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

Discussion Paper #1578

Many-to-Many Matching and Price Discrimination

Renato Gomes
Toulouse School of Economics

Alessandro Pavan
Northwestern University

June 2014

JEL classification: D82

Keywords: vertical matching markets, many-to-many matching, asymmetric information, mechanism design, cross-subsidization

Many-to-Many Matching and Price Discrimination[*]

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

Alessandro Pavan
Northwestern University
alepavan@northwestern.edu

June 2014

Abstract

We study centralized many-to-many matching in markets where agents have private information about (vertical) characteristics that determine match values. Our analysis reveals how matching patterns reflect cross-subsidization between sides. Agents are *endogenously* partitioned into *consumers* and *inputs*. At the optimum, the costs of procuring agents-inputs are compensated by the gains from agents-consumers. We show how such cross-subsidization can be achieved through matching rules that have a simple threshold structure and are assortative in the weak-order (set inclusion) sense. We then deliver testable predictions relating the optimal matching rules and price schedules to the distribution of the agents' characteristics. The analysis has implications for the design of large matching intermediaries, such as advertising exchanges, business-to-business platforms, and online job-matching agencies.

JEL classification: D82

keywords: vertical matching markets, many-to-many matching, asymmetric information, mechanism design, cross-subsidization

[*]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 and ECARE-ULB, as well as conference participants at Columbia-Duke-Northwestern IO Conference, SED, Csef-Igier Conference, Stony Brook Game Theory Festival, ESEM, EARIE, ESWM, and ESSET. We are grateful to Jacques Crémer, Jan Eeckhout, Johannes Horner, Doh-Shin Jeon, Bruno Jullien, Patrick Legros, 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

Since Becker (1973), matching models have been used to study a variety of markets, including marriage, labor, and education, in which agents are heterogeneous in some *vertical* characteristics that determine the value of the matches (e.g., attractiveness, or skill). A robust insight from this literature is that, when matches are *one-to-one*, the matching pattern is positive assortative provided that the match value function satisfies appropriate supermodularity conditions, which depend on the presence, and nature, of frictions, and on the possibility of transfers (see, e.g., Shimer and Smith (2000), Legros and Newman (2002, 2007), Eeckhout and Kircher (2010) and the references therein).

The theoretical literature, however, devotes scant attention to the study of vertical markets in which matches are *many-to-many*. This is surprising, given the central role played in modern economies by many-to-many intermediaries, such as advertising exchanges, business-to-business platforms, and online job-matching agencies. This paper takes a step in this direction by studying centralized (and frictionless) many-to-many matching. Our analysis provides answers to the following questions: What matching patterns should arise when agents are *privately informed* about the vertical characteristics that determine match values? How does the private (profit-maximizing) provision of matching services compare with the public (welfare-maximizing) provision? How are matching allocations affected by shocks that alter the distribution of the agents' characteristics?

In the spirit of Becker (1973), we assume that agents on each side of the market are heterogeneous in some vertical characteristic. This vertical characteristic is responsible for an agent's *valuation* for *matching intensity* (defined below). This valuation can be either positive or negative. For example, while advertisers derive a positive utility from reaching a greater number of viewers, the latter may derive a negative utility from being exposed to a larger amount of advertising. Importantly, we allow the "consumer value" and the "input value" of each agent to differ. The consumer value, v, coincides with the agent's valuation for matching intensity. The input value, instead, coincides with the agent's *salience*, and is given by $\sigma(v)$. The latter determines the utility (or, alternatively, disutility) that the agent provides to those agents from the other side of the market he is matched to.[1] We consider both the possibility that an agent's salience (his input value) increases with his valuation (his consumer value) as well as the opposite case where it decreases. Importantly, while to ease the exposition we focus on the case in which salience is a deterministic function of valuations, all our results extend to settings in which salience varies stochastically with valuations and agents have private information about both dimensions.

