CMS-EMS
Center for Mathematical Studies in Economics
And Management Science

Discussion Paper #1559

Cross-Subsidization and Matching Design

Renato Gomes
Toulouse School of Economics

Alessandro Pavan
Northwestern University

January 2013

JEL classification: D82

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

Cross-Subsidization and Matching Design*

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

Alessandro Pavan
Northwestern University
alepavan@northwestern.edu

January 2013

Abstract

We develop a theory of price discrimination in many-to-many matching markets in which agents' preferences are vertically and horizontally differentiated. The optimal plans induce negative assortative matching at the margin: agents with a low value for interacting with other agents are included in the matching sets of only those agents from the opposite side whose value for matching is sufficiently high (cross-subsidization). We deliver testable predictions relating the optimal matching plans and price schedules to the distribution of the agents' preferences and attractiveness. The analysis has implications for the design of business-to-business platforms, advertising, and cable TV packages.

JEL classification: D82

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

*Previous versions circulated under the titles "Price Discrimination in Many-to-Many Matching Markets" and "Many-to-Many-Matching Design". For comments and useful suggestions, we thank seminar participants at UC Berkeley, Bocconi, Bonn, Cambridge, Universitat Carlos III, Duke, École Polytechnique, Getulio Vargas Foundation (EPGE and EESP), Northwestern, NYU, Pompeu Fabra, Stanford, Toulouse, Yale, UBC, UPenn; as well as conference participants at 2011 Columbia-Duke-Northwestern IO Conference, SED 2011, 2011 Csef-Igier Conference, 2011 Stony Brook Game Theory Festival, ESEM 2011, EARIE 2011, ESWM 2012, and ESSET 2012. We are grateful to Jacques Crémer, Jan Eeckhout, 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 suggestions. André Veiga provided excellent research assistance. Pavan also thanks the Haas School of Business and the Economics Department at UC Berkeley for their hospitality during the 2011-2012 academic year, as well as the National Science Foundation (grant SES-60031281) and Nokia for financial support. 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 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 the defining feature of mediated many-to-many matching, which is the focus of this paper.[2]

We start by studying markets in which preferences are purely vertically differentiated. In such markets, all agents agree on the salience of the dimension that is responsible for each agent's attractiveness. Because of this concordance over salience, two agents from the same side may disagree on the relative attractiveness of any two agents from the opposite side only when one of the two agents positively values interacting with agents from the other side, while the other agent negatively values such interactions. Consider, for example, the market for nontargeted advertising. While all users may agree that one media outlet offers a larger quantity of advertising than the other (the relevant salience dimension in this example), different users may exhibit different degrees of tolerance towards advertising. In particular, while many users dislike advertising, some actually enjoy it. Vertical differentiation is believed to be an important dimension in B2B matching, in markets for nontargeted advertising, and in credit cards markets.

We then extend the analysis to markets in which preferences exhibit elements of both vertical and horizontal differentiation. In such markets, 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

[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.

with certain agents positively but with others negatively. For example, in the market for 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 horizontal differentiation is believed to play an important role include online targeted advertising, online dating sites, and the market for the provision of employment matching services.

For both types of markets (with and without horizontal differentiation), we consider the problem of a monopolistic platform that operates on two sides. Agents from each side have private information both about their own preferences and about personal characteristics that determine their attractiveness to those agents they are matched to. For example, in the market for B2B matching, each procurer has private information not only about his willingness to pay for meeting business partners, but also about his expected budget, his purchasing history, and various other traits of his profile that determine his attractiveness to the vendors. Similarly, in the cable TV example, each viewer possesses private information not only about the importance that he assigns to cable TV (the vertical dimension in his preferences) but also about his ideal channel profile (the horizontal dimension), which determines the viewer's attractiveness to the channels.

Our analysis explores the implications of such heterogeneity in preferences and in attractiveness for such questions as: What matching allocations are likely to emerge under profit maximization (private provision of matching services) and under welfare maximization (public provision)? How are these allocations affected by shocks that alter the joint distribution of the agents' preferences and attractiveness? What price schemes sustain such allocations?

We answer these questions using a mechanism design approach. The platform's problem consists in choosing a *matching rule* along with a *pricing rule* for each side of the market that jointly maximize either welfare or profits. We will 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 this is not needed. A matching rule assigns each agent to a set of agents from the other side. We require only 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.

Pure Vertical Differentiation. Our first result shows that, when the conditions described below hold, the optimal matching rules (i) discriminate only along the willingness-to-pay dimension (that is, any two agents with the same value for matching are matched to the same set of agents, irrespective of possible differences in their salience),[3] and (ii) have a threshold structure, according

[3]Hereafter, by *value for matching,* we mean the value an agent assigns to interacting with agents from the other side (the vertical dimension in the agent's preferences); this can be either positive or negative. By *salience,* we mean a combination of various dimensions in the agent's profile that determine the agent's attractiveness. A high salience implies a high attractiveness when evaluated from the perspective of someone from the other side whose value for matching is positive, but a low attractiveness when evaluated from the perspective of someone whose value for matching is negative.

to which each agent is matched to all agents from the other side whose value for matching is above a certain threshold.[4] The thresholds are decreasing in the reported values, thus implying *negative assortativeness at the margin*: those agents with a low value for matching are matched only to those agents from the other side whose value is sufficiently high.

We can divide into two sets the primitive conditions that imply that the optimal matching rules satisfy the above two properties. The first set of conditions requires that (1a) each agent's salience be either independent of or positively correlated with the agent's value for matching and (1b) those agents with a positive value for matching (i.e., who like being matched to other agents) have weakly diminishing marginal utility for the quality of their matching set, while those agents with a negative value for matching (i.e., who dislike being matched to other agents) have weakly diminishing marginal disutility. In the context of B2B matching, condition (1a) captures the idea that those procurers and vendors who are willing to pay the most to find a partner are, in general, seen as better matches. In turn, condition (1b) captures the idea that both procurers and vendors have scarce resources to evaluate the business potential of the proposed partners. Together with the fact that each agent has private information both about his salience and his value for matching, conditions (1a) and (1b) above imply that the cost-minimizing way to provide a matching set of high quality to those agents with a positive value for matching is to match those agents to all agents from the other side whose value for matching is sufficiently high, irrespective of their salience. This is because (a) the latter agents are, on average, the most attractive ones and (b) because using an agent with a negative value for matching intensively is less costly than using different agents with a negative value.

The second set of conditions that guarantee that the optimal rules both discriminate only along the willingness-to-pay dimension and have a threshold structure are the mirror image of conditions (1a) and (1b) above. Namely, they require that (2a) each agent's salience be independent of or negatively correlated with the agent's value for matching, and (2b) those agents with a positive value for matching have weakly increasing marginal utility for the quality of their matching set, while those agents with a negative value for matching have weakly increasing marginal disutility. Together with asymmetric information, conditions (2a) and (2b) above imply that the most effective way of using those agents with a negative value for matching is to match them to those agents with the highest positive values. This is because the latter agents are, on average, the least salient ones (and hence the most attractive ones to those agents with a negative value for matching) and because these agents have the largest matching sets (by virtue of incentive compatibility) and hence the highest positive marginal utility for increasing the quality of their matching set.

The threshold structure described above implies two interesting properties of optimal matching rules. First, the matching sets of any two agents from the same side are either identical, or nested, in the sense that one is a superset of the other. Second, for any two matching sets, the "marginal"

[4] Note that this property is not a consequence of incentive compatibility. The latter property requires only that the quality of the matching sets be invariant over characteristics that are irrelevant for individual preferences. It does not require that the composition of the matching sets also be invariant.

agents included only in the larger set are always those with the lowest value for matching among those included in either set. This implies that, as matching sets expand, the marginal agents are either the most or the least attractive ones, depending on whether matching is positively or negatively valued and on the correlation between values and salience.

Building on these results, we then identify the crucial condition that determines when a single complete network in which all agents from each side are matched to all agents from the other side is optimal. Specifically, such a network is optimal *if and only if* welfare (or, alternatively, profits) decrease when one removes from the complete network the link between the two agents from each side with the lowest values.[5] When, instead, welfare (or, alternatively, profits) increase by deleting such link, the optimal matching rule entails *separation*: those agents with a higher value for matching receive matching sets that are supersets of those offered to agents with a lower value for matching.

We also show that, when separation is optimal, the thresholds that define the optimal matching sets are given by an Euler equation that equalizes (i) the marginal gains in efficiency (or, alternatively, in profits) from expanding the matching sets on one side to (ii) the marginal losses in efficiency (or, alternatively, in profits) that, by reciprocity, arise on the other side of the market. Intuitively, this equation endogenously separates agents from each side into two groups. The first group consists of agents who play the role of *consumers*. These agents contribute positively to the platform's objective by "purchasing" sets of agents from the other side of the market. The second group consists of agents who play the role of *inputs*. These agents contribute negatively to the platform's objective, but serve to "feed" the matching sets of those agents from the opposite side who play the role of consumers.

As in the standard price discrimination problem analyzed in Mussa and Rosen (1978) and Maskin and Riley (1984), we identify conditions that ensure that a profit-maximizing platform separates types as finely as possible. It turns out that the familiar *regularity condition* (Myerson, 1981), according to which virtual values increase in true values, is not the right condition in our matching environment. In a standard setting, the marginal cost is independent of the agent's type; as a result, the monotonicity of the virtual values then implies the monotonicity of the trades. In contrast, in a matching environment, by virtue of reciprocity, the marginal cost of increasing the quality of the matching set of those agents who play the role of consumers is the profits loss that results from adding these agents to the matching sets of those agents from the opposite side who play the role of inputs. Since salience and values are potentially correlated, the marginal cost of expanding a matching set is thus also a function of the agent's value for matching. For the optimal matching rule to separate types as finely as possible, one must then require that the virtual values increase with the true values faster than their corresponding marginal cross-side effects do. In analogy to Myerson (1981), we refer to this condition as *Match Regularity*. Under this condition, bunching can occur only at "the top" (i.e., for the agents with the highest value for matching) due to capacity constraints, that is, because the stock of agents from the other side of the market has been exhausted; or at "the bottom" (i.e., for those agents with the lowest values) due to complete exclusion.

[5] Note that, although the "only if" part of the result is trivial here, the "if" part is not.

Building on the work of Wilson (1997), we then 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 so that the marginal gains in revenue from expanding the matching sets sold to consumer-agents equal the marginal costs of procuring input-agents from the opposite side. Interestingly, these marginal costs are endogenous and depend on the entire network structure of the matching allocations.

Our analysis also delivers testable predictions about the effects on profits of shocks that alter the joint distributions of values and salience. In particular, we show that a shock that (i) increases (stochastically) the salience of every agent from side A (albeit not necessarily uniformly across agents) while (ii) leaving unchanged the marginal distribution of the values for matching, (a) increases the quality of the matching sets sold to side-A agents with a low value for matching, but (b) reduces the quality of the matching sets sold to side-A agents with a high value. In terms of surplus, these shocks make low-value agents from side A better off at the expense of high-value ones. For example, in the B2B application, a positive shock to the salience of the procurers may induce the platform to respond by cutting the quality of the matching sets offered to high-value procurers.

Vertical and Horizontal Differentiation. Building on the above insights, we then extend the analysis to markets in which preferences combine elements of vertical and horizontal differentiation. To isolate the novel effects, we suppress the salience dimension and instead introduce a new spatial dimension that captures horizontal differentiation. Assuming that agents from each side are located on a circle, we then let the utility that each agent i from side A obtains from being matched to each agent j from side B continue to be an increasing function of agent i's value for matching (as in the model with only vertical differentiation), but now assume it is a decreasing function of the *distance* between the two agents' locations. In the cable TV example, a channel's location should be interpreted as the channel's profile (say, a "news" channel), while a viewer's location should be interpreted as the viewer's most preferred type of channel (say, a "sports" channel).

We show that, under fairly reasonable conditions that depend on the observability of the locations[6], the optimal matching rules (i) actively discriminate along the horizontal dimension (i.e., they assign different matching sets to agents with the same value for matching but with different locations), and (ii) 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 optimal matching rule specifies a threshold $t(x_B|x_A, v_A)$ for each location x_B on side B such that an agent from side A located at x_A and with value for matching v_A is matched to all agents from side B located at x_B whose value for matching v_B is above this threshold. Figure 1 below illustrates this structure.

As in the model with purely vertically differentiated preferences, the optimal rules thus induce a form of negative assortative matching at the margin according to which agents with a low value for matching are matched only to those agents from the other side with a sufficiently high value.

[6] Our analysis accommodates both the case where locations are public information (e.g., a channel's profile) and the case where they are private information (as in the case of viewers' tastes for news, sports, and movies).

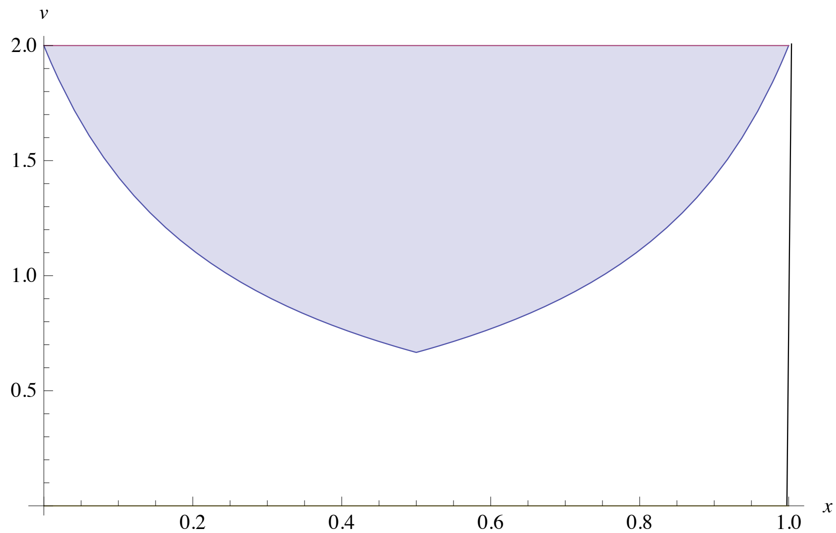

Figure 1: Optimal matching rule with horizontal differentiation. The shaded area in the figure describes the matching set for an agent from side A located at $x_A = 1/2$.

The difference is that, with horizontal differentiation, this form of negative assortativeness takes into account the agents' mutual attractiveness. As a result, in markets where horizontal differentiation plays a prominent role, the matching sets of any two agents are nested only if the two agents share the same location. This distinction has implications, for example, for the provision of cable TV services where the market has gradually moved from menus of packages with a nested structure (basic, 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 the utility functions 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).[7] In other words, the composition of the matching sets naturally respects the agents' idiosyncratic preferences: those agents from the opposite side whose profile makes them more attractive are present in the matching set in larger numbers.[8]

As in the case of vertical differentiation, we show how the optimal thresholds can be characterized by an Euler equation. In contrast to the case of vertical differentiation, however, the separation of agents between consumers and inputs now depends on their *joint* location: the same agent from side A may play the role of a consumer when matched to an agent from side B with a certain location, but the role of an input when matched to another agent with a different location.

