Many-to-Many Matching Design and Price Discrimination*

Renato Gomes
Toulouse School of Economics
renato.gomes@tse-fr.eu

Alessandro Pavan
Northwestern University
alepavan@northwestern.edu

February 2011
(preliminary and incomplete draft – do not circulate without permission)

Abstract

This paper studies second-degree price discrimination in matching markets, that is, in markets where the product sold by the monopolist is access to other agents. In order to investigate the optimality of a large variety of pricing strategies, we allow for any many-to-many matching rule that satisfies a weak reciprocity condition. In this context, we derive necessary and sufficient conditions for the welfare and profit-maximizing mechanisms to employ a single network or to offer a menu of non-exclusive networks (*multi-homing*). We characterize the matching schedules that arise under a wide range of preferences, and deliver testable comparative statics results that relate the pricing strategies of a profit-maximizing platform to conditions on demand and the distribution of match qualities. Our analysis sheds light on the distortions brought by the private provision of broadcasting, health insurance and job matching services.

JEL classification: D82

Keywords: matching, two-sided markets, networks, adverse selection, incentives.

*We are grateful to Johannes Horner, Doh-Shin Jeon, Bruno Jullien, Volker Nocke, Marco Ottaviani and Jean Tirole for very helpful conversations.

1 Introduction

This paper studies second-degree price discrimination in matching markets, that is, in markets where the product sold by the monopolist is access to other agents. In such markets, platforms engage in second-degree price discrimination by offering menus of matching plans that agents can choose from. For concreteness, consider the problem of a Cable TV provider that contracts with TV channels to offer packages of channels to home customers. The Cable company's problem can be seen from two perspectives. The more familiar one is that of customers who are offered a menu of packages of channels to choose from. In its mirror image, channels make transfers to the Cable company which are contingent on the number of viewers they reach (note that more viewers imply greater advertising revenues to channels). Importantly, by the very nature of matching markets, the menu of channels offered to customers pins down the quantity schedule faced by channels! As such, when designing its profit-maximizing plan, the Cable company has to internalize the cross-side effects of the schedules offered to customers and channels.

Such interdependency is ubiquitous in two-sided matching markets. Health care providers offer menus of health plans that differ on the access that patients (on one side of the market) have to doctors (on the other side of the market). Analogously to the cable TV example, the design of health plans on the patient side determines what services the platform has to procure on the doctor side of the market. As such, market conditions on the doctor side greatly matter for the profitability of different price-discriminatory strategies on the patient side.

In this paper, we study second-degree price discrimination by a monopolistic matching platform that offers many-to-many matching schemes. For greater generality, we consider the problem of a platform that operates in a market with two sides (alternatively, this problem can be reduced to a one-sided "group design" problem). Each agent on each side of the market has a multi-dimensional type, which is private information. The first component captures the agent's willingness to pay for the quality of the matching set he receives (i.e., for the attributes of the agents from the other side he/she to whom he is matched). In the Cable TV example, viewers value the quality of their TV packages and channels value extra viewers (as advertising rents increase). All other components are idiosyncratic characteristics that determine the agent's attractiveness from the eyes of agents on the other side (gender, education, income, etc for viewers; and the quality of shows, amount of advertising, etc for channels). All agents on a given side agree on the quality of agents on the other side, although their payoffs may differ due to differences in preferences for quality. The platform's problem consists in choosing a matching rule together with a pricing rule so as to maximize profits (respectively, welfare). A matching rule assigns each agent on each side to a set of agents on the other side of the market. We only impose that these rules satisfy a minimally restrictive feasibility constraint, which we call reciprocity condition, which requires that if agent i from side A is matched to agent j from side B, then agent j is matched to agent i. In the cable TV example, if viewer John

is matched to BBC News, then BBC News is matched to John.

Using standard techniques from the mechanism design literature, we then show that the problem of designing a profit-maximizing mechanism can be recasted entirely in terms of designing an optimal matching rule. This formulation renders the platform's profit-maximization problem analogous to welfare-maximization, by replacing agents' willingness to pay for the quality of their matching sets with their virtual-valuation counterparts.

Our first result pins down important properties that pertain to both the profit-maximizing and the welfare-maximizing matching rules. First, these rules discriminate only on the basis of willingness to pay, and not on the agent's idiosyncratic characteristics that determine his/her attractiveness for agents on side B. Intuitively, since these characteristics do not enter the agent's payoff directly, the only way to screen them would be to assign different matching sets of similar quality to agents with similar willingness to pay. This cannot be optimal however, since incentive compatibility prevents the platform from excluding agents with undesirable idiosyncratic characteristics while admitting others with similar willingness to pay and more desirable attributes.

Secondly, we show that the welfare and profit-maximizing matching rules assume a threshold structure, where each agent is matched to all agents on the other side whose willingness to pay is greater than some threshold. This result hinges on one important assumption of our model; namely, that the willingness to pay of agents and the match quality they provide to their matching partners are positively affiliated. As a consequence, it follows that the cost-minimizing way to provide a certain matching quality to a given agent is to design his matching set to include the highest number of high-valuation agents (which have higher average match quality) from the other side of the market.

Our main result shows that both the welfare-maximizing and the profit-maximizing matching rule belongs to one of the following two classes, and identifies necessary and sufficient conditions for each of the two classes to be optimal. The first class includes matching rules that employ a *single network*. Here, any two agents from the same side of the market have access to the same set of agents from the other side of the market. A single network is the analogue in matching environments to a single price (that is, the absence of quantity/quality discrimination) by a single-product monopolist. The second class is that of *nested multi-homing* matching rules. These rules work as follows: the platform offers a menu of non-exclusive networks that agents can join, and sets prices in a way that agents join an increasing number of networks. Under a nested multi-homing rule, any two agents from the same side of the market are assigned matching sets that either coincide or are nested in the set-theoretic sense. Nested multi-homing is the equivalent in matching environments to active quantity/quality price discrimination by a standard monopolist.

We prove that a single network is (profit-) welfare-maximizing if, starting from a complete network, receding the link between the two agents from each side with the lowest (virtual) valuations, while leaving all other links untouched, decreases (profits) welfare. If this is not the case, then the (profit-)

welfare-maximizing matching rule exhibits nested multi-homing. Because valuations are always larger than their virtual analogues, this condition predicts that matching rules that employ a single network are more often associated to welfare-maximizing platforms, while multi-homing matching rules are more often employed by profit-maximizing platforms. This prediction is consistent with casual empiricism: the public provision of broadcasting, health insurance and job matching services tend to employ a single network structure, while their private counterparts more often offer discriminatory menus (that is, multi-homing matching rules).

To understand the intuition behind this result, consider the platform's profit-maximization problem (the welfare problem is analogous) and assume that virtual valuations for quality are positive for all agents. In this simple case, optimality requires that all agents shall belong to the same complete network.

Things are different when virtual valuations are negative for certain types on one or both sides of the market. In this case, it is still true that all agents with positive virtual valuations shall be matched to all agents with positive virtual valuations on the other side of the market. However, the platform can still increase profits by adding to the matching sets of agents with high valuations on one side some agents on the other side with negative (but small) virtual valuations. This cross-subsidization strategy is a general feature of matching markets, in which the platform might be willing to accept revenue losses in one side to boost rent extraction on the other side. The size of the matching set of each agent at the optimum (that is, the extent of cross-subsidization) depends on his willingness to pay for matching quality. Whether this leads to single or multi-homing is determined by the marginal impact of linking the two agents with the lowest virtual valuations on both sides. If this quantity is positive, all agents receive similar matching sets (including the lowest types), that is, the optimum is a single (complete) network. If this quantity is negative, the platform does better by separating agents based on their valuations. In this case, agents with high valuations are assigned matching sets which are super sets of those assigned to agents with lower valuations, that is, the matching rule exhibits nested multi-homing.

Next, we offer a complete characterization of the (profit-) welfare-maximizing matching rule when nested multi-homing is optimal. We show that the thresholds associated to the matching sets of each agent solve an Euler equation that equalizes the marginal (revenue) efficiency gains from expanding his matching set on a given side to the marginal (revenue) efficiency losses that, by reciprocity, arise on the other side of the market. This endogenous cost structure is the fundamental feature of price discrimination in matching markets.

Similarly to the standard price discrimination problem analyzed in Mussa and Rosen (1978) and Maskin and Riley (1983), we need conditions that ensure that the platform is willing to separate types as finely as possible. However, the familiar monotonicity of virtual values is no longer the right condition in our matching environment, since now we also have to account for the cross-side effects

associated to each type. As it turns out, this is accomplished by requiring that virtual values increase *faster* than the marginal cross-side impact of each type (in analogy to Myerson (1981), we refer to this condition as Strong Regularity). If the platform's problem is strongly regular, bunching can only occur due to capacity constraints, that is, because the stock of agents on the other side of the market has already been exhausted.

As a by-product of the characterization result described above, our analysis reveals that profit-maximization leads to two distortions relative to efficiency. The first distortion, the *isolation effect*, states that under profit-maximization, each agent is matched to a subset of his efficient matching set. Unlike the standard results in nonlinear pricing, this distortion is present even for the highest type on each side of the market. The reason is that moving from welfare to revenue-maximization raises the endogenous marginal costs of matching agents (as virtual values are always smaller than values). As a consequence, matching sets are strictly smaller for all agents.

The second distortion, the *exclusion effect*, states that too many agents are excluded from the pool served by the platform. This distortion is analogous to the usual monopolistic distortion, and is present in both the single and the multi-homing solutions.

Our analysis delivers testable empirical implications that associate shocks that alter the cross-side effects of matches to changes in the platform's pricing strategies. In the context of the Cable TV application, let's consider a positive shock on the income distribution of households that leaves their valuations for TV packages unchanged. How would the platform's update its menu of TV packages? Take a consumer with high valuation and consider the marginal channel that integrates his TV package (that is, the channel with the lowest willingness to pay for an extra viewer). If nested multi-homing is optimal, the virtual valuation of this marginal channel is negative. Therefore, the platform's cost of cross-subsidizing viewers with high valuations goes up when they become more valuable to channels, since incentive compatibility requires that the platform decreases charges to all channels with higher valuations than the marginal channel. In contrast, the benefits of matching viewers with low valuations goes up, since the marginal channel on their TV package has a positive virtual value. As consequence, the platform's optimal response to positive shocks on the wealth of households is to improve the standard packages (which target low-valuation viewers) and worsen the premium packages (which target high-valuation viewers), which make viewers with high valuations worse off and with low valuations better off.

A special case of our model is that of a principal operating in a single-sided market characterized by peer effects that has to divide agents into non-exclusive groups. We show that this one-sided group formation problem is equivalent to a two-sided matching problem where both sides have symmetric primitives and the platform is constrained to choose among symmetric matching rules. As it turns out, in symmetric two-sided markets, the efficient (as well as the profit-maximizing) matching rules are naturally symmetric. As such, all results in the paper can be recasted if one replaces "single

network" by "single group" and "nested multi-homing matching rule" by "mutually non-exclusive groups". In particular, our results can be applied to problems in organization economics (e.g., the design of working groups).[1]

The rest of the paper is structured as follows. Below we complete this introductory section with a quick review of the pertinent literature. Section 2 presents the model. Section 3 derives the main results: first, it identifies necessary and sufficient conditions for the efficient (or profit-maximizing) matching rule to employ a single network or exhibit nested multi-homing. Next, it characterizes properties of optimal multi-homing rules and discusses the distortions brought in by profit maximization relative to efficiency. We then derive testable comparative statics results that relate the pricing strategies of a profit-maximizing platform to the underlying distributions of valuations and match qualities. Section 4 considers a few extensions. Finally, Section 5 concludes. All proofs omitted in the main text are in the Appendix at the end of the document.

Related Literature

This paper studies second-degree price discrimination (a la Mussa and Rosen (1978) and Maskin and Riley (1983)) in matching markets, that is, in markets where the product traded by the monopolist is access to other agents.[2] In this environment, the analogs of quantity (or quality) schedules are matching rules that assign to each agent a set of agents on the other side of the market.

While the literature that studies monopolistic pricing in two-sided markets restricts attention to a single network or to mutually exclusive networks (e.g., Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2008), Caillaud and Jullien (2001, 2003), Ambrus and Argenziano (2009) and Weyl (2010)), here we assume that platforms can design arbitrary matching functions and provide conditions for the optimality of a single network relative to more sophisticated price discriminatory schemes (multi-homing matching).

In the context of one-to-one matching, Damiano and Li (2007) and Johnson (2010) derive conditions on primitives for a profit-maximization platform to employ positive assortative matching. In turn, Hoppe, Moldovanu and Sela (2009) derive one-to-one positive assortative matching as the equilibrium outcome of a costly signaling game. In contrast, our paper studies second-degree price discrimination in many-to-many matching environments.

[1] In a recent paper on one-sided group design, Board (2009) allows the platform to assign agents to mutually exclusive groups (that is, single-homing matching). In this context, he shows that the partition induced by a profit-maximizing rule will never be coarser than the one induced by the efficient rule (note that this result, however, does not imply that the partition induced by the profit-maximizing rule will be finer). Relative to this paper, we extend the analysis of group design to two-sided matching environments and allow for more general matching rules. This gain in generality leads to stronger results, as the isolation effects discussed above implies that the profit-maximizing partition is finer than the efficient partition.

[2] For models of second-degree price discrimination on quality, see Deneckere and McAfee (1996), Ellison and Fudenberg (2000) and Anderson and Dana (2009).

In a decentralized economy, Shimer and Smith (2000), Shimer (2005), Smith (2006), Atakan (2006) and Eeckhout and Kircher (2010) consider extensions of the assignment model of Becker (1973) to a setting with search/match frictions. These papers show that the resulting one-to-one matching allocation is positive assortative provided that the match value function satisfies strong forms of supermodularity. Relative to this literature, we abstract from search frictions, but allow the matching rule to be many-to-many and consider significantly more general information structures (in particular, we allow for the information to be asymmetrically distributed).

