Matching Plans for Agents with Vertically and Horizontally Differentiated Preferences*

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

Alessandro Pavan
Northwestern University and CEPR
alepavan@northwestern.edu

30th December 2015

Abstract

We develop a location model of price discrimination in many-to-many matching markets in which agents' preferences are both vertically and horizontally differentiated. The optimal plans induce negative assortative matching at the margin: at any given location, agents with a low value for interacting with agents from the opposite side (the vertical dimension) are included in the matching sets of only those agents from the opposite side whose value for matching is sufficiently high (cross-subsidization). The thresholds for inclusion naturally respect the agents' mutual attractiveness and joint locations. The analysis has implications for the design of cable TV packages, business-to-business platforms, and advertising.

JEL classification: D82

keywords: many-to-many matching, two-sided markets, networks, asymmetric information, incentives, mechanism design

*Pavan thanks the NSF for financial support under the grant SES-60031281. The usual disclaimer applies.

1 Introduction

Matching intermediaries play a central role in modern economies. In electronic commerce, for example, business-to-business (B2B) platforms match vendors with procurers in search of business opportunities.[1] These platforms often engage in price-discriminatory practices by offering different "matching plans"to each side of the market. The matching plans offered to the procurers determine the matching plans faced by the vendors, while the matching plans offered to the vendors determine the matching plans the platform can offer to the procurers. As a consequence of this interdependency, when designing their matching plans, B2B platforms have to internalize the cross-side effects on profits that each side induces on the other side.

Another example of mediated many-to-many matching design is the provision of cable TV services. 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 the viewers. The mirror image of this problem consists in designing a menu of packages of viewers to offer to the channels (reaching more viewers yields larger advertising revenues for the channels, but may also imply larger expenses in terms of broadcasting rights). As in the case of B2B platforms, 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. The presence of such cross-side effects is one of the defining features of mediated many-to-many matching, which is the focus of this paper.[2]

We consider markets in which preferences for matching exhibit elements of both vertical and horizontal differentiation. The vertical differentiation reflects the fact that certain agents may value interacting with agents from the other side uniformly more than other agents. The horizontal differentiation reflects the possibility that two agents from the same side may disagree on the relative attractiveness of any two agents from the opposite side, even when both agents positively (or negatively) value interacting with such agents. Furthermore, the same agent may value interacting with certain agents positively but with others negatively. For example, in the case of cable TV, viewers typically differ not only in the overall importance that they assign to cable TV (the vertical dimension in their preferences) but also in the attractiveness that they assign to different channels (e.g., some viewers prefer sports channels, while others prefer movie channels). Other examples of markets where a combination of horizontal and vertical differentiation is believed to play an important role include online targeted advertising, and the market for the provision of employment matching services.

We consider the problem of a monopolistic platform that operates on two sides. Agents from each side differ both in their preferences as well as in personal traits that determine their attractiveness

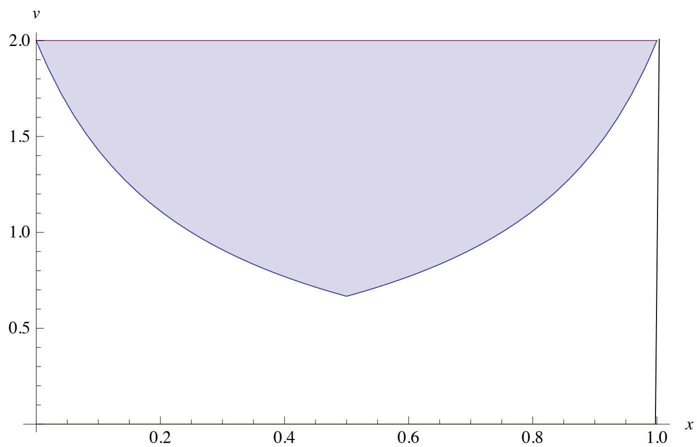
[1]According to the U.S Census Bureau, electronic commerce accounted for 47 percent of manufacturing total value in the US in 2010 (http://www.census.gov/econ/estats/2010/2010reportfinal.pdf).

[2]What makes the platform's problem nonseparable between the two sides is the fact that the cost of expanding the matching sets on each side depends on the entire matching schedule offered on the other side, which is part of the design.

to those agents they are matched to. For example, procurers may differ in their budget, purchasing history, and various other traits that determine their attractiveness to the vendors. Similarly, in the case of cable TV, viewers may differ not only in the importance that he assigns to cable TV (the vertical dimension) and in their ideal channel profile (their horizontal dimension), but also in various personal traits such as location, gender, and education, that contribute to their attractiveness to the channels. To capture the above possibilities, we consider a model in which agents from each side are located on a circle, and where the utility that each agent i from side A obtains from each agent j from side B is increasing in the agent's value for matching (the vertical dimension), and decreasing in the *distance* between the two agents' locations. In addition, we let the utility that each agent i derives from each agent j depend on agents j's personal traits other than agent j's location. For example, in the cable TV application, a channel's location should be interpreted as the channel's broadcasting profile (say, a "news"channel), while a viewer's location should be interpreted as the viewer's most preferred type of programming (say, "sports"). Holding fixed the channel's and the viewer's location, the utility the viewer derives from the channel may then depend also on various channel's traits such as the quality of its shows and the amount of its advertising. On the other hand, holding locations fixed, a channel's vertical dimension could simply reflect the channel's profit margin, per viewer.

We consider the possibility that agents possess private information about dimensions that may be relevant for their preferences as well as dimensions that are responsible for their attractiveness. The platform's problem then consists in choosing a menu of matching plans for each of the two sides of the market that jointly maximize either welfare or profits. We describe the allocations induced under such menus by means of a *matching rule* and a *pricing rule.* A matching rule assigns each agent to a set of agents from the other side. A pricing rule in turn describes the total payment made by the agent to the platform.

We start by identifying primitive conditions under which, irrespective of whether the platform maximizes profits or welfare, the optimal matching rules display a location-specific threshold structure. These rules work as follows: for any given location x_A on side A and any given value v_A for matching, the rule specifies a threshold $t(x_B, |x_A, v_A; \sigma_A, \sigma_B)$ for each location x_B on side B (which may also depend on personal traits, σ_A and σ_B) such that each agent from side A located at x_A with value for matching v_A and traits σ_A is matched to all agents from side B located at x_B, and with traits σ_B, whose value for matching v_B is above this threshold. Figure 1 below illustrates this structure, focusing for simplicity on the case where the traits σ are uniform across agents (and hence dropped from the picture).



Optimal matching rule. The shaded area in the figure describes the matching set for an agent from side A located at $x_A = 1/2$.

The optimal rules thus induce a form of negative assortative matching at the margin according to which, holding the locations and the traits fixed, agents with a low value for matching are matched only to those agents from the other side whose value for matching is sufficiently high. This form of negative assortativeness naturally takes into account any pair of agents' mutual attractiveness, which comes from a combination of location and personal traits. As a result, the matching sets of any two agents are nested only if the two agents share the same location and personal traits. This result has implications, for example, for the design of cable TV services where the market has gradually moved from menus of packages with a nested structure (basic package, premium, premium plus) to a non-nested structure whereby users can now personalize their packages by adding individual channels (see the discussion in Section 5).

We also show that, when preferences are supermodular in the vertical dimension and in the distance between locations, then the thresholds increase with distance (therefore reducing the mass of agents included in the matching set).[3] In other words, the composition of the matching sets naturally respects the agents' idiosyncratic preferences: those agents from the opposite side whose combination of location and personal traits makes them particularly attractive are present in the matching sets in the largest proportions.[4]

We then show how the optimal matching rules can be implemented by offering each agent a menu of *matching plans*. In the cable TV application, a plan is indexed by its category (e.g., movies, sports, news, etc.) and comes with a baseline price and a baseline configuration (the group of channels included in the baseline package). In addition, a plan specifies the price that the subscriber has to pay to add channels from each possible category. Agents then select the plan that best fits their

[3] In the cable TV application, supermodular preferences on the viewers' side mean that viewers with a high value for cable TV are, in general, more likely to also watch channels whose programming is distant from their ideal type.

[4] Importantly, this property, while natural, need not hold without the supermodularity assumption, which guarantees that the benefit of permitting two agents with given (vertical) characteristics to interact increases with their proximity.

preferences (e.g., viewers who like sports choose the sports package) and then personalize the package by adding a few additional channels of their choice. We show how the marginal prices in each plan can be conveniently expressed by means of a Lerner-Wilson formula that uses the (location-specific) elasticities of the aggregate demands to equalize the marginal gains of expanding the matching sets on each side to the cross-subsidization losses on the opposite side.

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 contains the main results. Section 5 discusses how the results relate to a few markets of interest. Section 6 concludes. All proofs are in the Appendix at the end of the document.

Related Literature

The paper is mostly 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 in which the product sold by the monopolist is access to other agents. The study of price discrimination in markets for many-to-many matching introduces novel features 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 serves as both a consumer and an input in the matching production function. This feature of matching markets implies that the cost of procuring an input is endogenous and depends in a nontrivial way on the entire matching rule. Lastly, relative to this literature, we consider a richer specification of the agents' preferences that combines elements of both vertical and horizontal differentiation.

Two-Sided Markets. Markets where agents purchase access to other agents are the focus of the literature on two-sided markets (e.g., Caillaud and Jullien (2003), Rochet and Tirole (2003, 2006), Armstrong (2006), Hagiu (2008), Ambrus and Argenziano (2009), Weyl (2010), and Jullien (2011)). This literature, however, restricts attention to a single network, or to mutually exclusive networks.[5] Our contribution relative to this literature, is in allowing for general matching rules and in introducing certain elements of private information on both the agents' preferences and their attractiveness.