Finally, we show how the optimal matching rules can be indirectly 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

[7]In the cable TV application, the supermodularity assumption means that viewers with a high value for cable TV are, in general, more likely to also watch channels that are far away from their ideal profile.

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

group of channels included in the baseline package), along with the specification of the price that the subscriber has to pay to add channels from each possible category. Agents then select the plan that best fits their preferences (e.g., viewers who like sports choose the sports package) and then personalize the package by adding a few additional channels. Using techniques similar to those in the case of vertical differentiation, 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.

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 with whom they interact. In this setting, the principal's problem consists in assigning the agents to nonexclusive groups. This one-sided problem is mathematically equivalent to a two-sided matching problem where both sides have symmetric primitives and 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. In particular, our results have implications for such problems in organization and personnel economics that pertain to the optimal 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 studies markets with purely vertically differentiated preferences. Section 3 shows how to extend the analysis to accommodate horizontal differentiation. Section 4 presents a few extensions, while Section 5 discusses how our predictions 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 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 two 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.

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.[9] Our contribution relative to this literature is in allowing for general matching rules and in introducing 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 exclusive groups (e.g., classes).[10]

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 study a setting with both vertically and horizontally differentiated preferences 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.[11] 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.

Decentralized Matching. In a decentralized matching economy, Eeckhout and Kircher (2010a)

[9]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 markets for cable TV, B2B matching, and online advertising.

[10]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.

[11]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 with agents from 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 possible and (b) agents can select their partner(s) deterministically, as in the case of cable TV or business-to-business matching. While we assume that the quality of a matching set is determined by the sum of the attractiveness of its members, we do accommodate the possibility that an agent's marginal utility for additional matches decreases with the quality of his matching set, thus potentially tempering the effect of match size on match quality.

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 matching, abstract from search frictions, and consider many-to-many matching rules.

2 Vertical Differentiation

2.1 Model and Preliminaries

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 type $\theta_k^i = (\sigma_k^i, v_k^i) \in \Theta_k \equiv \Sigma_k \times V_k$. The first component $\sigma_k^i \in \Sigma_k \equiv [\underline{\sigma}_k, \overline{\sigma}_k] \subseteq \mathbb{R}_+$ parametrizes the "salience" of agent i; the latter should be interpreted as a combination of various personal characteristics that determines agent i's attractiveness. 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 value that agent i assigns to interacting with agents from the other side. Formally, given any (Lebesgue measurable) set $\mathbf{s}$ of agents from side $l \neq k$ and any complete type profile $\theta \equiv (\theta_k^i)_{k=A,B}^{i \in [0,1]}$, the payoff that agent $i \in [0,1]$ from side $k \in \{A, B\}$ obtains from being matched, at a price p, to the set $\mathbf{s}$ is given by

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

where $g_k(\cdot)$ is a positive, strictly increasing, and continuously differentiable function satisfying $g_k(0) = 0$, and where

$$|\mathbf{s}|_l \equiv \int_{j \in \mathbf{s}} \sigma_l^j d\lambda(j)$$

is the total salience of $\mathbf{s}$ ($\lambda(\cdot)$ denotes the Lebesgue measure).

The case where agent i dislikes interacting with agents from the other side is thus captured by a negative value for matching $v_k^i < 0$. To avoid the uninteresting case where no agent from either side benefits from interacting with agents from the opposite side, we assume that $\bar{v}_k > 0$ for some $k \in \{A, B\}$. Importantly, note that any two agents from the same side who both benefit (or, alternatively, suffer) from interacting with agents from the other side rank any two agents from the opposite side in the same way. In this sense the model is one of pure vertical differentiation. Also note that, fixing the sign of the value for matching v_k^i, the attractiveness of any set $\mathbf{s}$ of agents from side $l \neq k$ coincides with its salience. Hereafter, we will thus often use the two terms interchangeably when there is no risk of confusion. The role of the functions $g_k(\cdot)$, $k = A, B$, is to capture increasing (or, alternatively, decreasing) marginal utility (or, alternatively, disutility) of expanding the salience of the matching sets.

Each agent's type $\theta_k^i = (\sigma_k^i, v_k^i)$ is the agent's own private information and is an independent draw from a distribution F_k with support Θ_k. We assume that F_k is absolutely continuous with respect to the Lebesgue measure, and denote by F_k^v the marginal distribution of F_k with respect to v_k (with density f_k^v), and by $F_k^\sigma(\cdot|v_k)$ the distribution of the salience conditional on v_k. We will assume that, for any $k = A, B$, the family of functions $\langle F_k^\sigma(\cdot|v_k)\rangle_{v_k \in V_k}$ is uniformly continuous in v_k in the L_1-norm. As is standard in the mechanism design literature, we also assume that each marginal distribution F_k^v is *regular* in the sense of Myerson (1981), meaning that the virtual values for matching $v_k - [1 - F_k^v(v_k)]/f_k^v(v_k)$ are continuous and nondecreasing.

The payoff formulation in (1) is fairly flexible and accommodates the following examples as special cases.

Example 1 ***(linear network externalities for quantity)*** Suppose that all agents have the same salience and that the marginal utility/disutility of higher salience are independent of the matching sets. This is equivalent to letting $\sigma_k^i \equiv 1$ for all $i \in [0,1]$, $k = A, B$, and to assuming that each agent i's payoff from each side $k = A, B$ is equal to

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

These preferences are the ones typically considered in the two-sided market literature (see Rysman (2009) for a survey) and in particular in the literature on B2B platforms (see, e.g., Jullien (2012)) and in the literature on credit-card payments (see, e.g., Rochet and Tirole (2003)).

Example 2 ***(supermodular matching values)*** Suppose that the match between agent $i \in [0,1]$ from side $k \in \{A, B\}$ and agent $j \in [0,1]$ from side $l \in \{A, B\}$, $l \neq k$, produces a surplus of $v_k^i \cdot v_l^j$ to each of the two agents, independently of the two agents' matching sets. These preferences can be re-conducted to the formulation in (1) by setting $\sigma_k^i \equiv v_k^i$ and $g_k(x) = x$ for all $i \in [0,1]$ and $k = A, B$. \\\\

This specification 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)). Applications of interest include online job matching agencies and online dating agencies.

Example 3 ***(limited attention/nuisance costs)*** Suppose that the payoff that each agent $i \in [0,1]$ from side $k \in \{A, B\}$ obtains from being matched to each agent $j \in [0,1]$ from side $l \in \{A, B\}$, $l \neq k$, is given by

$$\frac{v_k^i \cdot \sigma_l^j}{(|\mathbf{s}|_l)^\alpha}$$

where $\mathbf{s}$ is agent i's matching set. If $\alpha \in (0,1)$ and $V_k \subset \mathbb{R}_+$, the specification above captures the idea that the positive value of meeting an extra agent decreases with the total quality of the matching set (e.g., due to limited attention). In turn, if $\alpha < 0$ and $V_k \subset \mathbb{R}_-$, the above specification captures the idea that nuisance costs (e.g., in advertising markets) are convex in the intensity of advertising.

In either case, agent's i's total payoff is equal to

$$\int_{j\in\mathbf{s}} \frac{v_k^i \cdot \sigma_l^j}{(|\mathbf{s}|_l)^\alpha} d\lambda(j) - p = v_k^i \cdot g_k(|\mathbf{s}|_l) - p,$$

with $g_k(x) = x^{1-\alpha}$. The function $g(\cdot)$ is concave when $\alpha \in (0,1)$ (limited attention) and convex when $\alpha < 0$ (nuisance costs). \\

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).[12]

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, 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, 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 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 included in the matching set of any agent from side k reporting type θ_k. Note that $p_k(\cdot)$ maps Θ_k into $\mathbb{R}$, 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 $|\mathbf{s}_k(\theta_k)|_l$ the total salience of the matching set of any agent i from side k reporting type θ_k.

[12] 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.

Denote by $\hat{\Pi}_k(\theta_k, \hat{\theta}_k; M) \equiv v_k^i \cdot g_k(|\mathbf{s}_k(\hat{\theta}_k)|_l) - p_k(\hat{\theta}_k)$ the payoff that type $\theta_k = (\sigma_k, v_k)$ obtains when reporting type $\hat{\theta}_k = (\hat{\sigma}_k^i, \hat{v}_k^i)$, 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 $\theta_k \in \Theta_k$, $k = A, B$, and is incentive compatible (IC) if $\Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$ for all $\theta_k, \hat{\theta}_k \in \Theta_k$, $k = A, B$. 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.[13]

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 θ, the welfare generated by the mechanism M is given by

$$\Omega^W(M) = \sum_{k=A,B} \int_{\Theta_k} v_k \cdot g_k\left(|\mathbf{s}_k(\sigma_k, v_k)|_l\right) dF_k(\sigma_k, v_k), \tag{3}$$

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

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

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

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 salience of the matching set is nondecreasing in the value for matching v_k, i.e., $|\mathbf{s}_k(\sigma_k, v_k)|_l \geq |\mathbf{s}_k(\sigma_k', v_k')|_l$ for any (σ_k, v_k) and (σ_k', v_k') such that $v_k \geq v_k'$;

(ii) the expected payoff of any two agents with the same value for matching v_k is the same, irrespective of their salience σ_k;

(iii) the equilibrium payoffs $\Pi_k((\sigma_k, \underline{v}_k); M)$ of the agents with the lowest values for matching are non-negative, for all $\sigma_k \in \Sigma_k$;

[13] 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 Crémer and McLean (1988) — see also Mezzetti (2007).

(iv) the pricing rule satisfies the envelope formula

$$p_k(\sigma_k, v_k) = v_k \cdot g_k\left(|\mathbf{s}_k(\sigma_k, v_k)|_l\right) - \int_{\underline{v}_k}^{v_k} g_k\left(|\mathbf{s}_k(\sigma_k, x)|_l\right) dx - \Pi_k((\sigma_k, \underline{v}_k); M). \tag{6}$$

It is immediate to see that in any mechanism that maximizes the platform's profits, the IR constraints of those agents with the lowest values for matching bind, i.e., $\Pi_k((\sigma_k, \underline{v}_k); M^P) = 0$, for all $\sigma_k \in \Sigma_k$, $k = A, B$. Using the expression for payments (6), we can rewrite the platform's profit maximization problem in a manner analogous to the welfare maximization problem. We simply replace the true values with their virtual analogs (i.e., with the values discounted for informational rents). Formally, for any $k = A, B$, any $v_k \in V_k$, let $\varphi_k^W(v_k) \equiv v_k$ and $\varphi_k^P(v_k) \equiv v_k - [1 - F_k^v(v_k)]/f_k^v(v_k)$. Using the superscript $h = W$ (or, alternatively, $h = P$) to denote welfare (or, alternatively, profits), the platform's problem then consists in finding a matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ that maximizes

$$\Omega^h(M) = \sum_{k=A,B} \int_{\Theta_k} \varphi_k^h(v_k) \cdot g_k\left(|\mathbf{s}_k(\sigma_k, v_k)|_l\right) dF_k(\sigma_k, v_k) \tag{7}$$

among all rules that, together with the corresponding payment rules given by (6) (with $\Pi_k((\sigma_k, \underline{v}_k); M) = 0$ all $\sigma_k \in \Sigma_k$) satisfy constraints (i)-(ii) above and the reciprocity condition (5). 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 $\{v_k \in V_k : \varphi_k^h(v_k) \geq 0\} \neq \varnothing$. Finally, we let $\mathbf{1}_h$ denote the indicator function that equals one if $h = P$ and zero if $h = W$.

2.2 Optimal Matching Rules

As anticipated in the Introduction, our first result provides fairly natural primitive conditions under which the optimal matching rules have a simple structure.

Condition 1 ***[TP] Threshold Primitives:*** *One of the following two sets of conditions holds for both $k = A$ and $k = B$:*

(1.a) the function $g_k(\cdot)$ is weakly concave, and (1.b) the random variables $\tilde{\sigma}_k$ and $\tilde{v}_k$ are weakly positively affiliated;

(2.a) the function $g_k(\cdot)$ is weakly convex, and (2.b) the random variables $\tilde{\sigma}_k$ and $\tilde{v}_k$ are weakly negatively affiliated.[14]

Condition TP covers two alternative scenarios. The first one is one where agents with a positive value for matching have diminishing marginal utility for meeting new agents and where agents with a negative value for matching have diminishing marginal disutility for meeting additional agents. In either case, condition TP also requires that those agents with the highest values for matching are, on average, the most salient ones. These agents are thus seen as the most attractive ones by those

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

agents from the other side with a positive value for matching and as the least attractive ones by those agents with a negative value for matching.

The second scenario covers a symmetrically opposite situation. It assumes that those agents with a negative value for matching have a convex disutility for meeting new agents, whereas those agents with a positive value for matching have a weakly increasing marginal utility for meeting additional agents.[15] In either case, condition TP also requires that those agents with the highest values for matching are, on average, the least salient ones. These agents are thus seen as the most attractive ones by those agent from the opposite side with a negative value for matching and as the least attractive ones by those with a positive value for matching.

Importantly, note that when marginal benefits and costs are constant (linear $g_k(\cdot)$, $k = A, B$), Condition TP accommodates either positive or negative affiliation between salience and values for matching (in particular, it accommodates for the possibility that v_k and σ_k are independent) . In this case, the only restriction imposed by Condition TP is that the sign of the correlation between the two dimensions v_k and σ_k be the same on either side. We then have the following result.

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

$$\mathbf{s}_k^h(v_k) = \begin{cases} \Sigma_l \times [t_k^h(v_k), \overline{v}_l] \text{ if } v_k \in [\omega_k^h, \overline{v}_k] \\ \oslash \text{ otherwise,} \end{cases} \tag{8}$$

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

To understand the result, consider an agent with type $\theta_k = (\sigma_k, v_k)$ such that $\varphi_k^h(v_k) \geq 0$. In case of welfare maximization ($h = W$), this is an agent who values positively interacting with agents from the other side. In case of profit maximization ($h = P$), this is an agent who contributes positively to profits, even when accounting for informational rents. 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 $\mathbf{s}_k(\sigma_k, v_k) \supset \{(\sigma_l, v_l) : \varphi_l^h(v_l) \geq 0\}$ that includes all types $\theta_l = (\sigma_l, v_l)$ from side $l \neq k$ whose φ_l^h-value is non-negative. This is because (i) irrespective of their salience σ_l, these latter types contribute positively to type θ_k's payoff and (ii) these latter types have a non-negative φ_l^h-value which implies that adding type θ_k to these latter types' matching sets never reduces the platform's payoff.

In turn, consider now an agent with type $\theta_k = (\sigma_k, v_k)$ such that $\varphi_k^h(v_k) < 0$. It is also easy to see that it is never optimal to assign to this type a matching set $\mathbf{s}_k(\sigma_k, v_k)$ that contains agents from

[15] An increasing marginal utility for meeting new agents may result from possible complementarities in socio-economic interactions, as in the case of a team where the productivity of each member increases with the average productivity of the members.

side $l \neq k$ whose φ_l^h-value is also negative. The reason is that matching two agents with negative values (or, alternatively virtual values) for matching can only decrease the platform's payoff.

