Many-to-Many Matching Design*

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

Alessandro Pavan
Northwestern University
alepavan@northwestern.edu

April 2012

Abstract

We study second-degree price discrimination in markets where the product traded by the monopolist is access to other agents. We derive necessary and sufficient conditions for the welfare- and the profit-maximizing mechanisms to employ a single network or a menu of non-exclusive networks. We characterize the optimal matching schedules under a wide range of preferences, derive implications for prices, and deliver testable predictions relating the structure of the optimal pricing strategies to conditions on the distribution of match qualities. Our analysis sheds light on the distortions associated with the private provision of broadcasting, health insurance and job matching services.

JEL classification: D82

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

*Previous versions circulated under the title "Many-to-Many Matching Design and Price Discrimination"and "Price Discrimination in Many-to-Many Matching Markets." For comments and useful suggestions, we thank seminar participants at UC Berkeley, Bocconi, Cambridge, Carlos III, Duke, École Polytechnique, Northwestern, NYU, Pompeu Fabra, Stanford, Toulouse, Yale, UBC, UPenn, FGV-EPGE, FGV-EESP, 2011 Columbia-Duke-Northwestern IO Converence, SED 2011, 2011 Csef-Igier Conference, 2011 Stony Brook Game Theory Festival, ESEM 2011, EARIE 2011, ESWM, 2012. We are grateful to Jacques Crémer, Johannes Horner, Doh-Shin Jeon, Bruno Jullien, Nenad Kos, Volker Nocke, Marco Ottaviani, Francisco Ruiz-Aliseda, Ron Siegel, Bruno Strulovici, Jean Tirole, Glen Weyl, and Mike Whinston for helpful conversations and useful comments. André Veiga provided excellent research assistance. Alessandro Pavan also thanks the Haas School of Business and the Economics Department at UC Berkeley for hospitality during the 2011-2012 academic year and Nokia for kind financial support. The usual disclaimer applies.

1 Introduction

This paper studies second-degree price discrimination in markets where the product sold by a platform is access to other agents. In such markets, platforms offer menus of matching plans at different prices. For concreteness, consider the problem of a Cable TV provider contracting with TV channels on one side of the market and with 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 viewers. The mirror image of this problem consists in designing a matching schedule for channels where prices are contingent on the viewers that each channel is able to reach (more viewers yield higher advertising revenue). The menu of packages of channels offered to viewers determines the matching schedule faced by channels, while the matching schedule offered to channels determines the packages that the platform can offer to viewers. As a consequence, when designing its menus on each side, the cable company has to internalize the cross-side effects on profits that each side induces on the other side.

Similar problems emerge in many other two-sided matching markets. Consider, for example, the provision of health care services. Health care providers 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). Here too, the plans that the platform offers to patients determine the services that the platform has to procure on the doctor side. As such, market conditions on the doctor side determine the profitability of price-discriminatory strategies on the patient side.

The presence of such cross-side effects is what distinguishes price discrimination in many-to-many matching markets from price discrimination in markets for standard products. Other examples of price discrimination in many-to-many matching markets can be found in online advertising, credit cards markets, as well as in the practices followed by online employment agencies and business directories.

This paper builds a model that examines the implications of such cross-side effects for questions such as: What matching allocations and network structures are likely to emerge under profit maximization (private provision of matching services) and which ones under welfare maximization (public provision)? How are these allocations affected by shocks that alter the distribution of valuations and/or the attractiveness of the two sides? What price schemes sustain such allocations?

Model preview. We consider the problem of a monopolistic platform that operates in a market with two sides. What prevents the platform from appropriating the entire surplus is the fact that each agent from each side has private information both about his willingness to pay for the quality of his matching set (his valuation) and about various idiosyncratic characteristics that determine the agent's attractiveness for the other side. Consider the Cable TV example. In addition to his willingness to pay for the various packages, each viewer is likely to possess private information about idiosyncratic characteristics (e.g., educational background, consumption habits, income, etc.) that determine the advertising revenue that the channels expect from reaching the viewer. Likewise, channels have private information both about the revenue that they expect from reaching the viewers, as well as about the attractiveness of the shows and of the advertisement that they broadcast. Similarly, in the context of health care provision, patients possess private information both about their willingness to pay to join different networks as well as about idiosyncratic characteristics (e.g., health status, life style, various risks) that determine the surplus that the physicians obtain from being matched to these patients (as these characteristics affect the difficulty of the treatment, for instance).

All agents from the same side agree on the quality of the agents from the other side, but differ in their willingness to pay for such quality (vertical differentiation). Importantly, we allow such valuations to be negative for some agents: For example, in the case of health care provision, a physician's (negative) willingness to pay for accepting additional patients may originate from the physician's opportunity cost of time.

In such an environment, the platform's problem consists in choosing a menu of matching plans for each side of the market. Each item in the menu specifies a set of agents of given characteristics from the other side, along with a price. The objective is to characterize the profit-maximizing menus and compare the network allocations that they induce with the ones induced under welfare-maximization. In order to address these issues in full generality (i.e., without imposing a priori restrictions on the possible matching and pricing strategies), we follow a mechanism design approach. We recast the platform's problem as designing a *matching rule* together with a *pricing rule* so as to maximize profits or welfare. A matching rule assigns each agent from each side to a set of agents from the other side. We only impose that these rules satisfy a minimal feasibility constraint, which we call *reciprocity*. This condition 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.

Main Results. As is standard, the problem of designing profit-maximizing menus can be recasted entirely in terms of designing the matching rules that they induce. This makes the profit-maximization problem analogous to welfare-maximization, except that the agents' valuations are replaced by their virtual counterparts, i.e., they are discounted by informational rents.[1] Hereafter, we will thus refer to both the profit-maximizing and the welfare-maximizing rules as the optimal rules, and will often refrain from distinguishing between the two, when not needed.

Our first result shows that, under two reasonable conditions, optimal rules discriminate only along the willingness-to-pay dimension. In other words, two agents with the same valuation are matched to the same group of agents, irrespective of differences in other unobservable characteristics. These conditions are (i) (weakly) decreasing marginal utility for match quality, and (ii) (weakly) positive affiliation between willingness to pay and attractiveness. The first condition is self-explanatory. The second condition implies that willingness to pay and attractiveness are either independent or positively correlated. In the context of Cable TV, this assumption means that the channels that are

[1] The virtual valuation of an agent coincides with the marginal revenue that the agent brings to the monopolist (see, e.g., Bulow and Roberts (1989)).

willing to pay the most for viewers (e.g., because their advertisers are willing to pay the most) offer, in general, better shows and more pleasant advertisement. Likewise, the viewers who are willing to pay the most for better packages are the ones that the channels view as the most profitable ones (e.g., because these are the viewers preferred by advertisers). In the context of health care provision, where valuations are typically negative on the physicians' side reflecting the latter's opportunity cost of time, this assumption means that those patients with the highest willingness to pay to join larger networks are, on average, those that demand the most intensive care from physicians.[2] Likewise, those physicians with the lowest opportunity cost for serving more patients (equivalently, with the highest valuation to join the network) are those that the patients value the most at the ex-ante stage, i.e., before learning their specific illnesses (these physicians are, typically, the generalists). Clearly, because the assumption only requires affiliation to be weak, the analysis also applies to all markets where, a priori, there are no good reasons to expect willingness to pay and attractiveness to be correlated.

Our second result shows that the optimal matching rules have a threshold structure, according to which each agent is matched to all agents from the other side whose marginal willingness to pay for match quality is above some threshold. To understand both results, note that, under positive affiliation, the expected attractiveness of each agent from each side increases with the agent's willingness to pay. Moreover, 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. These two properties, along with the fact that asymmetric information poses restrictions on the way the match quality can vary with both attractiveness and willingness to pay, imply that the cost-minimizing way to provide a given match quality to an agent with a given willingness to pay is to match him to all agents from the other side whose willingness to pay is high enough, irrespective of all other dimensions that determine attractiveness.

Building on these results, we then show that the optimal rules 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 consists of 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 set. A single network is thus the analog in matching environments to a single price (that is, the absence of quantity/quality discrimination) in the context of a single-product monopolist. The second class is that of *nested multi-homing* matching rules. These rules are implemented by offering a menu of non-exclusive networks (this is known as *multi-homing* in the literature on two-sided markets) and setting prices so that agents with a higher willingness to pay join all networks that lower willingness-to-pay agents join and possibly a few more (therefore the qualification *nested*). Nested multi-homing is the equivalent in matching environments

[2] In general, having more patients in the network with a high willingness to pay may also mean a higher revenue for the physicians (the precise characterization depending on the sharing rule between the physicians and the health care provider). The statement in the main text is for given expected revenue. Holding constant this revenue, having more patients in the network with a higher wilingness to pay then implies a higher opportunity cost of time under the assumption that willingness to pay is positively affiliated with the required intensity of care.

to active quantity/quality price discrimination by a single-product monopolist.

We prove that a single network is optimal *if and only if*, starting from a complete network (i.e., a network that includes all agents from both sides) removing 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 optimal matching rule exhibits nested multi-homing. To understand the intuition for this result, consider the platform's profit-maximization problem (the welfare problem is analogous). Let us analyze first the situation where virtual valuations are positive for all agents from both sides. In such a case, the platform's marginal revenue of expanding the matching set of any agent is positive. Accordingly, profit-maximization clearly requires that each agent be matched to all other agents from the opposite side, i.e., a *single complete network*.

Things are different (and more interesting) when virtual valuations are negative for certain types. In this case, by the same reasoning as above, all agents with positive virtual valuations should be matched to all agents with positive virtual valuations from the other side. However, the platform can increase profits by adding to the matching sets of those agents with high positive virtual valuations some agents from the other side with negative virtual valuations. This cross-subsidization strategy is a general feature of matching markets, in which the platform is 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. If the effect is positive, then the optimal matching rule consists in creating a single complete network (that is, a network that includes all agents from both sides). If the effect is negative, then instead of creating a single, but incomplete network, the platform can do strictly better by separating agents based on their virtual valuations. Those agents with high virtual valuations are assigned matching sets which are supersets of those assigned to agents with lower virtual valuations; that is, the optimal matching rule induces nested multi-homing.

We then proceed by offering a complete characterization of the optimal matching rules. When nested multi-homing is optimal (this is the most interesting case), we show that the thresholds that define the matching sets of each agent are given by an 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. Intuitively, this optimality condition endogenously separate agents from each side into two groups. The first group is that of agents who play the role of *consumers*. These agents generate positive marginal revenues to the platform by "purchasing" sets of agents from the other side of the market. The second group is that of agents who play the role of *inputs*. These agents generate negative marginal revenues to the platform, but serve to "feed" the matching demand of agents-consumers on the other side of the market.

Building on the work of Wilson (1997), we derive a pricing formula that relates (observable) marginal prices to the elasticities of the demand for matching services on both sides of the market.

Intuitively, this formula derives the optimal price schedule by setting marginal prices for each additional match quality so that the marginal revenue gains from expanding the matching sets sold to agents-consumers on one side of the market equals the marginal costs of procuring agents-inputs from the opposite side. Interestingly, these marginal costs are endogenous and depend on the entire network structure of the matching allocations.

Similarly to the standard price discrimination problem analyzed in Mussa and Rosen (1978) and Maskin and Riley (1984), 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 valuations are monotonically increasing, is not the right condition in our matching environment. As pointed out by Bulow and Roberts (1989), this condition implies that the marginal effect on revenue of increasing trade is greater for agents with higher valuations. In a standard setting, because the marginal cost is independent of the agent's type, the monotonicity of the virtual valuations then implies the monotonicity of the trades. In contrast, in a matching environment, by virtue of reciprocity, the marginal cost of increasing the trade of some agent is the revenue loss of adding the agent to the matching sets of other agents from the opposite side. Since attractiveness and valuation are potentially correlated, the marginal cost of increasing trade is also a function of the agent's valuation. For the optimal matching rule to separate types as finely as possible, one must then require that the virtual valuations increase *faster* with the true valuations than their corresponding marginal cross-side effects. In analogy to Myerson (1981), we refer to this condition as *Strong Regularity.* Under this condition, bunching can occur only at "the top" (i.e., for the highest valuation agents) due to capacity constraints, that is, because the stock of agents from the other side of the market has been exhausted.

