Price Discrimination in Many-to-Many Matching Markets*

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

Alessandro Pavan
Northwestern University
alepavan@northwestern.edu

This Version: March 2011
(preliminary version)

Abstract

This paper studies second-degree price discrimination in matching markets, that is, in markets where the product sold by a platform is access to other agents. In order to investigate the optimality of a large variety of pricing strategies, we tackle the problem from a mechanism design approach and allow the platform to offer *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 the 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 in 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 the access to other agents. In such markets, platforms engage in second-degree price discrimination by offering menus of matching plans. For concreteness, consider the problem of a Cable TV provider contracting with TV channels on one side of the market and with home customers (i.e., viewers) on the other side. The Cable company's problem can be seen from two perspectives. The more familiar one is that of designing a menu of packages of channels to offer to its customers. The mirror image of this problem consists in designing a price schedule for the channels whereby prices are contingent on the number of viewers the channel will be able to reach (more viewers yields higher advertising revenue). By the very nature of the matching problem, the menu of channels offered to the viewers pins down the quantity schedule faced by the channels and the price schedule offered to the channels pins down the packages that the platform can offer to the viewers. As such, when designing its profit-maximizing menus, the Cable company has to internalize the cross-side effects of the schedules offered to both the viewers and the channels.

Such interdependency is ubiquitous in two-sided matching markets. Health care providers, for example, offer menus of health plans that differ in 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.

As is often the case in this type of environments, what prevents the platform from appropriating the entire surplus is the fact that agents on both sides have private information both about their willingness to pay for the quality of the matching set they receive as well as about idiosyncratic characteristics that determine their attractiveness to those agents they are matched with.

The Model. To examine the problem described above in full generality (i.e., without imposing a priori restrictions on the possible matching schedules and/or on the admissible pricing strategies), we tackle the problem from a mechanism design approach. We consider the problem of a monopolistic platform that operates in a market with two sides (as explained below, the "group design" problem of a platform operating on a single side consisting in assigning agents to non-exclusive groups is a special case of the more general problem considered in the paper). Each agent on each side of the market has a multi-dimensional type, which is his/her private information. The first component captures the agent's willingness to pay for the quality of the matching set he/she receives (i.e., for the attributes of the agents from the other side he/she is matched to). In the Cable TV example, viewers have private information on their willingness to pay for different TV packages and channels have private information on the extra advertisement revenue they expect from reaching more viewers. All other components are idiosyncratic characteristics that determine the agent's attractiveness from the eyes of the agents on the other side (education, consumption habits, income, for viewers; the attractiveness of the shows and the quality of the advertisement, for channels[1]). All agents on the same side agree on the quality of the agents on the other side but may have different values (and hence different willingness to pay) for such quality. Importantly, we allow such values to be negative for some agents: for example, in the case of health care provision, the (negative) willingness to pay that certain doctors may have for accepting the patients included in a given HMO (Health Maintenance Organization) or PPO (Preferred Provider Organization) may originate from the opportunity costs of the doctors' time; in this example, the patients' relevant idiosyncratic characteristics are their medical conditions.[2]

In this environment, the platform's problem consists in choosing a matching rule together with a pricing rule so as to maximize profits (alternatively, 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 minimal feasibility constraint, which we call *reciprocity condition.* The latter 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.

The Main Results. At the theoretical level, what distinguishes the problem described above from a standard monopolistic screening problem (e.g., Mussa and Rosen (1978) and Maskin and Riley (1983)) is twofold. First, by the very nature of the matching problem, the platform faces feasibility constraints with no equivalent in the adverse selection and price discrimination literatures. Second, each agent is both a customer and an input in the matching production function. The "customer" role of an agent is summarized in his/her willingness to pay while the"input" role is captured by the idiosyncratic characteristics that determine the agent's attractiveness for the other side. This feature of matching markets implies that the cost of procuring an input is endogenous (it depends on the entire matching rule) and incorporates nontrivial strategic considerations.

As standard in the mechanism design literature, the problem of designing a profit-maximizing mechanism can be recasted entirely in terms of designing an optimal matching rule. This makes the profit-maximization problem analogous to welfare-maximization, except that the agents' willingness to pay for the quality of their matching sets are replaced with their virtual-valuation counterparts. The novelty and the intricacies here come from the characterization of the properties of the optimal (i.e., welfare- or profit-maximizing) matching rule.

First, we show that under two fairly natural conditions, namely (i) (weakly) decreasing marginal

[1] While it is easy for the platform to verify what shows and advertisement the channels offer, it is believed that the channels possess superior information about the attractiveness of their shows and advertisement for the viewers. Such superiority may originate for example from market research as well as past experiences.

[2] In settings richer than the one considered in our baseline model, the same agent may derive positive profits/utility from being matched to certain agents from the other side but negative profits/utility from being matched to other agents. This is a possibility we discuss in an extension and that we plan to examine further in future work.

utility for the quality of the matching set, and (ii) (weak) positive affiliation between willingness to pay and attractiveness, then both the profit-maximizing and the welfare-maximizing matching rules discriminate only on the basis of willingness to pay. In other words, two agents with the same willingness to pay are matched to the same group of agents, irrespective of any other unobservable characteristics that may differentiate the two agents in terms of their attractiveness for those agents they are matched to.

Second, we show that the welfare- and profit-maximizing matching rules have a threshold structure, according to which each agent is matched to all agents on the other side whose willingness to pay is greater than some threshold. This result also hinges on the assumptions of positive affiliation between willingness to pay and attractiveness and diminishing marginal utility for match quality. To understand this result, first note that, under positive affiliation, the expected matching quality of an agent increases with his willingness to pay for quality. Second, under diminishing marginal utility, using the same agent as an input to provide match quality to many agents is less costly than using different agents. As a consequence, the cost-minimizing way to provide a given matching quality to an agent with a given willingness to pay is to match him with the highest willingness-to-pay agents from the other side.

Building on the aforementioned results, we then show that both the welfare-maximizing and the profit-maximizing matching rule belong to one of the following two classes, and identify 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 whose matching set is non-empty have the same matching sets. A single network is thus the analog in matching environments to a single price (that is, the absence of quantity/quality discrimination) in the contest of 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 with a higher willingness to pay join an increasing number of networks. Under a nested multi-homing rule, the matching sets of any two agents from the same side 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 single-product monopolist.

We prove that a single network is (profit-) welfare-maximizing *if and only 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. When this is not the case, then the (profit-) welfare-maximizing matching rule exhibits nested multi-homing. Because valuations are always larger than their virtual analogs, this result implies that matching rules that employ a single network are more often associated to welfare-maximizing platforms, while multi-homing matching rules are more often associated to profit-maximizing platforms. This prediction appears consistent with casual empiricism: the public provision of broadcasting, health insurance, and job-matching services tends to employ a single network structure, while their private counterparts 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) in the following situation. Assume that virtual valuations for quality are positive for all agents and that the total quality q of a given matching set is determined by the sum of the interaction qualities provided by each of its members, i.e., that $q = \sum_j \sigma(\mathbf{u}_j)$, where $\sigma(\mathbf{u}_j)$ is the interaction quality provided by an agent with attributes $\mathbf{u}_j$. Lastly assume that the payoff that each agent with virtual valuation $\varphi(v)$ obtains from being matched to a set of agents with total quality q is given by $\varphi(v) \cdot g(q)$ where g is a weakly concave function. That is, $\varphi \cdot g(q)$ is the agent's total willingness to pay for quality q. In this case, optimality clearly requires that each agent be matched to all other agents, i.e., a complete network.

Things are different (and more interesting) when virtual valuations are negative for certain types on one or both sides of the market. In this case, it remains 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 increase profits by adding to the matching sets of those agents with the highest positive virtual 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 on one side to boost rent extraction on the other side. Whether, at the optimum, this cross-subsidization leads to a single network or to multi-homing is then determined by the marginal effect on profits of linking the two agents with the lowest virtual valuations on each side. If this effect is positive, then the optimal matching rule consists in creating a single (complete) network where each agent is matched to any other agent. If this effect is negative, the optimal rule separates agents based on their virtual valuations. Those agents with the highest virtual valuations are assigned matching sets which are supersets of those assigned to agents with lower virtual valuations; that is, the optimal 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 a nice Euler equation that equalizes the marginal (revenue) efficiency gains from expanding the matching set on one side to the marginal (revenue) efficiency losses that, by reciprocity, arise on the other side of the market. This endogenous cost structure is one of the fundamental features 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 identify conditions that ensure that the platform is willing to separate types as finely as possible. It turns out that the familiar *regularity condition* (Myerson, 1981), according to which virtual values $\varphi(v)$ are monotonically increasing, is not the right condition in our matching environment. As pointed out by Bulow and Roberts (1989), this condition controls for the marginal effect on revenue of providing a higher quality to an agent with value v. In a standard setting, because the marginal cost is independent of the agent's type, the monotonicity of the virtual values then implies the monotonicity of the trades. In contrast, in a matching environment, by virtue of reciprocity, the marginal cost is also a function of the agent's type, through the effect of adding the agent to the matching sets of other agents on the opposite side. For the optimal matching rule to be maximally separating, one must then require that the virtual values $\varphi(v)$ increase *faster* (with type v) than their corresponding marginal cross-side effects. The latter are given by the marginal effect of adding type v from side A to the matching set of any agent from side B who is currently matched to any other agent from side A with type above v. In analogy to Myerson (1981), we refer to this condition as *Strong Regularity*. Under this condition, bunching can only occur at "the top," i.e., for the highest v due to capacity constraints, that is, because the stock of agents on the other side of the market has been exhausted.

Public and Private Provision in Matching Markets. As a by-product of the characterization results described above, our analysis reveals that profit-maximization leads to two distortions relative to welfare maximization. The first distortion, the *exclusion effect*, comes from the fact that too many agents are completely excluded by the platform. For example, in the context of health care provision, too many patients are left without any insurance. The second effect, the *isolation effect*, comes from the fact that, under profit-maximization, each agent (who is not excluded) is matched only to a subset of his efficient matching set. Unlike in standard mechanism design problems, this distortion applies also to the agents with the highest valuations on each side of the market. To see why, note that the size of an agent's matching set depends on his virtual valuation for quality as well as on the cross-side effect of adding this agent to the matching sets of other agents on the opposite side. Although the virtual valuations of those agents with the highest valuations coincide with the true valuations, the cross-side effect for such agents is always lower under profit maximization than under welfare maximization. Indeed, while such cross-side effect is proportional to the true value of the marginal agent on the opposite side under welfare maximization, it is proportional to the virtual value of the marginal agent under profit maximization. Because virtual values are lower than true values on both sides, this implies that the matching sets are strictly smaller under profit maximization than under welfare maximization for all agents, including those at the top of the distribution.

Comparative Statics. The model offers a convenient framework for studying various comparative statics. In particular, our analysis delivers testable predictions about the effects on the platform's pricing strategy of shocks that alter the distribution of valuations and/or the distribution of the cross-side effects of the match qualities. For example, in the context of the Cable TV application, consider a positive shock to the viewers income distribution that leaves unchanged the distribution of the viewers' willingness to pay for the TV packages but that raises the profits that the channels expect from reaching these viewers (via an increase in advertisement revenue). Take a viewer with a high willingness to pay for TV packages. If nested multi-homing was optimal before the shock, then the package of channels offered to this viewer must necessarily include channels with a negative virtual valuation. These are channels whose value $v_{channel}$ for extra viewers is positive but low enough (possibly because of low advertisement revenue) that its virtual value $\varphi_{channel}$ is negative. In other words, with these channels, the platform makes losses on the channel side but profits on the viewer side. Now, a positive income shock that leaves the distributions of values unaltered on each side but that raises the attractiveness of the viewers has the perverse effect of increasing the cost of adding low-willingness-to-pay channels to the packages offered to high-willingness-to-pay viewers. This is because these shocks leave the positive marginal revenue on the viewer side $\varphi_{viewer} \cdot q_{channels}$ unaltered while increasing the (negative) marginal revenue on the channel side $\varphi_{channel} \cdot q_{viewers}$ ($\varphi_{channel} < 0$ is unaltered but $q_{viewers}$ has gone up as a consequence of the shock). Putting it differently, holding constant the matching function, the platform's cost of cross-subsidizing viewers with high valuations goes up when they become more valuable to the channels.[3] In contrast, the benefit of expanding the packages of those viewers with a low willingness to pay goes up, since these viewers are always matched to channels with a positive virtual value. As consequence, our model predicts that the platform's optimal response to a positive income shock to the viewers (more generally, to any shock that increases the viewers' attractiveness) is to improve the quality of the "basic" packages (those targeted to low-valuation viewers) and worsen the quality of the "premium" packages (those targeted to the high-valuation viewers). In terms of consumer surplus, these shocks thus make low-end viewers better off at the expenses of high-end ones.