Building on the observations above, the proof of Proposition 1 establishes the optimality of a threshold structure for either one of the two scenarios covered by Condition TP. The arguments in the two scenarios are symmetric. Consider first the case where g is weakly concave on both sides and pick a type θ_k from side k with $\varphi_k^h(v_k) > 0$. Imagine that the platform wanted to assign to this type a matching set whose total salience

$$q = |\mathbf{s}|_l > \int_{\Sigma_l \times [r_l^h, \overline{v}_l]} \sigma_l dF_l(\sigma_l, v_l)$$

exceeds the salience of those agents from side l with a non-negative φ_l^h-value (i.e., for whom $v_l \geq r_l^h$). The combination of the assumptions that (i) salience and values are weakly positively affiliated, (ii) g_l are weakly concave and (iii) both salience and values for matching are private information, then implies that the least costly way to do so is to match type θ_k to all agents from side l whose $\varphi_l^h(v_l)$ is the least negative, irrespective of their salience. This is because (a) these latter types are, on average, the most salient ones, (b) using the same agents with a negative φ_l^h-value intensively is less costly than using different agents with a negative φ_l^h-value, and (c) asymmetric information poses restrictions on how the matching sets can vary with the agents' reported types[16]. This, in turn, means that type θ_k's matching set takes the form $\Sigma_l \times [t_k(v_k), \overline{v}_l]$, where the threshold $t_k(v_k)$ is computed so that

$$\int_{\Sigma_l \times [t_k(v_k), \overline{v}_l]} \sigma_l dF_l(\sigma_l, v_l) = q. \tag{9}$$

In other words, starting from any incentive compatible matching rule, one can construct a threshold rule that weakly improves upon the original one. The idea is that threshold rules *minimize the costs of cross-subsidization* by delivering to those agents who play the role of consumers (i.e., whose φ_k^h-value is nonnegative) matching sets of high quality in the most economical way. The formal proof in the Appendix uses results from the theory of stochastic orders (in particular, the monotone concave order) to verify the heuristics above.

Next, consider the case where g is weakly convex on both sides and pick a type θ_k from side k with $\varphi_k^h(v_k) < 0$. Imagine that the platform wanted to assign to this type a matching set $\mathbf{s}$ of salience $|\mathbf{s}|_l > 0$. The combination of the assumptions that (i) salience and values are weakly negatively affiliated, (ii) $g_l(\cdot)$ are weakly convex and (iii) both salience and value for matching are private information, then implies that the most profitable way of using type θ_k as an input in the matching process is to match him to those agents from side l with the highest positive φ_l^h-value, irrespective of their salience. This is because (a) these latter types are the ones that benefit the most from interacting with type θ_k (they have the highest $\varphi_l^h(v_l) > 0$ values and, by virtue of incentive compatibility, the largest matching sets and hence, by the convexity of g_l, the highest marginal utility)

[16] In particular, it imposes that the total salience of the assigned set can vary with each agent's own salience at most over a countable set of reports about values for matching.

and (b) these latter types are, on average, the least salient ones and hence exert the lowest negative externalities on type θ_k (recall that $\varphi_k^h(v_k) < 0$). Once again, this in turn means that type θ_k's matching set takes the form $\Sigma_l \times [t_k(v_k), \overline{v}_l]$, where the threshold $t_k(v_k)$ is computed according to (9). Note that, contrary to the first scenario considered above, the reason why a threshold structure is optimal in the second scenario is that it *maximizes the profits of cross-subsidization.*

Lastly note that the result that the optimal matching rules discriminate only along the v-dimension is not a mere consequence of incentive compatibility: the latter property only requires that the salience of the matching set be nondecreasing in v, thus permitting the *composition* of the matching set to depend on salience. Likewise, the optimality of threshold rules does not follow directly from incentive compatibility, for there exist multiple rules that are monotone (and hence implementable) and that do not have a threshold structure (for example, partitional rules).

Given the result in Proposition 1, hereafter we restrict attention to mechanisms whose matching rule takes the form given in (8). Letting $\hat{g}_k : V_l \to \mathbb{R}_+$ denote the functions defined by

$$\hat{g}_k(v_l) \equiv g_k\left(\int_{v_l}^{\overline{v}_l} \int_{\Sigma_l} \sigma dF_l(\sigma, v)\right),$$

$k, l = A, B$, $l \neq k$, the platform's problem can then be recasted in the following simplified manner. The platform's problem consists in choosing a pair of exclusion thresholds $(\omega_k^h)_{k\in\{A,B\}}$ and a pair of non-increasing threshold functions $(t_k^h(\cdot))_{k\in\{A,B\}}$ so as to maximize the objective

$$\Omega^h(M) = \sum_{k=A,B} \int_{\omega_k^h}^{\overline{v}_k} \hat{g}_k(t_k^h(v_k)) \cdot \varphi_k^h(v_k) dF_k^v(v_k) \tag{10}$$

subject to the reciprocity constraint that

$$t_k^h(v_k) = \inf\{v_l : t_l^h(v_l) \leq v_k\} \tag{11}$$

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

In order to further investigate the properties of optimal matching rules, the next definition extends to our two-sided matching setting the notion of separating schedules, as it appears, for example, in Maskin and Riley (1984).

Definition 1 ***(separation)*** *The h-optimal matching rule entails separation if there exists a (positive measure) set $\hat{V}_k \subset V_k$ of values such that, for any $v_k, v_k' \in \hat{V}_k$ $t_k^h(v_k) \neq t_k^h(v_k')$. The rule is maximally separating if $t_k^h(\cdot)$ is strictly decreasing over $[\omega_k^h, t_l^h(\omega_l^h)]$ (which, hereafter, we call the "separating range"). It entails exclusion at the bottom on side k if $\omega_k^h > \underline{v}_k$ and bunching at the top on side k if $t_l^h(\omega_l^h) < \bar{v}_k$.*

We are now ready to characterize the optimal matching rules for both welfare- and profit-maximizing platforms. This characterization is obtained by assuming that the following condition holds, which extends the standard monotonicity of virtual values to two-sided matching environments.

Condition 2 ***[MR] Match Regularity:*** *The functions* $\psi_k^h : V_k \to \mathbb{R}$ *given by*

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

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

To understand the condition above, take the case of profit-maximization, $h = P$. The numerator in $\psi_k^h(v_k)$ accounts for the effect on the platform's revenue of an agent from side k with value v_k as a *consumer* (as his virtual value $\varphi_k^h(v_k)$ is proportional to the marginal revenue produced by this agent). In turn, the denominator accounts for the effect on the platform's revenue 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 from side k with value above v_k). Therefore, the above match regularity condition requires that the value of an agent as a consumer (as captured by his virtual value) increases faster than his contribution as an input.

Note that, when $g_k(\cdot)$ is linear, we have that $\psi_k^h(v_k) = \varphi_k^h(v_k)/\mathbb{E}\left[\tilde{\sigma}_k | \tilde{v}_k = v_k\right]$. Therefore, condition MR is immediately satisfied when $\mathbb{E}\left[\tilde{\sigma}_k | \tilde{v}_k = v_k\right]$ is nonincreasing in v_k (which is true, for example, when $\tilde{\sigma}_k$ and $\tilde{v}_k$ are weakly negatively affiliated). When $\tilde{\sigma}_k$ and $\tilde{v}_k$ are strictly positively affiliated, MR holds provided that expected salience does not increase "too fast" with an agent's value for matching relative to his virtual value.

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

$$\triangle_k^h(v_k, v_l) \equiv -\hat{g}_k'(v_l) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l'(v_k) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l), \tag{12}$$

for $k, l \in \{A, B\}$, $l \neq k$. Note that $\triangle_A^h(v_A, v_B) = \triangle_B^h(v_B, v_A)$ represents the marginal effect on the platform's objective 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 value v_A is matched to all agents from side B with value above v_B and each agent from side B with value v_B is matched to all agents from side A with value above v_A.

Importantly, condition MR implies that $\triangle_k^h(v_k, v_l)$ satisfies the following single-crossing property: whenever $\triangle_k^h(v_k, v_l) \geq 0$, we have that $\triangle_k^h(v_k, \hat{v}_l) > 0$ for all $\hat{v}_l > v_l$ and $\triangle_k^h(\hat{v}_k, v_l) > 0$ for all $\hat{v}_k > v_k$. As established in Proposition 2 below, this single-crossing property rules out nonmonotonicities in the optimal matching rule. In this sense, MR is the analog of Myerson's standard regularity condition in two-sided matching problems.

Proposition 2 ***(optimal rules under vertical differentiation)*** *Assume Conditions TP and MR hold. Then, for both* $h = W$ *and* $h = P$*, the h-optimal matching rules are such that* $\mathbf{s}_k^h(v_k) = \Theta_l$ *for all* $v_k \in V_k$, $k = A, B$ *(that is, each agent from each side of the market is matched to any other agent from the opposite side) if* $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$. *When, instead,* $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$, *the h-optimal matching rule*

(i) is maximally separating;

(ii) entails bunching at the top on side k and no exclusion at the bottom on side l if $\triangle_k^h(\bar{v}_k, \underline{v}_l) > 0$;

(iii) entails exclusion at the bottom on side l and no bunching at the top on of side k if $\triangle_k^h(\bar{v}_k, \underline{v}_l) < 0$;[17]

(iv) is characterized by a threshold function $t_k^h(\cdot)$ that satisfies the following Euler equation

$$\triangle_k^h(v_k, t_k^h(v_k)) = 0, \tag{13}$$

for any v_k in separating range $[\omega_k^h, t_l^h(\omega_l^h)]$ (from this equation, one obtains $t_k^h(v_k) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v_k)\right)$).

To illustrate, 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 from side k with positive virtual values are matched to all agents from side l with positive virtual values, plus a measure of agents with negative virtual values (cross-subsidization). The optimal level of cross-subsidization is then determined by the Euler equation (13). As explained above, this equation equalizes the marginal benefit $-\hat{g}_k'(t_k^P(v_k)) \cdot \varphi_k^P(v_k) \cdot f_k^v(v_k)$ of expanding the matching set of an agent from side k with virtual value $\varphi_k^P(v_k) > 0$ who is already matched to all agents from side l with value above $t_k^P(v_k)$, with the marginal cost $-\hat{g}_l'(v_k) \cdot \varphi_l^P(t_k^P(v_k)) \cdot f_l^v(t_k^P(v_k))$ of expanding the matching set of any agent from side l with value $v_l = t_k^P(v_k)$ who is already matched to all agents from side k with value above $v_k = t_l^P(v_l)$, as required by reciprocity (recall that $\varphi_l^P(t_k^P(v_k)) < 0$).

Intuitively, agents from each side of the market are endogenously partitioned in two groups. Those agents with positive virtual values (equivalently, with values $v_k \geq r_k^P$) 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 negative virtual values (equivalently, with value $v_k < r_k^P$) play the role of *inputs* in the matching process, providing utility to those agents from the opposite side 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 $l \neq k$ has been exhausted. This is illustrated in the next example.

Example 4 ***(linear network externalities for quantity)*** Consider the case of linear network externalities, as described in Example 1 above, and assume that values v_k are uniformly distributed over $[\underline{v}_k, \bar{v}_k]$ (independently across sides). The welfare-maximizing rule matches each agent from each side to any other agent from the opposite side if $\triangle_A^W(\underline{v}_A, \underline{v}_B) = \underline{v}_A + \underline{v}_B \geq 0$, and is maximally separating otherwise. In turn, the profit-maximizing rule matches each agent from each side to any other agent from the opposite side if $\triangle_A^P(\underline{v}_A, \underline{v}_B) = \triangle_A^W(\underline{v}_A, \underline{v}_B) - [(\bar{v}_A - \underline{v}_A) + (\bar{v}_B - \underline{v}_B)] \geq 0$ and is

[17] In the knife-edge case where $\triangle_k^h(\bar{v}_k, \underline{v}_l) = 0$, the h-optimal rule entails neither bunching at the top on side k nor exclusion at the bottom on side l.

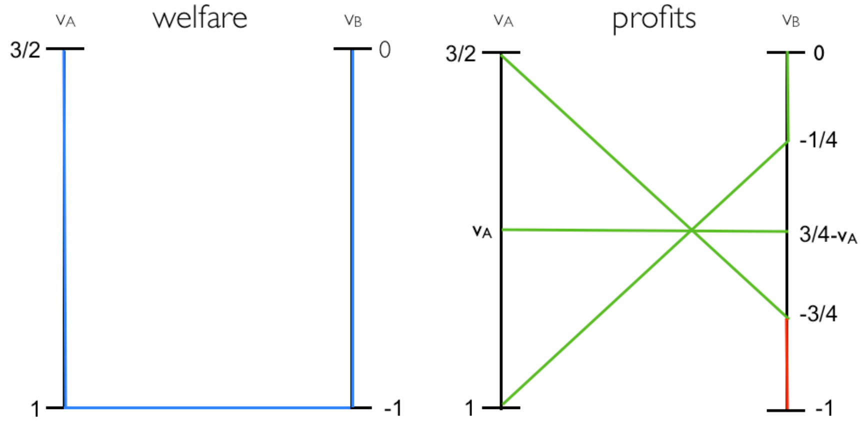

Figure 2: The optimal matching rules under linear network externalities for quantity when each agent from side A has a value drawn from a uniform distribution on $[1, 3/2]$, while each agent from side B has a value drawn from a uniform distribution on $[-1, 0]$. Notice that, under profit-maximization, there is bunching at the top and exclusion at the bottom on side B.

maximally separating otherwise. When the h-optimal rule entails separation, the threshold function is given by

$$t_k^h(v_k) = \mathbf{1}_h \cdot \frac{\overline{v}_k + \overline{v}_l}{2} - v_k.$$

\\

Example 5 ***(supermodular matching values)*** Consider the environment with supermodular matching values, as in Example 2. Each agent from each side has a value drawn independently 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 rule matches each agent to any other agent from the opposite side, while the profit-maximizing rule entails separation and is described by the threshold function $t_k^P(v_k) = \frac{v_k \cdot \bar{v}}{4 \cdot v_k - \bar{v}}$ 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 efficient matching set.

\\

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 subset of his efficient set. As reported in the next corollary, these distortions are general properties of profit-maximizing matching mechanisms (the proof follows directly from Proposition 2).

Corollary 1 ***(distortions)*** *Assume Conditions TP and MR hold. Relative to the welfare-maximizing rule, the profit-maximizing rule*

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

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* $v_k \geq \omega_k^P$, $k = A, B$.