Matching Design with Transfers. Rayo (2010) studies second-degree price discrimination by a monopolist selling a menu of conspicuous goods that serve as signals of consumers'hidden characteristics. Rayo's model can be interpreted as a one-sided matching model with purely vertically differentiated preferences where the utility of a matching set is proportional to the average quality of its members. Allowing for more general peer effects, Board (2009) studies the design of groups by a profit-maximizing platform (e.g., a school) that can induce agents to self-select into mutually

[5]Mutually exclusive networks correspond to matching rules with the following property: any two agents from the same side whose matching sets overlap, have the same matching sets. This property is not satisfied in the case of cable TV, B2B matching, and online advertising.

exclusive groups (e.g., classes).[6]

In a two-sided matching environment with purely vertically differentiated preferences, McAfee (2002) shows that partitioning agents on each side in two categories ("high" and "low"), and performing random one-to-one matching within category generates at least half of the welfare produced by one-to-one positive assortative matching. Hoppe, Moldovanu and Ozdenoren (2010) (i) sharpen McAfee's lower bounds in the case of welfare-maximization, and (ii) obtain lower bounds in the case of profit-maximization. Damiano and Li (2007) identify primitive conditions for a profit-maximizing platform to match agents in one-to-one positive assortative way. Johnson (2010) studies indirect implementations of one-to-one positive assortative matching through positions auctions. 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 to these papers, we allow for matching rules that assign agents to nonexclusive groups and in which the quality of a matching set is determined by the *sum* of the attractiveness of its members as opposed to the average attractiveness.[7] As a result, our predictions are fundamentally different from those derived in the above papers: the optimal rules induce many-to-many matching and are characterized by a threshold structure that implies a form of negative assortative matching at the margin, as described above. In this respect, the closest paper is Gomes and Pavan (2015). While that paper confines attention to markets in which preferences for matching exhibit only elements of vertical differentiation, the present paper studies markets with both vertically and horizontally differentiated preferences.

Decentralized Matching. In a decentralized matching economy, Eeckhout and Kircher (2010a) study price discrimination by principals who are randomly matched with agents. They show that, for partially rival meeting technologies, principals offer a distribution of posted prices, and agents with higher valuations choose principals with higher prices (ex-ante sorting).

Another strand of the literature (see, among others, Shimer and Smith (2000) and Eeckhout and Kircher (2010b)) extends the assignment model of Becker (1973) to settings 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 appropriate forms of supermodularity. Relative to this literature, we study mediated many-to-many matching, abstract from search frictions, and consider flexible discriminatory pricing rules.

---

[6] See also Arnott and Rowse (1987), Epple and Romano (1998), Helsley and Strange (2000), and Lazear (2001) for models of group design under complete information.

[7] The assumption that agents care only about the *average* quality of their partners offers an appropriate description of those markets in which (i) agents are interested in a single interaction/transaction with the other side and (ii) the identity of the partner is, to a large extent, random. It is less appropriate for those markets in which (a) multiple interactions are positively valued and/or (b) agents can select their partner(s) deterministically, as in the case of cable TV, or business-to-business matching.

2 Model

A monopolistic platform matches agents from two sides of a market. Each side $k \in \{A, B\}$ is populated by a unit-mass continuum of agents, indexed by $i \in [0, 1]$. Each agent $i \in [0, 1]$ from each side $k \in \{A, B\}$ has a three-dimensional type $\theta_k^i = (x_k^i, v_k^i, \sigma_k^i) \in \Theta_k \equiv X_k \times V_k \times \Sigma_k$. The first component, x_k^i, is the "location"of agent i; for convenience, we assume that agents are located on a circle of perimeter one, in which case $X_k = [0, 1]$, $k = A, B$. The second component $v_k^i \in V_k \equiv [\underline{v}_k, \overline{v}_k] \subseteq \mathbb{R}$ parametrizes agent i's "value for matching"; that is, the overall importance that agent i assigns to interacting with agents from the other side (the vertical dimension of the agent's preferences). The third component, $\sigma_k^i \in \Sigma_k \equiv [\underline{\sigma}_k, \overline{\sigma}_k] \subseteq \mathbb{R}$, parametrizes various traits of agent i other than his location that are irrelevant for the agent's own preferences, but contribute to the agent's attractiveness, as perceived from agents from the opposite side.

For example, as alluded to in the Introduction, in the cable TV application, the vertical parameter v_k^i may capture the overall importance that a viewer assigns to cable TV, or the overall importance that a channel assigns to reaching a representative viewer (e.g., reflecting the channel's expected advertising revenue as well as possible costs stemming from broadcasting rights). In turn, the location parameter x_k^i may capture a viewer's (horizontal) tastes for different types of programming. Finally, the traits σ_k^i may capture the quality, or quantity, of a channel's advertising, or, in case of a viewer's, a combination of demographic and socio-economic traits that determine the profitability of reaching the viewer from the prespective of the channel.

We assume that the vertical parameters v_k^i are the agents' own private information. As for the other parameters (x_k^i, σ_k^i), we will discuss cases where they are publicly observable as well as cases where they are the agents' private information. It the cable TV application, for example, it seems appropriate to assume that each viewer's ideal type of broadcasting is his own private information, whereas each channel's broadcasting profile is publicly observable.

The utility enjoyed by agent $i \in [0, 1]$ from side $k \in \{A, B\}$ from being matched to agent $j \in [0, 1]$ from side $l \neq k$ is given by

$$u_k(v_k^i, |x_k^i - x_l^j|, \sigma_l^j)$$

where $|x_k^i - x_l^j|$ is the distance between the two agents' locations. The function u_k is Lipschitz continuous, strictly increasing, continuously differentiable, and weakly concave in v_k^i, weakly increasing in σ_l^j, and weakly decreasing in $|x_k^i - x_l^j|$. The following example illustrates the type of preferences covered by the aforementioned specification.

Example 1 ***(log utility)*** The utility that each agent $i \in [0, 1]$ from each side $k \in \{A, B\}$ obtains from being matched to each agent $j \in [0, 1]$ from side $l \neq k$ is given by

$$u_k(v_k^i, |x_k^i - x_l^j|, \sigma_l^j) = \log \left[v_k^i \cdot \left(1 - \frac{\lambda_k}{\sigma_l^j} \cdot |x_k^i - x_l^j| \right) \right],$$

where $\sigma_l^j \in \Sigma_l \equiv [\underline{\sigma}_l, \overline{\sigma}_l]$, with $\underline{\sigma}_l > 1$, measures the attractiveness of agent j and where $\lambda_k \in [0, 1]$ parametrizes the importance that the side-k agents assign to the locations of the side-l agents. If

$v_k^i < 1$, agent i derives a negative utility from being matched to any agent j from side l. In turn, if $v_k^i > 1$, agent i derives a positive utility from being matched to agent j from side l if and only if, given j's location, j's personal traits σ_l^j make him sufficiently attractive, that is, if and only if $\frac{\sigma_l^j}{\lambda_k |x_k^i - x_l^j|} \geq \frac{v_k^i}{v_k^i - 1}$. For example, in the cable TV application, viewers (on side k) are certainly heterogeneous in their tastes for channels, and therefore $\lambda_k > 0$. In contrast, channels (on side l) are sometimes best viewed as homogenous (to a first approximation) in the profits they derive from reaching viewers, in which case $\lambda_l = 0$. \\\\

The type $\theta_k^i = (x_k^i, v_k^i, \sigma_k^i)$ of each agent $i \in [0,1]$ from each side $k = A, B$ is an independent draw from the distribution F_k with support Θ_k. We assume that F_k is absolutely continuous with respect to the Lebesgue measure. We denote by F_k^z (with density f_k^z) the marginal distribution of F_k with respect to the z dimension, by $F_k^{z,q}$ the marginal distribution of F with respect to the z and q dimensions, by $F_k^z(\cdot|q,r)$ the conditional distribution of the z dimension given the other two dimensions (with density $f_k^z(\cdot|q,r)$), and by $F_k^z(\cdot|q)$ the conditional distribution of the z dimension given the q dimension, for $z, q, r = v, x, \sigma$, $z \neq q, r$.

The total utility that each agent $i \in [0,1]$ from each side $k = A, B$ with type $\theta_k^i = (x_k^i, v_k^i, \sigma_k^i)$ obtains from being matched, at a price p, to any (Lebesgue measurable) set $\mathbf{s}$ of agents from side $l \neq k$ with type profile $(\theta_l^j)_{j \in \mathbf{s}}$ is given by

$$\pi_k^i(\mathbf{s}, p; \theta) = \int_{\mathbf{s}} u_k\left(v_k^i, |x_k^i - x_l^j|, \sigma_l^j\right) d\lambda(j) - p. \tag{1}$$

where $\theta \equiv (\theta_k^i)_{k=A,B}^{i \in [0,1]}$ and where $\lambda(\cdot)$ is the Lebesgue measure on $[0,1]$.

Remarks. The above specification assumes that the utility that each agent $i \in [0,1]$ from each side $k \in \{A, B\}$ obtains from each individual match is independent of who else the agent is matched to. Our result below about the optimality of (location-specific) threshold rules extends to more general payoffs of the form $\pi_k^i(\mathbf{s}, p; \theta) \equiv g_k(\langle \mathbf{s}, \theta \rangle_i) - p$, where $\langle \mathbf{s}, \theta \rangle_i \equiv \int_{\mathbf{s}} u_k\left(v_k^i, |x_k^i - x_l^j|, \sigma_l^j\right) d\lambda(j)$ and where $g_k(\cdot)$ is an increasing and weakly concave function that captures diminishing marginal utility for match quality. The characterization of the optimal thresholds is, however, more convoluted than in the case where $g_k(\cdot)$ is linear, as assumed in (1). On the other hand, it is important that $u_k\left(v_k^i, |x_k^i - x_l^j|, \sigma_l^j\right)$ is allowed to be nonlinear in its arguments; the special case where u_k is linear is uninteresting, for, in this case, the optimal thresholds can be shown to be location-independent.

As anticipated above, the payoff specification in (1) clearly accommodates the possibility that two agents from the same side may disagree on the relative attractiveness of any two agents from the opposite side even when both former agents like (or, alternatively, dislike) interacting with such latter agents. It also accommodates the possibility that the same agent may derive a positive utility from being matched to certain agents, while a negative utility from being matched to others. For example, while a viewer may derive a positive utility from adding to his package a news channel with low advertising, he may derive a negative utility from adding a channel that broadcasts primarily movies with a high degree of violence and nudity, or with a high volume of advertising.

Finally, note that allowing for a second dimension of attractiveness (here parametrized by the personal traits σ) in addition to location, adds flexibility to the model. To see this, consider, for example, the design of online advertising packages. Let web browsers belong to side k and advertisers to side l, and interpret $X_k = X_l = [0,1]$ as the spectrum of possible interests shared by browsers and advertisers. In the absence of a second dimension of attractiveness, the following property holds. Suppose browser i's location is x_k^i and advertisers j and $\hat{\jmath}$ have the same vertical dimension (i.e., $v_l^j = v_l^{\hat{\jmath}}$) but different locations (i.e., $x_l^j \neq x_l^{\hat{\jmath}}$). Then if browser i prefers j to $\hat{\jmath}$, then advertiser j's profits from reaching browser i are (weakly) higher than advertiser $\hat{\jmath}$'s. This property may be reasonable in certain applications, but is clearly restrictive in others.