Extensions. In order to implement its optimal matching rule, platforms offer price schedules to agents on both sides of the market who simultaneously choose their matching sets. Given such price schedules, agents play a coordination game. By design, this game has one equilibrium that implements the desired matching rule. This equilibrium, however, needs not be unique (*weak implementation,* in the mechanism design parlance, or the *chicken-and-egg problem*, in the two-sided markets parlance). We build on Weyl (2010) and construct *insulating payment rules* that, by conditioning payments on each side on the participation of agents from the opposite side, implement the optimal matching rule as a unique equilibrium (in dominant strategies).

In our baseline model, the platform matches agents on opposite sides without incurring any explicit costs (all "costs" are endogenous and stem from cross-side effects). We develop two extensions that introduce explicit costs to the platform's problem. First, we consider the effect of quasi-fixed costs, which are costs incurred by the platform for each agent that receives a non-empty matching set (in the Cable TV example, this is the cost of connecting the household to the underground cable system).

[3] This is because incentive compatibility requires that the platform offers larger discounts to those channels with a higher willingness to pay than the marginal channel.

Second, we introduce menu costs and study the coarse matching rules that appear in environments where the number of non-exclusive networks offered by the platform's menu is finite.

A special case of our model is that of a principal operating in a *single-sided*, rather than in a two-sided, market populated by multiple agents who experience differentiated peer effects from the other agents they interact with. In this setting, the principal's problem consists in assigning the agents to non-exclusive *groups* (rather than networks). We show that this one-sided group formation problem is equivalent to a two-sided matching problem where both sides have symmetric primitives and where the platform is constrained to selecting a symmetric matching rule. As it turns out, in two-sided markets with symmetric primitives, the efficient (as well as the profit-maximizing) matching rules are naturally symmetric. As such, all our results apply to this problem as well. It suffices to replace "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).

Outline of the Paper. The rest of the paper is organized as follows. Below, we close the introduction by briefly reviewing 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 to 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. It then uses these results to derive testable predictions about the effects of shocks to the distributions of valuations and to the distribution of the interaction qualities on prices and matching outcomes. Section 4 considers a few extensions. Section 5 concludes. All proofs omitted in the main text are in the Appendix at the end of the document.

Related Literature

As discussed above, this paper contributes to the literature on second-degree price discrimination (e.g., Mussa and Rosen (1978) and Maskin and Riley (1983)) by considering a setting where the product sold by the monopolist is access to other agents.[4] In addition, the paper is related to the following literatures.

Two-Sided Markets. Markets where agents buy access to other agents are the focus of the literature that studies monopolistic pricing in two-sided markets. This literature, however, restricts attention to a single network or to mutually exclusive networks (e.g., Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2008), Ambrus and Argenziano (2009), and Weyl (2010)).[5] In contrast, here we assume that platforms can design arbitrary matching functions and provide conditions for

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

[5]See Rysman (2009) for a recent survey of the two-sided markets literature.

the optimality of a single network relative to more sophisticated pricing schemes consistent with multi-homing matching rules.

Matching Design with Transfers. In the context of one-to-one matching, Damiano and Li (2007) and Johnson (2010) derive conditions on primitives for a profit-maximizing platform to induce 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. The key difference with respect to this literature is that we study second-degree price discrimination in many-to-many matching environments.

Group Design. As mentioned above, our two-sided matching model can be applied to solve (one-sided) group design problems with peer effects. Arnott and Rowse (1987) and Lazear (2001) study the problem of a school that, under complete information, wants to allocate students to disjoint classes. Besides restricting attention to mutually exclusive groups, these papers disregard the incomplete information issues that lie at the core of the present work.

Under incomplete information, Helsley and Strange (2000) analyze an economy with peer effects where agents can choose to stay in the public sector or secede to a private community. They allow for a single private community and disregard the matching design issues which are the focus of our analysis.[6]

More recently, Board (2009) and Rayo (2010) study profit-maximization by a monopolistic platform that can induce agents to self-select into mutually exclusive groups. Relative to these papers, we extend the analysis of matching design to two-sided environments and allow for matching rules that assign agents to non-exclusive groups.[7]

Cooperative Matching Theory. Our paper considers a many-to-many matching market in which agents have common preferences for agents on the other side of the market. In contrast, the matching theory surveyed in Roth and Sotomayor (1990) (for a more recent treatment, see Hatfield and Milgrom (2005)) and the recent literature on the school assignment problem (see, for example, Abdulkadiroglu, Pathak and Roth (2005a, 2005b) and Abdulkadiroglu and Sonmez (2003)) study one-to-one (or many-to-one) matching in a setting where agents have different rankings over agents on the other side. Moreover, these literatures are methodologically distinct from this paper, in that they focus on solution concepts such as stability and do not allow for transfers.

Decentralized Matching. 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/matching 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

[6] See also Epple and Romano (1998).

[7] In subsection 4.4, we specialize our model to solve the (one-sided) group design problem, and discuss in further detail the relation to the works of Board and Rayo.

and consider many-to-many matching rules. In a decentralized economy with perfect information, Farrell and Scotchmer (1998), Ellickson, Grodal, Scotchmer and Zame (1999) and Scotchmer (2002) characterize the core allocations of equilibrium models of group formation. Instead, we analyze centralized matching markets and 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

$$\pi_k^i(\mathbf{s}, p; \theta) \equiv v_k^i \cdot g_k(|\mathbf{s}|_k) - p, \tag{1}$$

where $g_k(\cdot)$ is a positive, strictly increasing, continuously differentiable, 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. Letting $F_k^{v,\sigma}$ denote the joint distribution of $(v_k, \sigma_l(\mathbf{u}_k))$, we then assume that such distribution is absolutely continuous and then denote by F_k^v the marginal distribution of $F_k^{v,\sigma}$ with respect to v_k (and by f_k^v its density) and by $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ the conditional distribution of the interaction quality $\sigma_l(\mathbf{u}_k)$ given v_k. We will assume that the function family $\left\langle F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)\right\rangle_{v_k}$ is uniformly continuous in v_k in the L_1-norm.

As is standard in the mechanism design literature, we also assume that the marginal distribution F_k^v of the willingness to pay is *regular* in the sense of Myerson (1981), meaning that the virtual values $v_k - [1 - F_k^v(v_k)]f_k^v(v_k)$ are strictly increasing.

In addition to the above technical conditions, we will assume that the following two key economic properties hold.

Condition 1 ***(Diminishing marginal utility)*** *The function g_k is weakly concave, $k = A, B$.*

Condition 2 ***(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, $\pi_k^i(\mathbf{s}, p; \theta) \equiv v_k^i \cdot \lambda(\mathbf{s}) - p$.*

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 $\pi_k^i(\mathbf{s}, p; \theta) = v_k^i \cdot \int_{j \in \mathbf{s}} v_l^j d\lambda(j) - p$.*

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* $\{\hat{\mathbf{s}}_k^i(\cdot)\}_{k=A,B}^{i\in[0,1]}$ along with a *payment rule* $\{\hat{p}_k^i(\cdot)\}_{k=A,B}^{i\in[0,1]}$ 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).[8]

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:

$$j \in \hat{\mathbf{s}}_A^i(\theta) \Leftrightarrow i \in \hat{\mathbf{s}}_B^j(\theta). \tag{2}$$

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

[8] 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.

truthfully. In our setting, a mechanism M is

$$\text{IR if } \Pi_k(\theta_k; M) \geq 0 \text{ for all } \theta_k \in \Theta_k, \tag{3}$$

$$\text{IC if } \Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M) \text{ for all } \theta_k, \hat{\theta}_k \in \Theta_k. \tag{4}$$

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 (3) and (4). 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

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

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 (3) and (4) 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* $g_k\left(\left.|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k\right) \geq g_k\left(\left.|\mathbf{s}_k(\mathbf{u}'_k, v'_k)\right|_k\right)$; *furthermore, except for a countable subset of* V_k,

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

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

3. *expected payments are given by*

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

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 $g_k\left(\left.|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k\right)$ 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 $\mathbb{E}[g_k\left(\left.|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)\right|_k\right)]$ 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 $g_k\left(\left.|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k\right)$ depends on $\mathbf{u}_k$, the expectation $\mathbb{E}[g_k\left(\left.|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)\right|_k\right)]$ 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

$$\Omega^h(M) = \sum_{k=A,B} \int_{\Theta_k} \varphi_k^h(v_k) \cdot g_k\left(\left.|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k) \tag{5}$$

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).[9] Hereafter, we will denote by $\{\mathbf{s}_k^h(\cdot)\}_{k=A,B}$ the matching rule that maximizes (5). 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 $\theta_k, \theta'_k \in \Theta_k$*

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

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 $\theta_k, \theta'_k \in \Theta_k$*

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

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.

[9] 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).

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* $\mathbf{s}^h_k(\cdot)$ *then has the following threshold structure*

$$\mathbf{s}^h_k(v_k) = \begin{cases} [t^h_k(v_k), \overline{v}_l] \text{ if } v_k \in [\omega^h_k, \overline{v}_k] \\ \oslash \text{ otherwise} \end{cases} \tag{6}$$

where the threshold $\omega^h_k \in [\underline{v}_k, \overline{v}_k]$ *determines which types are excluded and where the nonincreasing function* $t^h_k(\cdot)$ *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 $\varphi^h_k(v_k) \geq 0$. 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 $\mathbf{s}_k(\mathbf{u}_k, v_k) \supset \{\theta_l = (\mathbf{u}_l, v_l) : \varphi^h_l(v_l) \geq 0\}$ that includes all types $\theta_l = (\mathbf{u}_l, v_l)$ whose φ^h_l-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^h_l-valuation, and therefore adding type θ_k to these types' matching sets never reduces the platform's payoff $\Omega^h(M)$, as shown by (5). 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^h_l-valuation is non-negative, i.e., such that

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

Because of reciprocity, adding an agent whose φ^h_l-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 (6). 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

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

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

$$\Omega^h(M) = \sum_{k=A,B} \int_{\omega_k^h}^{\overline{v}_k} \hat{g}_k(t_k^h(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k). \tag{7}$$

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. This is obtained by assuming that the following condition holds, which strengthens the standard monotonicity of virtual values, as required in a two-sided matching environment.

Condition 3 ***(Strong Regularity)*** *The function*

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

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 in $\psi_k^h(v_k)$ 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 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 already matched to any other agent with valuation above v_k). 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 strictly increasing in v_k. The strong regularity condition above then requires that the value of an agent as a consumer increases faster than his value as an input as marginal willingness to pay 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". As we will see in Proposition 3 below, the key role played by strong regularity is to guarantee that bunching occurs only at the top or at the bottom of the value distribution. In this sense, it is the analog of Myerson standard regularity condition in two-sided matching problems. It also plays a role in the following proposition, but only in the special case where virtual valuations are always positive on one-side and both positive and negative on the other side.

Proposition 2 *Assume Condition 3 holds. 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 (7), that efficiency (respectively, 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 $\hat{\omega}_k^h \leq r_k^h$ 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 $k \in \{A, B\}$, $\hat{\omega}_k^h > r_k^h$ so that $\varphi_k^h(\hat{\omega}_k^h) > 0$. 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 $v_k \geq \hat{\omega}_k^h$ the same matching set as the original matching rule while it assigns to each agent with valuation $v_k \in [r_k^h, \hat{\omega}_k^h]$ the matching set $[\hat{v}_l^{\#}, \overline{v}_l]$, where $\hat{v}_l^{\#} \equiv \max\{r_l^h, \hat{\omega}_l^h\}$. 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 $\hat{\omega}_k^h \leq r_k^h$ for $k = A, B$.

Now suppose that $\hat{\omega}_k^h < r_k^h$ 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 $\mathbf{s}_k^{h\#}(\cdot)$ such that

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

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 $\hat{\omega}_k^h = r_k^h$ for $k \in \{A, B\}$ whereas $\hat{\omega}_l^h \leq r_l^h$. The platform could then do better by lowering the threshold type on side k and switching to the following multi-homing matching rule

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

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

$$-\hat{g}_k(r_l^h) \cdot \varphi_k^h(r_k^h) f_k^v(r_k^h) - \hat{g}_l'(r_k^h) \int_{r_l^h}^{\overline{v}_l} \varphi_l^h(v_l) dF_l^v(v_l) = -\hat{g}_l'(r_k^h) \int_{r_l^h}^{\overline{v}_l} \varphi_l^h(v_l) dF_l^v(v_l) > 0 \tag{8}$$

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 $\varphi_k^h(\underline{v}_k) \geq 0$. Note that this is the only case in which the strong regularity condition plays a role.