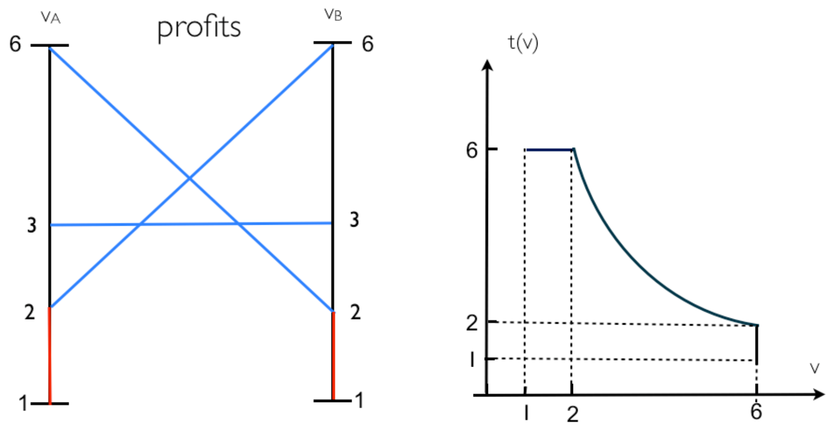

Figure 3: The profit-maximizing matching rule under supermodular matching values when each agent from each side has his value drawn independently from a uniform distribution on $[1, 6]$.

The intuition for both effects can be seen from the Euler condition (13): under profit-maximization, the platform only internalizes the cross-effects on marginal revenues (which are proportional to the virtual values), rather than the cross-effects on welfare (which are proportional to the true values). Contrary to the standard mechanism design problem, profit-maximization in a matching market may result in inefficiently small matching sets for *all* agents, including those with the highest values for matching. The reason is that, although the virtual values coincide with the true values for these latter agents, the cost of cross-subsidizing these agents is higher under profit maximization than under welfare maximization due to the inframarginal losses implied by reciprocity on the opposite side.

2.3 Pricing: The Lerner-Wilson Formula Under Vertical Differentiation

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 helped us identify 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 of matching plans. For example, business-to-business platforms typically offer menus of plans that differ by the number of matches available to each firm. Accordingly, we now show how the characterization from Proposition 2 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 markets where agents care only about the total number of agents they are matched to, which amounts to letting $\sigma_k(\cdot) \equiv 1$ for $k = A, B$. For any $q_k \in [0, 1]$, then let $\rho_k^h(q_k)$ denote the total price that each agent from side k has to pay for a matching set of size q_k under the h-optimal mechanism M^h. By optimality, the tariff $\rho_k^h(\cdot)$ has to satisfy

$$\rho_k^h(q_k) = p_k^h(\sigma_k, v_k) \text{ for all } (\sigma_k, v_k) \text{ such that } |\mathbf{s}_k^h(v_k)|_l = q_k. \tag{14}$$

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

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

denote the *individual demand* of each agent from side k whose value for matching is v_k. At any point $y_k^h(v_k)$ of differentiability of the tariff $\rho_k^h(\cdot)$, the following first-order condition must hold:

$$v_k \cdot g_k'(y_k^h(v_k)) - \frac{d\rho_k^h(y_k^h(v_k))}{dq_k} = 0. \tag{15}$$

Given the monotonicity of the individual demands 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:[18]

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

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

$$\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right) \equiv -\frac{\partial D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)}{\partial\left(\frac{d\rho_k^h}{dq_k}(q_k)\right)} \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)} = \frac{f_k^v\left(\frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}\right)}{1 - F_k^v\left(\frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}\right)} \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{g_k'(q_k)}. \tag{16}$$

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 values $v_k > 0$ (observe that $\frac{d\rho_k^h(q_k)}{dq_k}\frac{1}{g_k'(q_k)} = \left(y_k^h\right)^{-1}(q_k))$ and negative for all agents with negative values $v_k < 0$.

The next proposition recasts the first-order Euler condition (13) 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 3 ***(Lerner-Wilson under vertical differentiation)*** *In addition to Conditions TP and MR, suppose that network effects depend only on quantities ($\sigma_k(\cdot) \equiv 1$ for $k = A, B\}$), and that the h-optimal rule entails separation (i.e., $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$). Then the optimal price schedules $\rho_k^h(\cdot)$ and $\rho_l^h(\cdot)$ are differentiable and the marginal prices satisfy the Lerner-Wilson formula*

$$\underbrace{\frac{d\rho_k^h}{dq_k}(q_k) - \mathbf{1}^h \cdot \frac{\frac{d\rho_k^h}{dq_k}(q_k)}{\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)}}_{\text{net effect on side-}k\text{welfare (profits)}} + \underbrace{\frac{d\rho_l^h}{dq_l}\left(q_l^h(q_k)\right) - \mathbf{1}^h \cdot \frac{\frac{d\rho_l^h}{dq_l}\left(q_l^h(q_k)\right)}{\varepsilon_l\left(q_l^h(q_k), \frac{d\rho_l^h}{dq_l}\left(q_l^h(q_k)\right)\right)}}_{\text{net effect on side-}l\text{welfare (profits)}} = 0, \tag{17}$$

where $q_l^h(q_k) \equiv D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)$ is the aggregate demand for the q_k unit on side k at marginal price $d\rho_k^h(q_k)/dq_k$.

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

Importantly, note that the Lerner-Wilson formula (17) only depends on the *aggregate* demand for matching services on the two sides of the market. It can therefore be used for a structural estimate of the demand for matching services.

To provide intuition for (17), let us consider the case of profit-maximization (welfare maximization is analogous). 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 (17) 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 to the marginal cost of "procuring" extra agents from side l. In many-to-many matching markets, this cost is endogenous and is given by the revenue loss of expanding the matching sets of input-agents from side l who select a matching set of size $q_l^h(q_k) = D_k\left(q_k, \frac{d\rho_k^h}{dq_k}(q_k)\right)$, as implied by the threshold structure of the optimal rule. That this cost in turn depends on the elasticity of the the side-l aggregate demand reflects the fact that the platform is a monopolist on both sides of the market.[19]

2.4 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 readers, for example, affect the pricing strategies of advertising platforms (such as newspapers), since the advertisers' profits change for the same population of readers.

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

Definition 2 ***(higher attractiveness)*** *Consider a market where all side-l agents benefit from interacting with side-k agents, i.e., such that* $\underline{v}_l \geq 0$*. Side k is more attractive under the distribution* F_k *than under the distribution* $\hat{F}_k$ *if the following two conditions jointly hold: (a), for all* $v_k \in V_k$*,* $F_k^\sigma(\cdot|v_k)$ *dominates* $\hat{F}_k^\sigma(\cdot|v_k)$ *in the sense of first-order stochastic dominance, and (b)* $F_k^v = \hat{F}_k^v$*.*

The next proposition describes how the profit-maximizing matching rule changes as side k becomes more attractive.

Proposition 4 ***(increase in attractiveness)*** *In addition to Conditions TP and MR, suppose that the following conditions jointly hold: (i) network effects are linear (i.e.,* $g_A(x) = g_B(x) = x$*, all* $x \in \mathbb{R}_+$*); (ii) the P-optimal rule entails separation (i.e.,* $\triangle_k^P(\underline{v}_k, \underline{v}_l) < 0$*); and (iii) all side-l agents benefit from interacting with side-k agents (i.e.,* $\underline{v}_l \geq 0$*). Then, if the attractiveness of side k increases*

[19] In contrast to commodity markets (e.g., Maskin and Riley (1984)), quantity discounts are not a natural feature of the optimum in matching markets. In the case of linear network externalities for quantities ($g_k(x) \equiv x$), the first-order condition (15) implies that marginal prices increase with quantities, meaning that the price schedule $\rho_k^h(\cdot)$ is a convex function of q_k (i.e., the platform charges a quantity premium to those agents who play the role of consumers). 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 the sense of Definition 2), the platform switches from a matching rule $\mathbf{s}_k^P(\cdot)$ *to a matching rule* $\hat{\mathbf{s}}_k^P(\cdot)$ *such that*

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

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

Perhaps surprisingly, agents from side k can suffer from a positive shock to their attractiveness. Intuitively, an increase in the attractiveness of side-k agents alters the costs of cross-subsidization between the two sides. Recall that agents with $v_k \geq r_k^P$ are valued by the platform mainly as consumers. As these agents become more attractive, the costs of cross-subsidizing their "consumption" using agents from side l with negative virtual values increases, whereas the revenue gains on side k are unaltered. As a consequence, the matching sets of these agents shrink. The opposite is true for those agents with value $v_k \leq r_k^P$. These agents are valued by the platform mainly as inputs; as they become better inputs, their matching sets expand.

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

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

meaning that all agents from side k with value $v_k \leq r_k^P$ are necessarily better off. On the other hand, since $|\hat{\mathbf{s}}_k(v_k)|_l \leq |\mathbf{s}_k(v_k)|_l$ for all $v_k \geq r_k^P$, then either payoffs increase for all agents from side k, or there exists a threshold $\hat{\nu}_k > r_k^P$ 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 side-l agents. On the one hand, the fact that side k is more attractive implies that the payoff that each agent from side l derives from interacting with each side-k agent increases. On the other hand, by virtue of reciprocity, the matching sets for all agents with value $v_l < r_l^P$ shrink, which contributes negatively to payoffs. The net effect on the payoffs of side-l agents can thus be ambiguous and non-monotone in v_l. However, using (6), one can show that if a type θ_l with value for matching $\hat{\nu}_l \geq r_l^P$ is better off, then the same is necessarily true true for all types θ_l for which $v_l > \hat{\nu}_l$.

Next, consider the effect of an increase in the attractiveness of side k on the price schedule $\rho_k^P(q_k)$ that implements the P-optimal matching rule, as defined in 14.

Corollary 2 ***(effect of an increase in attractiveness on prices)*** *In addition to Conditions TP and MR, suppose that the following conditions jointly hold: (i) network effects are linear (i.e.,* $g_A(x) = g_B(x) = x$*, all* $x \in \mathbb{R}_+$*); (ii) the P-optimal rule entails separation (i.e.,* $\triangle_k^P(\underline{v}_k, \underline{v}_l) < 0$*); and (iii) all side-l agents benefit from interacting with side-k agents (i.e.,* $\underline{v}_l \geq 0$*). Then, if the attractiveness of side k increases (in the sense of Definition 2), the platform switches from a price* *schedule $\rho_k^P(\cdot)$ to a price schedule $\hat{\rho}_k^P(\cdot)$ such that $\hat{\rho}_k^P(q_k) \leq \rho_k^P(q_k)$ for any matching set of quality $q_k \leq \hat{q}_k$, where $\hat{q}_k > \left|\mathbf{s}_k^P(r_k^P)\right|_l = \left|\hat{\mathbf{s}}_k^P(r_k^P)\right|_l$.*

An increase in the attractiveness of side-k agents thus triggers an increase in the price that the platform charges to side-k agents for matching sets of high quality, and a decrease in the price that the platform charges to side-k agents for low-quality matching sets.

3 Vertical and Horizontal Differentiation

We now turn to markets where the agents' preferences exhibit elements of both vertical and horizontal differentiation. In such markets, 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 value positively (or, alternatively, negatively) interacting with such latter agents. Examples of markets in which horizontal differentiation is believed to play a prominent role include the market for cable TV and for online targeted advertising.

3.1 Preliminaries

The model is the same as in the previous section, except for the following adjustments. Each agent $i \in [0,1]$ from each side $k \in \{A, B\}$ has a bidimensional type $\theta_k^i = (x_k^i, v_k^i) \in \Theta_k \equiv X_k \times V_k$. 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 $X_k = [0,1]$, $k = A, B$. As in the model with vertical differentiation, the second component $v_k^i \in V_k \equiv [\underline{v}_k, \bar{v}_k]$ parametrizes the intensity of an agent's preferences for interacting with agents from the other side. For example, in the cable TV application, v_k^i captures the importance that a viewer assigns to cable TV or the importance that a channel assigns to reaching the viewers, reflecting the channel's expected advertising revenue as well as possible costs stemming from broadcasting rights. In turn, the location parameter x_k^i captures a viewer's (horizontal) tastes for different types of programming as well as a channel's broadcasting profile.

We assume that the vertical parameters v_k^i are the agents' own private information. As for the horizontal parameters x_k^i, the analysis covers both the case where they are publicly observable as well as the case 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

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

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, and weakly decreasing in $|x_k^i - x_l^j|$. The following example illustrates the type of preferences covered by the aforementioned specification.

Example 6 ***(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

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

where the parameter $\lambda_k \in [0, 1)$ measures the intensity of horizontal differentiation in the preferences of side-k 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 $v_k^i > 1$, then agent i derives a positive utility from being matched to agent j from side l if and only if j's location is sufficiently close to i's, that is, if and only if $|x_k^i - x_l^j| \leq \frac{v_k^i - 1}{\lambda_k \cdot v_k^i}$.

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 their tastes for viewers, in which case $\lambda_l = 0$. \\

The type $\theta_k^i = (x_k^i, v_k^i)$ 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, and denote by F_k^v (with density f_k^v) the marginal distribution of F_k with respect to the preference parameter v_k, and by F_k^x the marginal distribution of F_k with respect to x_k. The conditional distribution of v_k given x_k is denoted by by $F_k^v(\cdot|x_k)$, with density $f_k^v(\cdot|x_k)$.

The net utility that each agent $i \in [0,1]$ from each side $k = A, B$ with type $\theta_k^i = (x_k^i, v_k^i)$ 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

$$\int_{\mathbf{s}} u_k \left(v_k^i, |x_k^i - x_l^j| \right) d\lambda(j) - p. \tag{18}$$

As in the previous section, the platform's problem consists in choosing a matching rule and a pricing rule that jointly maximize welfare (or, alternatively, profits) subject to the reciprocity constraints and the individual rationality and incentive compatibility constraints considered in the previous section (note, however, that these constraints now depend on the observability of the location parameters x_k).

Remarks. In contrast to what assumed in the model of pure vertical differentiation of the previous section, here 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 $\theta \equiv (\theta_k^i)_{k=A,B}^{i \in [0,1]}$, $\langle \mathbf{s}, \theta \rangle_i \equiv \int_{\mathbf{s}} u_k \left(v_k^i, |x_k^i - x_l^j| \right) d\lambda(j)$ and where $g_k(\cdot)$ is an increasing and weakly concave function that possibly 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 (18). On the other hand, it is important that

$u_k\left(v_k^i, |x_k^i - x_l^j|\right)$ be 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 (18) 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 (this is in contrast to what assumed in the previous section). 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 (this possibility too was absent in the model of pure vertical differentiation of the previous section). For example, while a viewer may derive a positive utility from adding a news channel to his package, he may derive a negative utility from adding a channel that broadcasts movies with a high degree of violence or nudity in case parental control is difficult to enforce.

Finally, note that the assumption that payoffs on each side are (weakly) decreasing in the "circular" distance $|x_k^i - x_l^j|$ implies a certain type of symmetry in the payoff structure. For example, in the context of online advertising, 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. The above assumption then implies that if 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 seems reasonable in the applications discussed in this paper, but is clearly not without loss of generality.

3.2 Optimal Matching Rules

We now show that, under certain conditions, the optimal matching rules continue to have a simple threshold structure. We start by describing these conditions.