2 The Model

A monopolistic platform is in the business of bringing together agents from two sides of a market. Each side $k, l \in \{A, B\}$ is populated by a unit-mass continuum of agents indexed by $i, j \in [0, 1]$. Each agent i from each side k has a type $\theta_k^i = (\mathbf{u}_k^i, v_k^i) \in \Theta_k \equiv \mathbf{U}_k \times V_k$ that has two components. The first component $\mathbf{u}_k^i$ is a vector of individual characteristics that determines the attractiveness of agent i as seen from the eyes of each agent on side $l \neq k$. The second component v_k^i is a parameter that controls for agent i's willingness to pay for the quality of the set of agents from side l he is matched to. The support of $\mathbf{u}_k^i$ is some arbitrary set $\mathbf{U}_k$ which can assume discrete or continuous values on each of its dimensions. In contrast, the support of v_k^i is the real interval $V_k \equiv [\underline{v}_k, \overline{v}_k] \subseteq \mathbb{R}$. To accommodate the case where agent i dislikes interacting with agents from side l (negative externalities), we allow the support of v_k^i to take negative values.

In the cable TV example, let viewers belong to side A and channels to side B. In this case, $\mathbf{u}_A^i$ contains information about demographics, income, and educational background of viewer $i \in A$, whereas $\mathbf{u}_B^j$ contains information about the quality of the shows as well as the type of advertisement offered by channel $j \in B$. In this example, v_A^i then captures viewer i's willingness to pay for a better package of channels, while v_B^j stands for the marginal value of an extra viewer to channel j (reflecting an increase in advertisement revenues, for example).

Let $\sigma_k(\mathbf{u}_l^j)$ denote the interaction quality that each agent from side k obtains from being matched to an agent from side l with characteristics $\mathbf{u}_l^j$. The function $\sigma_k : \mathbf{U}_l \to \mathbb{R}_{++}$ thus maps the characteristics of an agent from side l to the interaction quality enjoyed by each agent on side k. For any given (Lebesgue measurable) set of agents $\mathbf{s}$ from side l with type profile $(\theta_l^j)_{j \in \mathbf{s}}$, we then denote by

$$|\mathbf{s}|_k = \int_{j \in \mathbf{s}} \sigma_k(\mathbf{u}_l^j) d\lambda(j),$$

the total *quality* of of the set from the eyes of each agent on side k, where $\lambda(\cdot)$ is the Lebesgue measure.

Given any type profile $\theta \equiv (\theta_k^i)_{k=A,B}^{i \in [0,1]}$, the payoff enjoyed by each agent i on each side k when

matched, at a price p, to a set $\mathbf{s}$ of agents from side l is given by

$$\pi_k^i(\mathbf{s}, p; \theta) \equiv v_k^i \cdot g_k\left(|\mathbf{s}|_k\right) - p,$$

where $g_k(\cdot)$ is a strictly increasing and weakly concave bounded function. Importantly, note that the parameter v_k^i summarizes all the information contained in agent i's type that is relevant for agent i's preferences.

The type $\theta_k^i = (\mathbf{u}_k^i, v_k^i)$ of each agent i from each side k is an independent drawn from the distribution F_k with support on Θ_k. Given F_k, we then denote by F_k^v the marginal distribution of F_k with respect to v_k and we assume that F_k^v is absolutely continuous with density f_k^v. As is standard in the mechanism design literature, we also assume that F_k^v is *regular* in the sense of Myerson (1981), meaning that the virtual values $v_k - \frac{1-F_k^v(v_k)}{f_k^v(v_k)}$ are strictly increasing.

Lastly, we assume that the distribution F_k satisfies the following affiliation condition.

Condition 1 ***(Affiliation)*** *The distribution F_k is such that $(\sigma_l(\tilde{\mathbf{u}}_k), \tilde{v}_k)$ are positively affiliated.*

In the cable TV example, the assumption of positive affiliation has two implications: first, channels that are willing to pay more for viewers (e.g., because their advertisers are willing to pay more) can afford better shows and more pleasant advertisement. Second, those viewers who are willing to pay more for the packages of channels they receive are the ones that the channels value the most (e.g., because these are the viewers preferred by the advertisers).

The following examples describe two important special cases of the preferences structure outlined above.

Example 1 ***(Linear Network Externalities for Quantity)*** *Suppose that agents from side k only care about the total mass of agents from side l they are matched to. In this case, $g_k(x) = x$ and $\sigma_k(\cdot) \equiv 1$ for $k \in \{A, B\}$, so that, $\pi_k^i(\mathbf{s}, p; \theta) \equiv v_k^i \cdot \lambda\left(\mathbf{s}\right) - p$.*

These preferences are the ones typically considered in the two-sided markets literature (e.g., Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2006) and Weyl (2010)).

Example 2 ***(Supermodular Matching Values)*** *Let $\mathbf{u}_k$ be a one-dimensional random variable identical to v_k, and suppose that $g_k(x) = x$ and $\sigma_k(\mathbf{u}_k) \equiv \sigma_k(v_k) = v_k$ for $k \in \{A, B\}$. The match between agent i from side k and agent j from side l produces a surplus of $v_k^i \cdot v_l^j$ to each of the two agents. As such, the total payoff that each agent i from side k obtains from being matched to a set $\mathbf{s}$ of agents from side l is given by $\pi_k^i(\mathbf{s}, p; \theta) = v_k^i \cdot \int_{j \in \mathbf{s}} v_l^j d\lambda\left(j\right) - p$.*

This production function appears, for example, in Damiano and Li (2007), Hoppe, Moldovanu and Sela (2009), as well as in the assignment/search literature (e.g., Becker (1973), Lu and McAfee (1996) and Shimer and Smith (2000)).

Matching mechanisms

A direct revelation mechanism consists of a *matching rule* $\left\{\hat{\mathbf{s}}_k^i(\cdot)\right\}_{k=A,B}^{i\in[0,1]}$ along with a *payment rule* $\left\{\hat{p}_k^i(\cdot)\right\}_{k=A,B}^{i\in[0,1]}$ such that, for any given type profile $\theta \equiv (\theta_k^i)_{k=A,B}^{i\in[0,1]}$, $\hat{\mathbf{s}}_k^i(\theta)$ represents the set of agents from side $l \neq k$ that are matched to agent i from side k, whereas $\hat{p}_k^i(\theta)$ denotes the payment made by agent i to the platform (i.e., to the match maker).[3]

A matching rule is feasible if and only if the following *reciprocity condition* holds: whenever agent j from side B belongs to the matching set of agent i from side A, then agent i belongs to j's matching set. Formally:

$$j \in \hat{\mathbf{s}}_A^i(\theta) \Leftrightarrow i \in \hat{\mathbf{s}}_B^j(\theta). \tag{1}$$

Because there is no aggregate uncertainty and because individual identities are irrelevant for payoffs, without any loss of optimality, hereafter we will restrict attention to *anonymous* mechanisms. In these mechanisms, the composition (i.e., the cross-sectional type distribution) of the matching set that each agent i from each side k receives, as well as the payment by agent i, depend only on agent i's reported type as opposed to the entire collection of reports θ by all agents (whose distribution coincides with F by the analog of the law of large numbers for a continuum of random variables). Furthermore, any two agents i and i' (from the same side) reporting the same type are matched to the same set and are required to make the same payments. Formally, an anonymous mechanism $M = \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ can be fully described by means of a pair of matching rules and a pair of payment rules such that, for any $\theta_k \in \Theta_k$, $p_k(\theta_k)$ is the payment made by each agent on side k reporting a type θ_k, whereas $\mathbf{s}_k(\theta_k) \subset \Theta_l$ is the set of types from side l to which each agent i from side k is matched to when reporting type θ_k. Note that $\mathbf{s}_k$ maps Θ_k into a sigma algebra over Θ_l. With some abuse of notation, hereafter we will then denote by $|\mathbf{s}_k(\theta_k)|_k$ the total quality of the matching set assigned to each agent i on side k reporting a type θ_k. Also note that, by virtue of reciprocity, $\theta_l \in \mathbf{s}_k(\theta_k)$ implies that $\theta_k \in \mathbf{s}_l(\theta_l)$. As such, a matching rule can be fully described by its side-k correspondence $\mathbf{s}_k(\cdot)$. By the Revelation Principle, we will restrict attention to direct revelation mechanisms which are *individually rational* (IR) and *incentive compatible* (IC). Denote by $\hat{\Pi}_k(\theta_k, \hat{\theta}_k; M) \equiv v_k^i \cdot g_k(|\mathbf{s}_k(\hat{\theta}_k)|_k) - p_k(\hat{\theta}_k)$ the payoff that type $\theta_k = (\mathbf{u}_k, v_k)$ obtains when reporting a type $\hat{\theta}_k = (\hat{\mathbf{u}}_k^i, \hat{v}_k^i)$, and by $\Pi_k(\theta_k; M) \equiv \hat{\Pi}_k(\theta_k, \theta_k; M)$ the payoff that type θ_k obtains by reporting truthfully. In our setting, a mechanism M is

$$\text{IR if} \qquad \Pi_k(\theta_k; M) \geq 0 \text{ for all } \theta_k \in \Theta_k, \tag{2}$$

$$\text{IC if} \qquad \Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M) \text{ for all } \theta_k, \hat{\theta}_k \in \Theta_k. \tag{3}$$

[3] To simplify notation, we do not allow the platform to randomize across matching sets, that is, we restrict attention to deterministic mechanisms. Deterministic mechanisms can be shown to be optimal in a more general model where stochastic matching rules are allowed. The proof is available upon request by the authors.

A matching rule $\mathbf{s}_k(\cdot)$ is *implementable* if there is a payment rule $\{p_k(\cdot)\}_{k=A,B}$ such that the mechanism $M = \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ satisfies the IR and IC constraints (2) and (3). Implicit in the aforementioned specification is the assumption that the platform must charge the agents before they observe their payoff. This seems a reasonable assumption in most applications of interest. Without such an assumption, the platform could extract all surplus and implement the efficient matching rule by using payments similar to those in Cremer and McLean (1988) – see also Mezzetti (2007).

3 Efficiency and Profit-Maximization

We start by defining what we mean by "efficient" and "profit-maximizing" mechanisms. For any given type profile θ, the welfare generated by the mechanism M is given by

$$\Omega^W(M) = \sum_{k=A,B} \int_0^1 v_k^i \cdot g_k(\left|\hat{\mathbf{s}}_k^i(\theta)\right|_k) d\lambda(i) = \sum_{k=A,B} \int_{\Theta_k} v_k \cdot g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k),$$

whereas the expected profits generated by the mechanism M are given by

$$\Omega^P(M) = \sum_{k=A,B} \int_0^1 \hat{p}_k^i(\theta) d\lambda(i) = \sum_{k=A,B} \int_{\Theta_k} p_k(\mathbf{u}_k, v_k) dF_k(\mathbf{u}_k, v_k)$$

Because there is no aggregate uncertainty, a mechanism M^W (respectively, M^P) is then said to be efficient (respectively, profit-maximizing) if it maximizes $\Omega^W(M)$ (respectively, $\Omega^P(M)$) among all mechanisms that are individually rational and incentive compatible, that is, among all mechanisms M that satisfy (2) and (3) above.

3.1 Preliminaries

Our first result provides necessary and sufficient conditions for a mechanism M to be individually rational and incentive compatible.

Lemma 1 *A mechanism M is individually rational and incentive compatible if and only if the following conditions jointly hold.*

1. *for all $\theta_k = (\mathbf{u}_k, v_k)$ and $\theta_k' = (\mathbf{u}_k', v_k')$, $v_k = v_k'$ implies that $\Pi_k(\theta_k; M) = \Pi_k(\theta_k'; M)$;*
2. *for all $\theta_k = (\mathbf{u}_k, v_k)$ and $\theta_k' = (\mathbf{u}_k', v_k')$, $v_k > v_k'$ implies that $g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) \geq g_k\left(|\mathbf{s}_k(\mathbf{u}_k', v_k')|_k\right)$; furthermore, except for a countable subset of V_k,*

$$g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) = g_k\left(\left|\mathbf{s}_k(\mathbf{u}_k', v_k)\right|_k\right)$$

for all $\mathbf{u}_k, \mathbf{u}_k' \in \mathbf{U}_k$;

3. expected payments are given by

$$p_k(\mathbf{u}_k, v_k) = v_k \cdot g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) - \Pi_k((\mathbf{u}_k, \underline{v}_k); M) - \int_{\underline{v}_k}^{v_k} g_k\left(|\mathbf{s}_k(\mathbf{u}_k, x)|_k\right) dx$$

for all $(\mathbf{u}_k, v_k) \in \Theta_k$, *with* $\Pi_k((\mathbf{u}_k, \underline{v}_k); M) \geq 0$.

To understand the result in Lemma 1, recall that agents' payoffs do not depend directly on their own characteristics $\mathbf{u}_k$, but only on the characteristics of those agents they are matched with. Therefore, incentive-compatible mechanisms have to deliver identical payoffs to all agents who share the same preference for quality v_k but have different characteristics $\mathbf{u}_k$. This result, however, does not mean that the platform cannot condition the value of the matching set $g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right)$ on individual characteristics $\mathbf{u}_k$. By designing the payment scheme appropriately, the platform can in fact preserve the indifference condition required by part 1 while letting the value of the matching set vary with $\mathbf{u}_k$. Condition 2 in the lemma establishes that this can be done at most over a countable subset of V_k. To see this, note that, because payoffs satisfy the increasing difference property between v_k and g_k, incentive compatibility requires that the quality of the matching set be nondecreasing in v_k. In turn, this implies that the expected quality $\mathbb{E}[g_k\left(|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)|_k\right)]$ must be nondecreasing in v_k, where the expectation is with respect to $\tilde{\mathbf{u}}_k$ given v_k Now at any point $v_k \in V_k$ at which $g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right)$ depends on $\mathbf{u}_k$, the expectation $\mathbb{E}[g_k\left(|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)|_k\right)]$ is necessarily discontinuous in v_k. Because monotone functions can be discontinuous at most over a countable set of points, this means that the value of the matching set may vary with the characteristics $\mathbf{u}_k$ only over a countable subset of V_k. Note, however, that this result pertains the value g_k of the matching set and not the matching set s_k itself (matching sets may vary with $\mathbf{u}_k$ over a continuous subset of V_k) implying that the platform may indeed find it useful to condition matching sets on characteristics other than willingness to pay.

The remaining condition on payments is the familiar envelope condition which pins down the payments from the matching rule, up to a scalar. An immediate implication of Lemma 1 is that a matching rule $\mathbf{s}_k(\cdot)$ is implementable if and only if it satisfies condition 2 above.