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

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{9}$$

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 (7) subject to the feasibility constraint (9) 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 $v_k < \bar{v}_k$.

The next proposition characterizes the h-optimal matching rule under the assumption that the strong regularity condition holds.

Proposition 3 *Let $\triangle^h < 0$ and assume Condition 3 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

$$\triangle_k^h(v_k, t_k^h(v_k)) = 0, \tag{10}$$

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 $t_k^h(v_k) \le r_l^h$ if and only if $v_k \ge r_k^h$. 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 (10).

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 (10) 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.

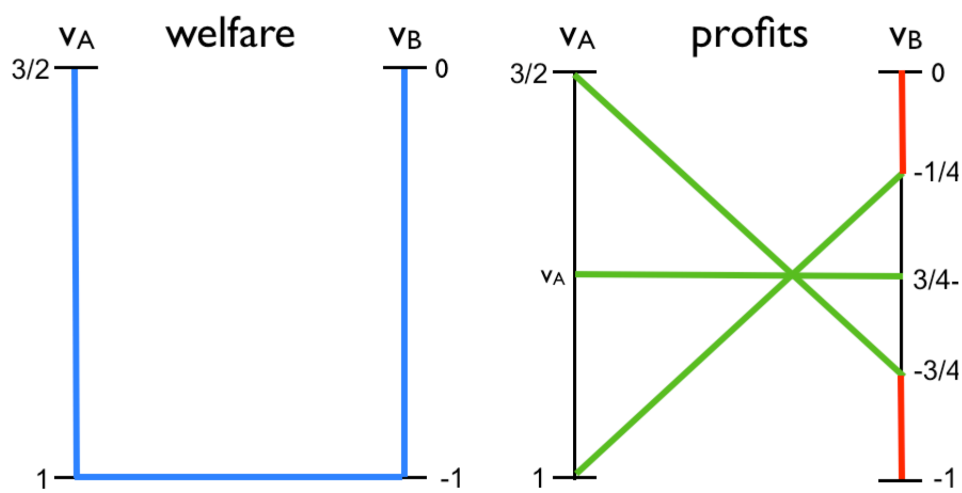

Figure 1: The health insurance plans offered by welfare-maximizing (public) and profit-maximizing (private) health-care providers when patients and doctors have valuations drawn from U[1,3/2] and U[-1,0], respectively.

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. *completely excludes a larger group of agents (**exclusion effect**); i.e.,* $\omega_k^P \geq \omega_k^W$ *for* $k = A, B$*, and*

2. *matches each agent on each side of the market to a subset of his efficient matching set (**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 intuition for both effects can be seen from equation (10): 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) and to exclusion of a larger group of agents.

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

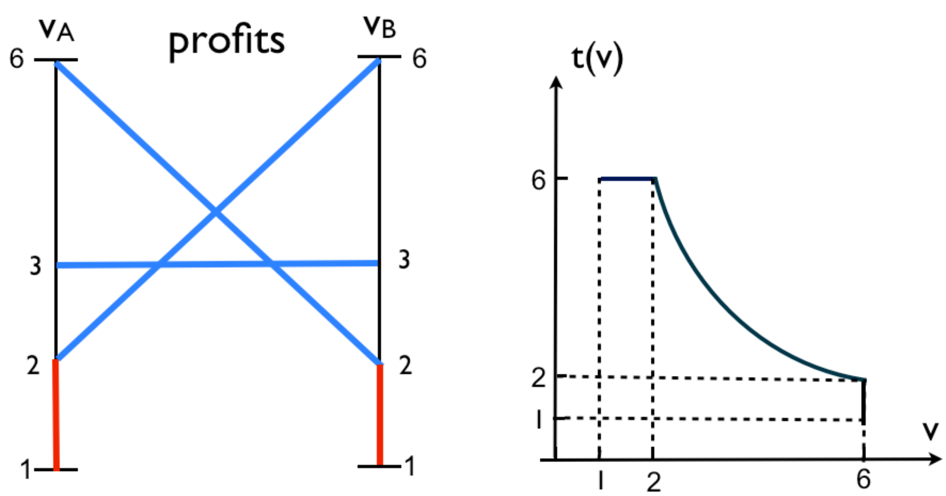

Figure 2: The menu offered by a profit-maximizing job employment agency when freelancers and firms have productivities drawn from U[1,6].

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* $2\underline{v} < \bar{v}$. *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 (10) 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* $[\underline{v}, \overline{v}] = [1, 6]$.

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 3 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. $\hat{\mathbf{s}}_l^P(v_l) \supseteq \mathbf{s}_l^P(v_l)$ *if and only if* $v_l \geq r_l^P$,

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. *suppose there exists a type* $\hat{\theta}_l$ *with* $\hat{\nu}_l \geq r_l^P$ *such that* $\Pi_l(\hat{\theta}_l; \hat{M}^P) \geq \Pi_l(\hat{\theta}_l; M^P)$; *then* $\Pi_l(\theta_l; \hat{M}^P) \geq \Pi_l(\theta_l; M^P)$ *for any type* θ_l *with* $v_l \geq \hat{\nu}_l$.

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 $v_k \leq r_k^P$ necessarily increase. On the other hand, since $|\hat{\mathbf{s}}_k(v_k)|_k \leq |\mathbf{s}_k(v_k)|_k$ for all $v_k \geq r_k^P$, then either payoffs increase for all agents on side k, or there exists a threshold type $\hat{\nu}_k > r_k^P$ such that the payoff of each agent on side k is higher under the new rule than under the original one if and only if $v_k \leq \hat{v}_k$. 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 effect on the payoffs of the agents on the opposite side is in general ambiguous. On the one hand, by virtue of reciprocity, the matching sets on side l expand for all agents with value $v_l \geq r_l^P$ and shrink for all agents with value $v_l < r_l^P$. On the other hand, the payoff that each agent on side k derives from interacting with side k increases as a result of the increase in popularity on side k. The net effect on payoffs can thus be ambiguous and nonmonotone in v_l. What remains true, though, is that, if there exists a type $\hat{\nu}_l \geq r_l^P$ who is better off, then necessarily the same is true for all types $v_l > \hat{\nu}_l$.

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 3 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. *suppose there exists a type $\hat{\theta}_l$ with $\hat{\nu}_l \geq r_l^P$ such that $\Pi_l(\hat{\theta}_l; \hat{M}^P) \geq \Pi_l(\hat{\theta}_l; M^P)$; Let $\hat{q}_l = \left|\hat{\mathbf{s}}_l^P(\hat{v}_l)\right|_l$. Then for any $q \geq \hat{q}_l$, $\hat{\rho}_l^P(q) \leq \rho_l^P(q)$.*

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 a decrease in the demand elasticity of side k.

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

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 3 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* $v_k \geq \omega_k^P$..

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 $v_l \geq \tilde{\nu}_l$.

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 $\tilde{F}_k^v$ dominates 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 3 holds. If the elasticity of side k decreases, then the platform moves from a price schedule $\rho_k^P(\cdot)$ to a price 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 and Robust Implementation

In order to implement an h-optimal mechanism $M^h = \{\mathbf{s}_k^h(\cdot), p_k^h(\cdot)\}_{k=A,B}$, the platform designs a coordination game which agents from both sides of the market simultaneously play. In the direct-revelation form of this game, an agent with type $(\mathbf{u}_k, v_k)$ sends a report $(\hat{\mathbf{u}}_k, \hat{v}_k)$, makes a payment $p_k^h(\hat{\mathbf{u}}_k, \hat{v}_k)$ as defined in Lemma 1, and is granted access to all agents on the opposite side who reported valuations above $t^h(\hat{v}_k)$. The construction of $\{p_k^h(\cdot)\}_{k=A,B}$ in Lemma 1 ensures that this game has one Nash equilibrium that implements the matching rule $\mathbf{s}_k^h(\cdot)$.

The implementation achieved by the payment rule $\{p_k^h(\cdot)\}_{k=A,B}$ is, however, partial since the coordination game described above has in general multiple equilibria. To see this, consider Example 6, in which the platform maximizes profits by matching free-lancers and firms according to the rule $\mathbf{s}_k^P(\cdot)$. In order to implement $\mathbf{s}_k^P(\cdot)$, the platform sets payments $\{p_k^P(\cdot)\}_{k=A,B}$ following Lemma 1. While reporting one's type $\theta_k = (\mathbf{u}_k, v_k)$ truthfully is a Nash equilibrium in the coordination game induced by $\{p_k^P(\cdot)\}_{k=A,B}$, this game has another (strict) Nash equilibrium in which free-lancers and firms report the lowest possible productivities $\underline{v}_A$ and $\underline{v}_B$, develop no projects and enjoy zero payoffs. Indeed, if all firms report $\underline{v}_B$, it is a strict best reply for every free-lancer to report $\underline{v}_A$ and pay nothing, since in equilibrium the platform offers no firms that free-lancers can work with. In the two-sided market literature, the multiplicity of equilibria (and the issue of partial implementation) is known as the "chicken and egg" problem (see Caillaud and Jullien (2001, 2003)) or the "failure to launch" problem (see Evans and Schmalensee (2009)).[10]

As pointed out by Weyl (2010) in the context of a monopolistic platform that employs a single network, the platform can circumvent this equilibrium multiplicity problem by designing *insulating tariffs* that condition the payment of each agent from side k on the joint reports of agents from side l (see also White and Weyl (2010)). In our model, consider an h-optimal matching rule $\mathbf{s}_k^h(\cdot)$ described by the threshold function $t_k^h(\cdot)$. Starting from $\mathbf{s}_k^h(\cdot)$, define the insulating payment rule $\varrho_k^h(v_k, (\hat{\theta}_l^i)^{i\in[0,1]})$ according to

$$\varrho_k^h(v_k, (\hat{\theta}_l^i)^{i\in[0,1]}) = v_k \cdot g_k\left(\left|\{i : \hat{v}_l^i \geq t_k(v_k)\}\right|_k\right) - \int_{\underline{v}_k}^{v_k} g_k\left(\left|\{i : \hat{v}_l^i \geq t_k(x)\}\right|_k\right) dx,$$

where $\{i : \hat{v}_l^i \geq t_k(v_k)\}$ is the set of all agents from side l who reported valuations above $t_k(v_k)$ and $\left|\{i : \hat{v}_l^i \geq t_k(v_k)\}\right|_k \equiv \int_{\{i:\hat{v}_l^i \geq t_k(v_k)\}} \sigma_k(\mathbf{u}_l^i) d\lambda(i)$.

The coordination game induced by $\left\{\varrho_k^h(v_k, (\hat{\theta}_l^i)^{i\in[0,1]})\right\}_{k=A,B}$ possesses a unique equilibrium; in this equilibrium, it is dominant for each agent to report truthfully. To see why this is the case, note that *for any* profile of reports $(\hat{\theta}_l^i)^{i\in[0,1]}$ by all agents on the other side, the quality of the matching set for each agent on side $k = A, B$ is increasing in his/her report. Along with the fact that the payments $\varrho_k^h(v_k, (\hat{\theta}_l^i)^{i\in[0,1]})$ satisfy the envelope formula, this ensures that reporting truthfully is optimal for each agent. Furthermore, the IR constraints are satisfied, since the payoff that each agent obtains by reporting truthfully is

$$\int_{\underline{v}_k}^{v_k} g_k\left(\left|\{i : \hat{v}_l^i \geq t_k(x)\}\right|_k\right) dx \geq 0,$$

which is positive, irrespective of the other agents' reports. The payment rule $\varrho_k^h(v_k, (\hat{\theta}_l^i)^{i\in[0,1]})$ thus implements the allocation rule $\mathbf{s}_k^P(\cdot)$ in dominant strategies.

4.2 Coarse Matching

The welfare and profit-maximizing matching rules when $\Delta^h < 0$ (that is, when multi-homing is optimal) employ menus with a continuum of matching sets that agents can choose from. In reality, however, matching platforms typically offer menus with a handful of alternatives. Cable companies, for example, offer a small number of packages of channels (although some of them allow viewers to add to these packages extra channels at an additional cost), and health care providers typically offer only a limited number of coverage plans. As pointed out by McAfee (2002) and Hoppe, Moldovanu and Ozdenoren (2010), the reason for such *coarse matching* is that platforms face costs for offering more complex menus (for example, menu costs for adding extra alternatives).[11]

[10]See also Ellison and Fudenberg (2003) and Ambrus and Argenziano (2009).

[11]See also Wilson (1989).