Condition 3 ***[LR] Location Regularity***: *For any $k, l \in \{A, B\}$, $l \neq k$, and any pair of locations $(x_k, x_l) \in X_k \times X_l$, the virtual values*

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

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

Condition 4 ***[I_k] Independence on side*** $k \in \{A, B\}$: *for any $(x_k, v_k) \in X_k \times V_k$, $F_k(v_k, x_k) = F_k^x(x_k) \cdot F_k^v(v_k)$.*

Condition 5 ***[S_k] Symmetry on side*** $k \in \{A, B\}$: *for any $(x_k, v_k) \in X_k \times V_k$, $F_k(x_k, v_k) = x_k \cdot F_k^v(v_k)$.*

Condition 6 ***[MS] Matching Supermodularity***: *The match value functions $u_k(\cdot, \cdot)$ are (weakly) supermodular, $k = A, B$.*

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

Condition I_k requires that the location parameter x_k and the vertical parameter v_k be independently distributed. In the cable TV application, this condition implies that knowing a viewer's "bliss point" (i.e., his preferred channel profile) 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. As shown below, this assumption simplifies the analysis by guaranteeing that the relevant incentive-compatibility constraints remain the one pertaining the vertical differentiation parameters, v_k.

Finally, Condition MS means that agents who value a lot interacting with agents from the other side (i.e., with a high v_k) suffer less from a deterioration in the attractiveness of their matching partners (that is, from an increase in the distance $|x_k - x_l|$). In the cable TV example, this assumption implies that those viewers who, in general, are keener on watching TV are also those who are more likely to watch channels whose profile is distant from their ideal one.

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

$$\begin{aligned}\triangle_k^h(\theta_k, \theta_l) &\equiv \varphi_k^h\left(v_k, |x_k - x_l|\right) + \varphi_l^h\left(v_l, |x_k - x_l|\right) \\ &= u_k\left(v_k, |x_k - x_l|\right) - \mathbf{1}_h \cdot \frac{1 - F_k^v(v_k|x_k)}{f_k^v(v_k|x_k)} \cdot \frac{\partial u_k}{\partial v}\left(v_k, |x_k - x_l|\right) \\ &+ u_l\left(v_l, |x_l - x_K|\right) - \mathbf{1}_h \cdot \frac{1 - F_l^v(v_l|x_l)}{f_l^v(v_l|x_l)} \cdot \frac{\partial u_l}{\partial v}\left(v_l, |x_l - x_k|\right)\end{aligned} \tag{19}$$

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 of adding a link between types θ_A and θ_B. We then have the following result.

Proposition 5 ***(optimal rules under horizontal differentiation)*** *Assume that, in addition to Condition LR, one of the following three conditions holds: (a) locations are public on both sides; (b) locations are private 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 the h-optimal matching rule $\mathbf{s}_k^h(\cdot)$ has the following threshold structure, $k = A, B$, $h = W, P$:

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

The threshold functions $t_k^h(\cdot|\cdot)$ are such that for any $\theta_k \in \Theta_k$, $(x_l, v_l) \in \Theta_l$, $k, l = A, B$, $l \neq k$, $h = W, P$:

1. *$t_k^h(x_l|\theta_k) = \underline{v}_l$ if $\triangle_k^h\left(\theta_k, (x_l, \underline{v}_l)\right) > 0$,*

2. *$t_k^h(x_l|\theta_k) = \bar{v}_l$ if $\triangle_k^h\left(\theta_k, (x_l, \bar{v}_l)\right) < 0$,*

3. *$t_k^h(x_l|\theta_k)$ is the unique solution to*

$$\triangle_k^h\left(\theta_k, (x_l, t_k^h(x_l|\theta_k))\right) = 0 \tag{20}$$

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

As in the case of pure vertical differentiation, when the conditions in Proposition 5 are satisfied, the optimal matching rules continue to have a threshold structure: agents with a low value for matching are matched only to those agents from the other side whose value for matching is sufficiently high. Contrary to the case of vertical differentiation, however, the thresholds are now location-specific: the optimal matching rules thus actively discriminate on the basis of mutual attractiveness (as captured by the joint location of any two agents).

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

Contrary to the case of vertical differentiation, however, the role of type $\theta_k = (x_k, v_k)$ as a consumer or as an input now varies across the x_l-locations. While type θ_k is a consumer of side-l agents located at x_l, he may be an input for side-l agents $\hat{\theta}_l = (\hat{x}_l, \hat{v}_l)$ located at $\hat{x}_l$ for whom $u_k(v_k, |x_k - \hat{x}_l|) < 0$ and $\triangle_l^W(\hat{\theta}_l, \theta_k) > 0$. That is, contrary to the case of pure vertical differentiation, the separation of agents between consumers and inputs now depends on the joint locations of any two agents.

As established in the proposition, the threshold $t_k^W(x_l|x_k, v_k)$ are weakly increasing in the distance $|x_k - x_l|$. To understand why, pick again a type $\theta_k = (x_k, v_k)$ from side k and a location x_l on side l such that $u_k(v_k, |x_k - x_l|) > 0$. Because *consumer values* $u_k(v_k, |x_k - x_l|) > 0$ go down as the partners' quality decreases (that is, as the distance $|x_k - x_l|$ increases), the *input costs* $u_l\left(t_k^W(x_l|\theta_k), |x_k - x_l|\right) < 0$ at the threshold $t_k^W(x_l|\theta_k)$ have to go down as well. In turn, this means that, as the distance $|x_k - x_l|$ increases, the marginal x_l-agent $t_k^h(x_l|\theta_k)$ in type θ_k's matching set must have a higher value v_l for matching. This logic extends to profit maximization under the supermodularity condition, which controls for how informational rents vary with the distance $|x_k - x_l|$.

The role of condition LR is to guarantee that, for any reported location x_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 the vertical dimension, v_k. Obviously, these conditions can be dispensed with when locations are public on both sides, for in this case

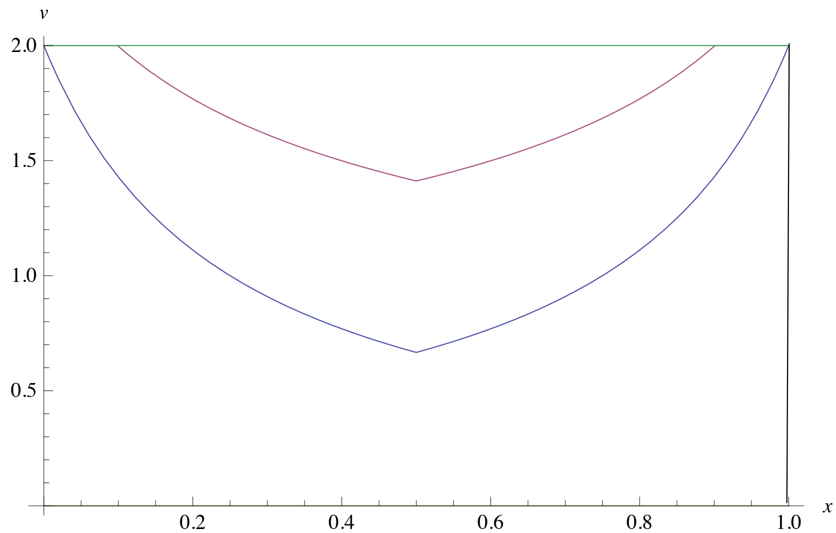

Figure 4: The welfare (blue) and profit-maximizing (red) matching rules under log utility preferences when $\lambda_A = \frac{4}{3}$, $\lambda_B = 0$, and agents from both sides have valuations and locations independently and uniformly drawn from $[0, 2]$ and $[0, 1]$, respectively.

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 the Euler condition (20), then imply that, for any report v_k about the vertical dimension, 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 v_k. This is because both the prices $p_k(x_k, v_k)$ and the sizes $\int_{\theta_l \in \mathbf{s}_k(x_k, v_k)} dF_l(\theta_l)$ of the matching sets are invariant in the reported x_l-dimension. Therefore, misreporting the bliss point (i.e., the horizontal dimension x_k) negatively affects the composition of the matching set, but does not change either its size or its price and hence is unprofitable. By the same logic, when locations are private on both sides, one has to replace Condition I_k with the stronger Condition S_k to guarantee that side-l agents find it optimal to report the locations x_l truthfully. We conjecture that the results in Proposition 5 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 following example illustrates the structure of the optimal matching rules when preferences are as in Example 6.

Example 7 (**optimal matching rules for log utility**) Suppose that preferences are as in Example 6 and that the conditions in Proposition 5 hold.[20] The welfare-maximizing matching rule is described by the following threshold function (at any point where $t_k^W(x_l|\theta_k) \in (\underline{v}_l, \bar{v}_l)$) :

$$t_k^W(x_l|\theta_k) = \frac{1}{v_k} \cdot \frac{1}{(1 - \lambda_k \cdot |x_k - x_l|) \cdot (1 - \lambda_l \cdot |x_k - x_l|)}.$$

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

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

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

The threshold functions $t_k^W(x_l|\theta_k)$ and $t_k^P(x_l|\theta_k)$ are illustrated in Figure 4. \\

3.3 Implementation: The Lerner-Wilson Formula Under Horizontal Differentiation

We now show how the characterization in Proposition 5 translates into properties of price schedules that indirectly implement the optimal matching rules.

To help the exposition, we describe the platform's pricing strategies in the context of the cable TV application. The platform offers to each viewer a menu of *packages* (also known 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 number. 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) incremental 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 then asked to choose a package and then personalize it by adding 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 number along with a baseline price and a collection of additional prices that the channel will have to pay to increase the number of viewers in each category to the desired level. Alternatively, and more realistically, 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.

We now use the results from Proposition 5 to relate the marginal prices in each package to the (category-specific) demand elasticities on each of the two sides. For the proposed indirect mechanism to implement the allocations and payments of the corresponding h-optimal direct revelation mechanism, we will let the baseline price $P_k(x_k)$ and the baseline configuration $\beta_k(x_k)$ coincide, respectively, with the equilibrium price and with the equilibrium matching set of any x_k-agents with the lowest value $\underline{v}_k$ for matching, in the corresponding h-optimal direct revelation mechanism. That is, we let

$$P_k^h(x_k) = p_k^h(x_k, \underline{v}_k) \tag{21}$$

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

$$\beta_k^h(x_k) = \mathbf{s}_k^h(x_k, \underline{v}_k). \tag{22}$$

Next, consider the additional 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 increase 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 selected 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:

$$u_k(v_k, |x_k - x_l|) - \frac{d\rho_k^h}{dq}(q_k^h(x_l|\theta_k), x_l|x_k) = 0. \tag{23}$$

Given the monotonicity of the individual demands in v_k, the x_k-*aggregate demand* for the q_k unit of x_l-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:

$$\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l|x_k), x_l|x_k\right) \equiv -\frac{\partial D_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l|x_k), x_l|x_k\right)}{\partial\left(\frac{d\rho_k^h}{dq}(q_k, x_l|x_k)\right)} \cdot \frac{\frac{d\rho_k^h}{dq}(q_k, x_l|x_k)}{D_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l|x_k), x_l|x_k\right)} \tag{24}$$

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

The next proposition uses the results in Proposition (5) 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 6 ***(Lerner-Wilson formula under horizontal differentiation)*** *Assume the conditions in Proposition 5 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 (21) and (22) respectively, along with a collection of incremental 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 raise 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*

$$\underbrace{\frac{d\rho_k^h}{dq}(q_k, x_l|x_k) - \mathbf{1}^h \cdot \frac{\frac{d\rho_k^h}{dq}(q_k, x_l|x_k)}{\varepsilon_k\left(q_k, \frac{d\rho_k^h}{dq}(q_k, x_l|x_k), x_l|x_k\right)}}_{\textit{net effect on side-kwelfare (profits)}} \tag{25}$$

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

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 formulas above is analogous to that for the formulas in Proposition 3: marginal prices are chosen so as to equalize the marginal revenue gains of expanding the number of x_l-agents in the matching set 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 are horizontally differentiated, 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.

4 Extensions

The analysis developed above can accommodate a few simple enrichments which we discuss hereafter.

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.[21]

Consider a market with purely vertically differentiated preferences. It is easy to see that the analysis developed in Section 2 extends to a setting where the platform can include no more than N plans in the menus offered to each side. Furthermore, as the number of plans increases (e.g., because menu costs decrease), the solution to the platform's problem uniformly converges to the h-optimal nested 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 matching rules of Proposition 2 are the limit as N grows large of those rules offered when the number of plans is finite. A similar conclusion applies to the model with both vertical and horizontal differentiation.

Quasi-Fixed Costs. Permitting an agent to interact with agents from the other side of the market typically involves a quasi-fixed cost. In the case of cable TV, for example, the platform must incur a cost to connect a house to the cable system. Likewise, in the case of credit cards, the platform must incur a cost to provide a merchant with the technology to operate its payment system. From the perspective of the platform, these costs are quasi-fixed, in the sense that they depend on whether or not an agent is completely excluded, but not on the composition of the agent's matching set.

The analysis developed above can easily accommodate such costs. Consider the model of pure vertical differentiation, and 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 through the following two-step procedure:

1. Step 1: Ignore quasi-fixed costs and maximize (10) among all weakly decreasing threshold functions $t_k^h(\cdot)$.

2. 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:

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

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 offering a menu of matching plans to offering a unique plan. Therefore, another testable prediction that the model delivers is that, ceteris paribus, discrimination should be more prevalent in matching markets with low quasi-fixed costs. A similar procedure can be used in the model with both vertical and horizontal differentiation (in this case, the exclusion types are location-specific).

[21] See also Wilson (1989).

Insulating Tariffs and Robust Implementation. Consider again the model with only vertical differentiation. In the direct revelation version of the matching game, 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 (6), and grants access to all agents from the other side of the market who reported a value above $t_k^h(v_k)$. This game admits one Bayes-Nash equilibrium implementing the h-optimal matching rule $\mathbf{s}_k^h(\cdot)$, along with other equilibria implementing different rules.[22]

As pointed out by Weyl (2010) in the context of a monopolistic platform offering a single plan, equilibrium uniqueness can however be guaranteed when network effects depend only on quantities (i.e., when $\sigma_k(\cdot) \equiv 1$ for $k = A, B$).[23] In the context of our model, it suffices to replace the payment rule $(p_k^h(\cdot))_{k=A,B}$ given by (6) with the payment rule

$$\varrho_k^h(v_k, (v_l^j)^{j\in[0,1]}) = v_k \cdot g_k\left(\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k\right) - \int_{\underline{v}_k}^{v_k} g_k\left(\left|\{j \in [0,1] : v_l^j \geq t_k(\tilde{v}_k)\}\right|_k\right) d\tilde{v}_k, \tag{26}$$

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 from side $l \neq k$ reporting a value 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 $(v_l^j)^{j\in[0,1]}$ by agents from the other side, the quality of the matching set for each agent from side k 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.[24]

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 the actual practices in many such markets, and draw attention to both the limitations of our analysis and, in some cases, to the somewhat puzzling implications of our results.