Using Lemma 1, we then have that, under any individually rational and incentive compatible mechanism M that gives zero surplus to each agent reporting the lowest willingness to pay, welfare, as well as the platform's profits, can be conveniently represented as follows

$$\Omega^h(M) = \sum_{k=A,B} \int_{\Theta_k} \varphi_k^h(v_k) \cdot g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k) \tag{4}$$

where $\varphi_k^W(v_k) = v_k$ if $h = W$ (i.e., in the case of welfare) and $\varphi_k^P(v_k) = v_k - \frac{1-F_k^v(v_k)}{f_k^v(v_k)}$ if $h = P$ (i.e., in the case of profits).[4] Hereafter, we will denote by $\{\mathbf{s}_k^h(\cdot)\}_{k=A,B}$ the matching rule that maximizes

[4]It is immediate to see that restricting attention to mechanisms that give zero surplus to each agent reporting the lowest willingness to pay is without loss of optimality both in case of profit maximization as well as in case of welfare maximization (subject to budget balance).

(4). We will then assume that $\bar{v}_k > 0$ for $k = A, B$, thus guaranteeing that an empty network is never optimal. For future reference, we also define the *reservation value* r_k^h as the unique solution to $\varphi_k^h(v_k) = 0$ whenever this equation has a solution.

3.2 Single network vs nested multi-homing

We now describe two important classes of matching rules: single- and multi-homing matching rules. Under a single-homing matching rule, agents on both sides of the market are assigned to mutually exclusive networks. As such, if two agents from side k are both matched to the same agent on side l (again, because identities play no role in our model, this formally means that they are matched to the same type θ_l on side l) then this means that their matching sets are the same, a property which is not imposed in case of multi-homing.

Definition 1 *A matching rule $\mathbf{s}_k(\cdot)$ exhibits single-homing if for all $\theta_k, \theta_k' \in \Theta_k$*

$$\mathbf{s}_k(\theta_k) \cap \mathbf{s}_k(\theta_k') \neq \oslash \Rightarrow \mathbf{s}_k(\theta_k) = s_k(\theta_k').$$

Single-homing matching rules can be implemented by offering the agents access to mutually exclusive networks and by charging appropriate fees for the different networks (we will come back to a precise description of the fees at the end of the document). Single-homing matching rules are at the heart of the two-sided markets literature (e.g., Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2006), Ambrus and Argenziano (2009) and Weyl (2010)). This literature studies optimal pricing by platforms that are restricted to use single-homing matching rules. Part of the contribution of the analysis here is to derive conditions under which such rules are optimal.

A particularly simple type of single-homing matching rule is one that employs a single network.

Definition 2 *A matching rule $\mathbf{s}_k(\cdot)$ employs a single network if for all $\theta_k, \theta_k' \in \Theta_k$*

$$\mathbf{s}_k(\theta_k), \mathbf{s}_k(\theta_k') \neq \oslash \Rightarrow \mathbf{s}_k(\theta_k) = s_k(\theta_k').$$

In contrast, under a multi-homing matching rule, the platform establishes a certain number of non-exclusive networks and allows agents on each side to join multiple networks. Of particular interest are nested multi-homing rules. Under these rules, if agents i_1 and i_2 from side k commonly meet agent j from side l, then the matching sets of agents i_1 and i_2 are nested.

Definition 3 *A matching rule $\mathbf{s}_k(\cdot)$ exhibits nested multi-homing if for all $\theta_k, \theta_k' \in \Theta_k$*

$$\mathbf{s}_k(\theta_k) \cap \mathbf{s}_k(\theta_k') \neq \oslash \Rightarrow \mathbf{s}_k(\theta_k) \subseteq \mathbf{s}_k(\theta_k') \text{ or } \mathbf{s}_k(\theta_k) \supseteq \mathbf{s}_k(\theta_k'),$$

where the inclusion is strict for some $\theta_k, \theta_k' \in \Theta_k$.

An example of a nested multi-homing matching rule is given by the cable TV application discussed above. In this case, the provider offers (mutually non-exclusive) packages that can be added to a "standard plan" and which grant access to extra channels. Multi-homing matching rules are also pervasive in online advertising (where advertisers can "buy" access to an increasing set of browsers) and health-care provision (where patients enroll in health plans that include different sets of doctors and hospitals). More generally, multi-homing is the equivalent in matching environments to quantity/quality price discrimination under standard (single-market) monopolistic screening.

The following proposition identifies important properties of optimal matching rules.

Proposition 1 *Let $h = P$ in case of profit-maximization and $h = W$ in case of welfare maximization. The h-optimal mechanism M^h does not depend on the vector of characteristics $\mathbf{u}_k$, $k = A, B$. Suppressing the dependence on $\mathbf{u}_k$, $k = A, B$, the h-optimal matching rule $\mathbf{s}_k^h(\cdot)$ then has the following threshold structure*

$$\mathbf{s}_k^h(v_k) = \begin{cases} [t_k^h(v_k), \overline{v}_l] \text{ if } v_k \in [\omega_k, \overline{v}_k] \\ \oslash \text{ otherwise} \end{cases} \tag{5}$$

where the threshold $\omega_k \in [\underline{v}_k, \overline{v}_k]$ determines which types are excluded and where the nonincreasing function $t_k^h(\cdot)$ determines the matching sets.

The result that optimal matching rules have the threshold structure outlined in the proposition hinges on two important assumptions: the weak concavity of the externality function $g_k(\cdot)$ and the positive affiliation of $(\sigma_l(\mathbf{u}_k), v_k)$. To understand the result, consider an agent with type $\theta_k = (\mathbf{u}_k, v_k)$ with $\varphi_k^h(v_k) \geq 0$. Ignoring for a moment the monotonicity constraints, it is easy to see that it is always optimal to assign to this type a matching set $\mathbf{s}_k(\mathbf{u}_k, v_k) \supset \{\theta_l = (\mathbf{u}_l, v_l) : \varphi_l^h(v_l) \geq 0\}$ that includes all types $\theta_l = (\mathbf{u}_l, v_l)$ whose φ_l^h-valuation is non-negative. This is because, (i) irrespective of their $\mathbf{u}_l$ characteristics, these types contribute positively to type θ_k's payoff (recall that $\sigma_k(\mathbf{u}_l) \geq 0$ for all $\mathbf{u}_l$) and (ii) these types have a non-negative φ_l^h-valuation, and therefore adding type θ_k to these types' matching sets never reduces the platform's payoff $\Omega^h(M)$, as shown by (4). Now imagine that the platform wants to assign to this type θ_k a matching set $\mathbf{s}_l$ whose intrinsic quality q is higher than the quality of the set of types on side l whose φ_l^h-valuation is non-negative, i.e., such that

$$|\mathbf{s}_l|_k = q > \int_{\{(\mathbf{u}_l, v_l): \varphi_l^h(v_l) \geq 0\}} \sigma_k(\mathbf{u}_l) dF_l(\mathbf{u}_l, v_l)$$

Because of reciprocity, adding an agent whose φ_l^h-valuation is negative to type θ_k's matching set now comes at a cost, through its negative effect on type θ_l's payoff. In this case, the assumption that $(\sigma_k(\mathbf{u}_l), v_l)$ are positively affiliated along with the assumption that the externality function $g_k(\cdot)$ is weakly concave imply that the least costly way to provide type θ_k with a matching set of quality q is by matching him with all types θ_l whose φ_l^h-valuation is the least negative, irrespective of their

$\mathbf{u}_l$ characteristics. This means that type θ_k's matching set takes the form $\mathbf{U}_l \cup [t_k^h(v_k), \overline{v}_l]$ where the threshold $t_k(v_k)$ is computed so that

$$\int_{\{(\mathbf{u}_l, v_l): v_l \in [t_k(v_k), \overline{v}_l]\}} \sigma_k(\mathbf{u}_l) dF_l(\mathbf{u}_l, v_l)_k = q.$$

Monotonicity of the matching quality in the willingness to pay v_k, as required by incentive compatibility, then implies that the threshold function $t_k^h(\cdot)$ is nonincreasing. The proof in the Appendix formalizes the heuristics described above.

The following corollary is a direct implication of Proposition 1.

Corollary 1 *The optimal mechanism M^h either employs a single network or exhibits nested multi-homing.*

Given the result in Proposition 1, from now on, we restrict attention to mechanisms whose matching rule takes the form given in (5). Under such mechanisms, the quality of the matching set that each agent i on each side k obtains when reporting a type $\theta_k = (\mathbf{u}_k, v_k)$ is given by

$$g_k\left(\left|\mathbf{s}_k^h(\mathbf{u}_k, v_k)\right|_k\right) = \hat{g}_k(t_k(v_k))$$

where the function $\hat{g}_k : V_l \to \mathbb{R}_+$ is given by

$$\hat{g}_k(v_l) \equiv g_k\left(\int_{v_l}^{\overline{v}_l} \int_{\mathbf{U}_l} \sigma_k(\mathbf{u}_l) \cdot dF_l(\mathbf{u}_l, v)\right).$$

The platform's objective then rewrites as

$$\Omega^h(M) = \sum_{k=A,B} \int_{\omega_k}^{\overline{v}_k} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k). \tag{6}$$

The next proposition provides a necessary and sufficient condition for the h-optimal mechanism to employ a single network or to exhibit nested multi-homing.

Proposition 2 *The h-optimal matching rule employs a single (complete) network if*

$$\triangle^h \equiv \sum_{k=A,B} g_k'(|\Theta_l|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l) | \tilde{v}_l = \underline{v}_l] \cdot \varphi_k^h(\underline{v}_k) \geq 0,$$

and exhibits nested multi-homing otherwise.

To see why the result above is true, consider first the case where $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$ implying that $\triangle^h \geq 0$. Because valuations (or virtual valuations) are all nonnegative, one can easily see from (6), that efficiency (resp. profits) are maximized by matching each agent on each side to all agents on the other side, meaning that the optimal matching rule employs a single network which includes all agents (i.e., a complete network).

Next, consider the case where $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, so that $\triangle^h < 0$. To see why in this case nested multi-homing is h-optimal, suppose instead that the platform were to use a single network and let $\hat{\omega}_k$ denote the threshold type on side k so that agents on this side are excluded if and only if $v_k < \hat{\omega}_k$. We will argue that for a single network to be optimal it must be that $\hat{\omega}_k \leq r_k^h$ for $k = A, B$, where the reservation value r_k^h is given by the unique solution to $\varphi_k^h(r_k^h) = 0$. Indeed, suppose, towards a contradiction, that for $k \in \{A, B\}$, $\hat{\omega}_k > r_k^h$ so that $\varphi_k^h(\hat{\omega}_k) > 0$. The platform could then improve upon its payoff by employing a nested multi-homing matching rule that assigns to each agent on side k with valuation $v_k \geq \hat{\omega}_k$ the same matching set as the original matching rule while it assigns to each agent with valuation $v_k \in [r_k^h, \hat{\omega}_k]$ the matching set $[\hat{v}_l^{\#}, \overline{v}_l]$, where $\hat{v}_l^{\#} \equiv \max\{r_l^h, \hat{\omega}_l\}$. The new matching rule is clearly monotone (and hence implementable) and gives the platform a higher h-payoff than the original mechanism. Hence, for a single network to be optimal, it must be that $\hat{\omega}_k \leq r_k^h$ for $k = A, B$.

Now suppose that $\hat{\omega}_k < r_k^h$ for $k = A, B$. Starting from this single network, the platform could then increase her payoff generated from side k by switching to a nested multi-homing rule $\mathbf{s}_k^{\#}(\cdot)$ such that

$$\mathbf{s}_k^{\#}(v_k) = \begin{cases} [\hat{\omega}_l, \overline{v}_l] & \Leftrightarrow & v_k \in [r_k^h, \overline{v}_k] \\ [r_l^h, \overline{v}_l] & \Leftrightarrow & v_k \in [\hat{\omega}_k, r_k^h] \\ \oslash & \Leftrightarrow & v_k \in [\underline{v}_k, \hat{\omega}_k] \end{cases}.$$

The new matching rule strictly improves upon the original one because it eliminates all matches between agents whose valuations (or virtual valuations) are both negative. Next, suppose that $\hat{\omega}_k = r_k^h$ for $k \in \{A, B\}$. The platform could then do better by lowering the threshold type and switching to the following multi-homing matching rule

$$\mathbf{s}_k^{\#}(v_k) = \begin{cases} [\hat{\omega}_l, \overline{v}_l] & \Leftrightarrow & v_k \in [r_k^h, \overline{v}_k] \\ [r_l^h, \overline{v}_l] & \Leftrightarrow & v_k \in [\hat{\omega}_k^{\#}, r_k^h] \\ \oslash & \Leftrightarrow & v_k \in [\underline{v}_k, \hat{\omega}_k^{\#}] \end{cases}$$

By setting the new exclusion threshold $\hat{\omega}_k^{\#}$ sufficiently close to r_k^h the platform increases her payoff. To see this, note that, starting from $\hat{\omega}_k^{\#} = r_k^h$, the marginal effect of decreasing the threshold $\hat{\omega}_k^{\#}$ by $\varepsilon > 0$ small enough is proportional to

$$-\hat{g}_k(\hat{\omega}_l) \cdot \varphi_k^h(r_k^h) f_k^v(r_k^h) - \hat{g}_l'(r_k^h) \int_{r_l^h}^{\overline{v}_l} \varphi_l^h(v_l) dF_l^v(v_l) = -\hat{g}_l'(r_k^h) \int_{r_l^h}^{\overline{v}_l} \varphi_l^h(v_l) dF_l^v(v_l) > 0 \tag{7}$$

where the equality follows from the fact that $\varphi_k^h(r_k^h) = 0$ whereas the inequality follows from the fact that $\hat{g}_l'(\cdot) < 0$. In words, the benefit of offering a matching set of higher quality to those agents on side l whose φ_l^h-valuation is positive more than offsets the cost of getting on board a few more agents on side k whose φ_k^h-valuation is negative. Note that for this network expansion to be profitable, it is essential that the platform assigns the newly added agents on side k only to those users on side l

with positive valuation, which requires employing a multi-homing matching rule. We conclude that when $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, the h-optimal mechanism necessarily exhibits nested multi-homing.

In the Appendix, we complete the proof by analyzing the remaining case where $\varphi_l^h(\underline{v}_l) < 0$ while $\varphi_k^h(\underline{v}_k) \geq 0$.

An interesting implication of Proposition 2 is that, since $\triangle^W > \triangle^P$, multi-homing matching rules are more often employed by profit-maximizing platforms, rather than by welfare-maximizing platforms. This is illustrated by the next two examples.