The analysis developed in the previous sections can be easily adapted to accommodate restrictions on the coarseness of the matching rules available to the platform. Denote by N the number of alternatives in the platform's menu. A matching rule implemented by a menu with N alternatives is referred to as an N-*coarse matching rule*. By the same logic of Proposition 1, the h-optimal rule in the class of all N-coarse matching rules does not depend on the vector of characteristics $\mathbf{u}_k$, and has a threshold structure. It therefore takes the following form:

$$\mathbf{s}_k(v_k) = \begin{cases} [\omega_l, \overline{v}_l] & \Leftrightarrow & v_k \in [t_k^1, \overline{v}_k] \\ [t_l^{N-1}, \overline{v}_l] & \Leftrightarrow & v_k \in [t_k^2, t_k^1) \\ \vdots & \vdots & \vdots \\ [t_l^2, \overline{v}_l] & \Leftrightarrow & v_k \in [t_k^{N-1}, t_k^{N-2}) \\ [t_l^1, \overline{v}_l] & \Leftrightarrow & v_k \in [\omega_k, t_k^{N-1}) \end{cases}, \tag{11}$$

where for $k \in \{A, B\}$

$$\underline{v}_k \leq \omega_k \equiv t_k^N \leq t_k^{N-1} \leq \ldots \leq t_k^2 \leq t_k^1 \leq t_k^0 \equiv \overline{v}_k \tag{12}$$

are the N thresholds that determine the matching sets on each side of the market. Incentive compatibility requires that $t_k^n \leq t_k^{n-1}$ for all $n \in \{1, ..., N\}$.

Such rules can be indirectly implemented by offering a menu of N mutually non-exclusive networks $\{[\omega_k, t_k^{N-1}), [t_k^{N-1}, t_k^{N-2}), ..., [t_k^2, t_k^1), [t_k^1, \overline{v}_k)\}$. All agents on side l who are not completely excluded by the platform obtain the "basic package" $[t_k^1, \overline{v}_k]$. On top of the basic package $[t_k^1, \overline{v}_k]$, agents can add "premium packages" represented by the lower intervals $[t_k^2, t_k^1), [t_k^3, t_k^2), ..., [\omega_k, t_k^{N-1})$. The price of each "extra package" typically depends on the list of packages already purchased by the agent, and can be obtained through the familiar integral pricing formula from Lemma 1.

The h-optimal N-coarse matching rule then solves the following $2N$-dimensional optimization problem:

$$\max_{\{t_k^N, t_k^{N-1}, ..., t_k^2, t_k^1\}_{k=A,B}} \quad \sum_{k=A,B} \sum_{n=1}^{N} \int_{t_k^n}^{t_k^{n-1}} \hat{g}_k(t_l^{N-n+1}) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) \tag{13}$$

subject to the monotonicity constraint (12). As the number of alternatives N increases (e.g., because menu costs decrease), the solution to (13) uniformly converges to the h-optimal nested multi-homing rule identified in Proposition 3. This follows from the fact that any weakly decreasing threshold function $t_k(\cdot)$ which is admissible under program (7) can be approximated arbitrarily well by step functions of the form (11) (in the sup-norm, i.e., in the norm of uniform convergence). As such, the maximally-separating nested multi-homing rules of Proposition 3 are the limit as N grows large enough of those offered when the number of non-exclusive networks is finite.

4.3 Quasi-Fixed Costs

Integrating an agent into a network structure typically involves a quasi-fixed cost. In the Cable TV example, the Cable company must connect an individual household to its underground cable system in order to permit the household to get access to its channels. Similarly, in the case of job matching services, firms and workers must incur the cost of setting up online profiles and building professional portfolios.[12] From the perspective of the platform, these costs are quasi-fixed, in the sense that they depend on the number of agents included in the network, but not on the matching sets offered to these agents.

In this subsection, we show how the analysis developed above can be extended to incorporate such costs. To this purpose, suppose that the platform has to incur a quasi-fixed cost $c_k > 0$ for each agent from side k who is included in the network (formally, for each agent who is matched to a non-empty set of agents from side l). The platform's problem of designing an h-optimal mechanism can then be solved in two steps.

Step 1 Ignore quasi-fixed costs and maximize (7) among all weakly decreasing threshold functions $t_k^h(\cdot)$.

Step 2 Given the optimal threshold function $t_k^h(\cdot)$ from Step 1, choose the h-optimal exclusion types ω_A^h, ω_B^h by solving the following problem:

$$\max_{\omega_A,\omega_B} \sum_{k=A,B} \int_{\omega_k}^{\overline{v}_k} \Big(\hat{g}_k(\max\{t_k^h(v_k),\omega_l\}) \cdot \varphi_k^h(v_k) - c_k\Big) \cdot dF_k^v(v_k).$$

At any interior solution (i.e., whenever $\omega_k^h > \underline{v}_k$), the exclusion types ω_A^h, ω_B^h therefore solve the following first-order conditions for $k, l \in \{A, B\}$:

$$\hat{g}_k(\max\{t_k^h(\omega_k^h),\omega_l^h\}) \cdot \varphi_k^h(\omega_k^h) - \hat{g}_l'(\omega_k^h) \cdot \int_{\max\{t_k^h(\omega_k^h),\omega_l^h\}}^{\overline{v}_l} \varphi_l^h(v_l) \cdot dF_l^v(v_l) = c_k, \tag{14}$$

where $t_k^h(\cdot)$ is the threshold rule characterized in Proposition 3 when $\Delta^h < 0$ (i.e., when nested multi-homing is optimal in the absence of quasi-fixed costs) and is given by $t_k^h(v_k) = \underline{v}_l$ for all $v_k \in [\underline{v}_k, \bar{v}_k]$ when $\Delta^h \geq 0$ (i.e., when a single network is optimal). The left-hand side of condition (14) reflects the effects on both sides k and l of marginally decreasing the exclusion type ω_k^h. At any interior optimum, these effects must offset the quasi-fixed cost c_k of adding type ω^k to the network.

The first-order condition above reveals that the exclusion types $\omega_k^h(c_A, c_B)$ are strictly increasing in the quasi-fixed costs c_A, c_B at any point where $\omega_k^h > \underline{v}_k$.[13] This observation leads to a testable empirical implication. Let $\Delta^h < 0$, that is, assume that, absent quasi-fixed costs, the h-optimal

[12] In the Cable TV example, the quasi-fixed cost is incurred by the platform, whereas in the job-matching example, the entry cost of uploading resumes and job profiles are paid by the firms and by the workers, i.e., by the agents on the two sides of the market. As long as these costs are common knowledge, this distinction is however irrelevant when it comes to the properties of the optimal mechanism.

[13] This follows from applying Cramer rule to the system defined by (14).

matching rule exhibits nested multi-homing. As quasi-fixed costs increase, so do the exclusion types ω_A^h, ω_B^h. For c_k sufficiently high, the exclusion types reach the reservation values r_k^h, in which case the platform switches from multi-homing to a single network. Therefore, our model predicts that, ceteris paribus, single networks are more often employed in matching markets with high quasi-fixed costs, while nested multi-homing is more prevalent in markets where the quasi-fixed costs of including agents into the network structure are low.

4.4 The Group Design Problem

The analysis so far studied the design of optimal many-to-many matching schemes in a two-sided market. As anticipated in the Introduction, the results apply also to one-sided matching problems. Consider, for example, the problem of a firm that has to design its internal communication system. This system determines which employees are directly connected in the firm's organizational chart. Young employees typically have greater willingness to interact (high v's) with other employees than senior employees (whose low, negative, v's, may stem from high opportunity costs). Furthermore, the value that an employee assigns to interacting with another employee typically depends on the latter's intrinsic characteristics (competence, work attitude, etc.) and these characteristics are often private information. The firm's problem then consists in designing the communication chart that maximizes the performance of the organization.

The problem described above is an example of the more general theme of how to assign agents to different groups in the presence of peer effects which is central to organization design. Such one-sided matching problem is a special case of the two-sided matching problems studied above. To understand why, note that the problem of designing non-exclusive groups in a one-sided matching setting is mathematically equivalent to the problem of designing an optimal matching rule in a two-sided matching setting with the additional requirement that the matching rule be symmetric across the two sides. Formally, this problem consists in choosing weakly decreasing threshold functions $t_A(\cdot), t_B(\cdot)$ so as to maximize (7) subject to the reciprocity constraint (9) and the additional symmetry constraint

$$t_A(v) = t_B(v). \tag{15}$$

Under this new constraint, maximizing (7) is equivalent to maximizing twice the objective associated to the one-sided matching problem (depending on the specific application, this objective could either be welfare, i.e., the sum of the payoffs of the individual agents forming the organization, or the platform's profits).

As it turns out, the symmetry constraint (15) in a two-sided matching problem is slack when the preferences and type distributions of sides A and B coincide (as is necessarily the case in the corresponding one-sided matching problem). This is immediate when $\triangle^h \geq 0$, that is, when a single complete network is h-optimal. Under nested-multi-homing, $\triangle^h < 0$, the characterization from

Proposition 3 reveals that, at any point where $t_k(\cdot)$ is strictly decreasing, because $\psi_l^h(\cdot) = \psi_k^h(\cdot)$,

$$t_k^h(v) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v)\right) = \left(\psi_k^h\right)^{-1}\left(-\psi_l^h(v)\right) = t_l^h(v).$$

Similarly, it is easy to see that (15) is also slack in the case the optimal rule exhibits bunching at the top. As a consequence, one can reinterpret all the results derived of the previous sections in terms of the group design problem. In this regard, a single complete network shall be interpreted as a *single complete group*, and nested multi-homing matching rules should be interpreted as *nested non-exclusive groups.*

In a recent paper on one-sided matching design, Board (2009) studies the problem of a monopolistic platform that assigns agents to *mutually exclusive* groups. In this context, Board 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 Board's paper, we extend the analysis of matching design to two-sided environments and allow for matching rules that assign agents to non-exclusive groups. By considering more general matching rules we obtain stronger results. For example, the isolation effects discussed above imply that the profit-maximizing rule indeed matches each agent to a subset of his/her efficient set. On the other hand, Board allows for more general preferences than the ones considered in this paper. Interestingly, in the class of preferences described by (1), our Proposition 2 reveals when Board's restriction to mutually exclusive groups entails any loss of generality.

Related is also the work by Rayo (2010) who considers a one-sided matching problem where the peer effect of a group is the average valuation of its members. As in Board (2010), Rayo restricts the platform to form mutually exclusive groups. In contrast to Board and the present paper, Rayo also characterizes the profit-maximizing group design problem when the hazard rate fails to be monotone.

5 Conclusions

In this paper we studied second-degree price discrimination in matching markets. We derived necessary and sufficient conditions for the platform to employ a single network or to engage in price-discriminatory matching (multi-homing). Interestingly, our model predicts that single networks are more often associated to the public (welfare-maximizing) provision of matching services, while multi-homing is more often employed by private (profit-maximizing) platforms. When multi-homing prevails, the platform designs its matching sets by weighting the efficiency (revenue) gains on one side of the market with the cross-subsidization losses on the other side. This endogenous cost structure leads to novel (sometimes perhaps counter-intuitive) comparative statics. In particular, we showed that high-valuation agents may lose from a positive shock to their popularity, since the endogenous cost of cross-subsidization goes up.

The analysis is worth extending in a number of important directions. First, in the present paper we assumed that the utility/profit that an agent assigns to any given matching set is independent of who else has access to it. In other words, we abstracted from "same-side" competitive effects. In advertising markets, for example, reaching a certain set of households becomes more valuable if competitors are excluded from reaching the same households. Given that "same-side" effects (e.g., congestion) are present in many matching markets, extending the analysis in this direction seems particularly promising.

Second, we assumed that each agent on a given side either benefits from being matched to *any* agent on the opposite side or dislikes being matched to *any* agent on the opposite side (the intensity varying with the particular agent he/she is matched to). In other words, we considered only "same-sign" externalities. In settings richer than the one analyzed here, the same agent may derive positive profits/utility from being matched to certain agents from the opposite side but negative profits/utility from being matched to other agents. More broadly, relaxing the assumption that agents agree on how to rank the attractiveness of agents on the opposite side also appear to be an interesting topic for future research.

Third, matching markets are often populated by a handful of competing matching platforms. Extending the analysis to the study of competition by matching platforms (in the spirit of Rochet and Stole (2002), for example) is likely to deliver new insights on the properties of these markets.[14]

6 Appendix

Proof of Proposition 1. If $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, then it is immediate from (5) that efficiency (respectively, 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 (12).

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 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 strictly 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 (12) 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:

[14]See Damiano and Li (2008) for a model in which two matchmakers compete through entry fees for agents on both sides of the market. They assume a one-to-one matching environment and restrict the platforms to establish single networks, where agents are randomly matched.