Matching Mechanisms

Appealing to the Revelation Principle, we focus on 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 $\theta \equiv (\theta_k^i)_{k=A,B}^{i\in[0,1]}$, $\hat{\mathbf{s}}_k^i(\theta)$ represents the set of agents from side $l \neq k$ that are matched to agent i from side k, whereas $\hat{p}_k^i(\theta)$ denotes the payment made by agent i to the platform (i.e., to the match maker).[8]

A matching rule is feasible if and only if the following *reciprocity condition* is satisfied: 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:

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

Because there is no aggregate uncertainty and because individual identities are irrelevant for payoffs (the latter depend only on the composition on the matching sets), without any loss of optimality, we 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 by all agents (whose distribution coincides with F_k on each side k). 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 thus 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 $l \neq k$ included in the matching set of any agent from side k reporting type θ_k. Note that $p_k(\cdot)$ maps Θ_k into $\mathbb{R}$, whereas $\mathbf{s}_k(\cdot)$ maps Θ_k into the Borel sigma algebra over Θ_l.

Denote by $\hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$ the payoff that type θ_k obtains when reporting type $\hat{\theta}_k$ 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 $\theta_k \in \Theta_k$, $k = A, B$, and is incentive compatible (IC)

[8] 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).

if $\Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$ for all $\theta_k, \hat{\theta}_k \in \Theta_k,\ k = A, B$. A matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ is *implementable* if there exists a payment rule $\{p_k(\cdot)\}_{k=A,B}$ such that the mechanism $M = \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ is individually rational and incentive compatible.[9]

Efficiency and Profit Maximization

Because there is no aggregate uncertainty, for any given type profile θ, the welfare generated by the mechanism M is given by

$$\Omega^W(M) = \sum_{k=A,B} \int_{\Theta_k} \int_{\mathbf{s}_k(\theta_k)} u_k\left(v_k^i, |x_k^i - x_l^j|, \sigma_l^j\right) dF_j(\theta_j) dF_k(\theta_k), \tag{3}$$

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

$$\Omega^P(M) = \sum_{k=A,B} \int_{\Theta_k} p_k(\theta_k) dF_k(\theta_k). \tag{4}$$

A mechanism M^W is efficient if it maximizes $\Omega^W(M)$ among all mechanisms that are individually rational, incentive compatible, and satisfy the **reciprocity condition**

$$\theta_l \in \mathbf{s}_k(\theta_k) \ \Rightarrow\ \theta_k \in \mathbf{s}_l(\theta_l). \tag{5}$$

Analogously, a mechanism M^P is profit-maximizing if it maximizes $\Omega^P(M)$ among all mechanisms that are individually rational, incentive compatible, and satisfy the above reciprocity condition.

Finally 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 $\mathbf{s}_k(\cdot)$.

3 Optimal Matching Rules

Our first result below shows that, under certain conditions, the optimal matching rules have a fairly natural structure. We start by describing these conditions. Let $\mathbf{1}_h$ be a dummy variable, taking value 1 if $h = P$ and zero if $h = W$.

Condition 1 ***[LR] Location Regularity:*** *For any $k, l \in \{A, B\}$, $l \neq k$, any $(\theta_k, \theta_l) \in \Theta_k \times \Theta_l$, the virtual values*

$$\varphi_k^h(\theta_k, \theta_l) \equiv u_k(v_k, |x_k - x_l|, \sigma_l) - \mathbf{1}_h \cdot \frac{1 - F_k^v(v_k|x_k, \sigma_k)}{f_k^v(v_k|x_k, \sigma_k)} \cdot \frac{\partial u_k}{\partial v}(v_k, |x_k - x_l|, \sigma_l)$$

are continuous and nondecreasing in v_k, $h = W, P$.

[9] 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 the entire surplus by using payments similar to those in Cremer and McLean (1988) — see also Mezzetti (2007).

Condition 2 ***[I$_k$] Independence on side*** $k \in \{A, B\}$*: for any* $(x_k, v_k, \sigma_k) \in X_k \times V_k \times \Sigma_k$, $F_k(v_k, x_k, \sigma_k) = F_k^{x,\sigma}(x_k, \sigma_k) \cdot F_k^v(v_k)$.

Condition 3 ***[S$_k$] Symmetry on side*** $k \in \{A, B\}$*: for any* $(x_k, v_k, \sigma_k) \in X_k \times V_k \times \Sigma_k$, $F_k(v_k, x_k, \sigma_k) = x_k \cdot F_k^\sigma(\sigma_k) \cdot F_k^v(v_k)$.

Condition 4 ***[MS] Matching Supermodularity****: The match value functions* u_k *are (weakly) submodular in* (v_k, σ_l) *and (weakly) supermodular in* $(v_k, |x_k - x_l|)$, $k, l = A, B$, $l \neq k$.

Condition LR extends the usual Myerson regularity condition to the conditional distribution of v_k, given the locations (x_k, x_l) and the traits (σ_k, σ_l).

Condition I$_k$ requires that the vertical parameters v_k be independently distributed from the location and trait parameters (x_k, σ_k). In the cable TV application, this condition implies that knowing a viewer's "bliss point"(i.e., his preferred channel profile) and personal traits σ_k carries little information about the overall importance that the viewer assigns to watching cable TV.

Condition S$_k$ strengthens the previous independence condition by further requiring that locations be uniformly distributed over $[0, 1]$, as typically assumed in models of horizontal differentiation, and independent of personal traits. As shown below, this assumption simplifies the analysis by guaranteeing that the relevant incentive-compatibility constraints are the ones pertaining to the vertical parameters, v_k.

Finally, Condition MS implies two properties. For any given σ_l, agents from side k who value more interacting with agents from the other side (i.e., whose v_k is higher) suffer less from an increase in the distance $|x_k - x_l|$. Second, for any given distance $|x_k - x_l|$, agents from side k whose v_k is higher benefit less from an increase in the partner's attractiveness σ_l. In the cable TV example, this property means that those viewers who, in general, are keener on cable TV are also those who are more likely to watch channels whose profile is distant from their bliss point and whose attractiveness is lower.

Now let $\triangle_k^h : \Theta_k \times \Theta_l \to \mathbb{R}$ denote the functions defined by

$$\triangle_k^h(\theta_k, \theta_l) \equiv \varphi_k^h(\theta_k, \theta_l) + \varphi_l^h(\theta_k, \theta_l)$$

for $k, l = A, B$, $l \neq k$. Note that $\triangle_A^h(\theta_A, \theta_B) = \triangle_B^h(\theta_B, \theta_A)$ represents the marginal effect on the platform's objective (be it welfare, $h = W$, or profit, $h = P$) of linking types θ_A and θ_B. Let $v \equiv (v_A, v_B) \in V_A \times V_B$, $x \equiv (x_A, x_B) \in X_A \times X_B$, and $\sigma \equiv (\sigma_A, \sigma_B) \in \Sigma_A \times \Sigma_B$. We then have the following result.

Proposition 1 ***(optimal matching rules)*** *Assume traits* σ *are publicly observable and that Condition LR holds. In addition, suppose one of the following three sets of conditions holds: (a) locations are publicly observable on both sides; (b) locations are private information on side* $k \in \{A, B\}$ *and publicly observable on side* $l \neq k$*, and Conditions* I_k*, and* S_l *hold; (c) locations are private information on both sides and Conditions* S_k *hold,* $k = A, B$.

Then there exist functions $t_k^h(\cdot)$ such that the h-optimal matching rules $\mathbf{s}_k^h(\cdot)$ have the following threshold structure, $k = A, B,\ h = W, P$:

$$\mathbf{s}_k^h(\theta_k) = \left\{(x_l, v_l, \sigma_l) \in \Theta_l : \ v_l > t_k^h(v_k, x, \sigma)\right\}.$$

The threshold functions $t_k^h(\cdot)$ are such that for $k, l = A, B,\ l \neq k,\ h = W, P$:

1. *$t_k^h(v_k, x, \sigma) = \underline{v}_l$ if $\triangle_k^h(\theta_k, (x_l, \underline{v}_l, \sigma_l)) > 0$,*
2. *$t_k^h(v_k, x, \sigma) = \bar{v}_l$ if $\triangle_k^h(\theta_k, (x_l, \bar{v}_l, \sigma_l)) < 0$,*
3. *$t_k^h(v_k, x, \sigma)$ is the unique solution to*

$$\triangle_k^h\left(\theta_k, (x_l, t_k^h(v_k, x, \sigma), \sigma_l)\right) = 0 \tag{6}$$

if $\triangle_k^h(\theta_k, (x_l, \underline{v}_l, \sigma_l)) < 0 < \triangle_k^h(\theta_k, (x_l, \bar{v}_l, \sigma_l))$. In this case, the threshold $t_k^h(v_k, x, \sigma)$ is locally strictly decreasing in v_k. When either (i) $h = W$, or (ii) $h = P$ and Condition MS holds, the threshold $t_k^h(v_k, x, \sigma)$ is also locally weakly increasing in the distance $|x_k - x_l|$ and weakly decreasing in $\sigma \equiv (\sigma_A, \sigma_B)$.

When the conditions in Proposition 1 are satisfied, the optimal matching rules thus have a *threshold structure*: agents with a low value for matching are matched only to those agents from the opposite side whose value for matching is sufficiently high. The thresholds are location- and trait-specific: the optimal matching rules thus actively discriminate on the basis of mutual attractiveness (as captured by the joint traits and locations of any two agents).

Consider the problem of welfare maximization (the problem of profit maximization is analogous). Take a type $\theta_k = (x_k, v_k, \sigma_k)$ from side k and a pair (x_l, σ_l) of location and traits from side l such that $u_k(v_k, |x_k - x_l|, \sigma_l) > 0$. This last condition makes type $\theta_k = (x_k, v_k, \sigma_k)$ a *consumer* of (x_l, σ_l)-agents. Type θ_k's matching set naturally includes all agents with traits σ_l located at x_l who like interacting with side-k agents located at x_k and with traits σ_k, i.e., for whom $u_l(v_l, |x_k - x_l|, \sigma_k) \geq 0$. It also includes some agents θ_l located at x_l with traits σ_l who dislike interacting with side-k agents located at x_k and with traits σ_k, provided that the cross-side effects on welfare generated by linking types θ_k and θ_l are positive, i.e., provided that $\triangle_k^W((x_k, v_k, \sigma_k), (x_l, v_l, \sigma_l)) > 0$. The latter agents included in type θ_k's matching set play the role of *inputs* in the W-optimal matching rule.