Business-to-business. B2B platforms work as brokers matching vendors with procurers for a fee (see, e.g., Lucking-Reyley 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

[22] 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).

[23] See also White and Weyl (2010).

[24] 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 (26) with $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_k \equiv \int_{\{j: v_l^j \geq t_k(v_k)\}} \underline{\sigma}_l d\lambda(j)$, where $\underline{\sigma}_l \equiv \min\{\sigma_l : \sigma_l \in \mathbf{\Sigma}_l\}$. However, these payments generate less revenue than the ones given in (6), implying that, in general, there is a genuine trade-off between robust (full) implementation and profit-maximization.

at increasing subscription fees. These practices appear broadly consistent with what is predicted by the version of our model with purely vertically differentiated preferences. 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 opened 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 an interesting line for future research.

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 the pure vertical-differentiation specification of our model. To see this, let values be positive on the viewers' side; assume that payoffs are linear, and that salience and willingness to pay 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, Direct TV offers five vertically differentiated (i.e., nested) English packages, four vertically differentiated Spanish packages, and eight international 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 (yet) reached the "extreme" form of customization predicted by our model with vertically and horizontally differentiated preferences, these 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 is 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 the cable TV providers. Extending the analysis to settings where (i) channels have bargaining power, and/or (ii) there are multiple providers is likely to provide new insights into the bundling practices of the cable TV industry and thus represents a promising 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 is predicted by the version of our model with purely vertically differentiated preferences, 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.[25]. Another interesting extension would consist in introducing richer forms of heterogeneity in the outside options, in line with those examined in Jullien (2000) 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

[25]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.

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 of horizontal differentiation offers a few insights in this direction, we derived our results under the (currently, unrealistic) 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 likely to introduce effects that can be relevant for this application.

Online Dating. Online dating websites typically offer menus of subscription plans that include a "join for free" option, along with (paid) plans providing a wider set of searching and interaction possibilities. These websites often suggest matches based on users' preference profiles and on geographic locations. Those users who subscribe to premium packages receive more suggestions for potential partners. Such features appear to be broadly in line with the predictions of our model with both vertical and horizontal differentiation. In particular, our prediction that, under supermodular utility functions, the thresholds increase with distance and decrease with price appears to be consistent with the practice of proposing matches whose profiles are more distant from the ideal one only if either the seeker or the proposed mate has subscribed to the premium service.

Credit Cards. Starting with Rochet and Tirole (2003), a vast literature has modeled the market for credit card payments as one where both users and merchants have linear network externalities for quantity (as in Example 1 above). The results from Propositions 1 and 2 suggest that credit card companies could reap higher profits by offering credit cards that differ in the access that they provide to merchants. To the best of our knowledge, this differentiation is not yet practiced. All major credit card companies offer cards that differ in their credit limits, mileage plans, and premium services. However, each of these cards is accepted by the same network of merchants. The presence of large quasi-fixed costs (the costs of including an agent in the network) discussed in the previous section, along with the pressure exerted by competition (which is missing in the model) might offer a partial explanation for this lack of discrimination.

The Group Design Problem. Consider now the problem of how to assign agents to different "teams" in the presence of peer effects, which is central to the theory of organizations and to personnel economics. As anticipated in the introduction, such a one-sided matching problem is a special case of the two-sided matching problems studied in this paper. To see this, note that the problem of designing nonexclusive 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) the preferences and type distributions of the two sides coincide, and (ii) the matching rule is required to be symmetric across sides, i.e., $\mathbf{s}_A(\theta) = \mathbf{s}_B(\theta)$ for all $\theta \in \Theta_A = \Theta_B$.

First, consider the case of purely vertically differentiated preferences. Under the new constraint that matching rules have to be symmetric across the two sides, maximizing (10) is equivalent to maximizing twice the objective function associated with the one-sided matching problem. As it turns out, the symmetry constraint is never binding in a two-sided matching market in which the two sides are symmetric (in which case $\psi_l^h(\cdot) = \psi_k^h(\cdot)$). This is immediate when $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$, that is, when a single network is optimal. When, instead, $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$, the characterization from Proposition 2 reveals that, at any point where the threshold rule $t_k^h(\cdot)$ is strictly decreasing, $t_k^h(v) = \left(\psi_l^h\right)^{-1}\left(-\psi_k^h(v)\right) = \left(\psi_k^h\right)^{-1}\left(-\psi_l^h(v)\right) = t_l^h(v)$. It is also easy to see that the symmetry condition is satisfied when the optimal rule entails bunching at the top.

Next, consider the case of preferences that are both vertically and horizontally differentiated. Again, from the analysis in Section 3 one can easily see that the h-optimal matching rules are naturally symmetric when the primitives on the two sides are symmetric, as one can verify from Proposition 5. All our results thus apply also to the group design problem.

6 Concluding Remarks

This paper has studied many-to-many matching in markets with both vertically and horizontally differentiated preferences. The analysis delivered two main results. First, under fairly reasonable assumptions, the optimal matching rules induce negative assortative matching at the margin. In the case of purely vertically differentiated preferences, this implies that, as matching sets expand, the marginal agents added to the matching sets are the ones with the lowest values for matching. Depending on whether values are positively or negatively correlated with salience, these marginal agents are, respectively, either the least or the most attractive ones. Similarly, when attractiveness is in the eyes of the beholders, as in the version of the model with both vertically and horizontally differentiated preferences, we find that, 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 ones of the agents under consideration. We believe that this particular form of 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 Euler 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, the Lerner-Wilson formulas that determine the marginal prices can be expressed in terms of observable market variables.

The above analysis 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. This was a reasonable starting point but is definitely inappropriate for certain markets. In advertising, for example, reaching a certain set of consumers is more profitable when competitors are blocked from reaching the same set. Extending the analysis to accommodate for congestion effects and other "same-side externalities" is nontrivial but worthwhile exploring.

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 aimed at inducing platforms "to get more agents on board" (for example, through subsidies, and in some cases the imposition of universal service obligations) are other important directions for future research.[26]

7 Appendix

Proof of Proposition 1. If $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, then it is immediate from (7) 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 (8).

Thus consider the situation where $\varphi_k^h(\underline{v}_k) < 0$ for some $k \in \{A, B\}$. Define $\Theta_k^{h+} \equiv \{\theta_k = (\sigma_k, v_k) : \varphi_k^h(v_k) \geq 0\}$ the set of types θ_k whose φ_k^h-value for matching is non-negative, and $\Theta_k^{h-} \equiv \{\theta_k = (\sigma_k, v_k) : \varphi_k^h(v_k) < 0\}$ the set of types with strictly negative φ_k^h-values.

Let $\mathbf{s}'_k(\cdot)$ be any implementable matching rule. We will show that when Condition TP holds, starting from $\mathbf{s}'_k(\cdot)$, one can construct another implementable matching rule $\hat{\mathbf{s}}_k(\cdot)$ that satisfies the threshold structure described in (8) and that weakly improves upon the original one in terms of the platform's objective $\Omega^h(M)$.

The proof proceeds as follows. First, it establishes a couple of lemmas that will be used throughout the rest of the proof. It then considers separately the two sets of primitive conditions covered by Condition TP.

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

Proof of Lemma 1. To see this, note that incentive compatibility requires that $|\mathbf{s}_k(\sigma_k, v_k)|_l \geq |\mathbf{s}_k(\sigma'_k, v'_k)|_l$ for any (σ_k, v_k) and (σ'_k, v'_k) such that $v_k \geq v'_k$. This in turn implies that $\mathbb{E}[|\mathbf{s}_k(\tilde{\sigma}_k, v_k)|_k]$ must be nondecreasing in v_k, where the expectation is with respect to $\tilde{\sigma}_k$ given v_k. Now at any point $v_k \in V_k$ at which $|\mathbf{s}_k(\sigma_k, v_k)|_l$ depends on σ_k, the expectation $\mathbb{E}[|\mathbf{s}_k(\tilde{\sigma}_k, v_k)|_l]$ is necessarily discontinuous in v_k. Because monotone functions can be discontinuous at most over a countable set of points, this means that the salience of the matching set may vary with σ_k only over a countable subset of V_k. Q.E.D.

The next lemma introduces a property for arbitrary random variables that will turn useful to establish the results.

[26] 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.

Definition 3 *[monotone concave/convex 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 any weakly increasing and weakly concave function $g : \mathbb{R} \to \mathbb{R}$. We say that z_2 is smaller than z_1 in the monotone convex 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 any weakly increasing and weakly convex function $g : \mathbb{R} \to \mathbb{R}$.*

Lemma 2 *Part (i). Suppose that $z_1, z_2 : [a,b] \to \mathbb{R}_+$ are nondecreasing and that z_2 is smaller than z_1 in the monotone concave order. Then for any weakly increasing and weakly concave function $g : \mathbb{R} \to \mathbb{R}$ and any weakly increasing and weakly negative function $h : [a,b] \to \mathbb{R}_-$, $\mathbb{E}\left[h(\tilde{\omega}) \cdot g\left(z_1(\tilde{\omega})\right)\right] \leq \mathbb{E}\left[h(\tilde{\omega}) \cdot g\left(z_2(\tilde{\omega})\right)\right]$.*

Part (ii). Suppose that $z_1, z_2 : [a,b] \to \mathbb{R}_+$ are nondecreasing and that z_2 is smaller than z_1 in the monotone convex order. Then for any weakly increasing and weakly convex function $g : \mathbb{R} \to \mathbb{R}$ and any weakly increasing and weakly positive function $h : [a,b] \to \mathbb{R}_+$, $\mathbb{E}\left[h(\tilde{\omega}) \cdot g\left(z_1(\tilde{\omega})\right)\right] \geq \mathbb{E}\left[h(\tilde{\omega}) \cdot g\left(z_2(\tilde{\omega})\right)\right]$.

Proof of Lemma 2. Consider first the case where z_2 is smaller than z_1 in the monotone concave order, g is weakly increasing and weakly concave and h is weakly increasing and weakly negative. Let $(h^n)_{n \in \mathbb{N}}$ be the family of weakly increasing and weakly negative step functions $h^n : [a,b] \to \mathbb{R}$, where n is the number of steps. Because z_2 is smaller than z_1 in the monotone concave order, the inequality in the lemma is obviously true for any one-step negative function h^1. Induction then implies that it is also true for any n-step negative function h^n, any $n \in \mathbb{N}$. Because the set of weakly increasing and weakly negative step functions is dense (in the topology of uniform convergence) in the set of weakly increasing and weakly negative functions, the result follows. Similar arguments establish part (ii) in the lemma. Q.E.D.

The rest of the proof considers separately the two sets of primitive conditions covered by Condition TP.

Case 1 *Consider markets in which the following primitive conditions jointly hold for $k = A, B$: (1a) the functions $g_k(\cdot)$ are weakly concave; (1b) the random variables $\tilde{\sigma}_k$ and $\tilde{v}_k$ are weakly positively affiliated.*

Let $\mathbf{s}'_k(\cdot)$ be the original rule and for any $\theta_k \in \Theta_k^{h+}$, let $\hat{t}_k(v_k)$ be the threshold defined as follows:

1. If $|\mathbf{s}'_k(\theta_k)|_l \geq \left|\Theta_l^{h+}\right|_l$, then let $\hat{t}_k(v_k)$ be such that
$$|\Sigma_l \times [\hat{t}_k(v_k), \bar{v}_l]|_l = \left|\mathbf{s}'_k(\theta_k)\right|_l .$$
2. If $|\mathbf{s}'_k(\theta_k)|_l \leq \left|\Theta_l^{h+}\right|_l = |\Theta_l|_l$, then $\hat{t}_k(v_k) = \underline{v}_l$.
3. If $0 < |\mathbf{s}'_k(\theta_k)|_l \leq \left|\Theta_l^{h+}\right|_l < |\Theta_l|_l$, then let $\hat{t}_k(v_k) = r_l^h$ (note that in this case $r_l^h \in (\underline{v}_l, \bar{v}_l)$).

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

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

By construction, $\hat{\mathbf{s}}_k(\cdot)_k$ is implementable. Moreover, $g_k\left(|\hat{\mathbf{s}}_k(\theta_k)|_l\right) \geq g_k\left(|\mathbf{s}'_k(\theta_k)|_l\right)$ for all $\theta_k \in \Theta_k^{h+}$, implying that for $k = A, B$,

$$\int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\sigma_k, v_k)|_l\right) dF_k(\sigma_k, v_k) \geq \int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l\right) dF_k(\sigma_k, v_k). \tag{27}$$

Below, we show that the matching rule $\hat{\mathbf{s}}_k(\cdot)$ also reduces the costs of cross-subsidization, relative to the original matching rule $\mathbf{s}'_k(\cdot)$. That is,

$$\int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l\right) dF_k(\sigma_k, v_k) \leq \int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(\sigma_k, v_k)|_l\right) dF_k(\sigma_k, v_k). \tag{28}$$

We start with 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 $z_1(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\mathbf{s}'_k(\tilde{\sigma}_k, v_k)|_l \,|v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\hat{\mathbf{s}}_k(\tilde{\sigma}_k, v_k)|_l \,|v_k\right]$, 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 values and salience, (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 $x \in [\underline{v}_k, r_k^h]$,

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

or, equivalently,

$$\int_{\underline{v}_k}^{x} z_1(v_k) dF_k^v(v_k) \geq \int_{\underline{v}_k}^{x} z_2(v_k) dF_k^v(v_k). \tag{29}$$

The result in the lemma clearly holds if for all $v_k \in [\underline{v}_k, r_k^h]$, $z_1(v_k) \geq z_2(v_k)$. Thus consider the case where $z_1(v_k) < z_2(v_k)$ for some $v_k \in [\underline{v}_k, r_k^h]$, and 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 $T \in \mathbb{N} \cup \{\infty\}$) subintervals of $[\underline{v}_k, r_k^h]$ 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 $[\underline{v}_k, r_k^h]$. 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

$$\int_{\underline{v}_k}^{r_k^h} \dot{z}_2(v_k)dF_k^v(v_k)/F_k^v(r_k^h) = \int_{\underline{v}_k}^{r_k^h} z_2(v_k)dF_k^v(v_k)/F_k^v(r_k^h). \tag{30}$$

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 (30), 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 (28). The results above imply that

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

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, $g_k\left(\left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l\right)$ is invariant in σ_k except over a countable subset of $[\underline{v}_k, r_k^h]$, as shown in Lemma 1. The first inequality follows from part (i) of 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 σ_k except over a countable subset of $[\underline{v}_k, r_k^h]$. The series of equalities and inequalities above establishes (28), as we wanted to show.

Combining (27) with (28) establishes the result that the threshold rule $\hat{\mathbf{s}}_k(\cdot)$ improves upon the original rule $\mathbf{s}'_k(\cdot)$ in terms of the platform's objective, thus proving the result in the proposition for the case of markets that satisfy conditions (1a) and (1b) in Condition TP.

Next, consider markets satisfying conditions (2a) and (2b) in Condition TP.