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

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

By construction, $g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, \cdot)|_k\right)$ is weakly increasing in v_k, for any $\mathbf{u}_k$, thus satisfying Condition 3 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

$$\int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k) \geq \int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s'}_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k), \;\; k = A, B \tag{16}$$

The rest of the proof shows that the matching rule $\hat{\mathbf{s}}_k(\cdot)$ reduces the "costs" of serving agents with negative φ_k^h-valuations relative to the original matching rule $\mathbf{s'}_k(\cdot)$; that is,

$$\int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s'}_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k) \leq \int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k). \tag{17}$$

Summing up (16) and (17) shows that the platform's objective is weakly greater under $\hat{\mathbf{s}}_k(\cdot)$ than under $\mathbf{s'}_k(\cdot)$, thus proving the result.

To establish (17), we start with the following definition.

Definition 7 *Take two random variables $z_1, z_2 : [a, b] \to \mathbb{R}$ and denote by F the probability measure on $[a, b]$. We say that z_2 is smaller than z_1 in the monotone concave order if $\mathbb{E}\left[g\left(z_2(\tilde{\omega})\right)\right] \leq \mathbb{E}\left[g\left(z_1(\tilde{\omega})\right)\right]$ for all weakly concave and weakly increasing functions $g : \mathbb{R} \to \mathbb{R}$.*

Since the function $g(x) = x$ is strictly increasing and weakly concave, it follows that if z_2 is smaller than z_1 in the monotone concave order, then $\mathbb{E}\left[z_2(\tilde{\omega})\right] \leq \mathbb{E}\left[z_1(\tilde{\omega})\right]$. Now let $z_1(v_k) \equiv \mathbb{E}_{\tilde{\mathbf{u}}_k}\left[|\mathbf{s'}_k(\tilde{\mathbf{u}}_k, v_k)|_k \,| v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\mathbf{u}}_k}\left[|\hat{\mathbf{s}}_k(\tilde{\mathbf{u}}_k, v_k)|_k \,| v_k\right]$, defined over the domain $[\underline{v}_k, r_k^h]$.

From the construction of $\hat{\mathbf{s}}_k(\cdot)$ and the assumption of positive affiliation, it follows that for all $\kappa \in [\underline{v}_k, r_k^h]$

$$\int_{\underline{v}_k}^{\kappa} \int_{\mathbf{U}_k} \left|\mathbf{s'}_k(\mathbf{u}_k, v_k)\right|_k dF_k(\mathbf{u}_k, v_k) \geq \int_{\underline{v}_k}^{\kappa} \int_{\mathbf{U}_k} |\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k \, dF_k(\mathbf{u}_k, v_k),$$

or, equivalently,

$$\int_{\underline{v}_k}^{\kappa} z_1(v_k) dF_k^v(v_k) \geq \int_{\underline{v}_k}^{\kappa} z_2(v_k) dF_k^v(v_k) \tag{18}$$

Incentive compatibility implies that $z_1(\cdot), z_2(\cdot)$ are weakly increasing in v_k. Denote by $[\dot{v}_k^1, \dot{v}_k^2]$, $[\dot{v}_k^3, \dot{v}_k^4]$, ..., $[\dot{v}_k^{2T-2}, \dot{v}_k^{2T}]$ the T (where $0 \leq T < \infty$) intervals in which $z_1(v_k) < z_2(v_k)$. Because $\int_{\underline{v}_k}^{r_k^h} z_1(v_k)dF_k^v(v_k) \geq \int_{\underline{v}_k}^{r_k^h} z_2(v_k)dF_k^v(v_k)$, it is clear that $\dot{T} \equiv \cup_{t=0}^{T-1}[\dot{v}_k^{2t+1}, \dot{v}_k^{2t+2}]$ is a proper subset of $[\underline{v}_k, r_k^h]$ whenever the inequality is strict.

Now construct $\dot{z}_2(\cdot)$ on the domain $[\underline{v}_k, r_k^h]$ so that:

1. $\dot{z}_2(v_k) = z_1(v_k) < z_2(v_k)$ for all $v_k \in \dot{T}$;

2. $z_2(v_k) \leq \dot{z}_2(v_k) = \alpha z_1(v_k) + (1-\alpha)z_2(v_k) \leq z_1(v_k)$, where $\alpha \in [0,1]$, for all $v_k \in [\underline{v}_k, r_k^h]\backslash\dot{T}$;

3. $\int_{[\underline{v}_k, r_k^h]\backslash\dot{T}} \{\dot{z}_2(v_k) - z_2(v_k)\} dF_k^v(v_k) = \int_{\dot{T}} \{z_2(v_k) - z_1(v_k)\} dF_k^v(v_k)$.

Because $\int_{\underline{v}_k}^{r_k^h} z_1(v_k)dF_k^v(v_k) \geq \int_{\underline{v}_k}^{r_k^h} z_2(v_k)dF_k^v(v_k)$, there always exists some $\alpha \in [0,1]$ such that 2 and 3 hold.

From the construction above, $\dot{z}_2(\cdot)$ is weakly increasing and $\int_{\underline{v}_k}^{r_k^h} \dot{z}_2(v_k)dF_k^v(v_k) = \int_{\underline{v}_k}^{r_k^h} z_2(v_k)dF_k^v(v_k)$. As a consequence, $z_2(\cdot)$ is a mean-preserving spread of $\dot{z}_2(\cdot)$. This implies that for all weakly concave and weakly increasing functions $g : \mathbb{R} \to \mathbb{R}$,

$$\int_{\underline{v}_k}^{r_k^h} g\left(z_2(v_k)\right) dF_k^v(v_k) \leq \int_{\underline{v}_k}^{r_k^h} g\left(\dot{z}_2(v_k)\right) dF_k^v(v_k) \leq \int_{\underline{v}_k}^{r_k^h} g\left(z_1(v_k)\right) dF_k^v(v_k),$$

where the first inequality follows from the weak concavity of $g(\cdot)$ and the second inequality follows from the fact that $\dot{z}_2(v_k) \leq z_1(v_k)$ for all $v_k \in [\underline{v}_k, r_k^h]$ and $g(\cdot)$ is weakly increasing. This shows that z_2 is smaller than z_1 in the monotone concave order.

The following lemma is useful.

Lemma 2 *Consider the positive random variables $z_1, z_2 : [a,b] \to \mathbb{R}_+$ which are weakly increasing functions of $\omega \in [a,b]$ and denote by F the probability measure on $[a,b]$. If z_2 is smaller than z_1 in the monotone concave order, then for any weakly concave and weakly increasing function $g(\cdot)$ and any weakly increasing function $h : [a,b] \to \mathbb{R}_-$*

$$\mathbb{E}\left[h(\tilde{\omega})g\left(z_1(\tilde{\omega})\right)\right] \leq \mathbb{E}\left[h(\tilde{\omega})g\left(z_2(\tilde{\omega})\right)\right].$$

Proof of Lemma 2. The proof argues that the inequality above is true for any weakly increasing step function $h^n : [a,b] \to \mathbb{R}_-$, where n is the number of steps. Because the set of weakly increasing step functions is dense (in the topology of uniform convergence) in the set of weakly increasing functions, the result follows.

Because z_2 is smaller than z_1 in the monotone concave order, the inequality above is obviously true for the one-step function h^1. Induction shows that this is true for all $n \in \mathbb{N}$. Q.E.D.

We are now ready to prove inequality (17). As it turns out,

$$\begin{aligned}
\int_{\Theta_k^-} \varphi_k^h(v_k)\cdot g_k\left(\left|\mathbf{s}'_k(\mathbf{u}_k,v_k)\right|_k\right) dF_k(\mathbf{u}_k,v_k) &= \int_{\underline{v}_k}^{r_k^h} \varphi_k^h(v_k)\cdot \mathbb{E}_{\tilde{\mathbf{u}}_k}\left[g_k\left(\left|\mathbf{s}'_k(\tilde{\mathbf{u}}_k,v_k)\right|_k\right)|v_k\right] dF_k^v(v_k) \\
&= \int_{\underline{v}_k}^{r_k^h} \varphi_k^h(v_k) g_k\left(z_1(v_k)\right) dF_k^v(v_k) \\
&= F_k^v(r_k^h)\cdot \mathbb{E}\left[\varphi_k^h(\tilde{v}_k) g_k\left(z_1(\tilde{v}_k)\right)|v_k \le r_k^h\right] \\
&\le F_k^v(r_k^h)\cdot \mathbb{E}\left[\varphi_k^h(v_k) g_k\left(z_2(v_k)\right)|v_k \le r_k^h\right] \\
&= \int_{\underline{v}_k}^{r_k^h} \varphi_k^h(v_k)\cdot g_k\left(\mathbb{E}_{\tilde{\mathbf{u}}_k}\left[|\hat{\mathbf{s}}_k(\tilde{\mathbf{u}}_k,v_k)|_k|v_k\right]\right) dF_k^v(v_k) \\
&= \int_{\Theta_k^-} \varphi_k^h(v_k)\cdot g_k\left(|\hat{\mathbf{s}}_k(\tilde{\mathbf{u}}_k,v_k)|_k\right) dF_k(\mathbf{u}_k,v_k).
\end{aligned}$$

The first equality follows from changing the order of integration. The second equality follows from the fact that, since $\mathbf{s}'_k(\cdot)$ is implementable, $g_k\left(|\mathbf{s}'_k(\mathbf{u}_k,v_k)|_k\right)$ is invariant in $\mathbf{u}_k$ except in a countable subset of $[\underline{v}_k, r_k^h]$. The first inequality follows from Lemma 2. The equality in the fifth line follows again from the fact that, by construction, $\hat{\mathbf{s}}_k(\cdot)$ is implementable, and hence invariant in $\mathbf{u}_k$ except in a countable subset of $[\underline{v}_k, r_k^h]$. The series of equalities and inequalities above establishes (17), as we wanted to show. Q.E.D.

Proof of Proposition 2: There are three cases to consider: (i) $\varphi_k^h(\underline{v}_k) \ge 0$ for $k = A, B$, (ii) $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, and (iii) $\varphi_l^h(\underline{v}_l) < 0$ while $\varphi_k^h(\underline{v}_k) \ge 0$. Cases (i) and (ii) are discussed in the main text, so here we consider case (iii). Below, we show that a single network is optimal if $\triangle^h > 0$, whereas a nested multi-homing is optimal if $\triangle^h < 0$.

From the same arguments as in case (ii) in the main text, if a single network is optimal, it must be that $\hat{\omega}_l^h < r_l^h$ and $\hat{\omega}_k^h = \underline{v}_k$. In turn, if nested multi-homing is optimal, it must be that $t_k^h(\underline{v}_k) \in (\underline{v}_l, r_l^h]$.

First, suppose that $\Delta^h > 0$ and, towards a contradiction, assume that the optimal mechanism employs a multi-homing rule. Take an arbitrary point $v_k \in [\underline{v}_k, \bar{v}_k]$ at which the function $t_k^h(\cdot)$ is strictly decreasing in a right neighborhood of v_k.Consider the effect of a marginal reduction in the threshold $t_k^h(v_k)$ around the point $v_l = t_k^h(v_k)$. This is given by

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

It is easy to see that

$$sign\{var(v_k, v_l)\} = sign\{Var(v_k, v_l)\}$$

where

$$Var(v_k, v_l) \equiv g'_k(|\mathbf{U}_l \times [v_l, \bar{v}_l]|_k)\cdot \mathbb{E}[\sigma_k(\tilde{\mathbf{u}}_l)|\tilde{v}_l = v_l]\cdot \psi_k^h(v_k) + \varphi_l^h(v_l)$$

Because (i) g_k is strictly positive, strictly increasing, and weakly concave, (ii) $\sigma_k(\mathbf{u}_l)$ and v_l are positively affiliated, (iii) $\psi_k^h(\cdot)$ is strictly increasing and nonnegative, we then have that $Var(v_k, v_l)$ is strictly increasing in v_k and v_l. Next, note that, given any interval $[v_k', v_k'']$ over which the function $t_k^h(\cdot)$ is constant and equal to v_l, the marginal effect of decreasing the threshold below v_l for any type $v_k \in [v_k', v_k'']$ is given by

$$\int_{v_k'}^{v_k''} [var(v_l, v_k)] dv_k$$

Finally note that $\Delta^h \geq 0$ implies $var(\underline{v}_k, \underline{v}_l) \geq 0$ and hence $var(v_k, v_l) > 0$ for all (v_k, v_l). The results above then imply that, starting from a multi-homing rule, the platform can strictly increase its objective by decreasing the threshold for any type for which $t_k^h(v_k) > \underline{v}_l$, proving that a single complete network is optimal.