The role of type $\theta_k = (x_k, v_k)$ as a consumer or as an input naturally varies across locations and traits. While type θ_k is a consumer of side-l agents with traits σ_l and location x_l, he may be an input for those side-l agents $\hat{\theta}_l = (\hat{x}_l, \hat{v}_l, \hat{\sigma}_l)$ for whom $u_k(v_k, |x_k - \hat{x}_l|, \hat{\sigma}_l) < 0$ and $\triangle_l^W(\hat{\theta}_l, \theta_k) > 0$. That is, the separation of agents between consumers and inputs depends on the joint locations and traits of any two agents.

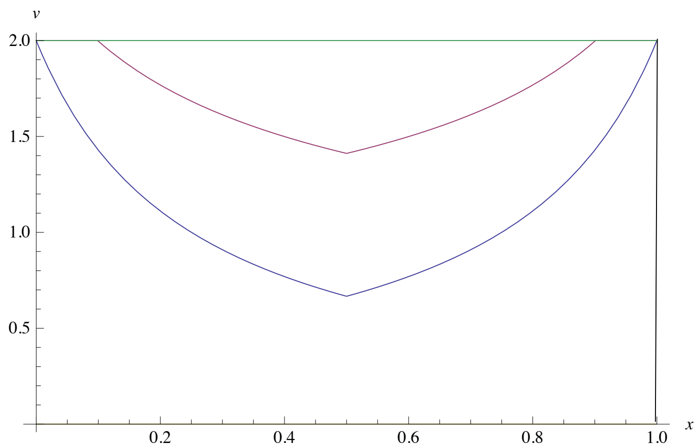
As established in the proposition, the threshold $t_k^W(v_k, x, \sigma)$ are weakly increasing in the distance $|x_k - x_l|$ and decreasing in the partner's quality σ_l. To understand why, pick again a type $\theta_k = (x_k, v_k, \sigma_k)$ from side k and a location-traits pair (x_l, σ_l) from side l such that $u_k(v_k, |x_k - x_l|, \sigma_l) >$

0. Because *consumer values* $u_k\left(v_k, |x_k - x_l|, \sigma_l\right) > 0$ go down as the partners' attractiveness decreases (that is, as the distance $|x_k - x_l|$ increases, or as the quality σ_l decreases), the *input costs* $u_l\left(t_k^W(v_k, x, \sigma), |x_k - x_l|, \sigma_k\right) < 0$ at the threshold $t_k^W(v_k, x, \sigma)$ have to go down as well. In turn, this means that, as the distance $|x_k - x_l|$ increases, or the quality σ_l decreases, the marginal x_l-agent $t_k^W(v_k, x, \sigma)$ in type θ_k's matching set must have a higher value for matching v_l. By the same token, the threshold $t_k^W(v_k, x, \sigma)$ decreases as an agent's own attractiveness σ_k increases, for again this implies a reduction in the agent's input cost.

This logic extends to profit maximization under the supermodularity condition MS, which controls for how informational rents vary with distance and personal traits.

The role of Condition LR is to guarantee that, for any reported location x_k and traits σ_k, the size of the matching sets increase in the value for matching v_k, as required by incentive compatibility. In turn, the role of conditions S_k and I_k, $k = A, B$, is to ensure that, under the optimal rules, the only binding incentive compatibility constraints are those that pertain to the vertical dimension, v_k. Obviously, these conditions can be dispensed with when not only the traits but also the locations are public on both sides, for in this case the only dimension that the agents can misreport is the vertical one. To understand the role of these two conditions, consider first the case where locations are private on side k but public on side l. Conditions I_k and S_l, together with Condition (6), then imply that, for any (v_k, σ_k), the matching sets associated with different reports about the horizontal dimension x_k are *parallel translations* of one another. As a consequence, misreporting the location x_k is never profitable, irrespective of whether or not the agent reports truthfully the vertical dimension. This is because both the prices $p_k(x_k, v_k, \sigma_k)$ and the quality-adjusted sizes $\int_{\theta_l \in \mathbf{s}_k(x_k, v_k, \sigma_k)} \sigma_l dF_l(\theta_l)$ of the matching sets are invariant in the reported x_k-dimension. Therefore, misreporting the bliss point (i.e., the x_k location) negatively affects the composition of the matching set, but does not change its price, or its quality-adjusted size, and hence is unprofitable. Conditions I_k and S_l then imply that an agent who prefers reporting truthfully to lying about the v_k dimension also prefers reporting truthfully to any arbitrary lie.

By the same logic, when locations are private on both sides, one has to replace Condition I_k with the stronger Condition S_k on side k to guarantee that side-l agents also find it optimal to report truthfully. We conjecture that the results in Proposition 1 extend to distributions F_k that are sufficiently "close"to the ones covered by the Symmetry and Independence conditions, but did not attempt to establish this formally.



The welfare (solid) and profit-maximizing (dashed) matching thresholds under log utility preferences, when $\lambda_A = \frac{4}{3}$, $\lambda_B = 0$, $\underline{\sigma}_l = \bar{\sigma}_l = 1$, $l = A, B$, and agents from both sides have valuations and locations independently and uniformly distributed over $[0, 2]$ and $[0, 1]$, respectively.

The following example illustrates the structure of the optimal matching rules when preferences are as in Example 1.

Example 2 (**optimal matching rules for log utility**) Suppose that preferences are as in Example 1, that traits are publicly observable, that locations are private on side $k \in \{A, B\}$ and public on side $l \neq k$, and that Conditions I_k, and S_l hold.[10] The welfare-maximizing matching rule is described by the following threshold function (at any point where $t_k^W(v_k, x, \sigma) \in (\underline{v}_l, \bar{v}_l))$:

$$t_k^W(v_k, x, \sigma) = \frac{1}{v_k} \cdot \frac{1}{\left(1 - \frac{\lambda_k}{\sigma_l} \cdot |x_k - x_l|\right) \cdot \left(1 - \frac{\lambda_l}{\sigma_k} \cdot |x_k - x_l|\right)}.$$

In turn, the profit-maximizing matching rule is described by the following threshold function (at any point where $t_k^P(v_k, x, \sigma) \in (\underline{v}_l, \bar{v}_l)$):

$$\frac{t_k^P(v_k, x, \sigma)}{\exp\left\{\frac{1 - F_l^v(t_k^P(v_k, x, \sigma))}{f_l^v(t_k^P(v_k, x, \sigma)) t_k^P(v_k, x, \sigma)}\right\}} = \frac{\exp\left\{\frac{1 - F_k^v(v_k)}{f_k^v(v_k) \cdot v_k}\right\}}{v_k} \cdot \frac{1}{\left(1 - \frac{\lambda_k}{\sigma_l} \cdot |x_k - x_l|\right) \cdot \left(1 - \frac{\lambda_l}{\sigma_k} \cdot |x_k - x_l|\right)}.$$

The threshold functions $t_k^W(v_k, x, \sigma)$ and $t_k^P(v_k, x, \sigma)$ are illustrated in Figure 4 for the special case where $\underline{\sigma}_l = \bar{\sigma}_l = 1$ (with the σ dropped to ease the exposition). \\

We conclude this section by showing that conclusions similar to those in the previous proposition obtain if one replaces the assumption that the traits σ are observable with the assumption that payoffs are separable in the traits and then add a second regularity condition that guarantees that the buyer-value of each agent increases with the vertical parameter v faster than the input value.

[10] Note that Condition MS is implied by the log-utility specification.

Condition 5 ***[PS] Payoffs Separable in*** σ*: For* $k = A, B$*, there exist functions* $g_k : \mathbb{R} \to \mathbb{R}$ *and* $\tilde{u}_k : \mathbb{R}^2 \to \mathbb{R}$*, with* $\tilde{u}_k$ *Lipschitz continuous, strictly increasing, continuously differentiable, and weakly concave in* v_k^i*, and weakly decreasing in* $|x_k - x_l|$*, and with* g_k *positive and non-decreasing, such that for all* $(v_k, x_k, x_l, \sigma_l)$*,*

$$u_k(v_k, |x_k - x_l|, \sigma_l) = \tilde{u}_k(v_k, |x_k - x_l|) \cdot g_k(\sigma_l).$$

Condition 6 (MR) ***Matching regularity****: Payoffs are separable in the sense of Condition PS and, in addition, the functions*

$$\psi_k^h(v_k|x_k, x_l) \equiv \frac{\tilde{u}_k(v_k, |x_k - x_l|) - \mathbf{1}_h \cdot \frac{1 - F_k^v(v_k|x_k)}{f_k^v(v_k|x_k)} \cdot \frac{\partial \tilde{u}_k}{\partial v}(v_k, |x_k - x_l|)}{\displaystyle\int g_l(\sigma_k) dF_k^\sigma(\sigma_k|x_k, v_k)}$$

are strictly increasing in v_k*, for all* $(x_k, x_l) \in X_k \times X_l$*,* $k, l = A, B$*,* $l \neq k$*.*

For all $(x_k, v_k) \in X_k \times V_k$, $(x_l, v_l) \in X_l \times V_l$, $k, l = A, B$, $l \neq k$, then let

$$\begin{aligned}\bar{\Delta}_k^h(v_k, v_l; x_k, x_l) \equiv & \int g_k(\sigma_l) dF_l^\sigma(\sigma_l|x_l, v_l) \left[\tilde{u}_k(v_k, |x_k - x_l|) - \mathbf{1}_h \cdot \frac{1 - F_k^v(v_k|x_k)}{f_k^v(v_k|x_k)} \cdot \frac{\partial \tilde{u}_k}{\partial v}(v_k, |x_k - x_l|)\right] \\ & + \int g_l(\sigma_k) dF_k^\sigma(\sigma_k|x_k, v_k) \left[\tilde{u}_l(v_l, |x_k - x_l|) - \mathbf{1}_h \cdot \frac{1 - F_l^v(v_l|x_l)}{f_l^v(v_l|x_l)} \cdot \frac{\partial \tilde{u}_l}{\partial v}(v_l, |x_k - x_l|)\right]\end{aligned}$$

denote the *expected* value of matching agents with characteristics (v_k, x_k) and (v_l, x_l), where the expectation is with respect to the traits σ_k and σ_l. The following result then extends the conclusions in the previous proposition to the case of unobservable traits.