Case 2 *Consider markets in which the following primitive conditions jointly hold for $k = A, B$: (2a) the functions $g_k(\cdot)$ are weakly convex; (2b) the random variables $\tilde{\sigma}_k$ and $\tilde{v}_k$ are weakly negatively affiliated.*

Again, let $\mathbf{s}'_k(\cdot)$ be an arbitrary (implementable) rule and for any $\theta_k \in \Theta_k^{h-}$, let $\hat{t}_k(v_k)$ be the threshold defined as follows:

1. If $|\Theta_l|_l > |\mathbf{s}'_k(\theta_k)|_l \geq \left|\Theta_l^{h+}\right|_l > 0$, then let $\hat{t}_k(v_k) = r_l^h$ (note that in this case $r_l^h \in (\underline{v}_l, \bar{v}_l)$);

2. If $|\mathbf{s}'_k(\theta_k)|_l \geq \left|\Theta_l^{h+}\right|_l = 0$, then let $\hat{t}_k(v_k) = \bar{v}_l$;

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

4. If $0 \leq |\mathbf{s}'_k(\theta_k)|_l < \left|\Theta_l^{h+}\right|_l$, then let $\hat{t}_k(v_k)$ be such that

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

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

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

By construction, $\hat{\mathbf{s}}_k(\cdot)_k$ is monotone and invariant in s_k and hence implementable. Moreover, we have that $|\hat{\mathbf{s}}_k(\theta_k)|_l \leq |\mathbf{s}'_k(\theta_k)|_l$ for all $\theta_k \in \Theta_k^{h-}$. This implies that, for $k = A, B$,

$$\int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot |\hat{\mathbf{s}}_k(\sigma_k, v_k)|_l \, dF_k(\sigma_k, v_k) \geq \int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot \left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l dF_k(\sigma_k, v_k). \tag{31}$$

The arguments below show that the new matching rule $\hat{\mathbf{s}}_k(\cdot)$, relative to $\mathbf{s}'_k(\cdot)$, also increases the surplus from the positive $\varphi_k^h(v_k)$-agents, $k = A, B$ (recall that, by assumption, there exists at least one side $k \in \{A, B\}$ for which $\varphi_k^h(v_k) > 0$ for v_k high enough, $h = P, W$). That is, for any side $k \in \{A, B\}$ for which $\Theta_k^{h+} \neq \varnothing$,

$$\int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot |\hat{\mathbf{s}}_k(\sigma_k, v_k)|_l \, dF_k(\sigma_k, v_k) \geq \int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot \left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l dF_k(\sigma_k, v_k) \tag{32}$$

We start with the following result.

Lemma 4 *Consider the two random variables $z_1, z_2 : [r_k^h, \bar{v}_k] \to \mathbb{R}_+$ given by $z_1(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\hat{\mathbf{s}}_k(\tilde{\sigma}_k, v_k)|_l \,|v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\mathbf{s}'_k(\tilde{\sigma}_k, v_k)|_l \,|v_k\right]$, where the distribution over $[r_k^h, \bar{v}_k]$ is given by $\frac{F_k^v(v_k) - F_k^v(r_k^h)}{1 - F_k^v(r_k^h)}$. Then z_2 is smaller than z_1 in the monotone convex order.*

Proof of Lemma 4. From (i) the construction of $\hat{\mathbf{s}}_k(\cdot)$, (ii) the assumption of negative affiliation between values and salience, (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 $x \in [r_k^h, \bar{v}_k]$,

$$\int_x^{\bar{v}_k} \int_{\mathbf{\Sigma}_k} |\hat{\mathbf{s}}_k(\sigma_k, v_k)|_l \, dF_k(\sigma_k, v_k) \geq \int_x^{\bar{v}_k} \int_{\mathbf{\Sigma}_k} \left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l dF_k(\sigma_k, v_k),$$

or, equivalently,

$$\int_x^{\bar{v}_k} z_1(v_k) dF_k^v(v_k) \geq \int_x^{\bar{v}_k} z_2(v_k) dF_k^v(v_k).$$

The result in the lemma clearly holds if for all $v_k \in [r_k^h, \bar{v}_k]$, $z_1(v_k) \geq z_2(v_k)$. Thus consider the case where $z_1(v_k) < z_2(v_k)$ for some $v_k \in [r_k^h, \bar{v}_k]$ and 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 $T \in \mathbb{N} \cup \{\infty\}$) subintervals of $[r_k^h, \bar{v}_k]$ in which $z_1(v_k) < z_2(v_k)$. Because $\int_{r_k^h}^{\bar{v}_k} z_1(v_k) dF_k^v(v_k) \geq \int_{r_k^h}^{\bar{v}_k} z_2(v_k) dF_k^v(v_k)$, 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 $[r_k^h, \bar{v}_k]$. Now construct $\dot{z}_2(\cdot)$ on $[r_k^h, \bar{v}_k]$ so that:

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

2. $\dot{z}_2(v_k) = z_2(v_k)$, for all $v_k \in \mathcal{T}$;

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

Because $\int_{r_k^h}^{\bar{v}_k} z_1(v_k) dF_k^v(v_k) \geq \int_{r_k^h}^{\bar{v}_k} z_2(v_k) dF_k^v(v_k)$, 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

$$\int_{r_k^h}^{\bar{v}_k} \dot{z}_2(v_k) dF_k^v(v_k) = \int_{r_k^h}^{\bar{v}_k} z_1(v_k) dF_k^v(v_k).$$

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

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

where the first inequality follows the fact that $z_2(v_k) \leq \dot{z}_2(v_k)$ for all $v_k \in [r_k^h, \bar{v}_k]$ and $g(\cdot)$ is weakly increasing, while the second inequality follows from the construction of $\dot{z}_2(v_k)$ and the weak convexity of $g(\cdot)$. Q.E.D.

We are now ready to prove inequality (32). The results above imply that

$$\begin{aligned}
\int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s'}_k(\sigma_k, v_k)\right|_l\right) dF_k(\sigma_k, v_k) &= \int_{r_k^h}^{\bar{v}_k} \varphi_k^h(v_k) \cdot \mathbb{E}_{\tilde{\sigma}_k}\left[g_k\left(\left|\mathbf{s'}_k(\tilde{\sigma}_k, v_k)\right|_l\right) | v_k\right] dF_k^v(v_k) \\
&= \int_{r_k^h}^{\bar{v}_k} \varphi_k^h(v_k) \cdot g_k\left(z_2(v_k)\right) dF_k^v(v_k) \\
&= \left(1 - F_k^v(r_k^h)\right) \cdot \mathbb{E}\left[\varphi_k^h(\tilde{v}_k) \cdot g_k\left(z_2(\tilde{v}_k)\right) | v_k \geq r_k^h\right] \\
&\leq \left(1 - F_k^v(r_k^h)\right) \cdot \mathbb{E}\left[\varphi_k^h(v_k) \cdot g_k\left(z_1(v_k)\right) | v_k \geq r_k^h\right] \\
&= \int_{r_k^h}^{\bar{v}_k} \varphi_k^h(v_k) \cdot g_k\left(z_1(v_k)\right) dF_k^v(v_k) \\
&= \int_{r_k^h}^{\bar{v}_k} \varphi_k^h(v_k) \cdot g_k\left(\mathbb{E}_{\tilde{\sigma}_k}\left[\left|\hat{\mathbf{s}}_k(\tilde{\sigma}_k, v_k)\right|_l | v_k\right]\right) dF_k^v(v_k) \\
&= \int_{r_k^h}^{\bar{v}_k} \varphi_k^h(v_k) \cdot g_k\left(\left|\hat{\mathbf{s}}_k(\sigma_k, v_k)\right|_l\right) dF_k(\sigma_k, v_k).
\end{aligned}$$

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, $g_k\left(\left|\mathbf{s'}_k(\sigma_k, v_k)\right|_l\right)$ is invariant in σ_k except over a countable subset of $[r_k^h, \bar{v}_k]$, as shown in Lemma 1. The first inequality follows from part (ii) of Lemma 2. The equality in the last line follows again from the fact that, by construction, $\hat{\mathbf{s}}_k(\cdot)$ is implementable, and hence invariant over σ_k, except over a countable subset of $[r_k^h, \bar{v}_k]$. The series of equalities and inequalities above establishes (32), as we wanted to show.

Combining (31) with (32) establishes that the threshold rule $\hat{\mathbf{s}}_k(\cdot)$ improves upon the original rule $\mathbf{s}'_k(\cdot)$ in terms of the platform's objective, thus proving the result in the proposition under the conditions in part 2 of Condition TP. Q.E.D.

Proof of Proposition 2. We start with the following lemma, which establishes the first part of the proposition.

Lemma 5 *Assume Conditions TP and MR hold. For $h = W, P$, the h-optimal matching rule is such that $t^h_k(v_k) = \underline{v}_l$ for all $v_k \in V_k$ if $\triangle^h_k(\underline{v}_k, \underline{v}_l) \geq 0$ and entails separation otherwise.*

Proof of Lemma 5. The proof considers separately the following three different cases.

- First, consider the case where $\varphi^h_k(\underline{v}_k) \geq 0$ for $k = A, B$, implying that $\triangle^h_k(\underline{v}_k, \underline{v}_l) \geq 0$. Because values (virtual values) 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 complete network.

- Next, consider the case where $\varphi^h_k(\underline{v}_k) < 0$ for $k = A, B$, so that $\triangle^h_k(\underline{v}_k, \underline{v}_l) < 0$. We then show that, starting from any non-separating rule, the platform can strictly increase its payoff by switching to a separating one. To this purpose, let $\hat{\omega}^h_k$ denote the threshold type corresponding to the non-separating rule so that agents from side k are excluded if $v_k < \hat{\omega}^h_k$ and are otherwise matched to all agents from side l whose value is above $\hat{\omega}^h_l$ otherwise.

  First, suppose that, for some $k \in \{A, B\}$, $\hat{\omega}^h_k > r^h_k$, where recall that $r^h_k \equiv \inf\{v_k \in V_k : \varphi^h_k(v_k) \geq 0\}$. The platform could then increase its payoff by switching to a separating rule that assigns to each agent from side k with value $v_k \geq \hat{\omega}^h_k$ the same matching set as the original matching rule while it assigns to each agent with value $v_k \in [r^h_k, \hat{\omega}^h_k]$ the matching set $[\hat{v}^{\#}_l, \overline{v}_l]$, where $\hat{v}^{\#}_l \equiv \max\{r^h_l, \hat{\omega}^h_l\}$.

  Next, suppose that $\hat{\omega}^h_k < r^h_k$ for both $k = A, B$. Starting from this non-separating rule, the platform could then increase its payoff by switching to a separating rule $\mathbf{s}^{\diamondsuit}_k(\cdot)$ such that, for some $k \in \{A, B\}$[27]

$$\mathbf{s}^{\diamondsuit}_k(v_k) = \begin{cases} [\hat{\omega}^h_l, \overline{v}_l] & \Leftrightarrow \quad v_k \in [r^h_k, \overline{v}_k] \\ [r^h_l, \overline{v}_l] & \Leftrightarrow \quad v_k \in [\hat{\omega}^h_k, r^h_k] \\ \oslash & \Leftrightarrow \quad v_k \in [\underline{v}_k, \hat{\omega}^h_k] \end{cases}.$$

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

  Finally, suppose that $\hat{\omega}^h_k = r^h_k$ for some $k \in \{A, B\}$ whereas $\hat{\omega}^h_l \leq r^h_l$ for $l \neq k$. The platform

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

could then do better by switching to the following separating 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 (but strictly below) 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-value is positive more than offsets the marginal cost of getting on board a few more agents from side k whose φ_k^h-value is negative, but sufficiently small.[28] 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-value is positive, which requires employing a separating rule.

- Finally, suppose that $\varphi_l^h(\underline{v}_l) < 0 \leq \varphi_k^h(\underline{v}_k)$. First, suppose that $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$ and that the matching rule is different from a single complete network (i.e., $t_k^h(v_k) > \underline{v}_l$ for some $v_k \in V_k$. 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 MR 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 employs a non-separating rule. First suppose that such rule entails full participation (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 separating rule that is obtained by increasing $t_k^h(\cdot)$ in a right neighborhood of $\underline{v}_k$ while leaving $t_k^h(\cdot)$ untouched elsewhere.

  Next consider the case where the original rule excludes some agents (but assigns the same matching set to each agent whose value is above $\hat{\omega}_k^h$). From the same arguments as above, for such rule 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

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

[28]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 $\varphi_k^h(r_k^h) = 0$.

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 MR 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 separating 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 a separating rule is optimal when $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. Q.E.D.

The rest of the proof shows that when, in addition to Conditions TP and MR, $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$ then the optimal separating rule satisfies properties (i)-(iv) in the proposition.

To see this, note that the h-optimal matching rule solves the following program, which we call the Full Program (P^F) :

$$P^F: \qquad \max_{\{\omega_k, t_k(\cdot)\}_{k=A,B}} \sum_{k=A,B} \int_{\omega_k}^{\overline{v}_k} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) \tag{33}$$

subject to the following constraints for $k, l \in \{A, B\}$, $l \neq k$

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{34}$$

$$t_k(\cdot) \text{ weakly decreasing,} \tag{35}$$

$$\text{and} \quad t_k(\cdot) : [\omega_k, \overline{v}_k] \to [\omega_l, \overline{v}_l] \tag{36}$$

with $\omega_k \in [\underline{v}_k, \overline{v}_k]$ and $\omega_l \in [\underline{v}_l, \overline{v}_l]$. Constraint (34) is the reciprocity condition, rewritten using the result in Proposition 1. Constraint (35) is the monotonicity constraint required by incentive compatibility. Finally, constraint (36) 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 $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$, it must be that $r_k^h > \underline{v}_k$ for some $k \in \{A, B\}$. Furthermore, from the arguments in the proof of Lemma 5 above, 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, $k = A, B$:

$$P_k^F: \qquad \max_{\omega_k, t_k(\cdot), t_l(\cdot)} \int_{\omega_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{37}$$

subject to $t_k(\cdot)$ and $t_l(\cdot)$ satisfying the reciprocity and monotonicity constraints (34) and (35), along with the following constraints:

$$t_k(\cdot) : [\omega_k, r_k^h] \to [r_l^h, \overline{v}_l], \qquad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\omega_k, r_k^h]. \tag{38}$$

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 (35) and fixes $\omega_k = \underline{v}_k$ and $\omega_l = \underline{v}_l$:

$$P_k^{Au}: \qquad \max_{t_k(\cdot),t_l(\cdot)} \int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{39}$$

subject to (34),

$$t_k(\cdot), t_l(\cdot) \text{ strictly decreasing}, \tag{40}$$

$$\text{and} \quad t_k(\cdot) : [\underline{v}_k, r_k^h] \to [r_l^h, \overline{v}_l], \quad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\underline{v}_k, r_k^h] \text{ are } \textit{bijections}. \tag{41}$$

By virtue of (40), (34) can be rewritten as $t_k(v_k) = t_l^{-1}(v_k)$. Plugging this into the objective function (39) yields

$$\int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) dv_k + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_k^{-1}(v_l)) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l) dv_l. \tag{42}$$

Changing the variable of integration in the second integral in (42) 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:

$$P_k^{Au}: \qquad \max_{t_k(\cdot)} \int_{\underline{v}_k}^{r_k^h} \left\{ \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l(v_k) \cdot \varphi_l^h(t_k(v_k)) \cdot f_l^v(t_k(v_k)) \cdot t_k'(v_k) \right\} dv_k \tag{43}$$

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

$$t_k(\underline{v}_k) = \overline{v}_l \quad \text{and} \quad t_k(r_k^h) = r_l^h. \tag{44}$$

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 (44).

