $\phi(0) \leq 0$.

  (b2) If $\phi(0) > 0 > \lim_{q_H^A \to +\infty} \phi(q_H^A)$, then let $\overline{q}_H^A$ be uniquely defined by $\phi(\overline{q}_H^A) = 0$. The set F has the shape of a sandglass, such that it includes all points in M such that $q_L \leq \overline{q}_H^A$ and $q_H \leq \overline{q}_H^A$, and some points in N if $q_L^A > \overline{q}_H^A$.

  (b3) If $\lim_{q_H^A \to +\infty} \phi(q_H^A) \geq 0$,then $F = M \cup D$.

 (ii) Suppose that $u_{12} < 0$ and $u_{112} \leq 0$ (not needed for part (a)).

  (a) when $q_H \leq q_L$, we have $F = M \cup D$.

  (b) When $q_H > q_L$, we have that (b1-b3) from part (i) hold.

Proof of part (i): Complements: $u_{12}^A(q^A, q^B) > 0$ and $u_{112}^A(q^A, q^B) \geq 0$ for each q^A, q^B

1. If $q_H^B \geq q_L^B$, then $u_1^A(q_H^A, q_H^B) \geq u_1^A(q_H^A, q_L^B)$ and $\phi(q_H^A) \geq (\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A)) u_1(q_H^A, q_L^B) > 0$. Therefore Φ^A is strictly increasing in q_H^A and $F = M \cup D$.

2. If $q_H^B < q_L^B$, then assume $u_{112}^A \geq 0$, that is u_{11}^A is increasing with respect to q^B, or equivalently u_{12}^A is increasing with respect to q^A. Then $\phi'(q_H^A) = \nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) u_{11}^A(q_H^A, q_H^B) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A) u_{11}^A(q_H^A, q_L^B) \leq (\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A)) u_{11}^A(q_H^A, q_H^B) < 0$. Therefore ϕ is strictly decreasing.

- If $\phi(0) \leq 0$, then $\phi(q_H^A) < 0$ for each $q_H^A > 0$. Therefore Φ^A is strictly decreasing in q_H^A and $F = N \cup D$.
- If $\phi(0) > 0 > \lim_{q_H^A \to +\infty} \phi(q_H^A)$, then let $\overline{q}_H^A$ be uniquely defined by $\phi(\overline{q}_H^A) = 0$.

  Now fix q_L^A, and consider $q_L^A < \overline{q}_H^A$. Then $\phi(q_H^A) > 0$ for $q_H^A \in (0, q_L^A)$ and $\Phi^A(q_H^A, q_L^A) < 0$ for each $q_H^A < q_L^A$. Conversely, $\Phi^A(q_H^A, q_L^A) > 0$ at least for $q_H^A \in (q_L^A, \overline{q}_H^A]$, because Φ^A is increasing in q_H^A for $q_H^A \in (q_L^A, \overline{q}_H^A)$. Since $\phi(q_H^A) < 0$ for $q_H^A > \overline{q}_H^A$, it is possible that $\Phi^A(q_H^A, q_L^A) < 0$ for q_H^A sufficiently larger than $\overline{q}_H^A$.

  Now consider $q_L^A > \overline{q}_H^A$. Then $\phi(q_H^A) < 0$ for each $q_H^A > q_L^A$, hence $\Phi^A(q_H^A, q_L^A) < 0$ for each $q_H^A > q_L^A$. Conversely, $\Phi^A(q_H^A, q_L^A) > 0$ at least for $q_H^A \in [\overline{q}_H^A, q_L^A)$ because Φ^A is decreasing in q_H^A for $q_H^A \in (\overline{q}_H^A, q_L^A)$. Since $\phi(q_H^A) > 0$ for $q_H^A < \overline{q}_H^A$, it is possible that $\Phi^A(q_H^A, q_L^A) < 0$ for q_H^A sufficiently smaller than $\overline{q}_H^A$.

  In this case the feasible set is non convex, and has vaguely the shape of a sandglass.
- If $\lim_{q_H^A \to +\infty} \phi(q_H^A) \geq 0$, then Φ^A is strictly increasing in q_H^A, hence $(\mathbf{I}^A)$ is satisfied if and only if $(q_H^A, q_L^A) \in M \cup D$.