Proposition 2 ***(optimal rules when personal traits are private information)*** *Assume that traits are the agents' private information and that, in addition to Conditions LR and PS, one of the following three sets of conditions holds: (a) locations are publicly observable on both sides and, for each* $x_k \in X_k$*, either* (v_k, σ_k) *are (weakly) positively affiliated,* $k = A, B$*, or* (v_k, σ_k) *are (weakly) negatively affiliated,* $k = A, B$*; (b) locations are private information on side* $k \in \{A, B\}$ *and public on side* $l \neq k$*, and Conditions* I_k*, and* S_l *hold; (c) locations are private on both sides and Conditions* S_k *hold,* $k = A, B$*.*

Then there exist functions $t_k^h(\cdot)$ *sich that the h-optimal matching rules* $\mathbf{s}_k^h(\cdot)$ *have the following threshold structure,* $k = A, B$*,* $h = W, P$ *:*

$$\mathbf{s}_k^h(\theta_k) = \Big\{(x_l, v_l, \sigma_l) \in \Theta_l : \ v_l > t_k^h(v_k, x)\Big\}.$$

If, in addition, Condition PS is replaced with the stronger Condition MR, then for any $v_k \in V_k$*,* $x = (x_k, x_l) \in X_k \times X_l$*,* $k, l = A, B$*,* $l \neq k$*,* $h = W, P$ *:*

1. $t_k^h(v_k, x) = \underline{v}_l$ *if* $\bar{\Delta}_k^h(v_k, \underline{v}_l; x_k, x_l) > 0$*,*

2. $t_k^h(v_k, x) = \bar{v}_l$ *if* $\bar{\Delta}_k^h(v_k, \bar{v}_l; x_k, x_l) < 0$*,*

3. $t_k^h(v_k, x)$ *is the unique solution to*

$$\bar{\Delta}_k^h \left(v_k, t_k^h(v_k, x); x \right) = 0$$

if $\bar{\Delta}_k^h (v_k, \underline{v}_l; x_k, x_l) < 0 < \bar{\Delta}_k^h (v_k, \bar{v}_l; x_k, x_l)$. *In this case, the threshold* $t_k^h(v_k, x)$ *is locally strictly decreasing in* v_k. *When* $F_k^\sigma(\cdot|x_k, v_k)$ *is invariant in* (x_k, v_k), $k = A, B$, *and either (i)* $h = W$, *or (ii)* $h = P$ *and* $\tilde{u}_k$ *are (weakly) supermodular in* $(v_k, |x_k - x_l|)$, *the threshold* $t_k^h(v_k, x)$ *is also locally weakly increasing in the distance* $|x_k - x_l|$.

In addition to the conditions in Proposition 1, there are two key assumptions that help retaining the optimality of threshold rules when personal traits are private. The first one is the separability of the payoffs in the personal traits (that is, Condition PS). The second one is the assumption that the affiliation between the values for matching v and the personal traits is of the same sign on both sides of the market. Along with the assumption that g is positive and non-decreasing, these assumptions imply that the most effective way of cross subsidizing the matches is to disregard the personal traits and follow a cut-off rule similar to the one in the previous proposition. The proof of this result follows from the combination of arguments similar to those that establish Proposition 1 above with arguments similar to those that establish Proposition 1 in Gomes and Pavan (2015) for settings in which preferences are only vertically differentiated.

We conjecture that the result that optimal rules do not depend on personal traits σ when the latter are the agents' private information extends to more general settings. In fact, note that, because the traits determine attractiveness but not the agents' own preferences, the only way a rule can discriminate along the traits dimension is to make the agents indifferent as to what traits to report. While, in principle, it is possible to play with the composition of the matching sets so as to guarantee such indifference, we conjectures that the cases where it is strictly optimal to do so are knife-edge. Relying on such indifference may also introduce into the mechanism an undesirable form of fragility, stemming from the possibility that the agents, when indifferent, play strategies other than the equilibrium ones, thus imposing losses on the platform. Formalizing these considerations is beyond the scope of the paper. Assuming that payoffs satisfy the separability property of Condition PS is an overkiller, but it guarantees in a simple way that optimal rules do not discriminate on personal traits that are irrelevant for preferences when the latter are not observable by the platform.

Another reason why optimal rules may be invariant in the traits even when condition PS is violated is that the agents themselves may be unable to observe such traits. This is likely to be the case in markets in which experimentation plays a key role, where the agents are in the process of learning their attractiveness through pairwise interactions, as in Fershtman and Pavan (2015).

Finally note that, under the specification of Condition PS, the value that type θ_k obtains from being matched to type θ_l need not increase with σ_l. Because of this possibility, under Condition PS, σ_l should be interpreted as a combination of personal traits that are responsible for the "*salience*" of type θ_l, as opposed to his attractiveness. However, observe that what matters for the results in

Proposition 2 is only that the sign of $u_k(v_k, |x_k - x_l|, \sigma_l)$ is independent of σ_l. In particular, the same conclusions as in Proposition 2 obtain if one replaces Condition PS with the following condition.

Condition 7 *[PSI] Payoffs of Sign Independent of* σ: *For* $k = A, B$, *there exist functions* $\tilde{u}_k : \mathbb{R}^2 \to \mathbb{R}$ *such that for all* $(v_k, x_k, x_l, \sigma_l)$, *the following are true: (a)* $sign\{u_k(v_k, |x_k - x_l|, \sigma_l)\} = sign\{\tilde{u}_k(v_k, |x_k - x_l|)\}$ *and (b)* u_k *is increasing in* σ_l *if* $\tilde{u}_k(v_k, |x_k - x_l|) \geq 0$ *and decreasing in* σ_l *if* $\tilde{u}_k(v_k, |x_k - x_l|) < 0$.

An example of a payoff function satisfying Condition PSI is

$$u_k(v_k, |x_k - x_l|, \sigma_l) = \tilde{u}_k(v_k, |x_k - x_l|) + sign\{\tilde{u}_k(v_k, |x_k - x_l|)\} \cdot g_k(\sigma_l)^{sign\{\tilde{u}_k(v_k, |x_k - x_l|)\}}$$

with $g_k(\cdot)$ non-decreasing and satisfying $g_k(\sigma_l) > 1$ for all σ_l.

4 Implementation: the Lerner-Wilson Pricing Formula for Matching Plans

We now show how the platform can implement the optimal matching rules by offering agents on both sides of the market menus of matching plans. We describe the platform's pricing strategies in the context of the cable TV application. To ease the exposition, we also abastract from personal traits by letting $\underline{\sigma}_k = \bar{\sigma}_k = 1$, $k = A, B$, and then dropping σ from all formulas.

The platform offers to each viewer a menu of *packages* (sometimes also referred to as *plans*)

$$\mathcal{M}_k \equiv \{\mathcal{P}(x_k) : x_k \in [0, 1]\}_k .$$

Each package $\mathcal{P}(x_k) = (\beta_k(x_k), P_k(x_k), \rho_k(\cdot, \cdot; x_k))$ is indexed by its *category* $x_k \in [0, 1]$ (sports, news, movies, etc.) which coincides with the profile of channels that are present in the largest amount. In addition, a package specifies a *baseline price* $P_k(x_k)$ and a *baseline configuration*

$$\beta_k(x_k) = \bigcup\nolimits_{x_l \in [0,1]} \underline{q}_{x_l}(x_k)$$

where $\underline{q}_{x_l}(x_k)$ denotes the quantity of x_l-channels included in the x_k-package. Finally, each package $\mathcal{P}(x_k)$ specifies the (possibly non-linear) price

$$\rho_k(q, x_l; x_k)$$

that the viewer has to pay to bring the total number of x_l-channels in the package to $q \in [0, 1]$, for every category $x_l \in [0, 1]$. Given the menu $\mathcal{M}_k$, each viewer i from side k is offered the possibility to choose a package from the menu and then personalize it by adding (or substracting) channels. Denoting by $q_{x_l}(x_k)$ the quantity of x_l-channels selected by a viewer who chooses the package $\mathcal{P}(x_k)$, we then have that the total price paid by the viewer is given by

$$P(x_k) + \int_0^1 \rho_k(q_{x_l}(x_k), x_l; x_k) dx_l.$$

Next, consider the channels' side. Here too the platform offers to each channel a menu of pricing plans, where each plan is again indexed by the channel's category. There are different ways one can describe such plans. By symmetry with the viewers' side, a plan could specify the type of viewers present in the package in the largest amount, along with a baseline price and a collection of additional prices that the channel has to pay to increase (or reduce) the number of viewers from each category. Alternatively, and more in line with actual practices, a plan can be described by the price the channel has to pay to be included in each of the viewers' packages. Because these distinctions are inconsequential to our results, we will not further pursue them and instead focus our discussion below on the viewers' side.

Below we show how the results from Proposition 1 can be used to relate the marginal prices in each package to the (category-specific) demand elasticities on each of the two sides. For the proposed menus of matching plans to implement the allocations and payments of the corresponding h-optimal direct revelation mechanism, we will let the baseline price $P_k^h(x_k)$ and the baseline configuration $\beta_k^h(x_k)$ coincide, respectively, with the equilibrium price and with the equilibrium matching set of any x_k-agent with the lowest value $\underline{v}_k$ for matching, in the corresponding h-optimal direct revelation mechanism. That is, we let

$$P_k^h(x_k) = p_k^h(x_k, \underline{v}_k) \tag{7}$$

and, for any $x_l \in [0,1]$, we let $\underline{q}_{x_l}^h(x_k) \equiv 1 - F_l^v(t_k^h(v_k, x_k, x_l))$ so that

$$\beta_k^h(x_k) = \mathbf{s}_k^h(x_k, \underline{v}_k). \tag{8}$$

Next, consider the price $\rho_k^h(q, x_l; x_k)$ that a viewer selecting the package $\mathcal{P}_k^h(x_k)$ has to pay to bring the number of x_l-channels to q. At any point of differentiability of the tariff $\rho_k^h(\cdot, x_l; x_k)$, we denote by $\frac{d\rho_k^h}{dq}(q, x_l; x_k)$ the marginal price for the q unit of x_l-channels under the package $\mathcal{P}_k^h(x_k)$.

Now consider the problem of a viewer with type $\theta^k = (x_k, v_k)$ who selects the plan $\mathcal{P}_k^h(x_k)$. His *individual demand* for x_l-channels then satisfies

$$q_k^h(x_l; \theta_k) \in \arg\max_{q \in [0,1]} \left\{ u_k(v_k, |x_k - x_l|) \cdot q - \rho_k^h(q, x_l; x_k) \right\}.$$

At any point $q_k^h(x_l; \theta_k)$ of differentiability of the tariff $\rho_k^h(\cdot, x_l; x_k)$, the following first-order condition must hold:

$$u_k(v_k, |x_k - x_l|) - \frac{d\rho_k^h}{dq}(q_k^h(x_l; \theta_k), x_l; x_k) = 0. \tag{9}$$

Given the monotonicity of the individual demands in v_k, the *aggregate demand* for the q_k unit of x_l-agents by the x_k-agents, at the *marginal price* $\frac{d\rho_k^h}{dq_k}(q_k, x_l, x_k)$ — in the cable TV application, the measure of viewers who demand q_k or more x_l-channels after selecting the package $\mathcal{P}_k^h(x_k)$ — is given by

$$D_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l; x_k), x_l; x_k\right) \equiv 1 - F_k^v\left(v_k(q_k, x_l; x_k); x_k\right),$$

where $v_k(q_k, x_l; x_k)$ solves $u_k(v_k(q_k, x_l; x_k), |x_k - x_l|) - \frac{d\rho_k^h}{dq}(q_k, x_l; x_k) = 0$.