Example 3 *Consider the case of linear network externalities, as described in Example 1 above, and assume that values v_k are uniformly distributed over $[\underline{v}_k, \overline{v}_k]$. Then, the welfare-maximizing matching rule employs a single network if*

$$\triangle^W = \underline{v}_A + \underline{v}_B \geq 0,$$

and exhibits nested multi-homing otherwise. In turn, the profit-maximizing matching rule employs a single network if

$$\triangle^P = 2\left(\underline{v}_A + \underline{v}_B\right) - \left(\overline{v}_A + \overline{v}_B\right) = \triangle^W - \left[\left(\overline{v}_A - \underline{v}_A\right) + \left(\overline{v}_B - \underline{v}_B\right)\right] \geq 0,$$

and exhibits nested multi-homing otherwise.

Example 4 *Consider the case of supermodular matching values, as described in Example 2 above, and assume that values v_k are uniformly distributed over $[\underline{v}_k, \overline{v}_k]$ with $\underline{v}_k > 0$. Then, the welfare-maximizing matching rule employs a single network, since*

$$\triangle^W = 2 \cdot \underline{v}_A \cdot \underline{v}_B > 0.$$

In turn, the profit-maximizing matching rule employs a single network if

$$\triangle^P = \sum_{k=A,B} \underline{v}_l \cdot \left(2\underline{v}_k - \overline{v}_k\right) = \triangle^W \cdot \frac{1}{2} \cdot \left[\left(2 - \frac{\overline{v}_A}{\underline{v}_A}\right) + \left(2 - \frac{\overline{v}_B}{\underline{v}_B}\right)\right] \geq 0,$$

and exhibits nested multi-homing otherwise.

3.3 Properties of optimal multi-homing rules

We now further investigate the properties of optimal matching rules when nested multi-homing is optimal, that is, when $\triangle^h < 0$. In order to do so, we use the characterization of Proposition 1 to rewrite the feasibility constraint in terms of the threshold functions that describe the optimal matching rule. Because $t_A(\cdot), t_B(\cdot)$ are weakly decreasing, reciprocity requires that

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{8}$$

that is, the threshold $t_k(v_k)$ of an agent with valuation v_k is the smallest value v_l on side l whose matching set contains v_k.

The platform's problem then consists of maximizing the objective (6) subject to the feasibility constraint (8) and the monotonicity condition, required by incentive compatibility, that $t_A(\cdot), t_B(\cdot)$ be weakly decreasing.

The next definition extends to the present two-sided matching model the notion of separating schedules, as it appears in Maskin and Riley (1984).

Definition 4 *The h-optimal matching rule $s_k^h(\cdot)$ is said to be maximally separating if $t_k^h(\cdot)$ is strictly decreasing in $[\omega_k^h, t_l^h(\omega_l^h)]$. If $\omega_k^h > \underline{v}_k$, we say that the h-optimal matching rule exhibits exclusion at the bottom on side k. In turn, if $t_l^h(\omega_l^h) < \bar{v}_k$ we say that the h-optimal matching rule exhibits bunching at the top on side k.*

When there is bunching at the top on side k, all agents with valuation $v_k \in [t_l(\omega_l^h), \bar{v}_k]$ obtain the same matching set $[\omega_l^h, \bar{v}_l]$. Reciprocity then implies that maximally separating matching rules can only exhibit bunching at the top when all agents on side l with valuation above ω_l^h are assigned to some agent on side k with valuation $v_k < \bar{v}_k$.

Unlike in one-sided price discrimination problems, assuming that virtual values are monotone is not enough to guarantee that the optimal matching rule is maximally separating. The next condition strengthens the standard monotonicity of virtual values, as required in a two-sided matching environment.

Condition 2 ***(Strong Regularity)*** *The function*

$$\psi_k^h(v_k) \equiv \frac{f_k^v(v_k) \cdot \varphi_k^h(v_k)}{-\hat{g}_l'(v_k)} = \frac{\varphi_k^h(v_k)}{g_l'\left(|\mathbf{U}_k \times [v_k, \bar{v}_k]|_l\right) \cdot \mathbb{E}\left[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k\right]}$$

is strictly increasing for all $v_k \in [\underline{v}_k, \overline{v}_k]$.

Take the case of profit-maximization, where $h = P$. The numerator on the expression above accounts for the effect as a *consumer* of an agent from side k with valuation v_k on the platform's revenue (as his virtual valuation $\varphi_k^h(v_k)$ is proportional to the marginal revenue produced by this agent). In turn, the denominator accounts for the expected effect of this agent as an *input* on the platform's revenue (as $-\hat{g}_l'(v_k)$ is proportional to the marginal utility brought by this agent to every agent on side l who is matched to him). The assumption of positive affiliation between v_k and $\sigma_l(\mathbf{u}_k)$, together with the assumption of weak concavity of $g_l(\cdot)$, imply that the denominator of $\psi_k^h(v_k)$ is a strictly increasing function of v_k. The strong regularity condition above then requires that the value of agents as consumers increases faster than their value as inputs, as their valuation v_k grows. In the linear model, $\psi_k^h(v_k) = \frac{\varphi_k^h(v_k)}{\mathbb{E}[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k]}$, which is strictly increasing provided that the positive affiliation between $\sigma_l(\mathbf{u}_k)$ and v_k is not "too strong".

The next proposition characterizes the h-optimal matching rule when the platform's problem is regular in the sense of Condition 2.

Proposition 3 *Let $\triangle^h < 0$ and assume Condition 2 holds. Then the h-optimal matching rule $s_k^h(\cdot)$ is maximally separating. Define*

$$\triangle_k^h(v_k, v_l) \equiv -\hat{g}_k'(v_l) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l'(v_k) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l).$$

The h-optimal matching rule $s_k^h(\cdot)$ is such that if $\triangle_k^h(\bar{v}_k, \underline{v}_l) > 0$, there is bunching at the top on side k and no exclusion at the bottom on side l. In turn, if $\triangle_k^h(\bar{v}_k, \underline{v}_l) < 0$, there is exclusion at the bottom on side l and no bunching at the top on of side k. Else, in the knife-edge case where $\triangle_k^h(\bar{v}_k, \underline{v}_l) = 0$, there is neither bunching at the top on side k nor exclusion at the bottom on side l.

Moreover, for all values v_k in the separating range, the h-optimal threshold $t_k^h(\cdot)$ satisfies the Euler equation

$$\triangle_k^h(v_k, t_k^h(v_k)) = 0, \tag{9}$$

which yields $t_k^h(v_k) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v_k)\right)$.

Assume $\underline{v}_k < 0$, $k = A, B$. An important feature of the maximally separating h-optimal rule described above is that $t_k^h(v_k) \leq r_l^h$ if and only if $v_k \geq r_k^h$. In the case of profit-maximization, this means that agents with positive virtual values on side k are matched to all agents with positive virtual values on side l, plus a measure of agents with negative virtual values on side l (cross-subsidization). The optimal level of cross-subsidization for an agent with $\varphi_k^h(v_k) > 0$ is determined by the Euler equation, which equalizes the marginal benefits and the marginal costs of enlarging the matching sets. The term $-\hat{g}_k'(t_k^h(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k)$ captures the marginal gain on side k (in terms of efficiency or revenue), while the term $-\hat{g}_l'(v_k) \cdot \varphi_l^h(t_k^h(v_k)) \cdot f_l^v(t_k^h(v_k))$ captures the associated marginal losses on side l from further decreasing $t_k^h(v_k)$ (recall that $\varphi_l^h(t_k^h(v_k)) < 0$ for any $v_k > r_k^h$). At the optimum, these two effects must balance each other, as implied by (9).

It is also worth noticing that optimality implies that there is bunching at the top on side k if and only if there is no exclusion at the bottom on side l. This means that bunching can only occur at the optimum due to binding capacity constraints, that is, when the "stock" of agents on side l has been exhausted. This is illustrated in the next example.

Example 5 ***(Patients and doctors with linear network externalities for quantity)*** *Let the platform be a health insurance company that provides patients on side A with access to doctors on side B. Each patient only cares about the number of doctors included in his health plan, while doctors only care about the number of patients they assist. Accordingly, the environment features linear network externalities, as defined in Examples 1 and 3. Patients have values for an extra doctor drawn from a uniform distribution on $[1, \frac{3}{2}]$, while doctors face costs (negative values) of treating an extra patient*

drawn from a uniform distribution on $[-1, 0]$. *Since* $\triangle^W = 0$, *the welfare-maximizing (say, public) provision of health insurance entails the adoption of a single (complete) network where all patients have access to all doctors. In turn, since* $\triangle^P = -\frac{3}{2}$, *the profit-maximizing (say, private) provision of health insurance entails the adoption of a nested multi-homing matching rule. It follows from (9) that this rule is described by the threshold function* $t_A^P(v_A) = \frac{3}{4} - v_A$ *defined over* $v_A \in [1, \frac{3}{2}]$. *Under profit-maximization, there is bunching at the top on side B (that is, cheap doctors are included in the network offered to all patients) and exclusion at the bottom on side B (that is, very expensive doctors are excluded from any health plan). Figure 1 depicts the situation described above.*

The example above shows that, as implied by Proposition 2, welfare-maximizing matching rules are more likely to implement single-homing matching rules. In turn, profit-maximizing platforms tend to employ nested multi-homing matching rules, since this allows for finer price discrimination. A salient feature of the health-care example is that, relative to welfare, the profit-maximizing matching rule (i) excludes more agents, and (ii) assigns matching sets which are strict subsets of the ones assigned by the efficient rule. As we discuss next, these two distortions generalize well beyond the case of linear network externalities and uniform valuations assumed in the example above.

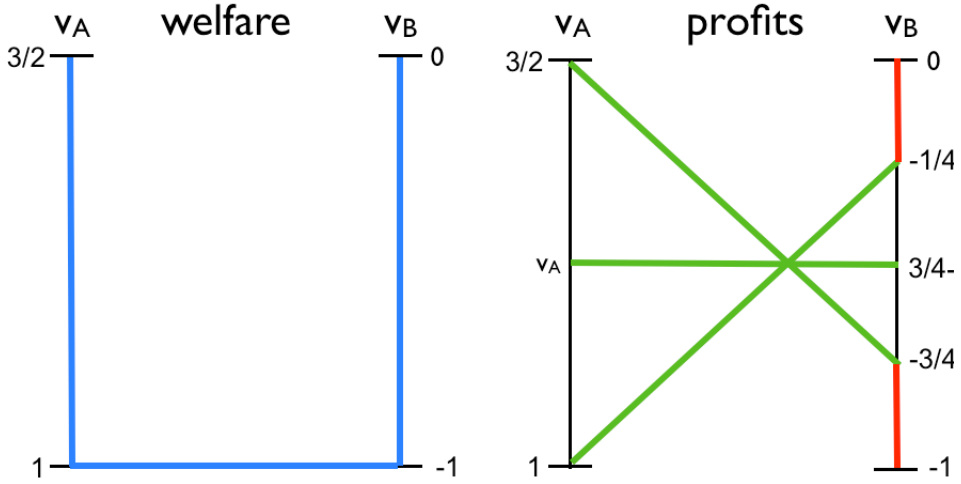

Figure 1: The health insurance plans offered by welfare-maximizing (public) and profit-maximizing (private) health-care providers when patients and doctors have valuations drawn from U[1,3/2] and U[-1,0], respectively.

3.4 Distortions Relative to Efficiency: The Exclusion and Isolation Effects

Relative to welfare maximization, profit-maximization leads to two distortions, as explained in the following proposition.

Proposition 4 *Relative to the welfare-maximizing matching rule* $\mathbf{s}_k^W(\cdot)$, *the profit-maximizing matching rule* $\mathbf{s}_k^P(\cdot)$*:*

1. *admits a weakly smaller group of agents (**exclusion effect**), i.e.,* $\omega_k^W \leq \omega_k^P$ *for* $k = A, B$*, and*

2. *matches each agent on each side of the market to a subset of agents on the other side of the market (**isolation effect**), i.e.,* $\mathbf{s}_k^P(v_k) \subseteq \mathbf{s}_k^W(v_k)$ *for all* $v_k > \omega_k^P$*,* $k = A, B$*.*

The first distortion, the *exclusion effect*, states that too many agents are excluded by the platform. This distortion is analogous to the usual downward distortion in standard monopolistic screening models.

The second distortion, the *isolation effect*, states that under profit-maximization, each agent is matched to a subset of his efficient matching set. The reason can be seen from equation (9): under profit-maximization, the platform only internalizes the cross-effects on marginal revenues (which are proportional to virtual values $\varphi_k^P(v_k)$), rather than the cross-effects on welfare (which are proportional to the true values v_k). Since virtual values are always smaller than the true values, the platform fails to internalize part of the marginal gains from new matches. This leads to smaller matching sets potentially for *all* types (including the highest types on each side of the market).

The next example illustrates the exclusion and isolation effects in the supermodular matching value case of Example 2.

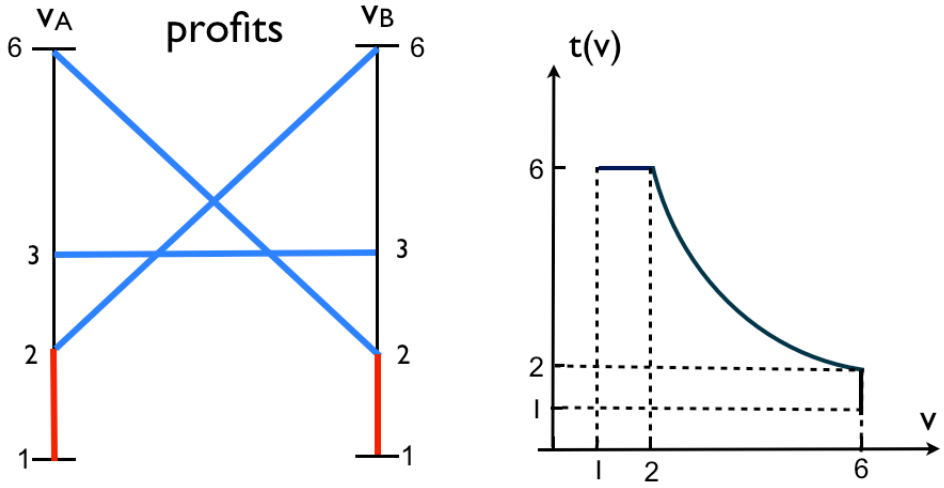

Figure 2: The menu offered by a profit-maximizing job employment agency when freelancers and firms have productivities drawn from U[1,6].