Next, suppose that $\Delta^h < 0$ and, towards a contradiction, suppose that a single network is optimal. First consider the case where such a network is complete (that is, $\hat{\omega}_l^h = \underline{v}_l$ or, equivalently, $t_k^h(\underline{v}_k) = \underline{v}_l$). The fact that $\Delta^h < 0$ implies that $var(\underline{v}_k, \underline{v}_l) < 0$ which in turn implies that the marginal effect of raising the threshold $t_k^h(\underline{v}_k)$ for the lowest type on side k, while leaving the threshold untouched for all other types is positive. By continuity of the marginal effects, the platform can then improve its objective by switching to a multi-homing rule that is obtained from the complete network by increasing $t_k^h(\cdot)$ in a right neighborhood of $\underline{v}_k$ while leaving $t_k^h(\cdot)$ untouched elsewhere, contradicting the optimality of a single network.

Then consider the case where $\hat{\omega}_l^h > \underline{v}_l$. For a single network to be optimal, it must then be that $\hat{\omega}_l^h$ satisfies the following first-order condition

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

which requires that the total effect of a marginal increase of the size of the network on side l (obtained by reducing the threshold $t_k^h(v_k)$ below $\hat{\omega}_l^h$ for all types v_k) be zero. This rewrites as

$$\int_{\underline{v}_k}^{\bar{v}_k} [var(v_k, \hat{\omega}_l^h)] dv_k = 0.$$

Because $sign\{var(v_k, \hat{\omega}_l^h)\} = sign\{Var(v_k, \hat{\omega}_l^h)\}$ and $Var(v_k, \hat{\omega}_l^h)$ is strictly increasing in v_k, this means that there exists a $v_k^\# \in (\underline{v}_k, \bar{v}_k)$ such that

$$\int_{v_k^\#}^{\bar{v}_k} [var(v_k, \hat{\omega}_l^h)] dv_k > 0.$$

This means that there exists a $\omega_l^\# < \hat{\omega}_l^h$ such that the platform could increase her payoff by switching to the following nested multi-homing rule

$$s_k^h(v_k) = \begin{cases} [\omega_l^\#, \bar{v}_l] & \Leftrightarrow \quad v_k \in [v_k^\#, \bar{v}_k] \\ [\hat{\omega}_l^h, \bar{v}_l] & \Leftrightarrow \quad v_k \in [\underline{v}_k, v_k^\#] \end{cases}$$

thus contradicting the optimality of a single (incomplete) network. We thus conclude that multi-homing is necessarily optimal when $\Delta^h < 0$. Q.E.D.

Proof of Proposition 3. Using the result in Proposition 1, the h-optimal matching rule solves the following program, which we call the Full Program (P_F) :

$$P_F : \qquad \max_{\{\omega_k, t_k(\cdot)\}_{k=A,B}} \sum_{k=A,B} \int_{\omega_k}^{\overline{v}_k} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) \tag{19}$$

subject to the following constraints for $k, l = A, B$, $l \neq k$

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{20}$$

$$t_k(\cdot) \text{ weakly decreasing} \tag{21}$$

$$\text{and} \quad t_k(\cdot) : [\omega_k, \overline{v}_k] \to [\omega_l, \overline{v}_l] \tag{22}$$

with $\omega_k \in [\underline{v}_k, \bar{v}_k]$ and $\omega_l \in [\underline{v}_l, \bar{v}_l]$. Constraint (20) is the reciprocity condition, rewritten using the result in Proposition 1. Constraint (21) is the monotonicity constraint of Lemma 1, also rewritten using the result in Proposition 1. Finally, constraint (22) is a domain-codomain restriction which requires the function $t_k(\cdot)$ to map each type from side k that is included in the network into the set of types from side l that is also included in the network.

Because $\triangle^h < 0$ (i.e., because multi-homing is optimal), it must be that $r_k^h > \underline{v}_k$ for some $k \in \{A, B\}$. Furthermore, from the arguments in the proof of Proposition 1, at the optimum, $\omega_k^h \in [\underline{v}_k, r_k^h]$ and, whenever r_l^h exists (which, given the assumption that $\bar{v}_l > 0$, is the case if and only if $\varphi_l^h(\underline{v}_l) \leq 0$), then $\omega_l^h \in [\underline{v}_l, r_l^h]$ and $t_k^h(r_k^h) = r_l^h$. Hereafter, we will assume that r_l^h exists. When this is not the case, then $\omega_l^h = \underline{v}_l$ and $t_k^h(v_k) = \underline{v}_l$ for all $v_k \geq r_k^h$, while the optimal ω_k^h and $t_k^h(v_k)$ for $v_k < r_k^h$ are obtained from the solution to program P_F^k below by replacing r_l^h with $\underline{v}_l$).

Thus assume $\varphi_k^h(\underline{v}_k) \leq 0$ for $k = A, B$. Program P_F can then be decomposed into the following two independent programs P_F^k, $k = A, B$:

$$P_F^k : \qquad \max_{\omega_k, t_k(\cdot), t_l(\cdot)} \int_{\omega_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{23}$$

subject to $t_k(\cdot)$ and $t_l(\cdot)$ satisfying the reciprocity and monotonicity constraints (20) and (21), along with the following constraints

$$t_k(\cdot) : [\omega_k, r_k^h] \to [r_l^h, \overline{v}_l], \qquad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\omega_k, r_k^h].^{15} \tag{24}$$

Program P_F^k is not a standard calculus of variations problem. As an intermediate step, we will thus consider the following Auxiliary Program (P_{Au}^k), which strengthens constraint (21) and fixes $\omega_k = \underline{v}_k$ and $\omega_l = \underline{v}_l$:

$$P_{Au}^k : \qquad \max_{t_k(\cdot), t_l(\cdot)} \int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{25}$$

subject to (20),[16]

$$t_k(\cdot), t_l(\cdot) \textit{ strictly } \text{decreasing} \tag{26}$$

$$\text{and} \quad t_k(\cdot) : [\underline{v}_k, r_k^h] \to [r_l^h, \overline{v}_l], \quad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\underline{v}_k, r_k^h] \text{ are bijections.} \tag{27}$$

By virtue of (26), (20) can be rewritten as $t_k(v_k) = t_l^{-1}(v_k)$. Plugging this into the objective function (25) yields

$$\int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) dv_k + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_k^{-1}(v_l)) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l) dv_l. \tag{28}$$

Changing the variable of integration in the second integral in (28) to $\tilde{v}_l \equiv t_k^{-1}(v_l)$, using the fact that $t_k(\cdot)$ is strictly decreasing and hence differentiable almost everywhere, and using the fact that $t_k^{-1}(r_l^h) = r_k^h$ and $t_k^{-1}(\overline{v}_l) = \underline{v}_k$, the auxiliary program can be rewritten as follows:

$$P_{Au}^k : \qquad \max_{t_k(\cdot)} \int_{\underline{v}_k}^{r_k^h} \left\{ \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l(v_k) \cdot \varphi_l^h(t_k(v_k)) \cdot f_l^v(t_k(v_k)) \cdot t_k^{'}(v_k) \right\} dv_k \tag{29}$$

subject to $t_k(\cdot)$ being continuous, strictly decreasing, and satisfying the boundary conditions

$$t_k(\underline{v}_k) = \overline{v}_l \quad \text{and} \quad t_k(r_k^h) = r_l^h. \tag{30}$$

Consider now the Relaxed Auxiliary Program (P_R^k) that is obtained from P_{Au}^k by dispensing with the condition that $t_k(\cdot)$ be continuous and strictly decreasing and instead allowing for any measurable control $t_k(\cdot) : [\underline{v}_k, r_k^h] \to [r_l^h, \overline{v}_l]$ with bounded subdifferential that satisfies the boundary condition (30).

Lemma 3 *P_R^k admits a piece-wise absolutely continuous maximizer $\tilde{t}_k(\cdot)$.*

Proof of Lemma 3. Program P_R^k is equivalent to the following optimal control problem $\mathcal{P}_R^k$:

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

subject to

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

where K is a large number. Program $\mathcal{P}_R^k$ satisfies all the conditions of the Filipov-Cesari Theorem (see Cesari (1983)). By that theorem, we know that there exists a measurable function $y(\cdot)$ that solves $\mathcal{P}_R^k$. By the equivalence of P_R^k and $\mathcal{P}_R^k$, it then follows that P_R^k admits a piece-wise absolutely continuous maximizer $\tilde{t}_k(\cdot)$. Q.E.D.

[16]That is, both one-to-one (i.e., injective)\ and onto (i.e., surjective).

Lemma 4 *Let* $\tilde{v}_k \equiv \inf\{v_k \in [\underline{v}_k, r_k^h]$:*(31) admits a solution*$\}$*, where (31) is given by*

$$\hat{g}_k'(\eta(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) + \hat{g}_l'(v_k) \cdot \varphi_l^h(\eta(v_k)) \cdot f_l^v(\eta(v_k)) = 0. \tag{31}$$

The solution to P_R^k *is given by*

$$\tilde{t}_k(v_k) = \begin{cases} \bar{v}_l & \text{if} \quad v_k \in [\underline{v}_k, \tilde{v}_k] \\ \eta(v_k) & \text{if} \quad v_k \in (\tilde{v}_k, r_k^h]. \end{cases} \tag{32}$$

Proof of Lemma 4. From Lemma 3, we know that P_R^k admits a piece-wise absolutely continuous solution. Standard results from calculus of variations then imply that such solution $\tilde{t}_k(\cdot)$ must satisfy the Euler equation at any interval $I \subset [\underline{v}_k, r_k^h]$ where its image $\tilde{t}_k(v_k) \in (r_l^h, \overline{v}_l)$. The Euler equation associated to program P_R^k is given by (31). Condition 3 ensures that (i) there exists a $\tilde{v}_k \in [\underline{v}_k, r_k^h)$ such that (31) admits a solution if and only if $v_k \in [\tilde{v}_k, r_k^h]$, (ii) that at any point $v_k \in [\tilde{v}_k, r_k^h]$ such solution is unique and given by $\eta(v_k) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v_k)\right)$, and (iii) that $\eta(\cdot)$ is continuous and strictly decreasing over $[\tilde{v}_k, r_k^h]$.

When $\tilde{v}_k > \underline{v}_k$, (31) admits no solution at any point $v_k \in [\underline{v}_k, \tilde{v}_k]$, in which case $\tilde{t}_k(v_k) \in \{r_l^h, \overline{v}_l\}$. Because $\varphi_k^h(v_k) < 0$ for all $v_k \in [\underline{v}_k, \tilde{v}_k]$ and because $\hat{g}_k(\cdot)$ is decreasing, it is then immediate from inspecting the objective (29) that $\tilde{t}_k(v_k) = \bar{v}_l$ for all $v_k \in [\underline{v}_k, \tilde{v}_k]$.

It remains to show that $\tilde{t}_k(v_k) = \eta(v_k)$ for all $v_k \in [\tilde{v}_k, r_k^h]$. Because the objective function in P_k^r is not concave in (t_k, t_k') for all v_k, we cannot appeal to standard sufficiency arguments. Instead, using the fact that the Euler equation is a necessary optimality condition for interior points, we will prove that $\tilde{t}_k(v_k) = \eta(v_k)$ by arguing that there is no function $\hat{t}_k^h(\cdot)$ that improves upon $\tilde{t}_k(\cdot)$ and such that $\hat{t}_k(\cdot)$ coincides with $\tilde{t}_k(\cdot)$ except on an interval $(v_k^1, v_k^2) \subseteq [\tilde{v}_k, r_k^h]$ over which $\hat{t}_k^h(v_k) \in \{r_l^h, \overline{v}_l\}$.

To see that this is true, fix an arbitrary $(v_k^1, v_k^2) \subseteq [\tilde{v}_k, r_k^h]$ and consider the problem that consists in choosing optimally a step function $\hat{t}_k(\cdot) : (v_k^1, v_k^2) \to \{r_l^h, \overline{v}_l\}$. Because step functions are such that $\hat{t}_k'(v_k) = 0$ at all points of continuity and because $\varphi_k^h(v_k) < 0$ for all $v_k \in (v_k^1, v_k^2)$, it follows that the optimal step function is given by $\hat{t}_k(v_k) = \overline{v}_l$ for all $v_k \in (v_k^1, v_k^2)$. Notice that the value attained by the objective (29) over the interval (v_k^1, v_k^2) under such step function is zero. Instead, an interior control $t_k(\cdot) : (v_k^1, v_k^2) \to (r_l^h, \overline{v}_l)$ over the same interval with derivative

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

for all $v_k \in (v_k^1, v_k^2)$ yields a strictly positive value. This proves that the solution to P_R^k must indeed satisfy the Euler equation (31) for all $v_k \in [\tilde{v}_k, r_k^h]$. Together with the property established above that $\tilde{t}_k(v_k) = \bar{v}_l$ for all $v_k \in [\underline{v}_k, \tilde{v}_k]$, this establishes that the unique piece-wise absolutely continuous function that solves P_R^k is the control $\tilde{t}_k(\cdot)$ that satisfies (32). Q.E.D.