Public and Private Provision of Matching Services. The above results have implications for the public and private provision of matching services. Because true valuations are always larger than their virtual analogs, our results imply that single networks are more often associated with welfare-maximizing platforms, while multi-homing is more often associated with 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).

Overall, profit-maximization leads to two distortions relative to efficiency. First, there is an *exclusion effect*, whereby too many agents are completely excluded from the market. For example, in the context of health care provision, too many patients are left without any insurance. The second distortion is in the form of an *isolation effect*, whereby 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 those agents with the highest valuations. The reason is that, although the virtual valuations of these agents coincide with the true ones, the cost of cross-subsidizing these agents is always higher under profit maximization than under welfare maximization. This is because such cost is proportional to the true valuation of the marginal agent from the opposite side under welfare maximization and to the virtual valuation of the marginal agent under profit maximization. Because virtual valuations are always lower than true valuations for the marginal agents, the matching sets are strictly smaller under profit maximization than under welfare maximization for all agents, including those at the top of the distribution.

Testable Predictions. The analysis also delivers various testable predictions about the effects of shocks that alter the distributions of individual characteristics affecting the attractiveness of the two sides and/or their willingness to pay to reach the other side.

In the context of the Cable TV application, consider, for example, a positive shock to the viewers' income (not necessarily uniform across viewers) that leaves unchanged the viewers' willingness to pay but raises the profits that the channels expect from reaching the viewers (e.g., via an increase in advertising revenue). Now consider a viewer with a positive virtual valuation. If nested multi-homing was optimal before the shock, then the package offered to this viewer includes channels with a negative virtual valuation. Because the losses that the platform incurs with these channels increase with the viewer's attractiveness, the platform's optimal response to such a shock is to reduce the viewer's matching set by taking off some of the negative virtual-valuation channels. Next, consider a viewer with a negative virtual valuation. Because the original matching set of such viewer includes only positive virtual-valuation channels, the optimal response to such a shock is to expand the viewer's matching set. In other words, the model predicts that the platform's optimal response to such shocks 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 high-valuation viewers). In terms of consumer surplus, these shocks make low-end viewers better off at the expenses of high-end ones.

Group-design Problem. A related problem is that of a principal operating in a single-sided market populated by multiple agents who experience differentiated peer effects from the agents they interact with. In this setting, the principal's problem consists in assigning the agents to non-exclusive groups (rather than networks). 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 optimal matching rules are naturally symmetric. Therefore, all our results naturally extend to single-sided matching problems. 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 teams).

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 optimal matching rule to employ a single network or to exhibit nested multi-homing. Next, it characterizes the optimal matching rules, and discusses the distortions caused by profit maximization relative to welfare maximization. It also derives testable predictions for the effects of shocks that alter the distribution of valuations and/or the distribution of attributes that determine the agents' attractiveness. Section 4 discusses various extensions and concludes. All proofs omitted in the main text are either in the Appendix at the end of the document or in the Online Supplementary Material.

Related Literature

The paper is related to the following literatures.

Price discrimination. The paper contributes to the literature on second-degree price discrimination (e.g., Mussa and Rosen (1978), Maskin and Riley (1983), Wilson (1997)) by considering a setting where the product sold by the monopolist is access to other agents.[3] The study of price discrimination in many-to-many matching markets brings two novelties relative to the standard monopolistic screening problem. First, the platform's feasibility constraint (namely, the reciprocity of the matching rule) has no equivalent in markets for commodities. Second, each agent is both a consumer and an input in the matching production function. The "consumer" role of an agent is summarized in his willingness to pay, while the "input" role is captured by the idiosyncratic characteristics that determine the agent's attractiveness to the other side. This feature of matching markets implies that the cost of procuring an input is endogenous and it depends in a nontrivial way on the entire matching rule.

Two-Sided Markets. Markets where agents purchase 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)).[4] In contrast, we assume that platforms can design arbitrary matching rules and provide conditions for the optimality of a single network relative to more sophisticated network structures (such as those associated 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. In contrast, we study second-degree price discrimination in many-to-many matching environments.

Group Design. As anticipated 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.[5] Under incomplete information, Board (2009) and Rayo (2010) study profit-maximization by a monopolistic platform

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

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

[5] See also Epple and Romano (1998) and Helsley and Strange (2000).

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.[6]

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 and consider many-to-many matching rules.

Double Auctions. A large literature starting with the seminal works of Myerson and Satterhwaite (1983) and Crampton, Gibbons and Klemperer (1987) (see for example McAfee (1991), McAfee (1992a), Fieseler, Kittsteiner and Moldovanu (2003), and Jehiel and Pauzner (2006)) study the (non-)existence of efficient trading mechanisms between a finite number of buyers and sellers. Alternatively, Gresik and Satterthwaite (1983), Satterthwaite and Williams (1989), McAfee (1992b), Rustichini, Satterthwaite and Williams (1994) and Cripps and Swinkels (2006) study the efficiency properties of double auctions in the limit as the number of buyers and sellers increase. In these models, quantities are rival goods across buyers and sellers: As buyers consume more, sellers consume less (or face higher production costs). In contrast, by virtue of reciprocity, in our model matching quality is nonrival across sides. This property is the key difference between matching and trading markets.

2 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 by each agent j from side $l \neq k$. The second component v_k^i is a scalar parameter that describes agent i's willingness to pay for the quality of the set of agents from side l to which i is matched. 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 avoid the uninteresting case where no agent from neither side is willing to interact with agents from the opposite side, we assume that $\bar{v}_k > 0$ for some side $k \in \{A, B\}$.

[6] Board shows that the partition induced under profit-maximization is never coarser than under welfare maximization (note that this result, however, does not imply that it is finer). By considering more general matching rules we show that the profit-maximizing rule indeed matches each agent to a subset of his efficient set. On the other hand, Board allows for more general preferences than the ones considered in this paper. Rayo (2010) considers a one-sided matching problem where the peer effect of a group is the average valuation of its members. In contrast to Board and the present paper, he characterizes the profit-maximizing group design problem when the hazard rate of the type distribution fails to be monotone.

Let $\sigma_k(\mathbf{u}_l^j)$ denote the interaction quality that each agent i from side k obtains from being matched to an agent j from side l with characteristics $\mathbf{u}_l^j$ (we call $\sigma_k(\mathbf{u}_l^j)$ the *attractiveness* of agent j). The function $\sigma_k : \mathbf{U}_l \to \mathbb{R}_+$ thus maps the characteristics of each agent j from side l to the interaction quality enjoyed by each agent i from side k.

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, educational background, consumption habits, of viewer i from side A, whereas $\mathbf{u}_B^j$ contains information about the shows and the advertisement offered by channel j from side B.[7] Accordingly, $\sigma_A(\mathbf{u}_B^j)$ is the quality (or attractiveness) of channel j as perceived by viewers, whereas $\sigma_B(\mathbf{u}_A^i)$ is the contribution of viewer i to the quality of the audience enjoyed by channels. In turn, v_A^i captures viewer i's willingness to pay for a higher quality package of channels, while v_B^j stands for channel j's willingness to pay for an audience of higher quality (e.g., having audiences of higher quality allows channels to increase advertising revenues).

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. For example, in the context of health care provision, let patients belong to side A and doctors to B. In this case, $\mathbf{u}_A^i$ describes patient i's medical condition, and v_A^i captures patient i's willingness to pay for a better physicians' network. In turn, $\mathbf{u}_B^j$ describes doctor j's skills, while v_B^j captures doctor j's disutility from meeting more patients (reflecting the doctors' opportunity cost of time). In this example, $\sigma_A(\mathbf{u}_B^j)$ is the contribution of doctor j to the quality of a given physicians' network, while $\sigma_B(\mathbf{u}_A^i)$ captures the intensity of care demanded by patient i.[8]

The quality of any set of agents from side l (as perceived by each agent from side k) is the sum of the interaction qualities (or attractiveness) of each of its side-l members. Accordingly, for any given (Lebesgue measurable) set of agents $\mathbf{s}$ from side l with type profile $(\theta_l^j)_{j\in\mathbf{s}}$, we denote by

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

the *quality* associated with the set $\mathbf{s}$ ($\lambda(\cdot)$ is the Lebesgue measure). Importantly, all agents from side k agree on the attractiveness of each agent j from side l, and hence on the quality of each set $\mathbf{s}$, which is what makes the model one of vertical differentiation.

Given any complete type profile $\boldsymbol{\theta} \equiv (\theta_k^i)_{k=A,B}^{i\in[0,1]}$, the payoff enjoyed by each agent i from each side k when matched, at a price p, to a set $\mathbf{s}$ of agents from side l is given by

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

[7] Assuming channels possess superior information than the Cable TV company about the impact of their shows and advertisement seems reasonable.

[8] Clearly, from a theoretical viewpoint, expressing attractiveness as a function of a vector $\mathbf{u}_k^i$ of individual attributes is redundant. This decomposition only serves the purpose of permitting us (and possibly the econometrician) to relate the abstract notion of attractiveness to more familiar (and in principle measurable) variables such as health status and idiosyncratic risks in the case of health care provision, or income, education and demographics, in the case of Cable TV provision.

where $g_k(\cdot)$ is a positive, strictly increasing, continuously differentiable, function such that $g_k(0) = 0$. 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 for quality, whereas $\sigma_l(\mathbf{u}_k^i)$ summarizes his attractiveness.

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

Example 1 ***(linear network externalities for quantity)*** *Suppose that the utility of each agent i from side k depends only on the total mass of agents from side l and is linear in this mass. In this case, $\sigma_k(\cdot) \equiv 1$ and $g_k(x) = x$ so that $\pi_k^i(\mathbf{s}, p; \boldsymbol{\theta}) \equiv v_k^i \cdot \lambda(\mathbf{s}) - p$.* \\

These preferences are the ones typically considered in the two-sided market literature (e.g., Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2008), Ambrus and Argenziano (2009), and Weyl (2010)).

Example 2 ***(supermodular match values)*** *Let $\mathbf{u}_k$ be a one-dimensional random variable almost surely equal 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. Then, $\pi_k^i(\mathbf{s}, p; \boldsymbol{\theta}) = v_k^i \cdot \int_{j \in \mathbf{s}} v_l^j d\lambda(j) - p$.* \\

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)).

We assume that the type $\theta_k^i = (\mathbf{u}_k^i, v_k^i)$ of each agent i from each side k is an independent draw from the distribution F_k with support Θ_k. Letting $F_k^{v,\sigma}$ denote the joint distribution of $(v_k, \sigma_l(\mathbf{u}_k))$, we then assume that $F_k^{v,\sigma}$ is absolutely continuous with respect to the Lebesgue measure and denote by F_k^v the marginal distribution of $F_k^{v,\sigma}$ with respect to v_k (with density f_k^v), and by $F_k^\sigma(\cdot|v_k)$ the distribution of the interaction quality $\sigma_l(\mathbf{u}_k)$ conditional on v_k. We will assume that the family of functions $\langle F_k^\sigma(\cdot|v_k)\rangle_{v_k}$ 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 valuations $v_k - [1 - F_k^v(v_k)]/f_k^v(v_k)$ are continuous and nondecreasing.

Matching Mechanisms

Appealing to the Revelation Principle, we focus on (deterministic) direct-revelation mechanisms, which consist 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 $\boldsymbol{\theta} \equiv (\theta_k^i)_{k=A,B}^{i\in[0,1]}$, $\hat{\mathbf{s}}_k^i(\boldsymbol{\theta})$ represents the set of agents from side $l \neq k$ that are matched to agent i from side k, whereas $\hat{p}_k^i(\boldsymbol{\theta})$ denotes the payment made by agent i to the platform (i.e., to the match maker).[9]

[9] Restricting attention to deterministic mechanisms is without loss of optimality under the assumptions in the paper. The proof is based on arguments similar to those in Strausz (2006) and is available upon request.

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 agent j's matching set. Formally:

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

Because there is no aggregate uncertainty and because individual identities are irrelevant for payoffs, without any loss of optimality, 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 $\boldsymbol{\theta}$ 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.

Suppressing superscripts, an anonymous mechanism $M = \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ is described by 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, and $\mathbf{s}_k(\theta_k) \subseteq \Theta_l$ is the set of types from side l associated with all agents from side k reporting type θ_k. Note that $p_k(\cdot)$ maps Θ_k into $\mathbb{R}$, and $\mathbf{s}_k(\cdot)$ maps Θ_k into the Borel 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 of each agent i from side k reporting type θ_k.