Example 6 ***(Free-lancers and firms with supermodular matching values)*** *Let the platform be an employment agency that matches free-lancers on side A to firms on side B. The match between a free-lancer and a firm generates a project whose output is* $2v_Av_B$*, which we assume is evenly split between the firm and the free-lancer. Accordingly, this environment features supermodular matching values, as defined in Examples 2 and 4. Free-lancers and firms have productivity values drawn from a uniform distribution on* $[\underline{v}, \overline{v}]$*, where* $\underline{v} > 0$ *and* $2\underline{v} < \overline{v}$*. Since* $\triangle^W = 2\underline{v}^2$*, the welfare-maximizing*

(say, public) employment agency creates a single (complete) network where all free-lancers have access to all firms. In turn, since $\triangle^P = 2\underline{v}(2\underline{v} - \bar{v}) < 0$*, the profit-maximizing (say, private) employment agency offers a menu of access plans that allows free-lancers to reach an increasing number of firms. It follows from (9) that the threshold function associated to this menu is* $t_k^P(v_k) = \frac{v_k \cdot \bar{v}}{4 \cdot v_k - \bar{v}}$ *defined over* $(\omega_k, \bar{v}) = (\frac{\bar{v}}{3}, \bar{v})$*. Under profit-maximization, there is exclusion at the bottom on both sides (that is, free-lancers and firms with low productivity develop no projects). Moreover, under profit-maximization, all free-lancers and firms develop a proper subset of their efficient number of projects. Figure 2 describes the profit-maximizing solution when* $[\underline{v}, \bar{v}] = [1, 6]$*.*

3.5 Comparative Statics under Profit-Maximization

In matching markets agents play the dual role of *consumers* ("purchasing"sets of agents from the other side of the market) and *inputs* (generating value to the agents on the other side of the market they are matched with). In our model, the consumer role of the agents on each side k is associated to the marginal distribution of the valuations F_k^v, whereas the input role is associated to the family of conditional distribution functions $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v)$ for the interaction qualities $\sigma_l(\mathbf{u}_k)$.

3.5.1 The Detrimental Effects of Becoming More Popular

Shocks that alter the cross-side effects of matches are common in two-sided markets. Changes in the income distribution of households, for example, shall affect the pricing strategies of Cable TV providers, since channels' profits change given the same set of viewers (e.g., because advertisers are willing to pay more for viewers with higher purchasing power).

The next definition uses a standard stochastic order relation to formalize the notion that the popularity (or attractiveness) of agents on a given side has changed.

Definition 5 *We say that side* k *is more popular under* $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ *than under* $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ *if, for all* v_k*,* $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ *dominates* $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ *in the usual stochastic order, and* $F_k^v = \hat{F}_k^v$*.*

The next proposition describes how the profit-maximizing matching rule changes as side k becomes more popular. Perhaps surprisingly, an increase in the popularity of agents from side k can be hurtful to agents from both sides of the market.

Proposition 5 *Consider the platform's profit-maximization problem when network effects are linear* $(g_A(x) = g_B(x) = x)$*, let* $\triangle^P < 0$*, and assume Condition 2 holds. If the popularity of side* k *increases, then the platform moves from a matching rule* $\mathbf{s}_k^P(\cdot)$ *to a matching rule* $\hat{\mathbf{s}}_k^P(\cdot)$ *such that*

1. $\hat{\mathbf{s}}_k^P(v_k) \supseteq \mathbf{s}_k^P(v_k)$ *if and only if* $v_k \leq r_k^P$*,*
2. $\hat{\mathbf{s}}_l^P(v_l) \subseteq \mathbf{s}_l^P(v_l)$ *if and only if* $v_l \leq r_l^P$*,*

3. there exists $\hat{\nu}_k \in (r_k^P, \bar{v}_k]$ *such that* $\Pi_k(\theta_k; \hat{M}^P) \geq \Pi_k(\theta_k; M^P)$ *if and only if* $v_k \leq \hat{\nu}_k$,

4. there exists $\hat{\nu}_l \in [\underline{v}_l, r_l^P)$ *such that* $\Pi_k(\theta_k; \hat{M}^P) \geq \Pi_l(\theta_l; M^P)$ *if* $v_l \geq \hat{\nu}_l$.

The proposition above shows that changes in the popularity of side k have heterogeneous effects on the matching sets and on the payoffs of agents from sides k and l. To see why this is true, note that as side k becomes more popular, the cost in terms of the negative revenue collected on side l from matching an agent with value $v_k \geq r_k^P$ to an agent with value $v_l \leq r_l^P$ increases, whereas the marginal benefit is unaltered. As a consequence, the matching sets of agents on side k with value $v_k \geq r_k^P$ shrink and, by reciprocity, the matching sets of agents on side l with $v_l \leq r_l^P$ also shrink.

In contrast, the benefit, in terms of the positive revenue collected from side l from matching agents with value $v_l \geq r_l^P$ to agents with value $v_k \leq r_k^P$ increases (since $\sigma_l(\mathbf{u}_k)$ is higher on average). As a consequence, the matching sets of agents on side l with $v_l \geq r_l^P$ expand and, by reciprocity, the matching sets of agents on side k with $v_k \leq r_k^P$ also expand.

To evaluate the impact on payoffs, we can use the characterization of Lemma 1 to obtain that for all $v_k \leq r_k^P$

$$\Pi_k(\theta_k; M^P) = \int_{\underline{v}_k}^{v_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k \leq \int_{\underline{v}_k}^{v_k} |\hat{\mathbf{s}}_k(\tilde{v}_k)|_k \, d\tilde{v}_k = \Pi_k(\theta_k; \hat{M}^P),$$

and therefore the payoffs of all agents on side k with $v_k \leq r_k^P$ increase. Since $|\mathbf{s}_k(v_k)|_k \geq |\hat{\mathbf{s}}_k(v_k)|_k$ for all $v_k \geq r_k^P$, incentive compatibility then implies that payoffs also increase for some agents on side k whose valuations are greater than r_k^P, implying that the threshold $\hat{\nu}_k$ is such that $\hat{\nu}_k > r_k^P$.

Intuitively, an increase in popularity on side k alters the costs of cross-subsidization across sides. Agents with $v_k \geq r_k^P$ are valued by the platform mainly by their role as consumers. As these agents become more popular, cross-subsidizing their "consumption" using agents from side l whose addition to the network is detrimental becomes more expensive, which explains why their matching sets shrink. The opposite is true for agents with $v_k \leq r_k^P$: these agents are valued by the platform mainly by their role as inputs. As they become better inputs, their matching sets expand and their payoffs increase.

The results from Proposition 5 offer testable predictions about the pricing strategies of many two-sided platforms. In the case of Cable TV providers, it implies that shocks on the wealth of households (which do not affect their valuation for channels) shall be accompanied by improvements on the standard packages and worsening of the premium packages offered by the platform.

Let q denote the quality of a given matching set $\mathbf{s}$, and let $\rho_k^P(q)$ denote the price that agents on side k have to pay for a matching set of quality q under the profit-maximizing mechanism M^P. Clearly, for any quality q that is attainable under the mechanism M^P

$$\rho_k^P(q) = p_k^P(\mathbf{u}_k, v_k) \quad \text{for all} \quad (\mathbf{u}_k, v_k) \quad \text{such that} \quad |\mathbf{s}_k^P(v_k)|_k = q.$$

The tariff $\rho_k^P(q)$ offers an indirect implementation of the mechanism M^P; it is designed so that each type $(\mathbf{u}_k, v_k)$ finds it optimal to choose the quality $|\mathbf{s}_k^P(v_k)|_k$ prescribed by the matching rule $\mathbf{s}_k^P(\cdot)$.

The next corollary translates Proposition 5 in terms of the tariff $\rho_k^P(q)$.

Corollary 2 *Consider the platform's profit-maximization problem when network effects are linear ($g_A(x) = g_B(x) = x$), let $\triangle^P < 0$, and assume Condition 2 holds. If the popularity of side k increases, then the platform moves from a price/quality schedule $\rho_k^P(\cdot)$ to a price/quality schedule $\hat{\rho}_k^P(\cdot)$ such that*

1. *there exists $\hat{q}_k > \left|\mathbf{s}_k^P(r_k^P)\right|_k = \left|\hat{\mathbf{s}}_k^P(r_k^P)\right|_k$ such that $\hat{\rho}_k^P(q) \leq \rho_k^P(q)$ if and only if $q \leq \hat{q}_k$,*

2. *there exists $\hat{q}_l < \left|\hat{\mathbf{s}}_l^P(r_l^P)\right|_l$ such that $\hat{\rho}_l^P(q) \leq \rho_l^P(q)$ if $q \geq \hat{q}_l$.*

In terms of quality/price schedules, an increase in the popularity of side k increases the prices for high quality matching sets on side k, and decreases the price for low quality matching sets. In turn, qualities above some threshold $\hat{q}_l$ become uniformly cheaper for agents on side l as side k becomes more popular.

3.5.2 The Cross-Side Effects of Changes in Demand Elasticity

The next definition formalizes the notion of an increase in the demand elasticity of side k.

Definition 6 *We say that side k is less elastic under F_k^v than under $\tilde{F}_k^v$ if F_k^v dominates $\tilde{F}_k^v$ in the hazard-rate order, and, for any v_k, $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k) = \tilde{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$.*

The seminal works of Rochet and Tirole (2006) and Armstrong (2006) argued that the structure of prices set by two-sided platforms that employ a single network shall depend on how demand elasticities compare across sides. The next proposition extends their analyses to the case of a two-sided platform that price-discriminates agents by offering them menus of networks.

Proposition 6 *Consider the platform's profit-maximization problem when network effects are linear ($g_A(x) = g_B(x) = x$), let $\triangle^P < 0$, and assume Condition 2 holds. If the elasticity of side k decreases, then the platform moves from a matching rule $\mathbf{s}_k^P(\cdot)$ to a matching rule $\tilde{\mathbf{s}}_k^P(\cdot)$ such that*

1. *$\tilde{\mathbf{s}}_k^P(v_k) \subseteq \mathbf{s}_k^P(v_k)$ for all v_k,*

2. *$\tilde{\mathbf{s}}_l^P(v_l) \subseteq \mathbf{s}_l^P(v_l)$ for all v_l,*

3. *$\Pi_k(\theta_k; \tilde{M}^P) < \Pi_k(\theta_k; M^P)$ for all v_k.*

Moreover, in the case of linear network externalities for quantity (where $\sigma_k(\cdot) \equiv 1$ for $k \in \{A, B\}$), there exists $\tilde{\nu}_l \in [\underline{v}_l, \bar{v}_l)$ such that $\Pi_l(\theta_l; \tilde{M}^P) < \Pi_l(\theta_l; M^P)$ if $\tilde{\nu}_l \leq v_l$.

The proposition above extends to a matching environment the familiar result for one-sided markets that quantities go down (and prices go up) as the demand of side k becomes less elastic. As a result, the payoffs of agents on side k uniformly decrease. In addition to these familiar results, in a matching market, reciprocity implies that the matching sets of agents from side l also have to shrink. Interestingly, while the quality of the matching sets offered to agents from side k decreases, the same might not be true to some agents on side l (since more agents from side k have higher valuations; recall that if F_k^v dominates $\tilde{F}_k^v$ in the hazard-rate order then it also does it in the usual first-order sense).

It is unambiguous however that in the case of linear network externalities for quantity the payoffs for agents with high valuations on side l have to decrease. To see why this is true, note that reciprocity implies the following aggregate condition:

$$\int_{\underline{v}_k}^{\bar{v}_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k = \int_{\underline{v}_l}^{\bar{v}_l} |\mathbf{s}_l(\tilde{v}_l)|_l \, d\tilde{v}_l,$$

which, heuristically, means that the total number of links departing from either side is the same. Since

$$\Pi_l((\bar{v}_l, \mathbf{u}_l); \tilde{M}^P) = \int_{\underline{v}_l}^{\bar{v}_l} |\tilde{\mathbf{s}}_l(\tilde{v}_l)|_l \, d\tilde{v}_l = \int_{\underline{v}_k}^{\bar{v}_k} |\tilde{\mathbf{s}}_k(\tilde{v}_k)|_k \, d\tilde{v}_k$$

$$< \int_{\underline{v}_k}^{\bar{v}_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k = \int_{\underline{v}_l}^{\bar{v}_l} |\mathbf{s}_l(\tilde{v}_l)|_l \, d\tilde{v}_l = \Pi_l((\bar{v}_l, \mathbf{u}_l); M^P)$$

we can conclude by continuity that $\Pi_l(\theta_l; \tilde{M}^P) < \Pi_l(\theta_l; M^P)$ if $v_l \geq \tilde{\nu}_l$ for some $\tilde{\nu}_l \in [\underline{v}_l, \bar{v}_l)$.

In the case of health care provision, Proposition 6 tells us that, under the profit-maximizing menu of health coverage plans, the number of patients that can access each doctor has to decrease as the costs of medical services go up. Moreover, patients with great need for doctors are certainly worse off as a result of this price increase.

The next corollary builds on Proposition 6 to determine the effects of a decrease in the elasticity of side k on the quality/price schedule $\rho_k^P(\cdot)$.

Corollary 3 *Consider the platform's profit-maximization problem when network effects are linear ($g_A(x) = g_B(x) = x$), let $\triangle^P < 0$, and assume Condition 2 holds. If the elasticity of side k decreases, then the platform moves from a price/quality schedule $\rho_k^P(\cdot)$ to a price/quality schedule $\tilde{\rho}_k^P(\cdot)$ such that $\tilde{\rho}^P(q) \geq \rho_k^P(q)$ for all q.*

The next example combines the lessons from Propositions 5 and 6 to analyze the effect of an increase in the productivity of firms in the profit-maximizing problem of a job employment agency.