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

$$\varepsilon_k\left(q_k,\frac{d\rho_k^h}{dq}(q_k,x_l;x_k),x_l;x_k\right)\equiv-\frac{\partial D_k\left(q_k,\frac{d\rho_k^h}{dq}(q_k,x_l;x_k),x_l;x_k\right)}{\partial\left(\frac{d\rho_k^h}{dq}(q_k,x_l;x_k)\right)}\cdot\frac{\frac{d\rho_k^h}{dq}(q_k,x_l;x_k)}{D_k\left(q_k,\frac{d\rho_k^h}{dq}(q_k,x_l;x_k),x_l;x_k\right)} \tag{10}$$

$$=\frac{f_k^v\left(v_k(q_k,x_l;x_k);x_k\right)}{1-F_k^v\left(v_k(q_k,x_l;x_k);x_k\right)}\cdot\left[\frac{\partial u_k}{\partial v}\left(v_k(q_k,x_l;x_k),|x_k-x_l|\right)\right]^{-1}\cdot\frac{d\rho_k^h}{dq}(q_k,x_l;x_k).$$

The next proposition uses the results in Proposition (1) to relate the marginal prices under the optimal menu of packages to the (location-specific) elasticities of the aggregate demands from each side of the market.

Proposition 3 ***(Lerner-Wilson formula for matching plans)*** *Assume the conditions in Proposition 1 hold. The h-optimal mechanism can be implemented by offering to each side a menu of matching plans $\mathcal{M}_k^h=\left\{\mathcal{P}_k^h(x_k):x_k\in[0,1]\right\}_{k=A,B}$. Each plan $\mathcal{P}_k^h(x_k)$ is defined by a baseline price $P_k^h(x_k)$ and a baseline matching set $\beta_k^h(x_k)$ given by (7) and (8) respectively, along with a collection of non-linear prices $\rho_k^h(q,x_l;x_k)$ that any agent from side k selecting the $\mathcal{P}_k^h(x_k)$ plan has to pay to bring the quantity of x_l-agents in his matching set to q. The h-optimal price schedules $\rho_k^h(\cdot,x_l;x_k)$ are differentiable and satisfy the following Lerner-Wilson formulas*

$$\underbrace{\frac{d\rho_k^h}{dq}(q_k,x_l;x_k)-\mathbf{1}^h\cdot\frac{\frac{d\rho_k^h}{dq}(q_k,x_l;x_k)}{\varepsilon_k\left(q_k,\frac{d\rho_k^h}{dq}(q_k,x_l;x_k),x_l;x_k\right)}}_{\text{net effect on side-k welfare (alternatively, profits)}} \tag{11}$$

$$+\underbrace{\frac{d\rho_l^h}{dq}(q_l^h(q_k,x_l;x_k),x_k;x_l)-\mathbf{1}^h\cdot\frac{\frac{d\rho_l^h}{dq}(q_l^h(q_k,x_l;x_k),x_k;x_l)}{\varepsilon_l\left(q_l^h(q_k,x_l;x_k),\frac{d\rho_l^h}{dq}(q_l^h(q_k,x_l;x_k),x_k;x_l)\right)}}_{\text{net effect on side-lwelfare (profits)}}=0,$$

where $q_l^h(q_k,x_l;x_k)\equiv D_k\left(q_k,\frac{d\rho_k^h}{dq}(q_k,x_l;x_k),x_l;x_k\right)$ is the aggregate demand for the q_k unit of x_l-agents by x_k-agents at the marginal price $\frac{d\rho_k^h}{dq}(q_k,x_l;x_k)$, $x_k,x_l\in[0,1]$, $k,l\in\{A,B\}$, $l\neq k$.

The intuition for the pricing formulas above is the following: marginal prices are chosen so as to equalize the marginal revenue gains of expanding the number of x_l-agents included in the matching plan of each x_k-agent to the marginal costs of "procuring"the extra x_l-agents from side l, taking into account the threshold rule used by the platform to minimize the cross-subsidizations costs. When preferences exhibit elements of both vertical and horizontal differentiation, as assumed here, these marginal prices are naturally pairwise location-specific; that is, they condition on both the characteristics of the agents added to the matching set as well as the preferences of the agent whose matching set is under consideration. Note that, in principle, the result in the proposition could be tested by collecting data on the prices that platforms charge to viewers to add channels to their packages along with the corresponding demands.

5 Discussion

Second-degree price discrimination is ubiquitous in markets for many-to-many matching. In what follows, we discuss how our results relate to practices in some of such markets.

Cable TV. Cable TV platforms are known to price-discriminate on the viewer side of the market by offering viewers different packages of channels at different prices. What is perhaps less understood is that they also price-discriminate on the channel side by setting transfers that depend on the audience level attained.

As reported by Crawford (2000) and Crawford and Yurukoglu (2012), before the 1990s, technological limitations in the available bandwidth were forcing cable TV providers to offer no more than two packages: a basic one targeted to viewers with a low willingness to pay and consisting primarily of "cheap" channels; and a premium package targeted to viewers with a high willingness to pay and including channels with higher costs per viewer in addition to the channels included in the basic package. This practice can be viewed as consistent with what is predicted by our model. To see this, let the values v be positive on the viewers' side; assume that the traits σ and the values v are independent on the viewers' side and negatively affiliated on the channels' side. That the channels' willingness to pay is negatively affiliated with the quality of their programs may reflect the fact that high-quality channels typically have outside options superior to the low-quality channels and/or have higher bargaining power vis-a-vis the cable providers. Our results then imply that low-willingness-to-pay viewers should be directed toward basic packages consisting primarily of cheap lower-quality channels, while high-willingness-to-pay viewers should be directed toward premium packages which include also higher-quality channels which are, however, less lucrative for the platform.

Advances in digital technology after the 1990s enabled cable TV providers to offer viewers customized packages to better respond to the heterogeneity in viewers' preferences. Many cable TV providers now offer a few (vertically differentiated) plans, and then allow viewers to add (horizontally differentiated) packages such as "sports", "news" and "foreign". For example, in the US, Direct TV offers five vertically differentiated (i.e., nested) English packages, four vertically differentiated Spanish packages, and eight foreign packages. It then allows viewers to add to these packages nine (horizontally differentiated) premium packages, which bundle together channels specialized in movies, sports, news, and games. In addition, viewers can choose among eighteen individual sports channels, specialized in golf, tennis, basketball, and other sports. Finally, viewers can purchase hundreds of individual pay-per-view movies and events and combine them with mobile applications and Internet services. Similar combinations of packages with different degrees of horizontal and vertical differentiation are offered by other providers. While the industry has not reached yet the "extreme"form of customization predicted by our model, recent developments seem to indicate a trend toward the practice of offering a combination of pre-designed packages and customization options in the spirit of what predicted in the paper.

Our model, however, has two important limitations when one applies it to the cable TV industry.

First, it abstracts from competition among providers. A second, related, limitation is that it assumes that the monopolistic platform can make take-it-or leave offers to the channels. In contrast, the empirical analysis of Crawford and Yurukoglu (2012) suggests that large channel conglomerates enjoy nontrivial bargaining power vis-a-vis cable TV providers. Extending the analysis to settings where (i) channels have bargaining power, and/or (ii) there are multiple providers is expected to offer new insights into the bundling practices of the cable TV industry and thus represents a promising line for future research.

Business-to-business platforms. B2B platforms work as brokers matching vendors with procurers for a fee (see, e.g., Lucking-Reiley and Spulber (2001) and Jullien (2012)). Typically, these platforms offer menus that include a "join for free"option along with plans that provide richer matching possibilities at increasing subscription fees. These practices appear broadly consistent with what predicted by our model. In particular, the presence in these menus of "join for free"options reflects cross-subsidization effects similar to those highlighted in the paper. On the other hand, B2B platforms have recently expanded their services to include e-billing and supply-management support. These recent developments open the door to more sophisticated price discriminatory practices that use instruments other than the composition of the matching sets. Extending the analysis to accommodate for such richer instruments represents another interesting line for future research.

In-Print Advertising. Many off-line advertising outlets, such as newspapers and magazines, offer different editions of the same outlet, combining different levels of advertising and content, at different prices (see, e.g., Ambrus and Argenziano (2009) and Kaiser and Wright (2006)). The *Washington Post*, for example, offers a tabloid edition for free, and a regular (paid) edition, with less advertising and more content. Advertisers typically face higher prices to place ads in the regular edition (which attracts readers with a lower tolerance for advertising and higher interest in content) than in the tabloid edition. As a consequence, advertisers with a high willingness to pay to reach readers advertise in both editions, while advertisers with a lower willingness to pay favor the tabloid edition. This structure appears broadly consistent with what predicted by our model, subject to one important qualification. Our model does not consider the possibility that either one or both sides of the market derive utility directly from the product provided by the platform, as in this application where readers derive a positive utility from content in addition to disliking advertising. By considering only the disutility from advertising, our model predicts negative prices on the viewers' side of the market. If one were to add to the model a direct utility for content which is negatively correlated with the readers' tolerance for advertising, the model would then predict a positive price for the regular edition and a lower (possibly zero) price for the tabloid. Introducing a direct utility for the platform's products is likely to add further realism to the model and bring more light to the pricing strategies in media markets.[11]. Another interesting extension would consist in introducing richer forms of heterogeneity in the outside options, in line with those examined in Jullien (2000)

[11] See also Kaiser and Song (2009), and Weyl (2010) for a discussion of how readers' preferences for content might be correlated with their tolerance for advertising.

and Rochet and Stole (2002).

Online Advertising. Several online outlets (such as web portals and online newspapers and magazines) offer different subscription plans with different ratios of advertising and content. The content provided for free is often accompanied by a large amount of advertising. In turn, paying subscribers have access to more content and face a smaller exposure to advertising. These stylized facts are similar to those discussed above for in-print advertising.

More generally, online advertising companies have recently improved their ability to offer targeted advertising, thereby exploring the heterogeneity in interests among web browsers and advertisers. While our model offers a few insights in this direction, we derived our results under the assumption that browsers can use payments to adjust the level of advertising they are willing to endure. Extending the analysis to accommodate the possibility that platforms may face constraints on their ability to use prices on one of the two sides of the market is expected to bring in additional effects that can be relevant for this application.