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 type $\hat{\theta}_k = (\hat{\mathbf{u}}_k^i, \hat{v}_k^i)$, and by $\Pi_k(\theta_k; M) \equiv \hat{\Pi}_k(\theta_k, \theta_k; M)$ the payoff that type θ_k obtains by reporting truthfully. A mechanism M is individually rational (IR) if $\Pi_k(\theta_k; M) \geq 0$ for all $\theta_k \in \Theta_k$, and it is incentive compatible (IC) if $\Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$ for all $\theta_k, \hat{\theta}_k \in \Theta_k$. A matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ 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.[10]

3 Properties of Optimal Mechanisms

3.1 Efficiency and Profit Maximization

We start by defining what we mean by "efficient" and "profit-maximizing" mechanisms. Because there is no aggregate uncertainty, for any given type profile $\boldsymbol{\theta}$, the welfare generated by the mechanism M is given by

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

[10] 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 by using payments similar to those in Crémer and McLean (1988) – see also Mezzetti (2007).

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

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

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, incentive compatible, and satisfy the **reciprocity condition**

$$\theta_l \in \mathbf{s}_k(\theta_k) \ \Rightarrow \ \theta_k \in \mathbf{s}_l(\theta_l). \tag{3}$$

Note that the reciprocity condition implies that the matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ can be fully described by its side-k correspondence $\mathbf{s}_k(\cdot)$.

It is standard to show that a mechanism M is individually rational and incentive compatible *if and only if* the following conditions jointly hold for each side $k = A, B$:

(i) the quality of the matching set is nondecreasing in the willingness to pay, i.e., $|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k \geq |\mathbf{s}_k(\mathbf{u}'_k, v'_k)|_k$ for any $(\mathbf{u}_k, v_k)$ and $(\mathbf{u}'_k, v'_k)$ such that $v_k \geq v'_k$;

(ii) the expected payoff of any two agents with the same willingness to pay v_k is the same, irrespective of their individual characteristics $\mathbf{u}_k$;

(iii) the equilibrium payoffs $\Pi_k((\mathbf{u}_k, \underline{v}_k); M)$ of the lowest willingness-to-pay agents is non-negative;

(iv) the pricing rule satisfies the envelope formula

$$p_k(\mathbf{u}_k, v_k) = v_k \cdot g_k\left(|\mathbf{s}_k(\mathbf{u}_k, v_k)|_k\right) - \int_{\underline{v}_k}^{v_k} g_k\left(|\mathbf{s}_k(\mathbf{u}_k, x)|_k\right) dx - \Pi_k((\mathbf{u}_k, \underline{v}_k); M). \tag{4}$$

It is immediate to see that in any mechanism that maximizes the platform's profits, the individual rationality constraint of each agent from each side with the lowest willingness to pay must bind, i.e., $\Pi_k((\mathbf{u}_k, \underline{v}_k); M) = 0$, all $\mathbf{u}_k \in \mathbf{U}_k$, $k = A, B$. Using the expression for payments (4), we can rewrite the platform's profit maximization problem in a manner analogous to the welfare maximization problem. We simply replace the true valuations with their virtual analogs (i.e., with the true valuations discounted for informational rents). Formally, for any $v_k \in V_k$, any $k = A, B$, let $\varphi_k^W(v_k) = v_k$ and $\varphi_k^P(v_k) = v_k - [1 - F_k^v(v_k)]/f_k^v(v_k)$. Using the superscript $h = W$ (resp. $h = P$) to denote welfare (resp. profits), the platform's problem is to find a matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ that maximizes

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

among all rules that, together with the price rule given by (4) with $\Pi_k((\mathbf{u}_k, \underline{v}_k); M) = 0$, satisfy constraints (i) and (ii) above and the reciprocity condition (3).

Hereafter, we will say that a matching rule $\{\mathbf{s}_k^h(\cdot)\}_{k=A,B}$ is h-optimal if it solves the above h-problem. For future reference, for both $h = W, P$, we also define the *reservation value* $r_k^h \equiv \inf\{v_k \in$

$V_k : \varphi_k^h(v_k) \geq 0\}$ when $\{v_k \in V_k : \varphi_k^h(v_k) \geq 0\} \neq \varnothing$.

3.2 Single Network vs Nested Multi-Homing

Before solving the platform's welfare and profit-maximization problems, we describe two important classes of matching rules.

Definition 1 ***(single-homing)*** *A matching rule* $\mathbf{s}_k(\cdot)$ *exhibits single-homing if for all* $\theta_k, \theta_k' \in \Theta_k$, $\mathbf{s}_k(\theta_k) \cap \mathbf{s}_k(\theta_k') \neq \oslash$ *implies that* $\mathbf{s}_k(\theta_k) = s_k(\theta_k')$.

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. As mentioned above, these rules are central to the two-sided markets literature. Part of the contribution of the present analysis 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 ***(single network)*** *A matching rule* $\mathbf{s}_k(\cdot)$ *employs a single network if for all* $\theta_k, \theta_k' \in \Theta_k$, $\mathbf{s}_k(\theta_k), \mathbf{s}_k(\theta_k') \neq \oslash$ *implies that* $\mathbf{s}_k(\theta_k) = s_k(\theta_k')$.

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

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

As anticipated in the Introduction, our first result shows that under the following two fairly natural conditions, the optimal matching rules have a simple structure.

Condition 1 ***[DMU]*** *Diminishing Marginal Utility: The function* $g_k(\cdot)$ *is (weakly) concave for all* $k \in \{A, B\}$.

Condition 2 ***[PA]*** *Positive Affiliation: The distribution* F_k *is such that* $(\sigma_l(\tilde{\mathbf{u}}_k), \tilde{v}_k)$ *are (weakly) positively affiliated for all* $k \in \{A, B\}$.[11]

[11] See Milgrom and Weber (1982) for a formal treatment of the concept of affiliation.

We then have the following result.

Proposition 1 ***(optimal rules)*** *Assume conditions DMU and PA hold. Then both the profit-maximizing ($h = P$) and the welfare-maximizing ($h = W$) rules discriminate only along the willingness-to-pay dimension (that is, $\mathbf{s}_k^h(v_k, \mathbf{u}_k) = \mathbf{s}_k^h(v_k, \mathbf{u}_k')$ for any $v_k, \mathbf{u}_k, \mathbf{u}_k'$, $h = W, P$). Suppressing the dependence on $\mathbf{u}_k$, $k = A, B$, the h-optimal matching rule $\mathbf{s}_k^h(\cdot)$ has the following threshold structure:*

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

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

Proof. See Appendix.

To understand the result, consider an agent with type $\theta_k = (\mathbf{u}_k, v_k)$ with $\varphi_k^h(v_k) \geq 0$. Ignoring for a moment the monotonicity constraints, it is easy to see that it is always optimal to assign to this type a matching set $\mathbf{s}_k(\mathbf{u}_k, v_k) \supset \{(\mathbf{u}_l, v_l) : \varphi_l^h(v_l) \geq 0\}$ that includes all types $\theta_l = (\mathbf{u}_l, v_l)$ whose φ_l^h-valuation is non-negative. This is because, (i) irrespective of their $\mathbf{u}_l$ characteristics, these types contribute positively to type θ_k's payoff (recall that $\sigma_k(\mathbf{u}_l) \geq 0$ for all $\mathbf{u}_l$) and (ii) these types have a non-negative φ_l^h-valuation, and therefore adding type θ_k to these types' matching sets never reduces the platform's payoff $\Omega^h(M)$, as implied 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_k is higher than the quality of the set of types on side l whose φ_l^h-valuation is non-negative, i.e., such that

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

Because of reciprocity, adding an agent whose φ_l^h-valuation is negative to type θ_k's matching set now comes at a cost. In the case of welfare-maximization, agents with negative valuations require a payment to accept larger matching sets. In the case of profit-maximization, this cost stems from the infra-marginal losses on revenue captured by negative virtual valuations.

The positive affiliation between $(\sigma_k(\mathbf{u}_l), v_l)$ along with the weak concavity of $g_k(\cdot)$ (which implies that using the same agent as an input is less costly than using different agents) and the limitations imposed by the asymmetry of information, imply that the least costly way to provide type θ_k with a matching set of quality q_k is by matching him to 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

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

Because the quality of the matching set in monotone in v_k, as required by incentive compatibility, the threshold function $t_k^h(\cdot)$ is nonincreasing. The proof in the Appendix uses results from the theory of monotone concave order of random variables to verify the heuristics above.

The following corollary is then a direct implication of Proposition 1.

Corollary 1 ***(optimal network structures)*** *Assume conditions DMU and PA hold. Then any h-optimal matching rule either employs a single network or induces nested multi-homing, $h = W, P$. All other single-homing structures are dominated.*

Given the result in Proposition 1, we can restrict attention to mechanisms whose matching rule takes the form given in (6). Letting $\hat{g}_k : V_l \to \mathbb{R}_+$ denote the function defined by

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

the platform's objective then consists in choosing a pair of exclusion thresholds $(\omega_k^h)_{k\in\{A,B\}}$ and a pair of nonincreasing threshold functions $(t_k^h(\cdot))_{k\in\{A,B\}}$ so as to maximize the objective

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

subject to the reciprocity constraint that

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

for all $v_k \in [\omega_k^h, \bar{v}_k]$, $k = A, B$.

The next two results provide necessary and sufficient condition for the h-optimal mechanism to employ either a single network or to exhibit nested multi-homing and characterize properties of optimal multi-homing rules. These results are obtained by assuming that the following condition holds, which strengthens the standard monotonicity of virtual valuations, as required in our two-sided matching environment.

Condition 3 ***[SR]*** *Strong Regularity:* *The functions $\psi_k^h : V_k \to \mathbb{R}$ given by*

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

are strictly increasing, $k = A, B$, $h = W, P$.

Take the case of profit-maximization, $h = P$. The numerator in $\psi_k^h(v_k)$ accounts for the effect on the platform's profits of an agent from side k with valuation v_k as a *consumer* (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 on the platform's profits of this agent as an *input* (as $-\hat{g}_l'(v_k)$ is proportional to the marginal utility brought by this agent to every agent from side l who is already matched to any other agent with valuation above v_k). The affiliation and concavity assumptions imply that the denominator of $\psi_k^h(v_k)$ is nondecreasing in v_k. Therefore, the strong regularity condition above requires that the value of an agent as a consumer (as captured by his virtual valuation) increases faster than his contribution as an input. In the linear model, $\psi_k^h(v_k) = \varphi_k^h(v_k)/\mathbb{E}\left[\sigma_l(\tilde{\mathbf{u}}_k)|\tilde{v}_k = v_k\right]$, which is 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 rule out nonmonotonicities in the schedule of matching qualities. In this sense, it is the analog of Myerson standard regularity condition in two-sided matching problems. It also plays a role in Proposition 2 below, but only in the special case where φ_l^h-valuations are always positive on one-side and both positive and negative on the other side.

Now let $\triangle_k^h : V_k \times V_l \to \mathbb{R}$ denote the function defined by

$$\triangle_k^h(v_k, v_l) \equiv -\hat{g}_k'(v_l) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l'(v_k) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l) \tag{9}$$

$k, l \in \{A, B\}$, $l \neq k$. Note that $\triangle_A^h(v_A, v_B) = \triangle_B^h(v_B, v_A)$ represents the marginal effect on the platform's payoff of decreasing the threshold $t_A^h(v_A)$ below v_B, while, by reciprocity, also reducing the threshold $t_B^h(v_B)$ below v_A. Equivalently, $-\triangle^h(v_A, v_B)$ represents the marginal effect of deleting the link between v_A and v_B starting from a network structure where each agent from side A with valuation v_A is matched to all agents from side B with valuation above v_B and each agent from side B with valuation v_B is matched to all agents from side A with valuation above v_A. For future reference, also note that $sign(\triangle_k^h(v_k, v_l)) = sign(\psi_k^h(v_k) + \psi_l^h(v_l))$. We then have the following result.

Proposition 2 ***(single- vs multi-homing)*** *Assume Conditions DMU, PA and SR hold. The h-optimal matching rule employs a single (complete) network if* $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$ *(that is, if starting from a complete network deleting the link between* $\underline{v}_A$ *and* $\underline{v}_B$ *reduces the platform's payoff) and exhibits nested multi-homing otherwise,* $h = W, P$.