Lemma 6 *P_k^R admits a piece-wise absolutely continuous maximizer $\tilde{t}_k(\cdot)$.*

Proof of Lemma 6. 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 7 *Consider the function $\eta(\cdot)$ implicitly defined by*

$$\Delta_k^h(v_k, \eta(v_k)) = 0. \tag{45}$$

Let $\tilde{v}_k \equiv \inf\{v_k \in [\underline{v}_k, r_k^h] :(45) \text{ admits a solution}\}$. The solution to P_k^R is given by

$$\tilde{t}_k(v_k) = \begin{cases} \overline{v}_l & \text{if } v_k \in [\underline{v}_k, \tilde{v}_k] \\ \eta(v_k) & \text{if } v_k \in (\tilde{v}_k, r_k^h]. \end{cases} \tag{46}$$

Proof of Lemma 7. From Lemma 6, 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 (45). Condition MR ensures that (i) there exists a $\tilde{v}_k \in [\underline{v}_k, r_k^h)$ such that (45) 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 $\tilde{v}_k > \underline{v}_k$, (45) 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 (43) that $\tilde{t}_k(v_k) = \bar{v}_l$ for all $v_k \in [\underline{v}_k, \tilde{v}_k]$.

It remains to show that $\tilde{t}_k(v_k) = \eta(v_k)$ for all $v_k \in [\tilde{v}_k, r_k^h]$. Because the objective function in P_k^R is not concave in (t_k, t_k') for all v_k, we cannot appeal to standard sufficiency arguments. Instead, using the fact that the Euler equation is a necessary optimality condition for interior points, we will prove that $\tilde{t}_k(v_k) = \eta(v_k)$ by arguing that there is no function $\hat{t}_k(\cdot)$ 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 (43) over the interval (v_k^1, v_k^2) under such step function is zero. Instead, an interior control $t_k(\cdot) : (v_k^1, v_k^2) \to (r_l^h, \overline{v}_l)$ over the same interval with derivative

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

for all $v_k \in (v_k^1, v_k^2)$ yields a strictly positive value. This proves that the solution to P_k^R must indeed satisfy the Euler equation (45) 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 (46). Q.E.D.

Denote by $\max\{P_k^R\}$ the value of program P_k^R (i.e., the value of the objective (43) evaluated under the control $\tilde{t}_k^h(\cdot)$ defined in Lemma 7). 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 8 $\sup\{P_k^F\} = \sup\{P_k^{Au}\} = \max\{P_k^R\}$.

Proof of Lemma 8. 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 (34), (35) and (38), 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 (34), (40) and (41). 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 7. 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 7. Finally, given $t_k^h(\cdot) : [\omega_k^h, r_k^h] \to [r_l^h, \bar{v}_l]$, let $t_l^k(\cdot) : [r_l^h, \bar{v}_l] \to [\omega_k^h, r_k^h]$ be the unique function that satisfies (34). It is clear that the tripe $\omega_k^h, t_k^h(\cdot), t_l^h(\cdot)$ constructed this way satisfies conditions (34), (35) and (38), and is therefore a feasible candidate for program P_k^F. It is also immediate that the value of the objective (37) 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 8, 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 $\left\{\omega_k^h, t_k^h(\cdot)\right\}_{k\in\{A,B\}}$ to program P_F.

By inspection, it is easy to see that the corresponding rule is maximally separating. Furthermore, from the arguments in Lemma 7, 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 MR 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 MR, 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 3. Let $y_k^h(v_k) \equiv \left|\mathbf{s}_k^h(v_k)\right|_l$ denote the size of the matching set that each agent with value v_k obtains under the mechanism M^h. Using (6), for any $q_k \in y_k^h(V_k)$, i.e., for any q_k induced by M^h,

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

where $\left(y_k^h\right)^{-1}(q_k) \equiv \inf\{v_k : y_k^h(v_k) = q_k\}$ is the generalized inverse of $y_k^h(\cdot)$. It follows from Proposition 2 that $\left(y_k^h\right)^{-1}(q_k)$ is strictly increasing and differentiable at any q_k in the image of the separating range, i.e., for any $q_k \in [\left|\mathbf{s}_k^h(\omega_k^h)\right|_l, \left|\mathbf{s}_k^h(t_l^h(\omega_l^h))\right|_l]$. 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

$$\frac{d\rho_k^h}{dq_k}(q_k) = \left(y_k^h\right)^{-1}(q_k) \cdot g_k'(q_k) = v_k \cdot g_k'\left(\left|\mathbf{s}_k^h(v_k)\right|_l\right), \tag{47}$$

where $|\mathbf{s}_k^h(v_k)|_l = q_k$. Substituting the elasticity formula (16) and the marginal price formula (47) into the Lerner-Wilson formula (17) and using the same formulas for side l and recognizing that

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

for v_k such that $|\mathbf{s}_k^h(v_k)|_l = q_k$, then leads to the Euler equation (13). Q.E.D.

Proof of Proposition 4. Hereafter, we use the notation "$\hat{}$"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 2, 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 $|\hat{\mathbf{s}}_k(v_k)|_l \geq |\mathbf{s}_k(v_k)|_l$ if and only if $v_k \leq r_k^P$. Using (6), note that for all types θ_k with value $v_k \leq r_k^P$

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

Furthermore, since $|\hat{\mathbf{s}}_k(v_k)|_l \leq |\mathbf{s}_k(v_k)|_l$ 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 2. Let $y_k(v_k) \equiv \left|\mathbf{s}_k^P(v_k)\right|_l$ denote the quality of the matching set that each agent with value v_k obtains under the original mechanism, and $\hat{y}_k(v_k) \equiv \left|\hat{\mathbf{s}}_k^P(v_k)\right|_l$ the corresponding quantity under the new mechanism. Using (6), for any $q \in y_k(V_k) \cap \hat{y}_k(V_k)$, i.e., for any q offered both under M^P and $\hat{M}^P$,

$$\rho_k^P(q) = y_k^{-1}(q)q - \int_{\underline{v}_k}^{y_k^{-1}(q)} y_k(v)dv \text{ and}$$
$$\hat{\rho}_k^P(q) = \hat{y}_k^{-1}(q)q - \int_{\underline{v}_k}^{\hat{y}_k^{-1}(q)} \hat{y}_k(v)dv,$$

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

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

From the results in Proposition 4, we know that $[y_k(v_k) - \hat{y}_k(v_k)][v_k - r_k^P] \geq 0$ with $y_k(r_k^P) = \hat{y}_k(r_k^P)$. Therefore, for all $q \in y_k(V_k) \cap \hat{y}_k(V_k)$, with $q \leq y_k(r_k^P) = \hat{y}_k(r_k^P)$,

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

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

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

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

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

Because $\hat{y}_k^{-1}(z) \geq y_k^{-1}(z)$ for $z > y_k(r_k^P)$, we then conclude that the price differential $\rho_k^P(q) - \hat{\rho}_k^P(q)$, which is positive at $q = y_k(r_k^P) = \hat{y}_k(r_k^P)$, declines as q grows above $y_k(r_k^P)$. Going back to the original notation, it follows that there exists $\hat{q}_k > \left|\mathbf{s}_k^P(r_k^P)\right|_l = \left|\hat{\mathbf{s}}_k^P(r_k^P)\right|_l$ (possibly equal to $\left|\hat{\mathbf{s}}_k^P(\bar{v}_k)\right|_l$) 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.

Proof of Proposition 5. By familiar envelope arguments, a necessary condition for each type $\theta_k = (x_k, v_k) \in \Theta_k$, $k = A, B$, to prefer to report truthfully rather than reporting the true location x_k but a different value v_k is that payments satisfy

$$p_k(\theta_k) = \int_{\mathbf{s}_k(\theta_k)} u_k\left(v_k, |x_k - x_l|\right) dF_l(\theta_l) - \int_{\underline{v}_k}^{v_k} \int_{\mathbf{s}_k(x_k, y)} \frac{\partial u_k}{\partial v}\left(y, |x_k - x_l|\right) dF_l(x_l, v_l)dy - \Pi(x_k, \underline{v}_k | M^h) \tag{48}$$

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

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

Using the indicator function $\mathbf{1}_h$ and letting $\Pi(x_k, \underline{v}_k | M^h) = 0$ for all $x_k \in X_k$, $k = A, B$, we can then conveniently combine welfare and profit maximization into the following objective function:

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

where

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

It is convenient to define the indicator function $m_k(\theta_k, \theta_l) \in \{0, 1\}$ that equals one if and only if $\theta_l \in \mathbf{s}_k(\theta_k)$. Now define the following measure on the Borel sigma-algebra over $\Theta_k \times \Theta_l$:

$$\nu_k(E) \equiv \int_E m_k(\theta_k, \theta_l) dF_k(x_k, v_k) dF_l(x_l, v_l). \tag{49}$$

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,

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

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

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

Notice that the function

$$\varphi_k^h(v_k, |x_k - x_l|)$$

is strictly increasing in v_k by Condition LR. Therefore, fixing θ_k, for any $x_l \in X_l$, there exists a threshold $t_k^h(x_l|\theta_k)$ such that $\triangle_k^h(\theta_k, \theta_l) \geq 0$ if and only if $v_l \geq t_k^h(x_l|\theta_k)$. Condition LR also implies that the threshold $t_k^h(x_l|\theta_k)$ is decreasing in v_k. Moreover, because u_k weakly decreases in $|x_k - x_l|$, the threshold $t_k^W(x_l|\theta_k)$ is weakly increasing in the distance $|x_k - x_l|$. The same is true for $t_k^P(x_l|\theta_k)$ under Condition MS.

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 (48) is incentive compatible (that the mechanism is individually rational follows directly from (48)):

(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 4 ***(nested matching)*** *A matching rule $\mathbf{s}_k(\cdot)$ is said to be nested if for any $\theta_k = (x_k, v_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k)$ such that $x_k = \hat{x}_k$, 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 5 ***(ICV)*** *A mechanism M satisfies incentive compatibility along the v dimension (ICV) if for any $\theta_k = (x_k, v_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k)$ such that $x_k = \hat{x}_k$, $\Pi_k(\theta_k; M) \geq \hat{\Pi}_k(\theta_k, \hat{\theta}_k; M)$.*

It is straight-forward to prove the following result.

Lemma 9 ***(ICV)*** *A nested mechanism M satisfies ICV if and only if the following conditions jointly hold:*

1. *for any $\theta_k = (x_k, v_k)$ and $\hat{\theta}_k = (\hat{x}_k, \hat{v}_k)$ such that $x_k = \hat{x}_k$ and $v_k > \hat{v}_k$, $\mathbf{s}_k(\theta_k) \supseteq \mathbf{s}_k(\hat{\theta}_k)$,*
2. *the envelope formula (48) holds.*

It is clear that the mechanism associated with the threshold function t_k^h is nested and satisfies the monotonicity condition 1 from 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 threshold function t_k^h is nested and satisfies the monotonicity condition 1 from the lemma above. In turn, incentive compatibility on side k requires that

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

or equivalently,

$$\begin{aligned} \int_{\underline{v}_k}^{v_k} \int_{\mathbf{s}_k(x_k,y)} \frac{\partial u_k}{\partial v}\left(y, |x_k - x_l|\right) dF_l(x_l, v_l)dy \geq & \int_{\underline{v}_k}^{\hat{v}_k} \int_{\mathbf{s}_k(\hat{x}_k,y)} \frac{\partial u_k}{\partial v}\left(y, |\hat{x}_k - x_l|\right) dF_l(x_l, v_l)dy \\ & + \int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k)} [u_k\left(v_k, |x_k - x_l|\right) - u_k\left(\hat{v}_k, |\hat{x}_k - x_l|\right)] dF_l(x_l, v_l). \end{aligned} \tag{50}$$

It is easy to see that, for any $v_k, v \in V_k$, any $x_k \in X_k$

$$\int_{\mathbf{s}_k(x_k,v_k)} \frac{\partial u_k}{\partial v}\left(v, |x_k - x_l|\right) dF_l(x_l, v_l) = \int_{d \in [0,1/2]} \frac{\partial u_k}{\partial v}\left(v, d\right) dW(d; x_k, v_k), \tag{51}$$

where $W(d; x_k, v_k)$ is the measure of agents whose distance from x_k is d included in the matching set of type (x_k, v_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 (51) is invariant in x_k. That is, $W(d; x_k, v_k) = W(d; x'_k, v_k)$ for any $d \in [0, 1/2]$ any $x_k, x'_k \in X_k$, any $v_k \in V_k$. This means that

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

By the same arguments,

$$\int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k)} u_k\left(\hat{v}_k, |\hat{x}_k - x_l|\right)] dF_l(x_l, v_l) = \int_{\mathbf{s}_k(x_k,\hat{v}_k)} u_k\left(\hat{v}_k, |x_k - x_l|\right)] dF_l(x_l, v_l),$$

and

$$\int_{\mathbf{s}_k(\hat{x}_k,\hat{v}_k)} [u_k\left(v_k, |x_k - x_l|\right) dF_l(x_l, v_l) < \int_{\mathbf{s}_k(x_k,\hat{v}_k)} [u_k\left(v_k, |x_k - x_l|\right) dF_l(x_l, v_l).$$

It follows that the right hand side of (50) is smaller than

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

which is the payoff that type (x_k, v_k) obtains by announcing $(x_k, \hat{v}_k)$. That the inequality in (50) holds then follows from the fact that the mechanism satisfies ICV.

Finally, consider scenario (c). Under conditions condition S_k and condition 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 6. 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 5, $q_k^h(x_l|\theta_k)$ 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

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

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

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

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 (24) and the marginal price formula (23) into the Lerner-Wilson formula (25) and using the same formulas for side-l agents and recognizing that

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

for $\theta_k = (x_k, v_k)$ such that $q_k^h(x_l|\theta_k) = q_k$, then leads to the Euler equation (20). 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] Cesari, L., 1983, "Optimization Theory and Applications: Problems with Ordinary Differential Equations." Springer-Verlag, New York.

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

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

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

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

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

[14] 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.

[15] 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.

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

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

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

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

[20] 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.

[21] Jonhson, T., 2010, "Matching Through Position Auctions." Working Paper.

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

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

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

[25] 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.

[26] 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.

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

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

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

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

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

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

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

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

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

[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, 2010, "Imperfect Platform Competition: A General Framework." Working Paper University of Chicago.

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

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

* Northwestern University
** Tel Aviv University
*** Northwestern University