Denote by $\max\{P_R^k\}$ the value of program P_R^k (i.e., the value of the objective (29) evaluated under the control $\tilde{t}_k^h(\cdot)$ defined in Lemma 4). Then denote by $\sup\{P_{Au}^k\}$ and $\sup\{P_F^k\}$ the supremum of programs P^k_{Au} and P^k_F, respectively. Note that we write sup rather than max as, a priori, a solution to these problems might not exist.

Lemma 5 $\sup\{P^k_F\} = \sup\{P^k_{Au}\} = \max\{P^k_R\}$.

Proof of Lemma 5. Clearly, $\sup\{P^k_F\} \geq \sup\{P^k_{Au}\}$, for P^k_{Au} is more constrained than P^k_F. Next note that $\sup\{P^k_F\} = \sup\{\hat{P}^k_F\}$ where $\hat{P}^k_F$ coincides with P^k_F except that ω_k is constrained to be equal to $\underline{v}_k$ and $t_k(\underline{v}_k)$ is constrained to be equal to $\bar{v}_l$. This follows from the fact that excluding types below a threshold ω'_k gives the same value as setting $t_k(v_k) = \bar{v}_l$ for all $v_k \in [\underline{v}_k, \omega'_k)$. That $\sup\{\hat{P}^k_F\} = \sup\{P^k_{Au}\}$ then follows from the fact any pair of measurable functions $t_k(\cdot), t_l(\cdot)$ satisfying conditions (20), (21) and (24), with $\omega_k = \underline{v}_k$ and $t_k(\underline{v}_k) = \bar{v}_l$ can be approximated arbitrarily well in the L^2-norm by a pair of functions satisfying conditions (20), (26) and (27). That $\max\{P^k_R\} \geq \sup\{P^k_{Au}\}$ follows from the fact that P^k_R is a relaxed version of P^k_{Au}. That $\max\{P^k_R\} = \sup\{P^k_{Au}\}$ in turn follows from the fact that the solution $\tilde{t}^h_k(v_k)$ to P^k_R can be approximated arbitrarily well in the L^2-norm by a function $t_k(\cdot)$ that is continuous and strictly decreasing. Q.E.D.

From the results above, we are now in a position to exhibit the solution to P^k_F. Let $\omega^h_k = \tilde{v}_k$, where $\tilde{v}_k$ is the threshold defined in Lemma 4. Next for any $v_k \in [\tilde{v}_k, r^h_k]$, let $t^h_k(v_k) = \tilde{t}_k(v_k)$ where $\tilde{t}_k(\cdot)$ is the function defined in Lemma 4. Finally, given $t^h_k(\cdot) : [\omega^h_k, r^h_k] \to [r^h_l, \bar{v}_l]$, let $t^h_l(\cdot) : [r^h_l, \bar{v}_l] \to [\omega^h_k, r^h_k]$ be the unique function that satisfies (20). It is clear that the tripe $\omega^h_k, t^h_k(\cdot), t^h_l(\cdot)$ constructed this way satisfies conditions (20), (21) and (24), and is therefore a feasible candidate for program P^k_F. It is also immediate that the value of the objective (23) in P^k_F evaluated at $\omega^h_k, t^h_k(\cdot), t^h_l(\cdot)$ is the same as $\max\{P^k_R\}$. From Lemma 5, we then conclude that $\omega^h_k, t^h_k(\cdot), t^h_l(\cdot)$ is a solution to P^k_F.

Applying the construction above to $k = A, B$ and combining the solution to program P^A_F with the solution to program P^B_F then gives the solution $\left\{\omega^h_k, t^h_k(\cdot)\right\}_{k\in\{A,B\}}$ to program P_F.

By inspection, it is easy to see that the corresponding rule is maximally separating. Furthermore, from the arguments in Lemma 4, one can easily verify that there is exclusion at the bottom on side k (and no bunching at the top on side l) if $\tilde{v}_k > \underline{v}_k$ and bunching at the top on side l (and no exclusion at the bottom on side k) if $\tilde{v}_k = \underline{v}_k$. By the definition of $\tilde{v}_k$, in the first case, there exists a $v'_k > \underline{v}_k$ such that

$$-\hat{g}'_k(\bar{v}_l) \cdot \varphi^h_k(v'_k) \cdot f^v_k(v'_k) - \hat{g}'_l(v'_k) \cdot \varphi^h_l(\bar{v}_l) \cdot f^v_l(\bar{v}_l) = 0.$$

Using the definition of $\psi^h_k(\cdot)$ this is equivalent to

$$\psi^h_k(v'_k) + \psi^h_l(\bar{v}_l) = 0$$

The strict monotonicity of $\psi^h_k(\cdot)$ then implies that, in this case, $\psi^h_k(\underline{v}_k) + \psi^h_l(\bar{v}_l) < 0$ or, equivalently, that $\triangle^h_k(\underline{v}_k, \bar{v}_l) = \triangle^h_l(\bar{v}_l, \underline{v}_k) < 0$. Hence, whenever $\triangle^h_k(\underline{v}_k, \bar{v}_l) = \triangle^h_l(\bar{v}_l, \underline{v}_k) < 0$, there is exclusion at the bottom on side k and no bunching at the top on side l. Symmetrically, $\triangle^h_l(\underline{v}_l, \bar{v}_k) = \triangle^h_k(\bar{v}_k, \underline{v}_l) <$

0, implies that there is exclusion at the bottom on side l and no bunching at the top on of side k, as stated in the proposition.

Next, consider the case where $\tilde{v}_k = \underline{v}_k$. In this case there exists a $\eta(\underline{v}_k) \in [r_l^h, \bar{v}_l]$ such that

$$-\hat{g}_k'(\eta(\underline{v}_k)) \cdot \varphi_k^h(\underline{v}_k) \cdot f_k^v(\underline{v}_k) - \hat{g}_l'(\underline{v}_k) \cdot \varphi_l^h(\eta(\underline{v}_k)) \cdot f_l^v(\eta(\underline{v}_k)) = 0$$

or equivalently

$$\psi_k^h(\underline{v}_k) + \psi_l^h(\eta(\underline{v}_k)) = 0.$$

Assume first that $\eta(\underline{v}_k) < \bar{v}_l$. By the strict monotonicity of $\psi_l^h(\cdot)$ it then follows that $\psi_k^h(\underline{v}_k) + \psi_l^h(\bar{v}_l) > 0$ or, equivalently, that $\triangle_k^h(\underline{v}_k, \bar{v}_l) = \triangle_l^h(\bar{v}_l, \underline{v}_k) > 0$. Hence, whenever $\triangle_k^h(\underline{v}_k, \bar{v}_l) = \triangle_l^h(\bar{v}_l, \underline{v}_k) > 0$, there is no exclusion at the bottom on side k and bunching at the top on side l. Symmetrically, $\triangle_l^h(\underline{v}_l, \bar{v}_k) = \triangle_k^h(\bar{v}_k, \underline{v}_l) > 0$, implies that there is bunching at the top on side k and no exclusion at the bottom on side l, as stated in the proposition.

Next, consider the case where $\eta(\underline{v}_k) = \bar{v}_l$. In this case $\omega_k^h = \underline{v}_k$ and $t_k^h(\underline{v}_k) = \bar{v}_l$. This is the knife-edge case where $\triangle_k^h(\underline{v}_k, \bar{v}_l) = \triangle_l^h(\bar{v}_l, \underline{v}_k) = 0$ in which there is neither bunching at the top on side l nor exclusion at the bottom on side k.

Finally, note that the Euler equation (31) is equivalent to $\triangle_k^h(v_k, \eta(v_k)) = 0$. Using the fact that $t_k^h(v_k) = \eta(v_k)$ for all v_k in the separating range together with the fact that $\triangle_k^h(v_k, t_k^h(v_k)) \equiv \psi_k^h(v_k) + \psi_l^h(t_k^h(v_k))$ then establishes the last claim in the proposition that $t_k^h(v_k) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v_k)\right)$. Q.E.D.

Proof of Proposition 4. The result trivially holds when $\triangle^E \geq 0$, for in this case the welfare-maximizing matching rule always employs a single complete network. Thus suppose that $\triangle^W < 0$. Because $\varphi_k^P(v_k) \leq \varphi_k^W(v_k)$ for all $v_k \in [\underline{v}_k, \bar{v}_k]$, with strict inequality for all $v_k < \bar{v}_k$, then $\triangle^P$ is also strictly negative. Furthermore, the same property implies that $\psi_k^P(v_k) \leq \psi_k^W(v_k)$ for all $v_k \in [\underline{v}_k, \bar{v}_k]$. Now recall, from the arguments in the proof of Proposition 3, that the h-optimal rule exhibits exclusion at the bottom on side k if and only if $\triangle_k^h(\underline{v}_k, \bar{v}_l) = \triangle_l^h(\bar{v}_l, \underline{v}_k) < 0$ or, equivalently, if and only if $\psi_k^h(\underline{v}_k) + \psi_l^h(\bar{v}_l) < 0$. In this case, the threshold ω_k^h is the unique solution to $\psi_k^h(\omega_k^h) + \psi_l^h(\bar{v}_l) = 0$. The fact that $\omega_k^P \geq \omega_k^E$ then follows directly from the ranking between $\psi_k^P(\cdot)$ and $\psi_k^W(\cdot)$ along with the strict monotonicity of these functions. This establishes the exclusion effect.

Next, take any $v_k > \omega_k^P (\geq \omega_k^W)$ and suppose that $t_k^W(v_k) > \underline{v}_l$. The threshold $t_k^W(v_k)$ then solves $\psi_k^W(v_k) + \psi_l^W(t_k^W(v_k)) = 0$. The same monotonicities discussed above then imply that $t_k^P(v_k) > t_k^W(v_k)$. This establishes the exclusion effect. Q.E.D.

Proof of Proposition 5. Hereafter, we denote by "ˆ" all the variables in the mechanism $\hat{M}^P$ corresponding to the new distribution $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|\cdot)$ and continue to denote the variables in the mechanism M^P corresponding to the original distribution $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|\cdot)$ without annotation. By definition, we have that $\hat{\psi}_k^P(v_k) \geq \psi_k^P(v_k)$ for all $v_k \leq r_k^P$ while $\hat{\psi}_k^P(v_k) \leq \psi_k^P(v_k)$ for all $v_k \geq r_k^P$. Recall, from the arguments in the proof of Proposition 3, that for any $v_k < \omega_k^P$, $\Delta_k^P(v_k, \bar{v}_l) < 0$ or, equivalently, $\psi_k^P(v_k) + \psi_l^P(\bar{v}_l) < 0$, whereas for any $v_k \in (\omega_k, r_k^P]$, $t_k^P(v_k)$ satisfies $\psi_k^P(v_k) + \psi_l^P(t_k^P(v_k)) = 0$. The ranking between $\hat{\psi}_k^P(\cdot)$ and $\psi_k^P\cdot)$, along with the strict monotonicity of these functions then implies that $\hat{\omega}_k^P \leq \omega_k^P$ and, for any $v_k > \omega_k^P$, $\hat{t}_k^P(v_k) \leq t_k^P(v_k)$. Symmetrically, because $\hat{\psi}_k^P(v_k) + \psi_l^P(v_l) < \psi_k^P(v_k) + \psi_l^P(v_l)$ for all $v_k > r_k^P$, all v_l, we have that $\hat{t}_k^P(v_k) \geq t_k^P(v_k)$ for all $v_k > r_k^P$. These properties together with the reciprocity condition that links $t_l^P(\cdot)$ to $t_k^P(\cdot)$ establish parts (1) and (2) in the proposition.

Next note that, because F_l is unchanged, parts (1) and (2) also imply $|\hat{\mathbf{s}}_k(v_k)|_k \geq |\mathbf{s}_k(v_k)|_k$ if and only if $v_k \leq r_k^P$. We can then use Lemma 1 to conclude that for all types θ_k with value $v_k \leq r_k^P$

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

Furthermore, since $|\hat{\mathbf{s}}_k(v_k)|_k \leq |\mathbf{s}_k(v_k)|_k$ for all $v_k \geq r_k^P$, there exists a threshold type $\hat{\nu}_k > r_k^P$ (possibly equal to $\bar{v}_k$) such that $\Pi_k(\theta_k; \hat{M}^P) \geq \Pi_k(\theta_k; M^P)$ if and only if $v_k \leq \hat{\nu}_k$, which establishes part (3) in the proposition.