Proof. Consider first the case where $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, implying that $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$. Because valuations (virtual valuations) are all nonnegative, welfare (profits) is (are) maximized by matching each agent from each side to all agents from 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 $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. Starting from any single network the platform can then increase its payoff by switching to multi-homing. To see this, let $\hat{\omega}_k^h$ denote the threshold type corresponding to the single network so that agents from side k are excluded if and only if $v_k < \hat{\omega}_k^h$. First, suppose that, for some $k \in \{A, B\}$, $\hat{\omega}_k^h > r_k^h$, where recall that $r_k^h \equiv \inf\{v_k \in V_k : \varphi_k^h(v_k) \geq 0\}$. The platform could then increase its payoff by switching to a nested multi-homing rule that assigns to each agent from 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 $\hat{v}_l^{\#} \equiv \max\{r_l^h, \hat{\omega}_l^h\}$.

Next, suppose that $\hat{\omega}_k^h < r_k^h$ for both $k = A, B$. Starting from this single network, the platform could then increase her payoff by switching to a nested multi-homing rule $\mathbf{s}_k^{\diamondsuit}(\cdot)$ such that, for some $k \in \{A, B\}$[12]

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

The new matching rule improves upon the original one because it eliminates all matches between agents whose valuations (virtual valuations) are both negative.

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

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

By setting the new exclusion threshold $\hat{\omega}_k^{\#}$ sufficiently close to r_k^h the platform increases its payoff. In fact, the marginal benefit of increasing the quality of the matching sets of those agents from side l whose φ_l^h-valuation is positive more than offsets the marginal cost of getting on board a few more agents from side k whose φ_k^h-valuation is negative, but sufficiently small.[13] Note that for this network expansion to be profitable, it is essential that the new agents from side k that are brought "on board" be matched only to those agents from side l whose φ_l^h-valuation is positive, which requires employing a multi-homing matching rule.

In the Appendix, we complete the proof by analyzing the remaining case where $\varphi_l^h(\underline{v}_l) < 0$ while $\varphi_k^h(\underline{v}_k) \geq 0$. This is the only case in which strong regularity plays a role. Q.E.D.

The following corollary then follows directly from Proposition 2 by noting that $\triangle_k^W(\underline{v}_k, \underline{v}_l) > \triangle_k^P(\underline{v}_k, \underline{v}_l)$.

Corollary 2 *Multi-homing matching rules are more often employed by profit-maximizing platforms than by welfare-maximizing platforms.*

The finding above are illustrated by the next two examples.

Example 3 Consider the case of linear network externalities, as described in Example 1 above, and assume that valuations v_k are uniformly distributed over $[\underline{v}_k, \overline{v}_k]$. Then, the welfare-maximizing rule employs a single network if $\triangle_A^W(\underline{v}_A, \underline{v}_B) = \underline{v}_A + \underline{v}_B \geq 0$, and exhibits nested multi-homing otherwise. In turn, the profit-maximizing rule employs a single network if $\triangle_A^P(\underline{v}_A, \underline{v}_B) = 2(\underline{v}_A + \underline{v}_B) - (\overline{v}_A + \overline{v}_B) = \triangle_A^W(\underline{v}_A, \underline{v}_B) - [(\overline{v}_A - \underline{v}_A) + (\overline{v}_B - \underline{v}_B)] \geq 0$, and exhibits nested multi-homing otherwise. $\backslash\backslash$

[12] The behavior of the rule on side l is then pinned down by reciprocity.

[13] To see this, note that, starting from $\hat{\omega}_k^{\#} = r_k^h$, the marginal benefit of decreasing the threshold $\hat{\omega}_k^{\#}$ is $-\hat{g}_l'(r_k^h) \int_{r_l^h}^{\overline{v}_l} \varphi_l^h(v_l) dF_l^v(v_l) > 0$, whereas the marginal cost is given by $-\hat{g}_k(r_l^h) \cdot \varphi_k^h(r_k^h) f_k^v(r_k^h) = 0$ since $\varphi_k^h(r_k^h) = 0$.

Example 4 Consider the case of supermodular matching values, as described in Example 2 above, and assume that valuations v_k are uniformly distributed over $[\underline{v}_k, \overline{v}_k]$ with $\underline{v}_k > 0$. Then, the welfare-maximizing rule employs a single network, since $\triangle_A^W(\underline{v}_A, \underline{v}_B) = 2 \cdot \underline{v}_A \cdot \underline{v}_B > 0$. In turn, the profit-maximizing rule employs a single network if $\triangle_A^P(\underline{v}_A, \underline{v}_B) = \sum_{k=A,B,l\neq k} \underline{v}_l \cdot (2\underline{v}_k - \overline{v}_k) = \triangle_A^W(\underline{v}_A, \underline{v}_B) \cdot \left[2 - \frac{1}{2} \cdot (\overline{v}_A/\underline{v}_A + \overline{v}_B/\underline{v}_B)\right] \geq 0$, and exhibits nested multi-homing *otherwise.* \\

3.3 Optimal Multi-Homing Rules

We now further investigate the properties of optimal matching rules when nested multi-homing is optimal, that is, when $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. The next definition extends to our two-sided matching setting the notion of separating schedules, as it appears in Maskin and Riley (1984).

Definition 4 ***(maximally separating rules)*** *The h-optimal matching rule is maximally separating if $t_k^h(\cdot)$ is strictly decreasing over $[\omega_k^h, t_l^h(\omega_l^h)]$ (which we call the separating range). It exhibits exclusion at the bottom on side k if $\omega_k^h > \underline{v}_k$ and bunching at the top on side k if $t_l^h(\omega_l^h) < \bar{v}_k$.*

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

Proposition 3 ***(optimal multi-homing rules)*** *Assume Conditions DMU, PA and SR hold and suppose that multi-homing is optimal (i.e., $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$). The following properties are true both for $h = W$ and for $h = P$:*

(i) The h-optimal matching rule is maximally separating.

(ii) 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.

(iii) 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 (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).

(iv) For all valuations v_k in the separating range $[\omega_k^h, t_l^h(\omega_l^h)]$, the h-optimal threshold function $t_k^h(\cdot)$ satisfies the Euler equation

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

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

Proof. See Appendix.

Assume $\underline{v}_k < 0$, $k = A, B$. An important feature of the maximally separating h-optimal rule described above is that $t_k^h(v_k) \leq r_l^h$ if and only if $v_k \geq r_k^h$. Consider the case of profit-maximization (The arguments for the case of welfare maximization are analogous). Agents with positive virtual valuations from side k are matched to all agents with positive virtual valuations on side l, plus a measure of agents with negative virtual valuations from side l (cross-subsidization). The optimal level of cross-subsidization for an agent with virtual valuation $\varphi_k^P(v_k) > 0$ is then determined by the

Euler equation (10). As explained above, this equation equalizes the marginal benefit $-\hat{g}'_k(t^P_k(v_k)) \cdot \varphi^P_k(v_k) \cdot f^v_k(v_k)$ of enlarging the matching set of an agent from side k who is already matched to all agents from side l with valuation above $t^P_k(v_k)$, with the marginal cost $-\hat{g}'_l(v_k) \cdot \varphi^P_l(t^P_k(v_k)) \cdot f^v_l(t^P_k(v_k))$ of enlarging the matching set of any agent from side l with valuation $v_l = t^P_k(v_k)$ who is already matched to all agents from side k with valuation above $v_k = t^P_l(v_l)$, as required by reciprocity (recall that $\varphi^P_l(t^P_k(v_k)) < 0$).

Intuitively, agents from each side of the market are endogenously partitioned in two groups. Those with positive virtual valuations (equivalently, with valuations $v_k \geq r^P_k$) play the role of *consumers*, "purchasing" sets of agents from the other side of the market (these agents contribute positively to the platform's profits). In turn, those agents with a negative virtual valuation (equivalently, with valuation $v_k < r^P_k$) play the role of *inputs*, generating utility to the agents from the other side of the market they are matched to (these agents contribute negatively to the platform's profits).

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. In other words, bunching can only occur at the top due to binding capacity constraints, that is, when the "stock" of agents from side l has been exhausted. This is illustrated in the next example.

Example 5 Consider the environment with linear network externalities as in Examples 1 and 3. Assume that each agent from side A has a valuation drawn from a uniform distribution on $[1, 3/2]$, while each agent from side B has a valuation drawn from a uniform distribution on $[-1, 0]$. Since $\triangle^W_A(3/2, 0) = 0$ and $\triangle^P_A(3/2, 0) = -3/2$, the welfare-maximizing provision of matching services employs a single (complete) network, while the profit-maximizing provision exhibits nested multi-homing with the threshold function $t^P_A(v_A) = 3/4 - v_A$ defined over $[1, 3/2]$. Under profit-maximization, there is bunching at the top on side B and exclusion at the bottom on side B, as illustrated in Figure 1.\\

The example above offers a stylized description of the market for health care services. The platform here is a health insurance company providing patients from side A access to physicians from side B. The physicians' negative valuations reflect their opportunity cost of treating additional patients. The welfare-maximizing (say, public) provision of health insurance adopts a single (complete) network where all patients have access to all doctors. In contrast, the profit-maximizing (say, private) provision of health insurance exhibits a nested multi-homing matching rule according to which patients and doctors are sorted into different nested network plans. Those physicians with a low opportunity cost are included in all plans, while the more expensive ones are included only in the plans offered to those patients with the highest willingness to pay for larger physician networks. The most expensive doctors are excluded from all plans.

Relative to what is efficient, the profit-maximizing matching rule thus (i) completely excludes more agents from the market, and (ii) provides to each agent who is not excluded a matching set that is a strict subset of his efficient set. As we show below, these two distortions are general properties of profit-maximizing matching mechanisms.

3.4 Distortions Relative to Efficiency: Exclusion and Isolation Effects

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

Proposition 4 ***(distortions)*** *Assume Conditions DMU, PA, and SR hold. Relative to the welfare-maximizing matching rule, the profit-maximizing matching rule:*

1. *completely excludes a larger group of agents (**exclusion effect**) — i.e., $\omega_k^P \geq \omega_k^W$ for $k = A, B$;*
2. *matches each agent from 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 $v_k \geq \omega_k^P$, $k = A, B$.*

Proof. See Online Supplementary Material.

The intuition for both effects can be seen from the Euler condition (10): Under profit-maximization, the platform only internalizes the cross-effects on marginal revenues (which are proportional to virtual valuations $\varphi_k^P(v_k)$), rather than the cross-effects on welfare (which are proportional to the true valuations v_k). Since virtual valuations are always smaller than the true valuations, the platform fails to internalize part of the marginal welfare gains from new matches. As explained in the Introduction, 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 context of the supermodular matching value case of Example 2.

Example 6 Consider the environment with supermodular matching values, as in Examples 2 and 4. Each agent from each side has a valuation drawn from a uniform distribution on $[\underline{v}, \bar{v}]$, where $\underline{v} > 0$ and $2\underline{v} < \bar{v}$. Since $\triangle_k^W(\underline{v}, \underline{v}) = 2\underline{v}^2$ and $\triangle_k^P(\underline{v}, \underline{v}) = 2\underline{v}(2\underline{v} - \bar{v}) < 0$, the welfare-maximizing provision of matching services entails the creation of a single complete network, while the profit-maximizing provision entails the adoption of a nested multi-homing structure with threshold function given by $t_k^P(v_k) = \frac{v_k \cdot \bar{v}}{4 \cdot v_k - \bar{v}}$ defined over $(\omega_k, \bar{v}) = (\frac{\bar{v}}{3}, \bar{v})$. Under profit-maximization, there is exclusion at the bottom on both sides and each agent who is not excluded is matched to a strict subset of his efficiency set. Figure 2 describes the profit-maximizing solution when $[\underline{v}, \bar{v}] = [1, 6]$. \\

This example offers a stylized description of a market where the platform is an employment agency that matches free-lancers on side A to firms on side B. The output of the match between each free-lancer and each firm is increasing in the quality/productivity of the two parties and is evenly split between the firm and the free-lancer. In this environment, a welfare-maximizing (say, public) employment agency creates a single (complete) network that gives each free-lancer access to all firms. In contrast, a profit-maximizing (say, private) agency offers a menu of access plans to each side which results in fewer matches and more exclusion on both sides.