Example 7 ***(Positive shock on the productivity of firms)*** *Let the platform be an employment agency that matches free-lancers on side A to firms on side B, as described in Example 6. Consider*

an increase in the productivity of firms, as captured by a shift from F_B^v to $\dot{F}_B^v$ such that $\dot{F}_B^v$ dominates F_B^v according to the hazard rate order. Since in this example $\mathbf{u}_k = v_k$, the productivity shock described above combines a change in the popularity with a change in the elasticity of side B. By combining Propositions 5 and 6, we can conclude that high productivity firms are hurt by a positive productivity shock, since the new matching rule $\dot{\mathbf{s}}_B^P(\cdot)$ is such that $\dot{\mathbf{s}}_B^P(v_B) \subseteq \mathbf{s}_B^P(v_B)$ for all $v_B \geq r_B^P$, and its associated payoff $\Pi_B(\theta_B; \dot{M}^P)$ is such that $\Pi_B(\theta_B; \dot{M}^P) \leq \Pi_B(\theta_B; M^P)$ for all $v_B \geq \dot{\nu}_B > r_B^P$. The effect on the payoffs of low productivity firms is ambiguous, though, since an increase in popularity makes them better off, while an increase in demand elasticity makes them worse off.

4 Extensions

4.1 Insulating Tariffs

4.2 Coarse Matching and Menu Costs

4.3 Quasi-Fixed Costs

4.4 The Group Design Problem

5 Conclusion

6 Appendix

Proof of Proposition 1. If $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, then it is immediate from (4) that efficiency (resp. profit) maximization requires that each agent on each side be matched to all agents on the other side, in which case $\mathbf{s}_k^h(\theta_k) = \Theta_l$ for all $\theta_k \in \Theta_k$. This rule trivially satisfies the threshold structure described in (5).

Thus consider the situation where $\varphi_k^h(\underline{v}_k) < 0$ for some $k \in \{A, B\}$. Denote by $\Theta_k^+ \equiv \{\theta_k = (\mathbf{u}_k, v_k) : \varphi_k^h(v_k) \geq 0\}$ the set of all types $\theta_k = (\mathbf{u}_k, v_k)$ whose φ_k^h-valuation is non-negative, and by $\Theta_k^- \equiv \{\theta_k = (\mathbf{u}_k, v_k) : \varphi_k^h(v_k) < 0\}$ the set of types with negative φ_k^h-valuation.

Let $s'_k(\cdot)$ be any implementable matching rule. We will show that, starting from $s'_k(\cdot)$, one can construct another implementable matching rule $\hat{s}_k(\cdot)$ that satisfies the threshold structure described in (5) and that weakly improves upon the platform's objective $\Omega^h(M)$.

In order to do so, for each $\theta_k \in \Theta_k^+$, let $t_k(v_k)$ be defined as follows:

1. If $|\mathbf{s}'_k(\theta_k)|_k \geq \left|\Theta_l^+\right|_k$, then let $t_k(v_k)$ be such that

$$|[t_k(v_k), \bar{v}_l) \times \mathbf{U}_l|_k = \left|\mathbf{s}'_k(\theta_k)\right|_k;$$

2. If $|\mathbf{s}'_k(\theta_k)|_k \leq \left|\Theta_l^+\right|_k = |\Theta_l|_k$, then $t_k(v_k) = \underline{v}_l$.

3. If $0 < |\mathbf{s}'_k(\theta_k)|_k \leq \left|\Theta_l^+\right|_k < |\Theta_l|_k$ (in which case $r_l^h \in (\underline{v}_l, \bar{v}_l)$), then let $t_k(v_k) = r_l^h$.

Now apply the construction above to $k = A, B$ and consider the matching rule $\hat{\mathbf{s}}_k(\cdot)$ such that

$$\hat{\mathbf{s}}_k(\theta_k) = \begin{cases} \mathbf{U}_l \cup [t_k(v_k), \bar{v}_l\}) & \Leftrightarrow \quad \theta_k \in \Theta_k^+ \\ \{(\mathbf{u}_l, v_l) \in \Theta_l^+ : t_l(v_l) \leq v_k\} & \Leftrightarrow \quad \theta_k \in \Theta_k^-. \end{cases}$$

By construction, $g_k\left(\hat{\mathbf{s}}_k(\mathbf{u}_k, \cdot)\right)$ is weakly increasing in v_k, for any $\mathbf{u}_k$, thus satisfying Condition 2 of Lemma 1 (i.e., $\hat{\mathbf{s}}_k(\cdot)$ is implementable). Moreover, $g_k\left(|\hat{\mathbf{s}}_k(\theta_k)|_k\right) \geq g_k\left(|\mathbf{s}'_k(\theta_k)|_k\right)$ for all $\theta_k \in \Theta_k^+$, implying that

$$\int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k) \geq \int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k), \ k = A, B \tag{10}$$

Similarly, one can show that

$$\int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k) \geq \int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k), \ k = A, B \tag{11}$$

holds. This is follows from the particular way by which $\hat{\mathbf{s}}_k(\cdot)$ has has been constructed from $\mathbf{s}'_k(\cdot)$, along with the following three properties: (i) $g(\cdot)$ is concave, (ii) $\sigma_l(\mathbf{u}_k)$ is positively affiliated with v_k and (iii) the quality of the matching sets $|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k$ can only vary in a countable set of v_k's. Summing up (10) and (11) shows that the platform's objective is weakly greater under $\hat{\mathbf{s}}_k(\cdot)$ than under $\mathbf{s}'_k(\cdot)$, as we wanted to show. QED

Proof of Proposition 2: There are three cases to consider: (i) $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, (ii) $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, and lastly (iii) $\varphi_l^h(\underline{v}_l) < 0$ while $\varphi_k^h(\underline{v}_k) \geq 0$. Cases (i) and (ii) appear in the main text, so let's consider case (iii).

In this case, either a single network or a nested multi-homing rule can be optimal depending on the sign of $\triangle^h$ in the proposition. From the same arguments used in case (ii), it must be that $\hat{\omega}_l < r_l^h$ and $\hat{\omega}_k = \underline{v}_k$ in the case of a single network. In turn, when nested multi-homing is optimal, it must be that $t_h(v_k) \leq r_l^h$ for all $v_k \in V_k$ and $t_h(\underline{v}_k) > \underline{v}_l$.

First, suppose that $\Delta^h \geq 0$ and, towards a contradiction, assume that the optimal mechanism employs a multi-homing rule. In such a case, optimality would require that $t_h(\cdot)$ is strictly decreasing in a right neighborhood of $\underline{v}_k$. Consider the marginal effect on the platform's objective of a reduction in the threshold $t_k(\underline{v}_k)$:

$$\begin{aligned} var &= -\hat{g}'_k(t_k(\underline{v}_k))\varphi_k^h(\underline{v}_k)f_k^v(\underline{v}_k) - \hat{g}'_l(\underline{v}_k)\varphi_l^h(t_k(\underline{v}_k))f_l^v(t_k(\underline{v}_k)) \\ &= g'_k(|\mathbf{U}_l \times [t_k(\underline{v}_k), \bar{v}_l]|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = t_k(\underline{v}_k)] \cdot f_l^v(t_k(\underline{v}_k)) \cdot \varphi_k^h(\underline{v}_k) \cdot f_k^v(\underline{v}_k) \\ &\quad + g'_l(|\Theta_k|_l) \cdot \mathbb{E}[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = \underline{v}_k] \cdot f_k^v(\underline{v}_k) \cdot \varphi_l^h(t_k(\underline{v}_k)) \cdot f_l^v(t_k(\underline{v}_k)). \end{aligned} \tag{12}$$

Because g_k is weakly concave and because $\sigma_k(\mathbf{u}_l)$ and v_l are positively affiliated, we have that

$$g_k'(|\mathbf{U}_l \times [t_k(\underline{v}_k), \bar{v}_l]|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = t_k(\underline{v}_k)] \geq g_k'(|\Theta_l|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = \underline{v}_l]. \tag{13}$$

Furthermore, because $\varphi_k^h(\cdot)$ is strictly increasing, $0 > \varphi_l^h(t_k(\underline{v}_k)) > \varphi_l^h(\underline{v}_l)$. Together with (13), this implies that $var > 0$. This contradicts the optimality of a multi-homing matching rule and shows that a single network is optimal. Furthermore, the same arguments imply that such a network must include all agents on each side, i.e., it must be a complete single network.

Second, suppose that $\Delta^h < 0$ and, again towards a contradiction, suppose that a single network is optimal. Let's first consider the case when such a network is complete (that is, $\hat{\omega}_l = \underline{v}_l$ or $t_k^h(\underline{v}_k) = \underline{v}_l$). We will argue that the fact that $\Delta^h < 0$ implies that the platform could improve upon its payoff by switching to a multi-homing rule that is obtained from the complete network by increasing $t_k(\cdot)$ in a right neighborhood of $\underline{v}_k$ while leaving $t_k(\cdot)$ untouched elsewhere. Indeed, the marginal effect on the platform's payoff from doing so equals $-var$, where var is the expression in (12). Since $\Delta^h < 0$, moving to multi-homing rule increases profits (respectively, welfare), contradicting the optimality of a single network.

Lastly, consider the case where $\hat{\omega}_l > \underline{v}_l$. At the optimal single network, the (interior) threshold $\hat{\omega}_l$ has to satisfy the following first-order condition:

$$\hat{g}_l(\underline{v}_k)\varphi_l^h(\hat{\omega}_l)f_l^v(\hat{\omega}_l) - \hat{g}_k'(\hat{\omega}_l)\int_{\underline{v}_k}^{\bar{v}_k} \varphi_k^h(v_k)dF_k^v(v_k) = 0, \tag{14}$$

where the left-hand-side of (14) is the total effect of a marginal increase of the size of the network on side l. Now note that the left hand side of (14) is equivalent to

$$\begin{aligned}
&\left[\int_{\underline{v}_k}^{\bar{v}_k} \left(-\hat{g}_l'(x)\right) dx\right] \varphi_l^h(\hat{\omega}_l)f_l^v(\hat{\omega}_l) - \hat{g}_k'(\hat{\omega}_l)\int_{\underline{v}_k}^{\bar{v}_k} \varphi_k^h(v_k)dF_k^v(v_k) \\
= &\left[\int_{\underline{v}_k}^{\bar{v}_k} \left(g_l'(|\mathbf{U}_k \times [x, \bar{v}_k]|_l) \cdot \mathbb{E}[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = x] \cdot f_k^v(x)\right) dx\right] \varphi_l^h(\hat{\omega}_l) \cdot f_l^v(\hat{\omega}_l) \\
&+g_k'(|\mathbf{U}_l \times [\hat{\omega}_l, \bar{v}_l]|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = \hat{\omega}_l] \cdot f_l^v(\hat{\omega}_l) \cdot \int_{\underline{v}_k}^{\bar{v}_k} \varphi_k^h(v_k)dF_k^v(v_k)
\end{aligned}$$

which is striclty greater than

$$\begin{aligned}
&g_k'(|\mathbf{U}_l \times [\hat{\omega}_l, \bar{v}_l]|_k) \cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = \hat{\omega}_l] \cdot f_l^v(\hat{\omega}_l) \cdot \varphi_k^h(\underline{v}_k) \\
&+g_l'(|\Theta_k|_l) \cdot \mathbb{E}[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = \underline{v}_k] \cdot \varphi_l^h(\hat{\omega}_l) \cdot f_l^v(\hat{\omega}_l),
\end{aligned} \tag{15}$$

because $\varphi_k^h(\cdot)$ is strictly increasing, g_k is weakly concave, and because $\sigma_k(\mathbf{u}_l)$ and v_l are positively affiliated. That the sign of (15) is negative in turn means that

$$\hat{g}_l'(\underline{v}_k)\varphi_l^h(\hat{\omega}_l)f_l^v(\hat{\omega}_l) + \hat{g}_k'(\hat{\omega}_l)\varphi_k^h(\underline{v}_k)f_k^v(\underline{v}_k) > 0.$$

However, as argued above, in this case the platform can, starting from such a single network, increase profits (alternatively, welfare) by switching to a multi-homing rule that is obtained from the original single network by increasing $t_k(\cdot)$ in a right neighborhood of $\underline{v}_k$ while leaving $t_k(\cdot)$ untouched elsewhere. Q.E.D.

Proof of Proposition 3: The h-optimal matching rule solves the following program, which we call the Full Program (P_F) :

$$P_F: \qquad \max_{\{\omega_k, t_k(\cdot)\}_{k\in\{A,B\}},} \sum_{k=A,B} \int_{\omega_k}^{\overline{v}_k} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) \tag{16}$$

subject to

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{17}$$

$$t_A(\cdot), t_B(\cdot) \text{ are weakly decreasing}, \tag{18}$$

$$\text{and} \quad t_k(\cdot) : [\omega_k, \overline{v}_k] \to [\omega_l, \overline{v}_l]. \tag{19}$$

Becase $\triangle^h < 0$ (i.e., multi-homing is optimal), $r_k^h > \underline{v}_k$ for some $k \in \{A, B\}$. The proof of Proposition 1 established that $t_k^h(r_k^h) = r_l^h$ whenever $r_k^h \geq \underline{v}_k$ and $r_l^h \geq \underline{v}_l$. Therefore, the program P_F can be dismembered in the following two independent programs P_F^k for $k \in \{A, B\}$:

$$P_F^k: \qquad \max_{\omega_k, \{t_k(\cdot)\}_{k\in\{A,B\}},} \int_{\omega_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{\max\{r_l^h, \underline{v}_l\}}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l)$$

subject to (17), (18) and

$$t_k(\cdot) : [\omega_k, r_k^h] \to [\max\{r_l^h, \underline{v}_l\}, \overline{v}_l], \qquad t_l(\cdot) : [\max\{r_l^h, \underline{v}_l\}, \overline{v}_l] \to [\omega_k, r_k^h]. \tag{20}$$

If $\varphi_l^h(\underline{v}_l) \geq 0$ (and therefore $\underline{v}_l > r_l^h = -\infty$), the solution to P_F sets $t_k^h(v_k) = \underline{v}_l$ for all $v_k \geq r_k^h$. In this case, we only have to solve P_F^k to learn the optimal thresholds $t_k^h(v_k)$ for all $v_k < r_k^h$ and $t_l^h(v_l)$ for all $v_l \in V_l$. To simplify notation, we will assume from now on that $r_l^h \geq \underline{v}_l$,

The program P_F^k is not a standard calculus of variations problem. It will be instrumental to consider the following Auxiliary Program (P_A^k), which strenghtens constraint (18) and fixes $\omega_k = \underline{v}_k$ and $\omega_l = \underline{v}_l$:

$$P_A^k: \qquad \max_{\{t_k(\cdot)\}_{k\in\{A,B\}}} \int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{21}$$

subject to (17),

$$t_A(\cdot), t_B(\cdot) \text{ are strictly decreasing}, \tag{22}$$

$$\text{and} \quad t_k(\cdot) : [\omega_k, r_k^h] \to [r_l^h, \overline{v}_l], \quad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\omega_k, r_k^h] \text{ are bijections}. \tag{23}$$

By virtue of (22), $t_A(\cdot), t_B(\cdot)$ are strictly monotone, in which case (17) can be rewritten as $t_k(v_k) = t_l^{-1}(v_k)$. Plugging that into the objective function (21) leads to

$$\int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) dv_k + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_k^{-1}(v_l)) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l) dv_l.$$

Let's change the variable of integration in the second equation from v_l to $\tilde{v}_l \equiv t_k^{-1}(v_l)$. This leads to

$$\int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) dv_k - \int_{\underline{v}_k}^{r_k^h} \hat{g}_l(\tilde{v}_l) \cdot \varphi_l^h(t_k(\tilde{v}_l)) \cdot f_l^v(t_k(\tilde{v}_l)) \cdot t_k'(\tilde{v}_l) d\tilde{v}_l.$$

After relabelling the variable of integration on the second integral, the Auxiliary Program can be recasted as:

$$P_A^k: \qquad \max_{t_k(\cdot)} \int_{\underline{v}_k}^{r_k^h} \left\{ \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l(v_k) \cdot \varphi_l^h(t_k(v_k)) \cdot f_l^v(t_k(v_k)) \cdot t_k'(v_k) \right\} dv_k \quad (24)$$

subject to (22) and

$$t_k(\underline{v}_k) = \overline{v}_l \quad \textit{and} \quad t_k(r_k^h) = r_l^h. \quad (25)$$

Consider now the Relaxed Auxiliary Program (P_{RA}^k), which is a relaxation of P_A^k that ignores constraint constraint (22) and allows for controls $t_k(\cdot) : [\underline{v}_k, r_k^h] \to [r_l^h, \overline{v}_l]$ satisfying constraint (25).