Proof of part (ii): Substitutes: $u_{12}^A(q^A, q^B) < 0$ and $u_{112}^A(q^A, q^B) \leq 0$ for each q^A, q^B

1. If $q_H^B \leq q_L^B$, then $u_1^A(q_H^A, q_H^B) \geq u_1^A(q_H^A, q_L^B)$ and $\phi(q_H^A) \geq (\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A))u_1^A(q_H^A, q_L^B) > 0$. Therefore $(\mathbf{I}^A)$ is equivalent to $q_H^A \geq q_L^A$.
2. If $q_H^B > q_L^B$, then assume $u_{112}^A \leq 0$, that is u_{11}^A is decreasing with respect to q^B, or equivalently u_{12}^A is decreasing with respect to q^A. Then $\phi'(q_H^A) = \nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) u_{11}^A(q_H^A, q_H^B) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A)u_{11}^A(q_H^A, q_L^B) \leq (\nu_H^B \left(\theta_{HH}^A - \theta_{LH}^A\right) + \nu_L^B(\theta_{HL}^A - \theta_{LL}^A))u_{11}^A(q_H^A, q_L^B) < 0$. Therefore ϕ is strictly decreasing and we obtain a feasible set similar to the case 2 above: (i) $N \cup D$ if $\phi(0) \leq 0$; (ii) a sandglass if $\phi(0) > 0 > \lim_{q_H^A \to +\infty} \phi(q_H^A)$; (iii) $M \cup D$ if $\lim_{q_H^A \to +\infty} \phi(q_H^A) \geq 0$. ■

A.4 Proof of Lemma 3

Then, the platform's price against advertisers is set to make L type advertisers earn zero net surplus:

$$p_L^B = p_H^B = \frac{1}{2}\left(\beta_{LH}(1-q_H) + \beta_{LL}(1-q_L)\right)a \tag{A.8}$$

A.4.1 Allocation $(q_H, q_L) = (0, 0)$

In this case IC_H^A and IC_L^A imply $p_L^A = p_H^A = p^A$, hence IR_H^A implies IR_L^A and $p^A = u_0 - \frac{1}{2}(\alpha_{HH} + \alpha_{HL})a$, with $a \leq \bar{a}_H \equiv \frac{2u_0}{\alpha_{HL}+\alpha_{HL}}$ in order to have $p^A \geq 0$. Since $p_L^B = p_H^B = \frac{1}{2}(\beta_{LH} + \beta_{LL})a$, the platform's profit is equal to

$$u_0 + \frac{1}{2}(\beta_{LH} + \beta_{LL})a - \frac{1}{2}(\alpha_{HH} + \alpha_{HL})a$$

and it should be maximized with respect to $a \in [0, \bar{a}_H]$. Assuming $\beta_{LH} + \beta_{LL} > \alpha_{HH} + \alpha_{HL}$, the optimal a is $\bar{a}_H$ (point C in Figure 1), hence the maximal value is

$$u_0 + (\frac{1}{2}\beta_{LH} + \frac{1}{2}\beta_{LL} - \frac{1}{2}\alpha_{HH} - \frac{1}{2}\alpha_{HL})\bar{a}_H.$$

A.4.2 Allocation $(q_H, q_L) = (0, 1)$

In this case IC_H^A and IC_L^A require $u_0 - \frac{1}{2}(\alpha_{HH} + \alpha_{HL})a - p_H^A \geq u_0 - p_L^A$ and $u_0 - p_L^A \geq u_0 - \frac{1}{2}(\alpha_{LH} + \alpha_{LL})a - p_H^A$. Combining the two inequality conditions, we obtain $(\alpha_{LH} + \alpha_{LL} - \alpha_{HH} - \alpha_{HL})a \geq 0$. Because of Assumption 2-(i), $(\alpha_{LH} + \alpha_{LL} - \alpha_{HH} - \alpha_{HL})a < 0$ and thus the derived condition holds only if $a = 0$. When $a = 0$, the platform extracts the full rent by charging $p_L^A = p_H^A = u_0$ and the profit is equal to u_0. As we assume that the platform is not viable without selling advertising, this situation is not optimal.