3.5 Implications for Prices

The analysis so far restricted attention to direct-revelation mechanisms, where the matching set and the payment of each agent depend on the reported type. While these mechanisms help us describe the allocations that are induced both under welfare and under profit maximization, in reality these allocations are typically obtained by letting agents choose from a menu. For example, in the case of health care provision, patients are typically offered menus of health plans where the price for each plan depends on the number of doctors included in the plan. Accordingly, we will now show how the characterization from Proposition 3 translates into properties of price schedules that indirectly implement the optimal mechanism M^h.

In order to express the optimal pricing formulas in terms of observable variables, in this subsection we will restrict attention to the case where agents care only about the total number of agents from the other side that they are matched to (that is, $\sigma_k(\cdot) \equiv 1$ for $k \in \{A, B\}$). For any $q_k \in [0, 1]$, then let $\rho_k^h(q_k)$ denote the total price that agents from side k have to pay for a matching set of size q_k under the h-optimal mechanism M^h. Accordingly, the tariff $\rho_k^h(\cdot)$ has to satisfy $\rho_k^h(q_k) = p_k^h(\mathbf{u}_k, v_k)$ for all $(\mathbf{u}_k, v_k)$ such that $|\mathbf{s}_k^h(v_k)|_k = q_k$.

At any point of differentiability of the tariff $\rho_k^h(\cdot)$, we will then denote by $\frac{d\rho_k^h}{dq_k}(q_k)$ the marginal price for the q_k unit. Now, given the tariff $\rho_k^h(\cdot)$, let

$$x_k^h(v_k) \in \arg\max_{q_k \in [0,1]} \; v_k \cdot g_k(q_k) - \rho_k^h(q_k)$$

denote the *individual demand* of each agent from side k with marginal willingness to pay v_k. At any point $x_k^h(v_k)$ of differentiability of the tariff $\rho_k^h(\cdot)$, the following first-order condition must hold:

$$v_k \cdot g_k'(x_k^h(v_k)) - \frac{d\rho_k^h(x_k^h(v_k))}{dq_k} = 0. \tag{11}$$

Given the monotonicity of the individual demand in v_k, the side-k *aggregate demand* for the q_k unit at the *marginal price* $\frac{d\rho_k^h}{dq_k}(q_k)$ is given by:[14]

$$D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right) \equiv 1 - F_k^v\left(\frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}\right).$$

Given the expression for aggregate demand above, we can compute the elasticity of the aggregate demand for the q_k unit with respect to its marginal price:

$$\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right) \equiv -\frac{\partial D_k}{\partial\left(\frac{d\rho_k^h}{dq_k}(q_k)\right)} \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{D_k} = \frac{f_k^v\left(\frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}\right)}{1 - F_k^v\left(\frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}\right)} \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}. \tag{12}$$

[14] This is the measure of agents whose matching set is of size greater or equal than q_k.

As usual, this elasticity measures the responsiveness of the aggregate demand for the q_k unit to variations of the marginal price of the q_k unit. The elasticity is positive (in the sense that an increase in the marginal price reduces demand) for all agents with positive valuations $v_k > 0$ (observe that $\frac{d\rho_k^h(q_k)}{dq_k}\frac{1}{g_k'(q_k)} = \left(x_k^h\right)^{-1}(q_k)$) and negative for all agents with negative valuations $v_k < 0$. For example, in the health care application where valuations are negative on the doctors' side, a negative elasticity reflects the idea that if marginal payments to doctors increase, then the doctors respond by accepting more patients.

The next proposition recasts the first-order Euler condition (10) in terms of demand elasticities and marginal prices. The expression below extends to matching markets the familiar Lerner-Wilson formula for second-degree price discrimination in commodity markets (see Wilson (1997)).

Proposition 5 ***(Lerner-Wilson formula for matching markets)*** *In addition to Conditions DMU and SR, suppose that network effects depend only on quantities ($\sigma_k(\cdot) \equiv 1$ for $k \in \{A, B\}$), and that multi-homing is optimal (i.e., $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$). Then the optimal price schedules $\rho_k^h(\cdot), \rho_l^h(\cdot)$ are differentiable and the marginal prices satisfy the Lerner-Wilson formula*

$$\frac{d\rho_k^h}{dq_k}(q_k) - \mathbf{1}^h \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)} = \tag{13}$$

$$-\left[\frac{d\rho_l^h}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right) - \mathbf{1}^h \cdot \frac{\frac{d\rho_l^h}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right)}{\varepsilon_l\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right), \frac{d\rho_l^h}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right)\right)}\right],$$

where the indicator function $\mathbf{1}^h$ equals one in case of profit-maximization ($h = P$) and zero in case of welfare maximization ($h = W$), and where $D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)$ is the aggregate demand for the q_k unit on side k at marginal price $d\rho_k^h(q_k)/dq_k$.

Proof. See Appendix.

First, consider welfare maximization. The result then says that the marginal price for the q_k agent on side k must equal the opposite of the marginal price for the $q_l = D_k\left(q_k, \frac{d\rho_k^W}{dq_k}(q_k)\right)$ agent on side l, where $D_k\left(q_k, \frac{d\rho_k^W}{dq_k}(q_k)\right)$ is the aggregate demand for the q_k agent on side k. In the health care application, this means that the marginal price that the platform charges to any patient who wants to add an additional doctor to a plan of size q_k is equal to the marginal payment that the platform makes to any physician whose plan includes $D_k\left(q_k, \frac{d\rho_k^W}{dq_k}(q_k)\right)$ patients.

Next consider profit-maximization. Take a matching set of size q_k sold to an agent from side k that plays the role of a consumer (i.e., for whom marginal revenue is positive, that is, $v_k > r_k^P$). The formula in (13) is the analog of the familiar Lerner formula

$$\frac{p - MC}{p} = \frac{1}{\varepsilon_d}$$

for optimal monopoly pricing. It equalizes the marginal revenue of expanding the matching set on side k (the left hand side) to the marginal cost of "procuring" extra agents from side l (the right hand side). As in standard monopoly pricing, at the optimum, the marginal price $d\rho_k^P(q_k)/dq_k$ for any quantity q_k sold to agents who play the role of consumers is set so that the aggregate demand is locally elastic at $\left(q_k, \frac{d\rho_k^h(q_k)}{dq_k}\right)$, i.e., $\varepsilon_k\left(q_k, \frac{d\rho_k^h(q_k)}{dq_k}\right) > 1$, as can be seen from (11) and (12) after plugging $v_k > r_k^P$.

The interesting part of the formula is the expression for the endogenous marginal cost of expanding a matching set of size q_k. This is the cost of procuring an additional agent from the opposite side. As shown in Proposition 3, this cost is minimized by picking an agent from side l whose matching set contains exactly $q_l = D_k\left(q_k, \frac{d\rho_k^P(q_k)}{dq_k}\right)$ agents. To see this, let $v_k = \left(x_k^P\right)^{-1}(q_k)$ denote the valuation of the marginal agent from side k who is induced to purchase a matching set of size q_k. Recall that, at the optimum, this agent is matched to all agents from side l with valuation above $v_l = t_k^P\left(\left(x_k^P\right)^{-1}(q_k)\right)$. The marginal cost of expanding the matching set of this agent under the optimal matching rule is thus equal to the cost of adding this agent to the matching set of an agent from side l with valuation $v_l = t_k^P\left(\left(x_k^P\right)^{-1}(q_k)\right)$ whose matching set contains all agents from side k with valuation above $v_k = \left(x_k^P\right)^{-1}(q_k)$. The measure of this set is $1 - F_k^v\left(\left(x_k^P\right)^{-1}(q_k)\right)$, which is exactly equal to the aggregate demand $D_k\left(q_k, \frac{d\rho_k^P(q_k)}{dq_k}\right)$ on side k for the q_k unit, at the marginal price $\frac{d\rho_k^P(q_k)}{dq_k}$. This means that the marginal cost of expanding a matching set of size q_k on side k is equal to the marginal price $\frac{d\rho_l^h}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right)$ that the platform pays on side l to those agents who select a matching set of size $D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)$, augmented by the (positive) term

$$\frac{\frac{d\rho_l^P}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right)}{\varepsilon_l\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right), \frac{d\rho_l^P}{dq_l}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)\right)\right)}$$

that reflects the monopsonistic role that the platform plays on side l under the impossibility of perfect (first degree) price discrimination.

Turning to the shape of the price schedules (and the existence of quantity discounts/premiums), note that, in the case of linear network externalities for quantities ($g_k(x) \equiv x$), the first-order condition (11) implies that marginal prices increase with quantities, meaning that the price schedule $\rho_k^h(\cdot)$ is a convex function of q_k. In other words, the platform charges a quantity premium to those agents who play the role of consumers for expanding the size of their matching sets. In the case of strictly diminishing marginal utility for match quality ($g_k(\cdot)$ strictly concave), the emergence of quantity discounts depends nontrivially on the interplay between the elasticities of demands on both sides. In general there does not appear to be a good reason to expect price schedules to exhibit quantity discounts for the full range of quantities sold.

Lastly, note that the Lerner-Wilson formula (13) only depends on the shape of the *aggregate* demand for matching services on the two sides of the market. It can be used for a structural estimate of the demand for matching services. It also provides testable implications about the effects of changes in elasticity on total and marginal prices. As an illustration, suppose that network effects are linear (i.e., $g_A(x) = g_B(x) = x$) and that the demand from side k becomes less elastic, meaning that the new distribution $\tilde{F}_k^v$ dominates the old distribution F_k^v in the hazard-rate order, while, for any v_k, $\tilde{F}_k^\sigma(\cdot|v_k) = F_k^\sigma(\cdot|v_k)$.[15] The platform's profit-maximizing response to such a shock is to reduce the matching sets of each agent from side k (with an unambiguous negative effect on payoffs). In turn, this is accomplished by increasing the total price for each size q_k of the matching set. One can then use the result in the preceding proposition to verify that marginal prices also go up (see Proposition A1 and Corollary A1 in the Supplementary Material for details). As for the effects of such shocks on the opposite side, these effects are in general ambiguous. On the one hand, by virtue of reciprocity, the matching sets of agents from side l also shrink. On the other hand, because of the positive affiliation assumption, agents from side k are on average more attractive. The reason is that, although the conditional distributions are held constant, more agents on side k exhibit 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). Clearly, agents from side l are worse off in the special case where network effects depend only on quantities (i.e., $\sigma_l(\cdot) \equiv 1$).

The next subsection provides further empirically testable predictions by studying the (less explored) cross-side effects of changes in attractiveness.

3.6 The Detrimental Effects of Becoming More Attractive

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

The next definition formalizes the notion of a change in attractiveness.

Definition 5 ***(higher attractiveness)*** *Side k is more attractive under the distribution F_k than under the distribution $\hat{F}_k$ if, for all v_k, $F_k^\sigma(\cdot|v_k)$ dominates $\hat{F}_k^\sigma(\cdot|v_k)$ in the sense of first-order stochastic dominance, and $F_k^v = \hat{F}_k^v$.*

The next proposition describes how the profit-maximizing matching rule changes as side k becomes more attractive (a similar analysis holds for the case of welfare-maximization). Perhaps surprisingly, agents from side k can be hurt by a positive shock to their attractiveness.

Proposition 6 ***(increase in attractiveness)*** *In addition to Conditions PA and SR, suppose that network effects are linear (i.e., $g_A(x) = g_B(x) = x$), and that multi-homing is optimal (i.e., $\triangle_k^P(\underline{v}_k, \underline{v}_l) < 0$). If the attractiveness of side k increases, then the platform switches from a matching rule $\mathbf{s}_k^P(\cdot)$ to a matching rule $\hat{\mathbf{s}}_k^P(\cdot)$ such that:*

[15] The distribution $\tilde{F}_k^v$ dominates F_k^v in the hazard-rate order if and only if $\frac{f_k^v(v_k)}{1-F_k^v(v_k)} \geq \frac{\tilde{f}_k^v(v_k)}{1-\tilde{F}_k^v(v_k)}$ for all $v_k \in V_k$.

1. *the matching sets on side k increase for low-valuation agents and decrease for high-valuation ones — $\hat{\mathbf{s}}_k^P(v_k) \supseteq \mathbf{s}_k^P(v_k)$ if and only if $v_k \leq r_k^P$;*

2. *low-valuations agents from side k are better off, whereas the opposite is true for high-valuation ones — 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 $v_k \leq \hat{\nu}_k$.*

Proof. See Online Supplementary Material.