Finally notice that, for all $v_l \geq r_l^P$, because $\hat{t}_l^P(v_l) \leq t_l^P(v_l)$ (and hence $\hat{\mathbf{s}}_l(v_l) \supseteq \mathbf{s}_l(v_l)$) and because $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ dominates $F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ while $\hat{F}_k^v = F_k^v$, then necessarily $|\hat{\mathbf{s}}_l(v_l)|_l \geq |\mathbf{s}_l(v_l)|_l$. In contrast, for $v_l < r_l^P$ the comparison between $|\mathbf{s}_l(v_l)|_l$ and $|\hat{\mathbf{s}}_l(v_l)|_l$ is ambiguous, On the one hand, these types are now matched to a smaller matching set, i.e., $\hat{\mathbf{s}}_l(v_l) \subseteq \mathbf{s}_l(v_l)$. On the other hand, the expected quality of each agent in the matching set is now higher given that $\hat{F}_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k) \succeq F_k^{\sigma_l(\mathbf{u}_k)}(\cdot|v_k)$ in the usual stochastic order. Nonetheless, if there exists a $\hat{v}_l \geq r_l^P$ who is better off, i.e., for whom

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

then necessarily $\Pi_l(\theta_l; \hat{M}^P) \geq \Pi_l(\theta_l; M^P)$ for each type θ_l whose valuation $v_l \geq \hat{\nu}_l$, which establishes part (4). Q.E.D.

Proof of Corollary 2. Let $x_k(v_k) \equiv \left|\mathbf{s}_k^P(v_k)\right|_k$ denote the quality of the matching set that each agent with value v_k obtains under the original mechanism, and $\hat{x}_k(v_k) \equiv \left|\hat{\mathbf{s}}_k^P(v_k)\right|_k$ the corresponding quantity under the new mechanism. Using Lemma 1, for any $q \in x_k(V_k) \cap \hat{x}_k(V_k)$, i.e., for any q offered both under M and $\hat{M}$

$$\begin{aligned} \rho_k^P(q) &= x_k^{-1}(q)q - \int_{\underline{v}_k}^{x_k^{-1}(q)} x_k(v)dv \\ \hat{\rho}_k^P(q) &= \hat{x}_k^{-1}(q)q - \int_{\underline{v}_k}^{\hat{x}_k^{-1}(q)} \hat{x}_k(v)dv \end{aligned}$$

where $x_k^{-1}(q) = \inf\{v_k : x_k(v_k) = q\}$ is the generalized inverse of $x_k(\cdot)$ and $\hat{x}_k^{-1}(q) = \inf\{v_k : \hat{x}_k(v_k) = q\}$ the corresponding inverse for $\hat{x}_k(\cdot)$. We thus have that

$$\rho_k^P(q) - \hat{\rho}_k^P(q) = \int_{\underline{v}_k}^{x_k^{-1}(q)} [\hat{x}_k(v) - x_k(v)]dv + \int_{x_k^{-1}(q)}^{\hat{x}_k^{-1}(q)} [\hat{x}_k(v) - q]dv$$

From the results in Proposition 5, we know that $[x_k(v_k) - \hat{x}_k(v_k)][v_k - r_k^P] \geq 0$ with $x_k(r_k^P) = \hat{x}_k(r_k^P)$. Therefore, for all $q \in x_k(V_k) \cap \hat{x}_k(V_k)$, with $q \leq x_k(r_k^P) = \hat{x}_k(r_k^P)$,

$$\begin{aligned}
\rho_k^P(q) - \hat{\rho}_k^P(q) &= \int_{\underline{v}_k}^{x_k^{-1}(q)} [\hat{x}_k(v) - x_k(v)]dv - \int_{\hat{x}_k^{-1}(q)}^{x_k^{-1}(q)} [\hat{x}_k(v) - q]dv \\
&= \int_{\underline{v}_k}^{\hat{x}_k^{-1}(q)} [\hat{x}_k(v) - x_k(v)]dv + \int_{\hat{x}_k^{-1}(q)}^{x_k^{-1}(q)} [q - x_k(v)]dv \\
&\geq 0
\end{aligned}$$

whereas for $q \geq x_k(r_k^P) = \hat{x}_k(r_k^P)$,

$$\begin{aligned}
\rho_k^P(q) - \hat{\rho}_k^P(q) &= \int_{\underline{v}_k}^{r_k^P} [\hat{x}_k(v) - x_k(v)]dv + \int_{r_k^P}^{x_k^{-1}(q)} [\hat{x}_k(v) - x_k(v)]dv + \int_{x_k^{-1}(q)}^{\hat{x}_k^{-1}(q)} [\hat{x}_k(v) - q]dv \\
&= \rho_k^P(x_k(r_k^P)) - \hat{\rho}_k^P(x_k(r_k^P)) + \int_{r_k^P}^{x_k^{-1}(q)} [\hat{x}_k(v) - x_k(v)]dv + \int_{x_k^{-1}(q)}^{\hat{x}_k^{-1}(q)} [\hat{x}_k(v) - q]dv \\
&= \rho_k^P(x_k(r_k^P)) - \hat{\rho}_k^P(x_k(r_k^P)) + \left(\int_{r_k^P}^{\hat{x}_k^{-1}(q)} \hat{x}_k(v)dv - \hat{x}_k^{-1}(q)q\right) - \left(\int_{r_k^P}^{x_k^{-1}(q)} x_k(v)dv - x_k^{-1}(q)q\right)
\end{aligned}$$

Integrating by parts, using the fact that $x_k(r_k^P) = \hat{x}_k(r_k^P)$, and changing variables we have that

$$\begin{aligned}
&\left(\int_{r_k^P}^{\hat{x}_k^{-1}(q)} \hat{x}_k(v)dv - \hat{x}_k^{-1}(q)q\right) - \left(\int_{r_k^P}^{x_k^{-1}(q)} x_k(v)dv - x_k^{-1}(q)q\right) \\
= &\left(r_k^P \hat{x}_k(r_k^P) - \int_{r_k^P}^{\hat{x}_k^{-1}(q)} v\frac{d\hat{x}_k(v)}{dv}dv\right) - \left(r_k^P x_k(r_k^P) - \int_{r_k^P}^{x_k^{-1}(q)} v\frac{dx_k(v)}{dv}dv\right) \\
= &-\int_{x_k(r_k^P)}^{q} (\hat{x}_k^{-1}(z) - x_k^{-1}(z))dz.
\end{aligned}$$

Because $\hat{x}_k^{-1}(z) \geq x_k^{-1}(z)$ for $z > x_k(r_k^P)$, we then conclude that the price differential $\rho_k^P(q) - \hat{\rho}_k^P(q)$, which is positive at $q = x_k(r_k^P) = \hat{x}_k(r_k^P)$, declines as q grows above $x_k(r_k^P)$. Going back to the original notation, 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 equal to $\left|\hat{\mathbf{s}}_k^P(\bar{v}_k)\right|_k$) such that $\hat{\rho}_k^P(q) \leq \rho_k^P(q)$ if and only if $q \leq \hat{q}_k$. This establishes part (1) in the proposition.

Next, consider part (2). Suppose there exists a type $\hat{\theta}_l \geq r_l^P$ such that $\Pi_l(\hat{\theta}_l; \hat{M}^P) \geq \Pi_l(\hat{\theta}_l; M^P)$. Using Lemma 1, this means that

$$\int_{\underline{v}_l}^{\hat{v}_l} [\hat{x}_l(v) - x_l(v)]dv \geq 0.$$

Let $\hat{q}_l \equiv \left|\hat{\mathbf{s}}_l^P(\hat{\nu}_l)\right|_l$. From Proposition 5, we then know that, for any $v_l \geq \hat{\nu}_l$, $\hat{x}_l(v_l) > x_l(v_l)$. This in turn implies that for any $q > \hat{q}_l$, with $q \in x_l(V_l) \cap \hat{x}_l(V_l)$,

$$\begin{aligned}
\rho_l^P(q) - \hat{\rho}_l^P(q) &= \int_{\underline{v}_l}^{x_l^{-1}(q)} [\hat{x}_l(v) - x_l(v)]dv - \int_{\hat{x}_l^{-1}(q)}^{x_l^{-1}(q)} [\hat{x}_l(v) - q]dv \\
&= \int_{\underline{v}_l}^{\hat{x}_l^{-1}(q)} [\hat{x}_l(v) - x_l(v)]dv + \int_{\hat{x}_l^{-1}(q)}^{x_l^{-1}(q)} [q - x_l(v)]dv \\
&\geq 0
\end{aligned}$$

which establishes the result. Q.E.D.

Proof of Proposition 6. Denote by $\tilde{\psi}_k^P(v_k)$ the psi-function associated to the new distribution $\tilde{F}_k^v(\cdot)$, and 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, because g_k and g_l are linear, $\tilde{\psi}_k^P(v_k) \leq \psi_k^P(v_k)$ for all $v_k \in V_k$. Therefore for all $v_k \in V_k$,

$$\tilde{\psi}_k^P(v_k) + \psi_l^P(t_k^P(v_k)) < \psi_k^P(v_k) + \psi_l^P(t_k^P(v_k)).$$

From the arguments in the proof of Proposition 3, we then have that $\tilde{\omega}_k^P \geq \omega_k^P$ and $\tilde{t}_k^P(v_k) \geq t_k^P(v_k)$ for v_k, which establishes part 1. Part 2 follows from reciprocity.

Because F_l is unchanged, it then follows that $|\tilde{\mathbf{s}}_k(v_k)|_k \leq |\mathbf{s}_k(v_k)|_k$ for all $v_k \geq \omega_k^P$. Furthermore, because necessarily $|\tilde{\mathbf{s}}_k(v_k)|_k < |\mathbf{s}_k(v_k)|_k$ for all $v_k < r_k^P$, we have that $\Pi_k(\theta_k; \tilde{M}^P) < \Pi_k(\theta_k; M^P)$ for all $v_k \geq \omega_k^P$, which proves part 3. Part (4) is proved in the main text. Q.E.D.

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

References

[1] Abdulkadiroglu, A., P. Pathak and A. Roth, 2005a, Boston Public School Matching, *American Economic Review*, Vol. 95(2), pp. 368–71.

[2] Abdulkadiroglu, A., P. Pathak and A. Roth, 2005b, The New York City High School Match, *American Economic Review*, Vol. 95(2), pp. 364–367.

[3] Abdulkadiroglu, A. and T. Sonmez, 2003, School Choice: A Mechanism Design Approach, *American Economic Review*, Vol. 93(3), pp. 729–47.

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

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

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

[7] Arnott, R. and J. Rowse, 1987, Peer Group Effects and Educational Attainment, *Journal of Public Economics*, Vol. 32, pp. 287-305.

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

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

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

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

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

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

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

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

[16] Damiano, E. and H. Li, 2008, Competing Matchmaking, *Journal of the European Economic Association*, Vol. 6, pp. 789-818.

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

[18] Ellickson, B., B. Grodal, S. Scotchmer and W. Zame, 1999, Clubs and the Market, *Econometrica*, Vol. 67(5), pp. 1185-1217.

[19] 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.

[20] Ellison, G. and D. Fudenberg, 2003, Knife-Edge or Plateau: When do Market Models Tip?, *Quarterly Journal of Economics*, Vol. 118(4), pp. 1249-1278.

[21] Epple, D. and R. Romano, 1998, Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects", *American Economic Review*, Vol. 88, pp. 33-62.

[22] Evans, D. and R. Schmalensee, 2009, Failure to Lauch: Critical Mass in Platform Businesses, *Working Paper*.

[23] Farrell, J. and S. Scotchmer, 1988, Partnerships, *Quarterly Journal of Economics*, Vol. 103, pp. 279-297.

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

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

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

[27] 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.

[28] Lazear, E., 2001, Educational Production, *Quarterly Journal of Economics*, Vol. 116, pp. 777-803.

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

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

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

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

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

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

[35] Rayo, L., 2010, Monopolistic Signal Provision, *B.E. Journal of Theoretical Economics.*

[36] Rochet, J.-C. and L. Stole, 2002, Nonlinear Pricing with Random Participation, *Review of Economic Studies*, Vol. 69(1), pp. 277-311.

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

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

[39] 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.

[40] Rysman, M., 2009, The Economics of Two-Sided Markets, *Journal of Economic Perspectives*, Vol. 23(3), pp. 125-143.

[41] Scotchmer, S., 2002, Local Public Goods and Clubs, In *Handbook of Public Economics. Volume 4* (Alan J. Auerbach and Martin Feldstein, eds.), 1997-2042, Elsevier, Amsterdam.

[42] 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.

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

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

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

[46] White, A. and G. Weyl, 2010, Imperfect Platform Competition: A General Framework, *Working Paper.*

[47] Wilson, R., 1989, Efficient and Competitive Rationing, *Econometrica*, Vol. 57, pp. 1-40.