A.4.3 Allocation $(q_H, q_L) = (1, 0)$

In this case the constraints are given by

$$\begin{aligned}
&(\text{IR}_H^A) && u_0 - p_H^A \geq 0 \\
&(\text{IR}_L^A) && u_0 - \frac{1}{2}(\alpha_{LH} + \alpha_{LL})a - p_L^A \geq 0 \\
&(\text{IC}_H^A) && u_0 - p_H^A \geq u_0 - \frac{1}{2}(\alpha_{HH} + \alpha_{HL})a - p_L^A \\
&(\text{IC}_L^A) && u_0 - \frac{1}{2}(\alpha_{LH} + \alpha_{LL})a - p_L^A \geq u_0 - p_H^A
\end{aligned}$$

and the optimal tariffs are $p_H^A = u_0$, $p_L^A = u_0 - \frac{1}{2}(\alpha_{LH}+\alpha_{LL})a$, with $a_L \leq \bar{a}_L \equiv \frac{2u_0}{\alpha_{LH}+\alpha_{LL}}$ for $p_L^A \geq 0$. Since $p_L^B = p_H^B = \frac{1}{2}\beta_{LL}a$, the profit is

$$u_0 - \frac{1}{4}(\alpha_{LH} + \alpha_{LL})a + \frac{1}{2}\beta_{LL}a$$

and it should be maximized with respect to $a \in [0, \bar{a}_L]$. Assuming $2\beta_{LL} > \alpha_{LH} + \alpha_{LL}$, the optimal a is $\bar{a}_L$ (point E in Figure 1), hence the maximal profit given $(q_H, q_L) = (1, 0)$ is

$$u_0 + (\frac{1}{2}\beta_{LL} - \frac{1}{4}\alpha_{LH} - \frac{1}{4}\alpha_{LL})\bar{a}_L.$$

A.5 Proof of Lemma 4

Given $(q_H, q_L) = (0, 0)$, the profit function is

$$\begin{aligned} \pi(0,0,a_H,a_L) &= 1 + \frac{1}{4}\left(1 + \frac{1}{6}\delta + \frac{101}{100}\right)a_H + \frac{1}{4}\left(2\cdot\frac{101}{100} + 2 - 1 - \frac{1}{6}\delta - \frac{101}{100}\right)a_L \\ &\quad -\frac{1}{2}\begin{cases} (1-\frac{1}{12}\delta)a_H + \frac{7}{9}a_L & \text{if } a_H \geq a_L \geq \frac{3\delta}{10}a_H \\ a_H + \frac{1}{2}a_L & \text{if } a_L < \frac{3\delta}{10}a_H \end{cases} \end{aligned}$$

Hence

$$\begin{aligned} \pi(0,0;A) &= \frac{201}{200} + \frac{1}{12}\delta \\ \pi(0,0;B) &= \frac{2309\delta + 6030 - 150\delta^2}{300(3\delta+20)} \\ \pi(0,0;C) &= \frac{1809}{25\,(64-3\delta)} \end{aligned}$$

Given $(q_H, q_L) = (0, 1)$, the profit function is

$$\pi(0,1,a_H,a_L) = 1+\frac{1}{4}\cdot\left(1+\frac{1}{6}\delta\right)a_H - \frac{1}{4}\cdot\left(1-\frac{1}{12}\delta\right)a_H + \left(\frac{1}{2}\cdot\frac{101}{100} - \frac{1}{4}\cdot(1+\frac{1}{6}\delta)\right)a_L - \frac{1}{4}\cdot\frac{7}{9}a_L$$

$$\begin{aligned} \pi(0,1;B) &= \frac{1}{150}\frac{934\delta - 75\delta^2 + 3000}{3\delta+20} \\ \pi(0,1;D) &= \frac{1}{2}\frac{\delta+24}{12-\delta} \end{aligned}$$

Given $(q_H, q_L) = (1, 0)$, the profit function is

$$\pi(1, 0, a_H, a_L) = 1 + \frac{1}{4}\left(\frac{101}{100}\right) a_H - \frac{1}{4} a_H + \left(\frac{1}{2} - \frac{1}{4} \cdot \frac{101}{100}\right) a_L - \frac{1}{4} \cdot \frac{1}{2} a_L$$

Hence,

$$\begin{aligned} \pi(1, 0; B) &= \frac{3}{100} \frac{149\delta + 670}{3\delta + 20} \\ \pi(1, 0; E) &= \frac{7}{6} \end{aligned}$$

Comparing these derived payoffs, we obtain the stated result. ■