Intuitively, an increase in the attractiveness of side k alters the costs of cross-subsidization between sides. Agents with valuation $v_k \geq r_k^P$ are valued by the platform mainly by their role as consumers. As these agents become more attractive, the costs of cross-subsidizing their "consumption" using agents from side l with negative virtual valuations increases, whereas the revenue gain on side k is unaltered. As a consequence, the matching sets of these agents shrink. The opposite is true for those agents with valuation $v_k \leq r_k^P$ — these agents are valued by the platform mainly by their role as inputs. As they become better inputs, their matching sets expand.

In terms of payoffs, for all $v_k \leq r_k^P$

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

meaning that all agents from side k with $v_k \leq r_k^P$ are necessarily better off. 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 from side k is higher under the new rule than under the original one if and only if $v_k \leq \hat{v}_k$.

Next, consider the effect of the increase in attractiveness on side k on the payoffs of agents from side l. On the one hand, the fact that side k becomes more attractive implies that the payoff that each agent from side l derives from interacting with side k increases. On the other hand, by virtue of reciprocity, the matching sets for all agents with valuation $v_l < r_l^P$ shrink, which contributes negatively to profits. The net effect on the payoffs of agents from side l can thus be ambiguous and nonmonotone in v_l. Still, using equation (4), one can show 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 6 offer testable predictions about the pricing strategies of many two-sided platforms. In the case of Cable TV providers, it implies that shocks to households' income or wealth (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.

Consider the profit-maximizing price schedule $\rho_k^P(\cdot)$ defined in Subsection 3.4. The next corollary translates Proposition 6 in terms of the tariff $\rho_k^P(\cdot)$.

Corollary 3 ***(effect of attractiveness on prices)*** *In addition to Conditions PA and SR, suppose that network effects are linear (i.e., $g_A(x) = g_B(x) = x$), and that multi-homing is optimal (i.e., $\triangle_k^P(\underline{v}_k, \underline{v}_l) < 0$). If the attractiveness of side k increases, then the platform moves from a price* *schedule* $\rho_k^P(\cdot)$ *to a price schedule* $\hat{\rho}_k^P(\cdot)$ *such that* $\hat{\rho}_k^P(q_k) \leq \rho_k^P(q_k)$ *if and only if* $q_k \leq \hat{q}_k$*, where* $\hat{q}_k > \left|\mathbf{s}_k^P(r_k^P)\right|_k = \left|\hat{\mathbf{s}}_k^P(r_k^P)\right|_k$.

Proof. See Online Supplementary Material.

In terms of price schedules, an increase in the attractiveness of side k increases the prices that agents on side k have to pay for high quality matching sets, and decreases the price for low quality matching sets.

4 Extensions and Conclusions

The analysis developed above is worth extending in a number of interesting directions. Below we first discuss how the results can accommodate a few simple enrichments and then conclude by discussing various lines for future research.

Insulating Tariffs and Robust Implementation. In the direct revelation version of the matching game described above, each agent from each side is asked to submit a report θ_k which leads to a payment $p_k^h(\theta_k)$, as defined in (4), and which grants access to all agents from the opposite side who reported valuations above $t_k^h(v_k)$. This game admits one Bayes-Nash equilibrium implementing the matching rule $\mathbf{s}_k^h(\cdot)$, together with other equilibria implementing different rules.[16]

As pointed out by Weyl (2010) (see also White and Weyl (2010)) in the context of a monopolistic platform designing a single network, equilibrium uniqueness can however be guaranteed when network effects depend only on quantities (i.e., when $\sigma_k(\cdot) \equiv 1$ for $k \in \{A, B\}$). In the context of our model, it suffices to replace the payment rule $(p_k^h(\cdot))_{k=A,B}$ given by (4) with the payment rule

$$\varrho_k^h(v_k, (v_l^j)^{j\in[0,1]}) = v_k \cdot g_k\left(\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k\right) - \int_{\underline{v}_k}^{v_k} g_k\left(\left|\{j \in [0,1] : v_l^j \geq t_k(\tilde{v}_k)\}\right|_k\right) d\tilde{v}_k, \tag{14}$$

where $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k \equiv \int_{\{j: v_l^j \geq t_k(v_k)\}} d\lambda(j)$ denotes the measure of agents reporting a valuation above $t_k(v_k)$. Given the above payment rule, it is *weakly dominant* for each agent to report truthfully (This follows from the fact that, *given any* profile of reports $(v_l^j)^{i\in[0,1]}$ by all agents from the other side, the quality of the matching set for each agent from side $k = A, B$ is increasing in his report, along with the fact that the payment rule $\varrho_k^h(\cdot; (v_l^j)^{j\in[0,1]})$ satisfies the familiar envelope formula with respect to v_k). In the spirit of the Wilson doctrine, this also means that the the optimal allocation rule can be robustly (fully) implemented in weakly undominated strategies.[17]

[16] In the implementation literature, this problem is referred to as "partial implementation", whereas in the two-sided market literature as the "chicken and egg" problem (e.g., Caillaud and Jullien (2001, 2003)) or the "failure to launch" problem (e.g., Evans and Schmalensee (2009)). See also Ellison and Fudenberg (2003) and Ambrus and Argenziano (2009).

[17] With more general preferences, it is still possible to robustly (fully) implement any monotone matching rule in weakly undominated strategies by replacing the definition of $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k$ in (14) with $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k \equiv \int_{\{j: v_l^j \geq t_k(v_k)\}} \underline{\sigma}_k, \lambda(j)$, where $\underline{\sigma}_k \equiv \min\{\sigma_k(\mathbf{u}_l) : \mathbf{u}_l \in \mathbf{U}_l\}$.However, these payments generate less revenue than the ones given in (4).

Coarse Matching. In reality, platforms typically offer menus with finitely many alternatives. As pointed out by McAfee (2002) and Hoppe, Moldovanu and Ozdenoren (2010), the reason for such *coarse matching* is that platforms may face costs for adding more alternatives to their menus.[18]

It is easy to see that the analysis developed in this paper extends to a setting where the platform can include no more than N alternatives in the menus offered to each side. Furthermore, as the number of alternatives increases (e.g., because menu costs decrease), the solution to the platform's problem uniformly converges to the h-optimal nested multi-homing rule identified in the paper (This follows from the fact that any weakly decreasing threshold function $t_k(\cdot)$ can be approximated arbitrarily well by a step function in the sup-norm, i.e., in the norm of uniform convergence). In other words, the maximally-separating nested multi-homing rules of Proposition 3 are the limit as N grows large of those offered when the number of non-exclusive networks is finite.

Quasi-Fixed Costs. Integrating an agent into a network structure typically involves a quasi-fixed cost. In the Cable TV example, a household must be connected to the underground cable system to get access to the Cable Company 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. From the perspective of the platform, these costs are quasi-fixed, in the sense that they depend on whether a given agent is included in *some* network, but not on the agent's matching set.

The analysis developed above can easily incorporate such costs. Let c_k denote the quasi-fixed cost that the platform must incur for each agent from side k whose matching set is non empty. The h-optimal mechanism can then be obtained by the following two-step procedure:

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:

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

As the quasi-fixed costs increase, so do the exclusion types $\omega_k^h(c_A, c_B)$, $k = A, B$. 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, another testable prediction that the model delivers is that, ceteris paribus, single networks should be more prevalent in matching markets with high quasi-fixed costs, while nested multi-homing in markets with low quasi-fixed costs.

The Group Design Problem. Consider the problem of how to assign agents to different groups in the presence of peer effects, which is central to the theory of organizations and personnel economics. As anticipated in the Introduction, such one-sided matching problem is a special case

[18] See also Wilson (1989).

of the two-sided matching problems studied in this paper. To see this, 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 where (i) preferences and type distributions of sides A and B coincide, and (ii) the matching rule is required to be symmetric across sides, i.e., $t_A(v) = t_B(v)$ all $v \in V_A = V_B$. Under this new constraint, maximizing (7) is equivalent to maximizing twice the objective associated with the one-sided matching problem. As it turns out, the symmetry constraint is never binding in a two-sided matching market where the two sides are perfectly symmetric. This is immediate when $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$, that is, when a single complete network is h-optimal. Under nested-multi-homing, $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$, 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 the symmetry condition is satisfied also when the optimal rule exhibits bunching at the top. As a consequence, one can reinterpret all our results in terms of the group design problem.

We now discuss lines of future research.

Same-side Externalities. The analysis developed above assumed that the utility/profit that each agent derives from any given matching set is independent of who else from the same side has access to the same set. In other words, we abstracted from "same-side" externalities. In advertising markets, for example, reaching a certain set of households is more profitable if competitors are precluded from reaching the same set. Similar congestion effects are present in other matching markets. Extending the analysis in this direction is promising and likely to introduce novel effects that complement those documented in the present paper.

Different-sign Externalities. Our analysis also assumed that each agent either benefits or suffers from being matched to the other side of the market (the intensity varying with the particular agents he/she is matched to). Allowing the same agent to derive positive utility from interacting with certain agents and negative utility from interacting with others is also likely to deliver new insights.

Horizontal Differentiation. The model considered in the present paper is one of pure vertical differentiation. While this was useful to isolate important effects, many applications of interest feature both vertical and horizontal differentiation. Extending the analysis to incorporate elements of horizontal differentiation is challenging but highly promising.

Platform Competition. Matching markets are often populated by competing platforms. Understanding to what extent the distortions identified in the present paper are affected by the degree of market competition and studying policy interventions (subsidies/taxes and in some cases the imposition of universal service obligations) aimed at boosting welfare by inducing platforms to get more agents "on board" is another important direction of future research.[19]

[19] Damiano and Li (2008) consider a model in which two matchmakers compete through entry fees on two sides. However, they restrict the analysis to one-to-one matching thus abstracting from many of the effects identified in the present paper.

While we do expect the above extensions to open the door to novel effects, we also expect the key insights identified in the present paper to remain valid also in these richer settings.[20]

5 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 h-optimality requires that each agent from each side be matched to all agents from 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 (6).

Thus consider the situation where $\varphi_k^h(\underline{v}_k) < 0$ for some $k \in \{A, B\}$. Define $\Theta_k^+ \equiv \{\theta_k = (\mathbf{u}_k, v_k) : \varphi_k^h(v_k) \geq 0\}$ the set of types $\theta_k = (\mathbf{u}_k, v_k)$ whose φ_k^h-valuation is non-negative, and $\Theta_k^- \equiv \{\theta_k = (\mathbf{u}_k, v_k) : \varphi_k^h(v_k) < 0\}$ 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 (6) and that weakly increases the platform's objective $\Omega^h(M)$.

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

1. If $|\mathbf{s'}_k(\theta_k)|_k \geq \left|\Theta_l^+\right|_k$, then let $\hat{t}_k(v_k)$ be such that

$$|\mathbf{U}_l \times [\hat{t}_k(v_k), \bar{v}_l]|_k = \left|\mathbf{s'}_k(\theta_k)\right|_k ;$$

2. If $|\mathbf{s'}_k(\theta_k)|_k \leq \left|\Theta_l^+\right|_k = |\Theta_l|_k$, then $\hat{t}_k(v_k) = \underline{v}_l$.

3. If $0 < |\mathbf{s'}_k(\theta_k)|_k \leq \left|\Theta_l^+\right|_k < |\Theta_l|_k$ (in which case $r_l^h \in (\underline{v}_l, \bar{v}_l)$), where r_l^h is implicitly defined by $\varphi_l^h(r_l^h) = 0$, then let $\hat{t}_k(v_k) = r_l^h$.

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

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

By construction, $\hat{\mathbf{s}}_k(\cdot)_k$ is monotone and hence 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 for any $k \in \{A, B\}$

$$\int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)|_k\right) dF_k(\mathbf{u}_k, v_k) \geq \int_{\Theta_k^+} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s'}_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k). \qquad (15)$$

[20] For example, we expect competition to reduce the amount of surplus captured by the platforms but not necessarily the distortions in the provision of the matching services identified in the present paper. Indeed, as indicated in the literature on competition in nonlinear prices, distortions may be even larger under (imperfect) competition than in the monopolist case. Furthermore, when contract offers are allowed to depend on the offers made by the competitors (aka "meet the competition clause"), it is often possible to sustain the monopolist outcome as a non-cooperative equilibrium, which suggests that the results in the present paper are likely to remain relevant also for the case of competing platforms.