6 Concluding Remarks

This paper studies many-to-many matching in markets in which agents have vertically and horizontally differentiated preferences. The analysis delivers two main results. First, it identifies primitive conditions under which the optimal matching rules induce negative assortative matching at the margin. As the matching sets expand, the marginal agents from each location are always those with the lowest value for matching. The composition of the pool of marginal agents, however, naturally respects horizontal differences in preferences, with most of the marginal agents coming from "locations"close to the bliss-point of the agents under consideration. We believe that this particular form of location-specific negative assortativeness at the margin is a general property of markets for many-to-many matching.

Second, the optimal matching sets are specified by a simple condition that equalizes the marginal gains in welfare (or, alternatively, in profits) with the cross-subsidization losses in welfare (or, alternatively, in profits) that the platform must incur on the other side of the market. This condition can be used to construct the matching plans and to derive the price schedules that implement the optimal matching rules. Importantly, using our results, one can express the Lerner-Wilson formulas for the marginal prices in terms of observable variables, such as elasticties, prices and quantities.

In future work it would be interesting to extend the analysis to accommodate for "within-side" externalities (e.g., congestion and limited attention) and competition among platforms.[12]

[12] 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.

7 Appendix

Proof of Proposition 1. By familiar envelope arguments, a necessary condition for each type $\theta_k = (x_k, v_k, \sigma_k) \in \Theta_k$, $k = A, B$, to prefer to report truthfully both x_k and v_k to lying with respect to v_k while truthful reporting x_k (recall that σ_k is assumed observable, so the agent does not need to report it) is that payments satisfy the familiar envelope conditions

$$p_k(\theta_k) = \int_{\mathbf{s}_k(\theta_k)} u_k\left(v_k, |x_k - x_l|, \sigma_l\right) dF_l(\theta_l) - \int_{\underline{v}_k}^{v_k} \int_{\mathbf{s}_k(x_k, y, \sigma_k)} \frac{\partial u_k}{\partial v}\left(y, |x_k - x_l|, \sigma_l\right) dF_l(\theta_l) dy \tag{12}$$
$$- \Pi(x_k, \underline{v}_k, \sigma_k | M^h).$$

Plugging the formula above into $\Omega^P(M)$ leads to

$$\Omega^P(M) = \sum_{k=A,B} \int_{\Theta_k} \left\{ \begin{array}{c} \int_{\mathbf{s}_k(\theta_k)} \left[u_k\left(v_k, |x_k - x_l|, \sigma_l\right) - \frac{1 - F_k^v(v_k | x_k, \sigma_k)}{f_k^v(v_k | x_k, \sigma_k)} \cdot \frac{\partial u_k}{\partial v}\left(v_k, |x_k - x_l|, \sigma_l\right)\right] dF_l(\theta_l) \\ -\Pi(x_k, \underline{v}_k, \sigma_k | M^h) \end{array} \right\} dF_k(\theta_k).$$

Because at the optimum $\Pi(x_k, \underline{v}_k, \sigma_k | M^h) = 0$ for all $(x_k, \sigma_k) \in X_k \times \Sigma_k$, $k = A, B$, we can then conveniently combine welfare and profit maximization into the following objective function:

$$\Omega^h(M) = \sum_{k=A,B} \int_{\Theta_k} \left\{ \int_{\mathbf{s}_k(\theta_k)} \varphi_k^h\left(\theta_k, \theta_l\right) dF_l(\theta_l) \right\} dF_k(\theta_k)$$

where $\varphi_k^h\left(\theta_k, \theta_l\right)$ is as defined in the main text.

Next, define the indicator function $m_k(\theta_k, \theta_l) \in \{0, 1\}$ taking value one if and only if $\theta_l \in \mathbf{s}_k(\theta_k)$, that is, if and only if the two types θ_k and θ_l are connected. Now define the following measure on the Borel sigma-algebra over $\Theta_k \times \Theta_l$:

$$\nu_k(E) \equiv \int_E m_k(\theta_k, \theta_l) dF_k(\theta_k) dF_l(\theta_l). \tag{13}$$

Reciprocity implies that $m_k(\theta_k, \theta_l) = m_l(\theta_l, \theta_k)$. As a consequence, the measures ν_k and ν_l satisfy $d\nu_k(\theta_k, \theta_l) = d\nu_l(\theta_l, \theta_k)$. Therefore,

$$\begin{aligned} \Omega^h(M) &= \sum_{k=A,B} \int_{\Theta_k \times \Theta_l} \varphi_k^h\left(\theta_k, \theta_l\right) d\nu_k(\theta_k, \theta_l) \\ &= \int_{\Theta_k \times \Theta_l} \triangle_k^h(\theta_k, \theta_l) m_k(\theta_k, \theta_l) dF_k(\theta_k) dF_l(\theta_l). \end{aligned}$$

By point-wise maximization of the integral above, it is then clear that the matching rule that maximizes $\Omega^h(M)$ is such that $m_k(\theta_k, \theta_l) = m_l(\theta_l, \theta_k) = 1$ if and only if

$$\triangle_k^h(\theta_k, \theta_l) \geq 0.$$

Next note that the function

$$\varphi_k^h\left(\theta_k, \theta_l\right) \equiv u_k\left(v_k, |x_k - x_l|, \sigma_l\right) - \mathbf{1}_h \cdot \frac{1 - F_k^v(v_k | x_k, \sigma_k)}{f_k^v(v_k | x_k, \sigma_k)} \cdot \frac{\partial u_k}{\partial v}\left(v_k, |x_k - x_l|, \sigma_l\right)$$

is strictly increasing in v_k by Condition LR. Therefore, fixing θ_k, for any $(x_l, \sigma_l) \in X_l \times \Sigma_l$, there exists a threshold $t_k^h(v_k, x, \sigma)$ such that $\triangle_k^h(\theta_k, \theta_l) \geq 0$ if and only if $v_l \geq t_k^h(v_k, x, \sigma)$. Condition LR also implies that the threshold $t_k^h(v_k, x, \sigma)$ is decreasing in v_k. Moreover, because u_k weakly decreases in $|x_k - x_l|$, when $h = W$, the threshold $t_k^W(v_k, x, \sigma)$ is weakly increasing in the distance $|x_k - x_l|$. The same is true, when $h = P$, for $t_k^P(v_k, x, \sigma)$ as long as Condition MS holds, along with one of the two conditions I_k or S_k, for $k = A, B$. Finally, observe that, because u_k weakly increases with σ_l, when $h = W$, the threshold $t_k^W(v_k, x, \sigma)$ is weakly decreasing in σ. The same property holds for $t_k^P(v_k, x, \sigma)$ when Condition MS holds along with one of the two conditions I_k or S_k, $k = A, B$.

That the threshold rule that maximizes $\Omega^h(M)$, $h = W, P$, satisfies conditions (1)-(3) in the proposition then follows directly from the properties above.

Below, we complete the proof by showing that under any of the three scenarios described below, the mechanism M^h where the matching rule is given by the threshold rule in the proposition and where the payment rule is the one in (12) is incentive compatible (that the mechanism is individually rational follows directly from (12)):

(a) Locations are public on both sides;

(b) Locations are private on side k and public on side $l \neq k$ and, in addition, Condition S_l and I_k hold;

(c) Locations are private on both sides and Conditions S_k holds, $k = A, B$.

Definition 1 ***(nested matching)*** *A matching rule $\mathbf{s}_k(\cdot)$ is said to be nested if for any $\theta_k = (x_k, v_k, \sigma_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k, \sigma_k)$ such that $x_k = \hat{x}_k$ and $\sigma_k = \hat{\sigma}_k$, either $\mathbf{s}_k(\theta_k) \subseteq \mathbf{s}_k(\hat{\theta}_k)$ or $\mathbf{s}_k(\theta_k) \supseteq \mathbf{s}_k(\hat{\theta}_k)$ is true. A mechanism that employs a nested matching rule is said to be nested.*

Definition 2 ***(ICV)*** *A mechanism M satisfies incentive compatibility along the v dimension (ICV) if for any $\theta_k = (x_k, v_k, \sigma_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k, \hat{\sigma}_k)$ such that $x_k = \hat{x}_k$ and $\sigma_k = \hat{\sigma}_k$, $\Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$.*

It is straight-forward to prove the following result.

Lemma 1 ***(ICV)*** *A nested mechanism M satisfies ICV if and only if the following conditions jointly hold:*

1. *for any $\theta_k = (x_k, v_k, \sigma_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k, \hat{\sigma}_k)$ such that $x_k = \hat{x}_k$ and $\sigma_k = \hat{\sigma}_k$, $v_k > \hat{v}_k$ implies that $\mathbf{s}_k(\theta_k) \supseteq \mathbf{s}_k(\hat{\theta}_k)$,*

2. *the envelope formula (12) holds.*

It is clear that the mechanism associated with the threshold function t_k^h is nested and satisfies the monotonicity condition 1 in the lemma above. Because the envelope formula holds by construction, it follows that this mechanism satisfies ICV.

It is then immediate that, under scenario (a), i.e., when locations are public on both sides, the mechanism M^h is incentive compatible.