Lemma 2 *P_{RA}^k has a piece-wise absolutely continuous maximizer $t_k^h(\cdot)$ among all measurable controls $t_k(\cdot)$ with bounded sub-gradients.*

Proof of Lemma 2. Program P_{RA}^k is equivalent to the following optimal control problem $\mathbb{P}_{RA}^k$:

$$\mathbb{P}_{RA}^k: \qquad \max_{y(\cdot)} \int_{\underline{v}_k}^{r_k^h} \left\{ \hat{g}_k(x(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l(v_k) \cdot \varphi_l^h(x(v_k)) \cdot f_l^v(x(v_k)) \cdot y(v_k) \right\} dv_k$$

subject to

$$x'(v_k) = y(v_k), \quad x(\underline{v}_k) = \overline{v}_l, \quad x(r_k^h) = r_l^h \quad y(v_k) \in [-K, +K] \;\; \textit{and} \;\; x(v_k) \in [r_l^h, \overline{v}_l],$$

where K is a large number. Program $\mathbb{P}_{RA}^k$ satisfies all the conditions of the Filipov-Cesari Theorem (see Cesari (1983)), and we can conclude that exists a measurable function $y(\cdot)$ that solves $\mathbb{P}_{RA}^k$. By the equivalence of P_{RA}^k and $\mathbb{P}_{RA}^k$, it then follows that P_{RA}^k has a piece-wise absolutely continuous maximizer $t_k^h(\cdot)$ among all measurable controls $t_k(\cdot)$ with bounded sub-gradients. Notice that $t_k^h(\cdot)$ has (at most) a unique discontinuity at $\underline{v}_k$. Q.E.D.

Lemma 3 *The solution to P_{RA}^k among all among all measurable controls $t_k(\cdot)$ with bounded sub-gradients is given by*

$$\tilde{t}_k^h(v_k) = \begin{cases} \eta(v_k) & \Leftrightarrow \quad v_k \in (\max\{\underline{v}_k, \eta^{-1}(\overline{v}_l)\}, r_k^h] \\ \overline{v}_l & \Leftrightarrow \quad v_k \in [\underline{v}_k, \max\{\underline{v}_k, \eta^{-1}(\overline{v}_l)\}], \end{cases} \quad (26)$$

where $\eta(\cdot)$ is implicitly defined by

$$\hat{g}_k'(\eta(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) + \hat{g}_l'(v_k) \cdot \varphi_l^h(\eta(v_k)) \cdot f_l^v(\eta(v_k)) = 0, \tag{27}$$

and $\eta^{-1}(\bar{v}_l)$ is set to $-\infty$ if $\eta(v_k) < \bar{v}_l$ for all $v_k \in [\underline{v}_k, r_k^h]$.

Proof of Lemma 3. From Lemma 2, we know that the solution to P_{RA}^k among all measurable controls $t_k(\cdot)$ with bounded sub-gradients is piece-wise absolutely continuous. As a consequence, we can apply standard calculus of variations to conclude that the function $\tilde{t}_k^h(\cdot)$ that solves P_{RA}^k has to satisfy the Euler equation at any interior point of its domain where the restriction that its co-domain is $[r_l^h, \bar{v}_l]$ is slack (that is, $\tilde{t}_k^h(v_k) \in (r_l^h, \bar{v}_l)$). The Euler equation associated to program P_{RA}^k is given by (27). The Strong Regularity Condition assures that (27) has a unique (striclty decreasing and continuous) solution given by $\eta(v_k) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v_k)\right)$ at every point where $\eta(v_k) \in (\underline{v}_l, \bar{v}_l)$.

If $\eta^{-1}(\bar{v}_l) > \underline{v}_k$, the restriction on the co-domain belonging to $[r_l^h, \bar{v}_l]$ is binding at the optimum for all $v_k \in [\underline{v}_k, \eta^{-1}(\bar{v}_l)]$. In such a case, the optimum at $v_k \in [\underline{v}_k, \eta^{-1}(\bar{v}_l)]$ is a step function that assumes values on $\{r_l^h, \bar{v}_l\}$. Becase step functions are such that $t_k'(v_k) = 0$ at all points of continuity and because $\varphi_k^h(v_k) < 0$ for all $v_k \in [\underline{v}_k, \eta^{-1}(\bar{v}_l)]$, it is immediate from the objective (24) that $\tilde{t}_k^h(v_k) = \bar{v}_l$ for all $v_k \in [\underline{v}_k, \eta^{-1}(\bar{v}_l)]$, as in (26).

It remains to be shown that $\tilde{t}_k^h(v_k) = \eta(v_k)$ for all $v_k \in [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$. It turns out that the objective function of P_A' is not concave in (t_k, t_k') for all v_k, in which case we cannot appeal to the standard sufficiency arguments. Instead, we will prove that the solution to P_A' is given by $\tilde{t}_k^h(v_k) = \eta(v_k)$ for all $v_k \in [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$ by arguing that there is no function $\hat{t}_k^h(\cdot)$ that improves upon $\tilde{t}_k^h(v_k)$ such that $\hat{t}_k^h(\cdot)$ coincides with $\tilde{t}_k^h(\cdot)$ except in a interval $(\hat{v}_k, \tilde{v}_k) \subseteq [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$ where $\hat{t}_k^h(v_k) \in \{r_l^h, \bar{v}_l\}$ for all $v_k \in (\hat{v}_k, \tilde{v}_k)$.

To see why this is true, fix $(\hat{v}_k, \tilde{v}_k) \subseteq [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$ and let's optimize over step functions such that $\hat{t}_k^h(v_k) \in \{r_l^h, \bar{v}_l\}$ for all $v_k \in (\hat{v}_k, \tilde{v}_k)$. Becase step functions are such that $t_k'(v_k) = 0$ at all points of continuity and because $\varphi_k^h(v_k) < 0$ for all $v_k \in (\hat{v}_k, \tilde{v}_k)$, it follows that the optimal step function sets $\hat{t}_k^h(v_k) = \bar{v}_l$ for all $v_k \in (\hat{v}_k, \tilde{v}_k)$. Notice that the value attained by the objective (24) restricted to $(\hat{v}_k, \tilde{v}_k)$ evaluated at such step function is zero.

However, the platform can choose an interior control $t_k(\cdot) \in (r_l^h, \bar{v}_l)$ and set its derivative

$$t_k'(v_k) < \frac{\hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k)}{\hat{g}_l(v_k) \cdot \varphi_l^h(t_k(v_k)) \cdot f_l^v(t_k(v_k))}$$

in such a way that the integrand in (24) is strictly positive for all $v_k \in (\hat{v}_k, \tilde{v}_k)$. Therefore, for all $v_k \in [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$ there is an interior control $t_k(\cdot) \in (r_l^h, \bar{v}_l)$ that improves upon corner solutions. This allows us to conclude that the Euler equation (27) has to hold for all $v_k \in [\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$.

Finally, notice that $\eta(r_k^h) = r_l^h$, what shows that (26) indeed satisfies the constraint (25). Q.E.D.

Denote by $\max\{P_{RA}^k\}$ the value attained by program P_{RA}^k at the optimum (26). Denote by $\sup\{P_A^k\}$ and $\sup\{P_F^k\}$ the supremum of programs P_A^k and P_F^k. Note that we write sup rather than max, as a priori a solution to these problems might not exist.

Lemma 4 $\sup\{P_F^k\} = \sup\{P_A^k\} \leq \max\{P_{RA}^k\}$.

Proof of Lemma 4. It is immediate that $\sup\{P_A^k\} \leq \max\{P_{RA}^k\}$, since P_{RA}^k is a relaxed version of P_A^k. Moreover, $\sup\{P_F^k\} = \sup\{P_A^k\}$ because any measurable function $\{t_k(\cdot)\}_{k\in\{A,B\}}$ satisfying conditions (17), (18) and (20) can be arbitrarily well approximated in the L^2-norm by a function satisfying conditions (17), (22) and (23). Q.E.D.

We will now exhibit the solution to P_F^k. Set $\omega_k^h = \max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}$ and define $t_k^h(\cdot)$ as the function $\tilde{t}_k^h(\cdot)$ restricted to $(\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$. It is clear that the value of the objective (24) of program P_{RA}^k evaluated at $t_k^h(\cdot)$ and integrated over $(\max\{\underline{v}_k, \eta^{-1}(\bar{v}_l)\}, r_k^h]$ is equal to $\max\{P_{RA}^k\}$. Now define $t_l^h(\cdot)$ using (17) evaluated at $t_k^h(\cdot)$. It is clear that $\omega_k^h, \left\{t_k^h(\cdot)\right\}_{k\in\{A,B\}}$ satisfy conditions (17), (18) and (20), and are therefore feasible under P_F^k. As a consequence, Lemma 4 allows us to conclude that $\omega_k^h, \left\{t_k^h(\cdot)\right\}_{k\in\{A,B\}}$ is a solution to P_F^k.

By the same token, let's extend $t_k^h(\cdot)$ to $[r_k^h, \bar{v}_k]$ by using (17) evaluated at the solution $t_l^h(\cdot)$ to program P_F^l. Now define $\omega_l^h = \max\{\underline{v}_l, \varpi^{-1}(\bar{v}_k)\}$, where $\varpi(\cdot)$ is the analogous of function $\eta(\cdot)$ for problem P_F^l. By construction, $\left\{\omega_k^h, t_k^h(\cdot)\right\}_{k\in\{A,B\}}$ satisfy the conditions (17), (18) and (19) from program P_F. Evauating the objective function (16) of program P_F at $\left\{\omega_k^h, t_k^h(\cdot)\right\}_{k\in\{A,B\}}$ yields

$$\max\{P_{RA}^k\} + \max\{P_{RA}^l\} \geq \sup\{P_F^k\} + \sup\{P_F^l\} = \sup\{P_F\},$$

where the inequality follows from Lemma 4 and $\sup\{P_F\}$ denotes the supremum of program P_F. We therefore conclude that P_F admits a solution given by $\left\{\omega_k^h, t_k^h(\cdot)\right\}_{k\in\{A,B\}}$.

Finally, there is exclusion at the bottom on side l and no bunching at the top of side k if $\varpi^{-1}(\bar{v}_k) > \underline{v}_l$, which is equivalent to $\triangle_l^h(\underline{v}_l, \overline{v}_k) = \triangle_k^h(\bar{v}_k, \underline{v}_l) < 0$. A similar reasoning shows that there is no exclusion at the bottom on side l and bunching at the top of side k if $\triangle_k^h(\bar{v}_k, \underline{v}_l) = \triangle_l^h(\underline{v}_l, \bar{v}_k) > 0$, and neither bunching at the top on side k nor exclusion at the bottom on side l if $\triangle_k^h(\bar{v}_k, \underline{v}_l) = \triangle_l^h(\underline{v}_l, \bar{v}_k) = 0$. Finally, note that the Euler equation (27) is equivalent to $\triangle_k^h(v_k, t_k^h(v_k)) = 0$ for all v_k in the separating range. Q.E.D.

Proof of Proposition 5. Denote by $\hat{\psi}_k^P(v_k)$ the psi-function associated to the distribution $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$, and denote by $\hat{\mathbb{E}}\left[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k\right]$ the expected value of $\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k$ computed according to $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$. Because $\mathbb{E}\left[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k\right] \leq \hat{\mathbb{E}}\left[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k\right]$, it follows that $\hat{\psi}_k^P(v_k) \geq \psi_k^P(v_k)$ for all $v_k \leq r_k^P$ and $\hat{\psi}_k^P(v_k) \leq \psi_k^P(v_k)$ for all $v_k \geq r_k^P$. As a consequence, for all $v_k \leq r_k^P$ in the separating range

$$\hat{t}_k^P(v_k) = \left(\psi_l^P\right)^{-1}\left(-\hat{\psi}_k^P(v_k)\right) \leq \left(\psi_l^P\right)^{-1}\left(-\psi_k^P(v_k)\right) = t_k^P(v_k),$$

and the reverse inequality holds for all $v_k \geq r_k^P$ in the separating range. For $v_k \geq r_k^P$ outside the separating range (that is, for $v_k \geq r_k^P$ in the top bunching interval, if one exists), the fact that $\left(-\psi_k^P\right)^{-1} \psi_l^P(\underline{v}_l) \leq \left(-\hat{\psi}_k^P\right)^{-1} \psi_l^P(\underline{v}_l)$ implies that $\hat{t}_k^P(v_k) \geq t_k^P(v_k)$, as desired.