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

$$\int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k) \leq \int_{\Theta_k^-} \varphi_k^h(v_k) \cdot g_k\left(\left|\hat{\mathbf{s}}_k(\mathbf{u}_k, v_k)\right|_k\right) dF_k(\mathbf{u}_k, v_k). \tag{16}$$

To establish (16), we start with the following result.

Lemma 1 *A mechanism M is incentive compatible only if, with the exception of a countable subset of V_k, $\left|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k = \left|\mathbf{s}_k(\mathbf{u}'_k, v_k)\right|_k$ for all $\mathbf{u}_k, \mathbf{u}'_k \in \mathbf{U}_k$, $k = A, B$.*

Proof of Lemma 1. To see this, note that incentive compatibility requires that $\left|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k \geq \left|\mathbf{s}_k(\mathbf{u}'_k, v'_k)\right|_k$ for any $(\mathbf{u}_k, v_k)$ and $(\mathbf{u}'_k, v'_k)$ such that $v_k \geq v'_k$. This in turn implies that $\mathbb{E}[\left|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)\right|_k]$ 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 $\left|\mathbf{s}_k(\mathbf{u}_k, v_k)\right|_k$ depends on $\mathbf{u}_k$, the expectation $\mathbb{E}[\left|\mathbf{s}_k(\tilde{\mathbf{u}}_k, v_k)\right|_k]$ is necessarily discontinuous in v_k. Because monotone functions can be discontinuous at most over a countable set of points, this means that the quality of the matching set may vary with the characteristics $\mathbf{u}_k$ only over a countable subset of V_k. Q.E.D.

The next lemma introduces a property for arbitrary random variables that we will use below to establish the result.

Definition 6 *[monotone concave order] Let F be a probability measure on the interval $[a, b]$ and $z_1, z_2 : [a, b] \to \mathbb{R}$ be two random variables defined over $[a, b]$. 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 $g : \mathbb{R} \to \mathbb{R}$.*

Lemma 2 *Suppose $z_1, z_2 : [a, b] \to \mathbb{R}_+$ are nondecreasing. If z_2 is smaller than z_1 in the monotone concave order, then for any weakly concave and weakly increasing function $g : \mathbb{R} \to \mathbb{R}$ and any weakly negative and 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 then have the following result.

Lemma 3 *Consider the two random variables $z_1, z_2 : [\underline{v}_k, r_k^h] \to \mathbb{R}_+$ given by $z_1(v_k) \equiv \mathbb{E}_{\tilde{\mathbf{u}}_k}\left[\left|\mathbf{s}'_k(\tilde{\mathbf{u}}_k, v_k)\right|_k |v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\mathbf{u}}_k}\left[\left|\hat{\mathbf{s}}_k(\tilde{\mathbf{u}}_k, v_k)\right|_k |v_k\right]$, where the distribution over $[\underline{v}_k, r_k^h]$ is given by $F_k^v(v_k)/F_k^v(r_k^h)$. Then z_2 is smaller than z_1 in the monotone concave order.*

Proof of Lemma 3. From (i) the construction of $\hat{\mathbf{s}}_k(\cdot)$, (ii) the assumption of positive affiliation between attractiveness and willingness to pay, (iii) the fact that the measure $F_k^v(v_k)$ is absolute continuous with respect to the Lebesgue measure and (iv) Lemma 1, we have that for all $x \in [\underline{v}_k, r_k^h]$,

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

or, equivalently,

$$\int_{\underline{v}_k}^{x} z_1(v_k) dF_k^v(v_k) \geq \int_{\underline{v}_k}^{x} z_2(v_k) dF_k^v(v_k). \tag{17}$$

Denote by $[\dot{v}_k^1, \dot{v}_k^2]$, $[\dot{v}_k^3, \dot{v}_k^4]$,$[\dot{v}_k^5, \dot{v}_k^6]$,... the collection of T (where $T \in \mathbb{N} \cup \{\infty\}$) 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 $\mathcal{T} \equiv \cup_{t=0}^{T-1} [\dot{v}_k^{2t+1}, \dot{v}_k^{2t+2}]$ 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 $v_k \in \mathcal{T}$;
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 $v_k \in [\underline{v}_k, r_k^h] \backslash \mathcal{T}$;
3. $\int_{[\underline{v}_k, r_k^h] \backslash \mathcal{T}} \{\dot{z}_2(v_k) - z_2(v_k)\} dF_k^v(v_k) = \int_{\mathcal{T}} \{z_2(v_k) - z_1(v_k)\} dF_k^v(v_k)$.

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

$$\int_{\underline{v}_k}^{r_k^h} \dot{z}_2(v_k) dF_k^v(v_k) / F_k^v(r_k^h) = \int_{\underline{v}_k}^{r_k^h} z_2(v_k) dF_k^v(v_k) / F_k^v(r_k^h). \tag{18}$$

This implies that for all weakly concave and weakly increasing functions $g : \mathbb{R} \to \mathbb{R}$,

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

where the first inequality follows from the weak concavity of $g(\cdot)$ along with (18), while 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. Q.E.D.

We are now ready to prove inequality (16). The results above imply that

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

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, $g_k\left(\left|\mathbf{s}'_k(\mathbf{u}_k, v_k)\right|_k\right)$ is invariant in $\mathbf{u}_k$ except over a countable subset of $[\underline{v}_k, r_k^h]$, as shown in Lemma 1. 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 (16), as we wanted to show.

Summing up (15) and (16) 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. Q.E.D.

Proof of Proposition 2 (case $\varphi_l^h(\underline{v}_l) < 0 \le \varphi_k^h(\underline{v}_k)$). We show that a single network is optimal if and only if $\triangle_k^h(\underline{v}_k, \underline{v}_l) \ge 0$, whereas nested multi-homing is optimal if and only if $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$.

First, suppose that $\triangle_k^h(\underline{v}_k, \underline{v}_l) \ge 0$ and that the matching rule is multi-homing. 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 $\triangle_k^h(v_k, v_l)$. 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 $\int_{v'_k}^{v''_k} [\Delta_k^h(v_k, v_l)] dv_k$. Lastly note that $sign\{\Delta_k^h(v_k, v_l)\} = sign\{\psi_k^h(v_k) + \psi_l^h(v_l)\}$. Under the SR condition, this means that $\triangle_k^h(v_k, v_l) > 0$ for all (v_k, v_l). The results above then imply that the platform can 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 $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$ and that the platform uses a single network. First suppose that 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 $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$ 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.

Next consider the case where the single network is incomplete. From the same arguments as in the main text, for such a network to be optimal, it must be that $\hat{\omega}_l^h < r_l^h$ and $\hat{\omega}_k^h = \underline{v}_k$, with $\hat{\omega}_l^h$ satisfying the following first-order condition

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

This condition 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 $\int_{\underline{v}_k}^{\overline{v}_k}[\Delta_k^h(v_k,\hat{\omega}_l^h)]dv_k = 0$. Because $sign\{\Delta_k^h(v_k,\hat{\omega}_l^h)\} = sign\{\psi_k^h(v_k)+\psi_l^h(\hat{\omega}_l^h)\}$, under Condition SR this means that there exists a $v_k^{\#} \in (\underline{v}_k, \overline{v}_k)$ such that $\int_{v_k^{\#}}^{\overline{v}_k}\Delta_k^h(v_k,\hat{\omega}_l^h)dv_k > 0$. This means that there exists a $\omega_l^{\#} < \hat{\omega}_l^h$ such that the platform could increase its payoff by switching to the following nested multi-homing rule:

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

We conclude that multi-homing is optimal when $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. 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 $k, l \in \{A, B\}$, $l \neq k$

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

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

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

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

Because $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$ (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 2, at the optimum, $\omega_k^h \in [\underline{v}_k, r_k^h]$. In addition, whenever $r_l^h > \underline{v}_l$, $\omega_l^h \in [\underline{v}_l, r_l^h]$ and $t_k^h(r_k^h) = r_l^h$. Hereafter, we will assume that $r_l^h > \underline{v}_l$. 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_k^F below by replacing r_l^h with $\underline{v}_l$).

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

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

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

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

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

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

subject to (20),

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

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

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:

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

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_k^R) that is obtained from P_k^{Au} 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 4 *P_k^R admits a piece-wise absolutely continuous maximizer $\tilde{t}_k(\cdot)$.*

Proof of Lemma 4. Program P_k^R is equivalent to the following optimal control problem $\mathcal{P}_k^R$:

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

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 $\mathcal{P}_k^R$ 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 $\mathcal{P}_k^R$. By the equivalence of P_k^R and $\mathcal{P}_k^R$, it then follows that P_k^R admits a piece-wise absolutely continuous maximizer $\tilde{t}_k(\cdot)$. Q.E.D.

Lemma 5 *Consider the function $\eta(\cdot)$ implicitly defined by*

$$\Delta_k^h(v_k, \eta(v_k)) = 0. \tag{31}$$

Let $\tilde{v}_k \equiv \inf\{v_k \in [\underline{v}_k, r_k^h]$:(31) admits a solution$\}$. The solution to P_k^R is given by

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

Proof of Lemma 5. From Lemma 4, we know that P_k^R 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 with program P_k^R is given by (31). Condition SR 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 $\tilde{t}_k(v_k) = \overline{v}_l$ 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 $\hat{t}_k(\cdot)$ 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 valuation. This proves that the solution to P_k^R 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_k^R is the control $\tilde{t}_k(\cdot)$ that satisfies (32). Q.E.D.

Denote by $\max\{P_k^R\}$ the value of program P_k^R (i.e., the value of the objective (29) evaluated under the control $\tilde{t}_k^h(\cdot)$ defined in Lemma 5). Then denote by $\sup\{P_k^{Au}\}$ and $\sup\{P_k^F\}$ 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 6 $\sup\{P_k^F\} = \sup\{P_k^{Au}\} = \max\{P_k^R\}$.

Proof of Lemma 6. 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 $\tilde{t}_k^h(\cdot)$ 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_F^k. Let $\omega_k^h = \tilde{v}_k$, where $\tilde{v}_k$ is the threshold defined in Lemma 5. Next for any $v_k \in [\tilde{v}_k, r_k^h]$, let $t_k^h(v_k) = \tilde{t}_k(v_k)$ where $\tilde{t}_k(\cdot)$ is the function defined in Lemma 5. Finally, given $t_k^h(\cdot) : [\omega_k^h, r_k^h] \to [r_l^h, \overline{v}_l]$, let $t_l^k(\cdot) : [r_l^h, \overline{v}_l] \to [\omega_k^h, r_k^h]$ be the unique function that satisfies (20). It is clear that the tripe $\omega_k^h, t_k^h(\cdot), t_l^h(\cdot)$ 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 $\omega_k^h, t_k^h(\cdot), t_l^h(\cdot)$ is the same as $\max\{P_k^R\}$. From Lemma 6, we then conclude that $\omega_k^h, t_k^h(\cdot), t_l^h(\cdot)$ 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 $\{\omega_k^h, t_k^h(\cdot)\}_{k\in\{A,B\}}$ to program P_F.

By inspection, it is easy to see that the corresponding rule is maximally separating. Furthermore, from the arguments in Lemma 5, 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 $\Delta_k^h(v_k', \bar{v}_l) = 0$, or equivalently $\psi_k^h(v_k') + \psi_l^h(\bar{v}_l) = 0$. Condition SR along with the fact that $sign\{\Delta_k^h(v_k, v_l)\} = sign\{\psi_k^h(v_k) + \psi_l^h(v_l)\}$ then implies 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 exclusion at the bottom on side k and no 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 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 $\triangle_k^h(\underline{v}_k, \eta(\underline{v}_k)) = 0$, 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 Condition SR, 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. Q.E.D.