Now consider scenario (b). Incentive compatibility on side l (which coincides with ICV) follows from the fact that the matching mechanism defined by the threshold rule t_k^h is nested and satisfies the monotonicity condition in the lemma above. In turn, incentive compatibility on side k requires that

$$\Pi_k((x_k, v_k, \sigma_k); M) \geq \hat{\Pi}_k((x_k, v_k, \sigma_k), (\hat{x}_k, \hat{v}_k, \sigma_k); M),$$

for all $(x_k, \hat{x}_k, v_k, v_k, \sigma_k) \in X_k^2 \times V_k^2 \times \Sigma_k$. Equivalently,

$$\int_{\underline{v}_k}^{v_k} \int_{\mathbf{s}_k(x_k,y,\sigma_k)} \frac{\partial u_k}{\partial v}(y, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l) dy \geq \int_{\underline{v}_k}^{\hat{v}_k} \int_{\mathbf{s}_k(\hat{x}_k,y,\sigma_k)} \frac{\partial u_k}{\partial v}(y, |\hat{x}_k - x_l|, \sigma_l)\, dF_l(\theta_l) dy \quad (14)$$
$$+ \int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k,\sigma_k)} [u_k(v_k, |x_k - x_l|, \sigma_l) - u_k(\hat{v}_k, |\hat{x}_k - x_l|, \sigma_l)] dF_l(\theta_l).$$

It is easy to see that, for any $\theta_k \in \Theta_k$,

$$\int_{\mathbf{s}_k(\theta_k)} \frac{\partial u_k}{\partial v}(v, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l) = \int_{\sigma_l \in \Sigma_l} \int_{d \in [0,1/2]} \frac{\partial u_k}{\partial v}(v, d, \sigma_l)\, dW(d; \theta_k, \sigma_l) dF_l^\sigma(\sigma_l), \quad (15)$$

where $W(d; \theta_k, \sigma_l)$ is the measure of agents with traits σ_l whose distance from x_k is at most d that are included in the matching set of type θ_k under the mechanism M^h. It is also easy to see that, under Conditions I_k and S_l, $l \neq k$, the expression in (15) is invariant in x_k. That is, $W(d; \theta_k, \sigma_l) = W(d; \theta'_k, \sigma_l)$ for any $d \in [0, 1/2]$ any $\theta_k, \theta'_k \in \Theta_k$ with $v_k = v'_k$.and $\sigma_k = \sigma'_k$. This means that

$$\int_{\underline{v}_k}^{\hat{v}_k} \int_{\mathbf{s}_k(\hat{x}_k,y,\sigma_k)} \frac{\partial u_k}{\partial v}(y, |\hat{x}_k - x_l|, \sigma_l)\, dF_l(\theta_l) dy = \int_{\underline{v}_k}^{\hat{v}_k} \int_{\mathbf{s}_k(x_k,y,\sigma_k)} \frac{\partial u_k}{\partial v}(y, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l) dy.$$

By the same arguments,

$$\int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k,\sigma_k)} u_k(\hat{v}_k, |\hat{x}_k - x_l|, \sigma_l)\, dF_l(\theta_l) = \int_{\mathbf{s}_k(x_k,\hat{v}_k,\sigma_k)} u_k(\hat{v}_k, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l),$$

and

$$\int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k,\sigma_k)} u_k(v_k, |\hat{x}_k - x_l|, \sigma_l)\, dF_l(\theta_l) < \int_{\mathbf{s}_k(x_k,\hat{v}_k,\sigma_k)} u_k(v_k, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l).$$

It follows that the right hand side of (14) is smaller than

$$\int_{\underline{v}_k}^{\hat{v}_k} \int_{\mathbf{s}_k(x_k,y,\sigma_k)} \frac{\partial u_k}{\partial v}(y, |x_k - x_l|, \sigma_l)\, dF_l(\theta_l) dy$$
$$\int_{\mathbf{s}_k(x_k,\hat{v}_k,\sigma_k)} [u_k(v_k, |x_k - x_l|, \sigma_l) - u_k(\hat{v}_k, |x_k - x_l|, \sigma_l)] dF_l(\theta_l),$$

which is the payoff that type (x_k, v_k, σ_k) obtains by announcing $(x_k, \hat{v}_k)$. That the inequality in (14) holds then follows from the fact that the mechanism satisfies ICV.

Finally, consider scenario (c). Under Condition S_k and S_l, that the proposed mechanism is incentive compatible follows from the same arguments above applied to side l as well. Q.E.D.

Proof of Proposition 2. Consider first the case in which locations are public information, as in part (a) in the proposition. Because the agents' payoffs π are the sum of the utilities u they derive from each of the matches, it is then easy to see that the h-optimal rule is separable in the locations. What we mean is that, for any $\theta_k = (x_k, v_k, \sigma_k)$, any x_l, $k, l = A, B$, $l \neq k$, the measure of agents from side l with location x_l that are included in the matching set of type θ_k under any optimal mechanism is independent of the measure of agents from side l with location x_l' included in θ_k's matching set. This turns the problem into a collection of independent optimization problems, one for each pair (x_k, x_l). Thus fix the pair (x_k, x_l). The assumption that g is positive and non-decreasing, along with the assumption that the affiliation between the values for matching v_k and the personal traits σ_k is of the same sign on both sides of the market then implies that it is optimal to include in θ_k's matching set all agent from side l with location x_l whose value for matching is weakly higher than $t_k^h(v_k, x)$, irrespective of the personal traits. The proof for this last claim follows from arguments similar to those that establish Proposition 1 in Gomes and Pavan (2015) and hence is omitted.

Next, consider the case where locations are private information on at least one of the two sides of the market. The additional conditions in parts (b) and (c) in the proposition guarantee that the agents have no incentives to misreport their locations. The proof for this claim is similar to the one in Proposition 1 above and hence also omitted. Finally, to see that the optimal thresholds satisfy the properties in the proposition, observe that the value of allowing matches between agents from side k with location x_k and value for matching v_k and agents from side l with location x_l and value for matching v_l, when the matching rule does not discriminate on the basis of the personal traits is given by

$$f_l^v(v_l|x_l) f_k^v(v_k|x_k) \bar{\Delta}_k^h (v_k, v_l; x_k, x_l)$$

Observe that, when Condition MR holds, then the function $\bar{\Delta}_k^h (v_k, v_l; x_k, x_l)$ satisfies the following single crossing condition: whenever $\bar{\Delta}_k^h (v_k, v_l; x_k, x_l) \geq 0$, then $\bar{\Delta}_k^h (v_k, \hat{v}_l; x_k, x_l) > 0$ for all $\hat{v}_l > v_l$ and $\bar{\Delta}_k^h (\hat{v}_k, v_l; x_k, x_l) > 0$ for all $\hat{v}_k > v_k$. It is then easy to see that the optimal thresholds satisfy the properties in the proposition. Q.E.D.

Proof of Proposition 3. Let $q_k^h(x_l; \theta_k)$ denote the measure of x_l-agents included in the matching set of any agent from side k reporting a type θ_k, under the h-optimal matching rule $\mathbf{s}_k^h(\cdot)$. From Proposition 1, $q_k^h(x_l; \theta_k)$ is weakly increasing in v_k (strictly increasing whenever $q_k^h(x_l; \theta_k) \in (0, 1)$). Now define the marginal price for the q_k-th unit of x_l-agents by those agent from side k located at x_k by

$$\frac{d\rho_k^h}{dq}(q_k, x_l; x_k) \equiv u_k \left(\left(V_k^h(x_l, x_k) \right)^{-1} (q_k), |x_k - x_l| \right),$$

where $V_k^h(x_l, x_k)^{-1}(q_k) \equiv \inf\{v_k : q_k^h(x_l; x_k, v_k) = q_k\}$. Now define the price schedule $\rho_k^h(\cdot, x_l; x_k)$ as follows

$$\rho_k^h(q_k, x_l; x_k) \equiv \left(1 - F_l^v(t_k^h(\underline{v}_k, x_k, x_l)\right) \cdot u_k\left(\underline{v}_k, |x_k - x_l|\right) + \int_0^{q_k} \frac{d\rho_k^h}{dq}(q, x_l; x_k) dq.$$

From the integral formula above, we get that the optimal price schedules $\rho_k^h(\cdot)$ are differentiable at any quantity $q_k \in (0, 1)$.

Finally, substituting the elasticity formula (10) and the marginal price formula (9) into the Lerner-Wilson formula (11) and using the same formulas for side-l agents and recognizing that

$$V_l^h(x_k, x_l)^{-1}\left(D_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l; x_k), x_l; x_k\right)\right) = V_l^h(x_k, x_l)^{-1}(q_l^h(q_k, x_k; x_l)) = t_k^h(v_k, x_k, x_l)$$

for types $\theta_k = (x_k, v_k)$ such that $q_k^h(x_l; \theta_k) = q_k$, permits us to establish Condition (6). Q.E.D.

References

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

[2] Armstrong, M., 2006, "Competition in Two-Sided Markets." *Rand Journal of Economics*, Vol. 37(3), pp. 668-691.

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

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

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

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

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

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

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

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

[11] Eeckhout, J. and P. Kircher, 2010a, "Sorting vs Screening: Search Frictions and Competing Mechanisms." *Journal of Economic Theory*, Vol. 145, pp. 1354-1385.

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

[13] 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.

[14] 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.

[15] Evans, D. and R. Schmalensee, 2009, "Failure to Lauch: Critical Mass in Platform Businesses." Working Paper MIT Sloan.

[16] Fershtman D. and A. Pavan (2015), "Re-Matching, Experimentation, and Cross-Subsidization." Working Paper Northwestern University.

[17] Gomes, R. and A. Pavan, 2015, "Many-to-Many Matching and Price Discrimination," *Theoretical Economics*, forthcoming.

[18] Hagiu, A., 2008, "Pricing and Commitment by Two-Sided Platforms." *Rand Journal of Economics*, Vol. 37(3), pp. 720-737.

[19] Helsley, R. and W. Strange, 2000, "Social Interactions and the Institutions of Local Government." *American Economic Review*, Vol. 90(5), pp. 1477-1490.

[20] Hoppe, H., B. Moldovanu and E. Ozdenoren, 2010, "Coarse Matching with Incomplete Information." *Economic Theory*.

[21] 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.

[22] Jonhson, T., 2010, "Matching Through Position Auctions." *Journal of Economic Theory*, Vol. 148, pp. 1700-1713.

[23] Jullien, B., 2000, "Participation Constraints in Adverse Selection Models." *Journal of Economic Theory*, Vol. 93(1), pp. 1-47.

[24] Jullien, B., 2011, "Competition in Multi-Sided Markets: Divide-and-Conquer." *American Economic Journal: Microeconomics*, Vol. 3(4), pp. 1-35.

[25] Jullien, B., 2012, "B2B Two-sided Platforms." In Oxford Handbook of Digital Economics (M. Peitz and J. Waldfogel eds.), Oxford University Press.

[26] Kaiser, U. and M. Song, 2009, "Do Media Consumers Really Dislike Advertising?: An Empirical Assessment of the Role of Advertising in Print Media Markets." *International Journal of Industrial Organization*, Vol. 27, pp. 292-301.

[27] Kaiser, U. and J. Wright, 2006, **?**Price structure in two-sided markets: Evidence from the magazine industry.**?** *International Journal of Industrial Organization*, Vol. 24, pp. 1-28.

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

[29] Lu, X. and P. McAfee, 1996, "The Evolutionary Stability of Auctions Over Bargaining." *Games and Economic Behavior*, Vol. 15, pp. 228-254.

[30] Lucking-Reyley, D. and D. Spulber, 2001, "Busisness-to-Business Electronic Commerce." *Journal of Economic Prespectives*, Vol. 15, pp. 55-68.

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

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

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

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

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

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

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

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

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

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

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

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

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

[44] White, A. and G. Weyl, 2014, "Imperfect Platform Competition: A General Framework." Working Paper Microsoft Research.

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