Note that because F_l is left unchanged, it follows that $|\mathbf{s}_k(v_k)|_k \leq |\hat{\mathbf{s}}_k(v_k)|_k$ if and only if $v_k \leq r_k^P$. As a consequence, we can use Lemma 1 to conclude that for all $v_k \leq r_k^P$

$$\Pi_k(\theta_k; M^P) = \int_{\underline{v}_k}^{v_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k \leq \int_{\underline{v}_k}^{v_k} |\hat{\mathbf{s}}_k(\tilde{v}_k)|_k \, d\tilde{v}_k = \Pi_k(\theta_k; \hat{M}^P).$$

Since $|\mathbf{s}_k(v_k)|_k \geq |\hat{\mathbf{s}}_k(v_k)|_k$ for all $v_k \geq r_k^P$, there exists $\hat{\nu}_k > r_k^P$ (possibly equal to $\bar{v}_k$) such that $\Pi_k(\theta_k; M^P) \leq \Pi_k(\theta_k; \hat{M}^P)$ for all $v_k \leq \hat{\nu}_k$.

Finally, notice that reciprocity implies that $\hat{t}_l^P(v_l) \leq t_l^P(v_l)$ (and therefore $\mathbf{s}_l(v_l) \subseteq \hat{\mathbf{s}}_l(v_l)$) if and only if $v_l \geq r_l^P$. Therefore, $|\mathbf{s}_l(v_l)|_l \leq |\hat{\mathbf{s}}_l(v_l)|_l$ for all $v_l \geq r_l^P$. If $v_l < r_l^P$ the comparison between $|\mathbf{s}_l(v_l)|_l$ and $|\hat{\mathbf{s}}_l(v_l)|_l$ is ambiguous (since $\hat{\mathbf{s}}_l(v_l) \subseteq \mathbf{s}_l(v_l)$ but $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k) \succeq F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ according to the usual stochastic order). Still, by continuity there exists $\hat{\nu}_l \in [\underline{v}_l, r_l^P)$ such that $|\mathbf{s}_l(v_l)|_l \leq |\hat{\mathbf{s}}_l(v_l)|_l$ if $v_l \geq \hat{\nu}_l$. Appealing to Lemma 1 as above, it then follows that $\Pi_l(\theta_l; M^P) \leq \Pi_k(\theta_k; \hat{M}^P)$ if $v_l \geq \hat{\nu}_l$. Q.E.D.

Proof of Corollary 2. We will now show that $\hat{\rho}_k^P(q) \leq \rho_k^P(q)$ for all $q \leq |\mathbf{s}_k^P(r_k^P)|_k = |\hat{\mathbf{s}}_k^P(r_k^P)|_k$. Applying integration by parts and using Lemma 1 leads to

$$\begin{aligned}
\rho_k^P(q) &= \underline{v}_k \, |\mathbf{s}_k(\underline{v}_k)|_k + \int_{\underline{v}_k}^{v} \tilde{v}_k \cdot \frac{\partial}{\partial \tilde{v}_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k \\
&= \underline{v}_k \, |\mathbf{s}_k(\underline{v}_k)|_k + \int_{|\mathbf{s}_k^P(\underline{v}_k)|_k}^{q} \mu_k(z) dz,
\end{aligned}$$

after changing the variable of integration from $\tilde{v}_k$ to z and defining

$$\mu_k(z) \equiv \inf \{v_k : |\mathbf{s}_k(v_k)|_k = z\}.$$

Denoting by $\hat{\mu}_k(\cdot)$ the function analogous to $\mu_k(\cdot)$ associated to $\hat{\mathbf{s}}_k^P(\cdot)$, it follows that $\hat{\mu}_k(z) \leq \mu_k(z)$ for all $z \leq |\mathbf{s}_k^P(r_k^P)|_k = |\hat{\mathbf{s}}_k^P(r_k^P)|_k$. Therefore, for all $q \leq |\mathbf{s}_k^P(r_k^P)|_k = |\hat{\mathbf{s}}_k^P(r_k^P)|_k$,

$$\begin{aligned}
\rho_k^P(q) &\geq \underline{v}_k \, |\mathbf{s}_k(\underline{v}_k)|_k + \int_{|\hat{\mathbf{s}}_k^P(\underline{v}_k)|_k}^{q} \hat{\mu}_k(z) dz + \int_{|\mathbf{s}_k^P(\underline{v}_k)|_k}^{|\hat{\mathbf{s}}_k^P(\underline{v}_k)|_k} \hat{\mu}_k(z) dz \\
&\geq \underline{v}_k \, |\mathbf{s}_k(\underline{v}_k)|_k + \int_{|\hat{\mathbf{s}}_k^P(\underline{v}_k)|_k}^{q} \hat{\mu}_k(z) dz + \underline{v}_k \int_{|\mathbf{s}_k^P(\underline{v}_k)|_k}^{|\hat{\mathbf{s}}_k^P(\underline{v}_k)|_k} dz \\
&= \underline{v}_k \, |\hat{\mathbf{s}}_k(\underline{v}_k)|_k + \int_{|\hat{\mathbf{s}}_k^P(\underline{v}_k)|_k}^{q} \hat{\mu}_k(z) dz = \hat{\rho}_k^P(q),
\end{aligned}$$

as we wanted to show. For $q > |\mathbf{s}_k^P(r_k^P)|_k = |\hat{\mathbf{s}}_k^P(r_k^P)|_k$, note that

$$\hat{\rho}_k^P(q) - \rho_k^P(q) = \hat{\rho}_k^P(r_k^P) - \rho_k^P(r_k^P) + \int_{|\mathbf{s}_k^P(r_k^P)|_k}^{q} (\hat{\mu}_k(z) - \mu_k(z)) dz,$$

where $\hat{\mu}_k(z) \geq \mu_k(z)$ for all $z \geq |\mathbf{s}_k^P(r_k^P)|_k = |\hat{\mathbf{s}}_k^P(r_k^P)|_k$. Because $\hat{\rho}_k^P(r_k^P) - \rho_k^P(r_k^P) \leq 0$ and $\int_{|\mathbf{s}_k^P(r_k^P)|_k}^{q}(\hat{\mu}_k(z) - \mu_k(z))dz$ is positive and increasing, it follows that there exists $\hat{q}_k > \left|\mathbf{s}_k^P(r_k^P)\right|_k = \left|\hat{\mathbf{s}}_k^P(r_k^P)\right|_k$ (possibly $\left|\hat{\mathbf{s}}_k^P(\bar{v}_k)\right|_k$) such that $\hat{\rho}_k^P(q) \leq \rho_k^P(q)$ if and only if $q \leq \hat{q}_k$.

Define $\hat{q}_l \equiv \left|\hat{\mathbf{s}}_l^P(\hat{\nu}_l)\right|$, where $\hat{\nu}_l$ is the threshold established in the proof of Proposition 5 such that $|\mathbf{s}_l(v_l)|_l \leq |\hat{\mathbf{s}}_l(v_l)|_l$ if $v_l \geq \hat{\nu}_l$. The same computations as above then show that $\hat{\rho}_l^P(q) \leq \rho_l^P(q)$ if $q \geq \hat{q}_l$. Q.E.D.

Proof of Proposition 6. Denote by $\tilde{\psi}_k^P(v_k)$ the psi-function associated to the distribution $\tilde{F}_k^v(\cdot)$, and denote by $\tilde{\varphi}_k^P(v_k)$ the virtual values associated to $\tilde{F}_k^v(\cdot)$. Since $\tilde{F}_k^v(\cdot)$ dominates $F_k^v(\cdot)$ in the hazard rate order, it follows that $\tilde{\varphi}_k^P(v_k) \leq \varphi_k^P(v_k)$ and therefore $\tilde{\psi}_k^P(v_k) \leq \psi_k^P(v_k)$ for all $v_k \in V_k$. Therefore for all $v_k \in V_k$ in the separating range

$$\tilde{t}_k^P(v_k) = \left(\psi_l^P\right)^{-1}\left(-\tilde{\psi}_k^P(v_k)\right) \geq \left(\psi_l^P\right)^{-1}\left(-\psi_k^P(v_k)\right) = t_k^P(v_k).$$

Becase $\left(-\psi_k^P\right)^{-1}\psi_l^P(\underline{v}_l) \leq \left(-\tilde{\psi}_k^P\right)^{-1}\psi_l^P(\underline{v}_l)$, we can safely conclude that $\tilde{t}_k^P(v_k) \geq t_k^P(v_k)$ for $v_k \geq r_k^P$ outside the separating range (that is, for $v_k \geq r_k^P$ in the top bunching interval, if one exists).

Because F_l is left unchanged, it follows that $|\mathbf{s}_k(v_k)|_k \geq |\tilde{\mathbf{s}}_k(v_k)|_k$ for all $v_k \in V_k$. By the same argument in the proof of Proposition 5, it then follows that

$$\Pi_k(\theta_k; M^P) = \int_{\underline{v}_k}^{v_k} |\mathbf{s}_k(\tilde{v}_k)|_k \, d\tilde{v}_k \geq \int_{\underline{v}_k}^{v_k} |\tilde{\mathbf{s}}_k(\tilde{v}_k)|_k \, d\tilde{v}_k = \Pi_k(\theta_k; \tilde{M}^P)$$

for all $v_k \in V_k$.

Finally, notice that reciprocity implies that $\tilde{t}_l^P(v_l) \geq t_l^P(v_l)$ (and therefore $\mathbf{s}_l(v_l) \supseteq \tilde{\mathbf{s}}_l(v_l)$) for all $v_l \in V_l$. Q.E.D.

Proof of Corollary 3. Because $|\mathbf{s}_k(v_k)|_k \geq |\tilde{\mathbf{s}}_k(v_k)|_k$ for all $v_k \in V_k$, we can follow the same steps from the proof of Corollary 2 to get the result. Q.E.D.

References

[1] Anderson, E. and James Dana, 2009, When Is Price Discrimination Profitable?, *Management Science*, Vol. 55(6), pp. 980-989.

[2] Ambrus, A. and R. Argenziano, 2009, Asymmetric Networks in Two-Sided Markets, *American Economic Journal*, Vol. 1(1), pp. 17-52.

[3] Armstrong, M., 1996, Multiproduct Nonlinear Pricing, *Econometrica*, Vol. 64(1), pp. 51-75.

[4] Atakan, A., 2006, Assortative Matching with Explicit Search Costs, *Econometrica*, Vol. 74(3), pp. 667-680.

[5] Becker, G., 1973, A Theory of Marriage: Part 1, *Journal of Political Economy*, pp. 813-846.

[6] Board, S., 2009, Monopoly Group Design with Peer Effects, *Theoretical Economics*, Vol. 4, pp. 89-125.

[7] Bulow, J. and J. Roberts, 1989, The Simple Economics of Optimal Auctions, *The Journal of Political Economy*, Vol. 97(5), pp. 1060-1090.

[8] Caillaud, B. and B. Jullien, 2001, Competing Cybermediaries, *European Economic Review*, Vol. 45, pp. 797-808.

[9] Caillaud, B. and B. Jullien, 2003, Chicken & Egg: Competition Among Intermediation Service Providers, *Rand Journal of Economics*, Vol. 34, pp. 309-328.

[10] Cremer, J. & R. McLean, 1988, Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions, *Econometrica*, vol. 56(6), 1247-57.

[11] Damiano, E. and H. Li, 2007, Price Discrimination and Efficient Matching, *Economic Theory*, Vol. 30, pp. 243-263.

[12] Eeckhout, J. and P. Kircher, 2010, Sorting and Decentralized Price Competition, *Econometrica*, Vol. 78(2), pp. 539-574.

[13] Ellison, G. and D. Fudenberg, 2000, The Neo-Luddite's Lament: Excessive Upgrades in the Software Industry, *Rand Journal of Economics*, Vol. 31(2), pp. 253-272.

[14] Hagiu, A., 2006, Pricing and Commitment by Two-Sided Platforms, *Rand Journal of Economics*, Vol. 37(3), pp. 720-737.

[15] Hatfield, J. and P. Milgrom, 2005, Matching with Contracts, *American Economic Review*, Vol. 95(4), pp. 913-935.

[16] Hoppe, H., B. Moldovanu and Emre Ozdenoren, 2010, Course Matching with Incomplete Information, *Economic Theory*.

[17] Hoppe, H., B. Moldovanu and A. Sela, 2009, The Theory of Assortative Matching Based on Costly Signals, *Review of Economic Studies*, Vol. 76, pp. 253-281.

[18] Maskin, E., and J. Riley, 1984, Monopoly with Incomplete Information, *Rand Journal of Economics*, Vol. 15, pp. 171-196.

[19] McAfee, R. P., 2002, Course Matching, *Econometrica*, Vol. 70, pp. 2025-2034.

[20] McAfee, R. P. and R. Deneckere, 1996, Damaged Goods, *Journal of Economics and Management Strategy*, Vol. 5(2), pp. 149-174.

[21] Mezzetti, C., 2007, Mechanism Design with Interdependent Valuations: Surplus Extraction, *Economic Theory*, 31(3), 473-488.

[22] Mussa, M. and S. Rosen, 1978, Monopoly and Product Quality, *Journal of Economic Theory*, Vol. 18, pp. 301-317.

[23] Myerson, R., 1981, Optimal Auction Design, *Mathematics of Operations Research*, Vol. 6(1), pp. 58-73.

[24] Rayo, L., 2010, Monopolistic Signal Provision, *B.E. J Theoretical Economics*.

[25] Rochet, J.-C. and J. Tirole, 2003, Platform Competition in Two-sided Markets, *Journal of the European Economic Association*, Vol. 1, pp. 990-1029.

[26] Rochet, J.-C. and J. Tirole, 2006, Two-sided Markets: A Progress Report, *Rand Journal of Economics*, Vol. 37(3), pp. 645-667.

[27] Roth, A. and M. Sotomayor, 1990, Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, *Econometric Society Monograph Series*, Cambridge University Press, Cambridge.

[28] Shimer, R., 2000, The Assignment of Workers to Jobs in an Economy with Coordination Frictions, *Journal of Political Economy*, Vol. 113(5), pp. 996-1025.

[29] Shimer, R. and L. Smith, 2000, Assortative Matching and Search, *Journal of Political Economy*, Vol. 114(6), pp. 996-1025.

[30] Smith, L., 2006, The Marriage Model with Search Frictions, *Journal of Political Economy*, Vol. 113(5), pp. 1124-1144.

[31] Weyl, G., 2010, A Price Theory of Two-Sided Markets, *American Economic Review*.