Proof of Proposition 5. Let $x_k^h(v_k) \equiv \left|\mathbf{s}_k^h(v_k)\right|_k$ denote the size of the matching set that each agent with valuation v_k obtains under the mechanism M^h. Using (4), for any $q_k \in x_k^h(V_k)$, i.e., for any q_k induced by M^h,

$$\rho_k^h(q_k) = \left(x_k^h\right)^{-1}(q_k) \cdot g_k(q_k) - \int_{\underline{v}_k}^{\left(x_k^h\right)^{-1}(q_k)} g_k(x_k^h(v))dv,$$

where $\left(x_k^h\right)^{-1}(q_k) \equiv \inf\{v_k : x_k^h(v_k) = q_k\}$ is the generalized inverse of $x_k^h(\cdot)$. It follows from Proposition 3 that $\left(x_k^h\right)^{-1}(q_k)$ is strictly increasing and differentiable at any q_k in the image of the separating range, i.e., for any $q_k \in [\left|\mathbf{s}_k^h(\omega_k^h)\right|_k, \left|\mathbf{s}_k^h(t_l^h(\omega_l^h))\right|_k]$. Therefore, from the integral formula above, we get that the optimal price schedules $\rho_k^h(\cdot)$ are differentiable at any quantity q_k in the image of the separating range, and

$$\frac{d\rho_k^h}{dq_k}(q_k) = \left(x_k^h\right)^{-1}(q_k) \cdot g_k'(q_k) = v_k \cdot g_k'\left(\left|\mathbf{s}_k^h(v_k)\right|_k\right), \tag{33}$$

where $|\mathbf{s}_k^h(v_k)|_k = q_k$. Substituting the elasticity formula (12) and the marginal price formula (33) into the Lerner-Wilson formula (13) and using the same formulas for side l and recognizing that

$$\left(x_l^h\right)^{-1}\left(D_k\left(q_k,\frac{d\rho_k^h}{dq_k}(q_k)\right)\right)=\left(x_l^h\right)^{-1}\left(1-F_k^v\left(\left(x_k^h\right)^{-1}(q_k)\right)\right)=t_k^h(\left(x_k^h\right)^{-1}(q_k))=t_k^h(v_k)$$

for v_k such that $|\mathbf{s}_k^h(v_k)|_k=q_k$, then leads to the Euler equation (10). Q.E.D.

References

[1]

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

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

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

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

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

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

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

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

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

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

[12] Cramton, P., R. Gibbons and P. Klemperer, 1987, Dissolving a Partnership Efficiency, *Econometrica*, vol. 55(3), 615-632.

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

[14] Cripps, M. and J. Swinkels, 2006, Efficiency of Large Double Auctions, *Econometrica*, vol. 74(1), 47-92.

[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] 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.

[19] 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.

[20] 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.

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

[22] Gresik, T. and M. Satterthwaite, The Rate at Which a Simple Market Converges to Efficiency as the Number of Traders Increases: An Asymptotic Result for Optimal Trading Mechanisms, *Journal of Economic Theory*, Vol. 48, pp. 304-332.

[23] Fieseler, K., T. Kittsteiner and B. Moldovanu, 2003, Partnerships, Lemmons, and Efficient Trade, *Journal of Economic Theory*, Vol. 113, pp. 223-234.

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

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

[26] 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.

[27] Jehiel, P. and A. Pauzner, 2006, Partnership Dissolution with Interdependent Values, *Rand Journal of Economics*, Vol. 37(1), pp. 1-22.

[28] Jonhson, T., 2010, Matching Through Position Auctions, *Working Paper*.

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

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

[31] McAfee, R. P., 1991, Efficient Allocation with Continuous Quantities, *Journal of Economic Theory*, Vol. 53, pp. 51-74.

[32] McAfee, R. P., 1992a, Amicable Divorce: Dissolving a Partnership with Simple Mechanisms, *Journal of Economic Theory*, Vol. 56, pp. 266-293.

[33] McAfee, R. P., 1992b, A Dominant Strategy Double Auction, *Journal of Economic Theory*, Vol. 56, pp. 434-450.

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

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

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

[37] Milgrom, R. and J. Weber, 1982, A Theory of Auctions and Competitive Bidding, Econometrica, Vol. 50(5), pp. 1089-1122.

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

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

[40] Myerson, R. and M. Satterthwaite, 1983, Efficient Mechanism for Bilateral Trading, *Journal of Economic Theory*, Vol. 28, pp. 265-281.

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

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

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

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

[45] Rustichini, M. A. Satterthwaite and S. Williams, 1994, Convergence to Efficiency in a Simple Market with Incomplete Information, *Econometrica*, Vol. 62(5), pp. 1041-1063.

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

[47] Satterthwaite, M. and S. Williams, 1989, The Rate of Convergence to Efficiency in the Buyer's Bid Double Auction as the Market Becomes Large, *The Review of Economic Studies*, Vol. 56(4), pp. 477-498.

[48] 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.

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

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

[51] Strausz, R., 2006, Deterministic versus Stochastic Mechanisms in Principal-Agent Models, *Journal of Economic Theory*, Vol. 127(1), pp. 306-314.

[52] Weyl, G., 2010, A Price Theory of Two-Sided Markets, *American Economic Review*, Vol. 100(4), pp. 1642–72

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

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

[55] Wilson, R., 1997, Nonlinear Pricing, Oxford University Press.

Online Supplementary Material

1. Proofs omitted in the main text

Proof of Proposition 4. The result trivially holds when $\triangle_k^W(\underline{v}_k, \underline{v}_l) \geq 0$, for in this case the welfare-maximizing matching rule always employs a single complete network. Thus suppose that $\triangle_k^W(\underline{v}_k, \underline{v}_l) < 0$. Because $\varphi_k^P(v_k) \leq \varphi_k^W(v_k)$ for all $v_k \in [\underline{v}_k, \overline{v}_k]$, with strict inequality for all $v_k < \bar{v}_k$, then $\triangle_k^P(\underline{v}_k, \underline{v}_l)$ is also strictly negative. Furthermore, the same property implies that $\psi_k^P(v_k) \leq \psi_k^W(v_k)$ for all $v_k \in [\underline{v}_k, \overline{v}_k]$. 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 $\omega_k^P \geq \omega_k^W$ 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 isolation effect. Q.E.D.

Proof of Proposition 6. Hereafter, we use the annotation "ˆ" for all variables in the mechanism $\hat{M}^P$ corresponding to the new distribution $\hat{F}_k^\sigma(\cdot|\cdot)$ and continue to denote the variables in the mechanism M^P corresponding to the original distribution $F_k^\sigma(\cdot|\cdot)$ 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 $v_k \in [\omega_k^P, r_k^P]$, $\hat{t}_k^P(v_k) \leq t_k^P(v_k)$. 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 $v_k > r_k^P$. This completes the proof of part (1) in the Proposition.

Next consider part (2). Note that, because F_l is unchanged, the result in part 1 implies that $|\hat{\mathbf{s}}_k(v_k)|_k \geq |\mathbf{s}_k(v_k)|_k$ if and only if $v_k \leq r_k^P$. Using (4), note that for all types θ_k with valuation $v_k \leq r_k^P$

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

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 2 in the proposition. Q.E.D.

Proof of Corollary 3. 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 valuation 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 (4), for any $q \in x_k(V_k) \cap \hat{x}_k(V_k)$, i.e., for any q offered both under M^P and $\hat{M}^P$,

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

where $x_k^{-1}(q) \equiv \inf\{v_k : x_k(v_k) = q\}$ 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 6, 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 $q \geq x_k(r_k^P) = \hat{x}_k(r_k^P)$,

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

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 the result. Q.E.D.

2. Effects of Changes in Demand Elasticity

In this part, we formally prove the claims in Section 3.5 in the main text about the effects of changes in demand elasticity on matching sets and prices, under profit maximization.

Definition 7 (higher elasticity) *Side k is less elastic under distribution $\tilde{F}_k$ than under distribution F_k if $\tilde{F}_k^v$ dominates F_k^v in the hazard-rate order, and, for any v_k, $\tilde{F}_k^\sigma(\cdot|v_k) = F_k^\sigma(\cdot|v_k)$.*[21]

The next proposition extends the results in Rochet and Tirole (2006) and Armstrong (2006) for single-networks to the case of a two-sided platform that price-discriminates by offering menus of networks.

Proposition A1 (effects of changes in elasticity on matching sets and payoffs). *In addition to Conditions PA and SR, suppose that network effects are linear (i.e., $g_A(x) = g_B(x) = x$), and that $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. If the elasticity of side k decreases, then the platform moves from a matching rule $s_k^P(\cdot)$ to a matching rule $\tilde{\mathbf{s}}_k^P(\cdot)$ such that:*

1. *$\tilde{\mathbf{s}}_k^P(v_k) \subseteq s_k^P(v_k)$ for all $v_k \in V_k$,*
2. *$\Pi_k(\theta_k; \tilde{M}^P) < \Pi_k(\theta_k; M^P)$ for all $v_k \geq \omega_k^P$.*

Proof of Proposition A1. Denote by $\tilde{\psi}_k^P(v_k)$ the ψ-function associated with the new distribution $\tilde{F}_k^v(\cdot)$, and by $\tilde{\varphi}_k^P(v_k)$ the virtual valuations associated with $\tilde{F}_k^v(\cdot)$. 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 $g_k(\cdot)$ and $g_l(\cdot)$ 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. For Part 2 note that, because F_l is unchanged, it then follows that

[21] The distribution $\tilde{F}_k^v$ dominates F_k^v in the hazard-rate order if and only if $\frac{f_k^v(v_k)}{1-F_k^v(v_k)} \geq \frac{\tilde{f}_k^v(v_k)}{1-\tilde{F}_k^v(v_k)}$ for all $v_k \in V_k$.

$|\tilde{\mathbf{s}}_k(v_k)|_k \leq |\mathbf{s}_k(v_k)|_k$ 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 $v_k \geq \omega_k^P$. Q.E.D.

The next corollary builds on the preceding Proposition A1 to determine the effects of a decrease in the elasticity on side k on the price schedule $\rho_k^P(\cdot)$.

Corollary A1 (effects of changes in elasticity on prices) *In addition to Conditions PA and SR, suppose that network effects are linear (i.e., $g_A(x) = g_B(x) = x$), and that $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. If the elasticity of side k decreases, then the platform switches 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. Furthermore, in the case where network effects depend only on quantities (i.e., $\sigma_l(\cdot) \equiv 1$), the marginal price for each quantity q that is offered both under the old and the new distribution increases, that is, $d\tilde{\rho}_k^P(q)/dq > d\rho_k^P(q)/dq$.*

Proof of Corollary A1. Because $|\tilde{\mathbf{s}}_k(v_k)|_k \leq |\mathbf{s}_k(v_k)|_k$ for all $v_k \in V_k$, the result that $\tilde{\rho}^P(q) \geq \rho_k^P(q)$ for all q follows from the same steps as in the proof of Corollary 3. The result for the marginal prices follows from Proposition 5. To see this, first note that, for each quantity q offered both under the old and the new rule, $\left(\tilde{x}_l^P\right)^{-1}\left(\tilde{D}_k\left(q, \frac{d\tilde{\rho}_k^P(q)}{dq}\right)\right) = \left(x_l^P\right)^{-1}\left(D_k\left(q, \frac{d\rho_k^P(q)}{dq}\right)\right)$. To see this, recall that, by construction $(x_l^P)^{-1}\left(D_k\left(q, \frac{d\rho_k^P(q)}{dq}\right)\right) = t_k^P(\left(x_k^P\right)^{-1}(q))$ and $\left(\tilde{x}_l^P\right)^{-1}\left(\tilde{D}_k\left(q, \frac{d\tilde{\rho}_k^P(q)}{dq}\right)\right) = \tilde{t}_k^P(\left(\tilde{x}_k^P\right)^{-1}(q))$. Because the distribution on side l has not changed, it must be that $t_k^P(\left(x_k^P\right)^{-1}(q)) = \tilde{t}_k^P(\left(\tilde{x}_k^P\right)^{-1}(q))$. Using the fact that, when network effects are linear,

$$(x_l^P)^{-1}\left(D_k\left(q, \frac{d\rho_k^P(q)}{dq}\right)\right) = \frac{d\rho_l^P\left(D_k\left(q, \frac{d\rho_k^P(q)}{dq}\right)\right)}{dq_l}$$

we then have that the right-hand-side in the formula for the marginal prices in Proposition 5 is unaffected by the shock. This result, together with the fact that, for any given marginal price $d\rho_k^P(q_k)/dq_k$ on side k, a decrease in elasticity on side k implies that the right-hand-side in the same formula is smaller when evaluated at the same marginal price and the fact that the right-hand-side is increasing in $d\rho_k^p(q_k)/dq_k$ then gives the result. Q.E.D.