The matching intensity experienced by an agent who is matched to a set $\mathbf{s}$ of agents from the other side is the *sum of the saliences* of all agents in the matching set. The utility of an agent with valuation v is then given by $v \cdot g(|\mathbf{s}|)$, where $|\mathbf{s}|$ is the intensity of the matching set and where the increasing

[1] We favor the expression salience to attractiveness because we allow for the possibility that certain agents may dislike interacting with agents from the other side, in which case a high salience may result in a lower attractiveness. For example, in advertising markets, more salient or, equivalently, more prominent, advertisers may be less attractive or, equivalently, more annoying, to consumers.

function $g(\cdot)$ captures the agent's preferences over matching intensity (as opposed to the agent's preferences over *individual* matches). For example, in advertising markets, an advertiser's valuation stands for its willingness to pay for viewers, whereas the salience of an advertiser captures the nuisance to viewers produced by its ads. In turn, a viewer's valuation measures the viewer's intolerance to advertising, whereas a viewer's salience captures his purchasing habits. In this example, viewers with convex nuisance costs are captured by a negative valuation v and a convex $g(\cdot)$ function.

We study the matching assignments that maximize either social welfare or profits (as many matching intermediaries are privately owned). In the tradition of Koopmans and Beckmann (1957), we consider a setting with transfers and let preferences be quasi-linear in money. Accordingly, for each side of the market, the platform chooses a *pricing rule* and a *matching rule.* The former specifies payments as a function of an agent's vertical characteristic, v. The latter, instead, specifies the composition of the matching set (i.e., the set of types from the opposite side the agent is granted access to). Along with incentive compatibility, we require only that matching mechanisms 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. The cases of welfare and profit maximization can then be treated similarly, after one replaces valuations by their "virtual"counterparts (which discount for informational rents). Bearing this observation in mind, and for the sake of clarity, the description below focuses on the case of welfare maximization.

The recurring theme of this paper is how matching patterns reflect optimal cross-subsidization between sides. Our first main result identifies primitive conditions under which optimal matching rules exhibit a *threshold structure*, according to which each agent from side A with valuation v is matched to all agents from the other side whose valuation is above a threshold $t_A(v)$. Because thresholds are decreasing in valuations, threshold rules are *assortative*, in the weak-order (set inclusion) sense:

$$\hat{v} \leq v \quad \Longrightarrow \quad \mathbf{s}(\hat{v}) \preceq \mathbf{s}(v),$$

where $\mathbf{s}(v)$ is the matching set of an agent with valuation v. Moreover, for any two matching sets, the "marginal" agents included only in the larger set are always those with the lowest valuations, among those included in either set. This implies that matching patterns exhibit *negative assortativeness at the margin*: Those agents with a low valuation are matched only to those agents from the other side whose valuation is sufficiently high.

Threshold rules are simple and quite natural. However, their optimality should not be taken for granted. Our theory identifies primitive conditions that, when combined with incentive compatibility, make threshold rules optimal. These conditions cover two alternative (and complementary) scenarios. The first scenario is one where, on each side $k = A, B$, (1a) salience is increasing in valuations (i.e., $\sigma_k(\cdot)$ is non-decreasing), and (1b) agents have diminishing marginal utility, or disutility, for matching intensity (i.e., $g_k(\cdot)$ is weakly concave). In the context of B2B matching, condition (1a) captures the idea that those firms that are willing to pay the most for business connections are, in general, seen as better partners. In turn, condition (1b) captures the idea that firms have scarce resources to evaluate the business potential of the proposed connections, in which case the marginal utility of expanding the matching sets is decreasing in matching intensity. Coupled with the monotonicity requirements of incentive compatibility, conditions (1a) and (1b) imply that the *cross-subsidization costs* (i.e., the welfare losses from matching an agent with a negative valuation to one with a positive valuation) are *minimized* under a threshold structure. This is because, under the above two conditions, among all agents with negative valuations, those with the highest valuations are (a) the most salient ones (and hence the most attractive), and (b) the ones with the lowest marginal disutility for matching intensity (and hence the "cheapest" to serve).

The second scenario is defined by conditions that are the mirror image of those above. Namely, it is one where, on each side $k = A, B$, (2a) salience is decreasing in valuations (i.e., $\sigma_k(\cdot)$ is non-increasing), and (2b) agents have increasing marginal utility (or disutility) for matching intensity (i.e., $g_k(\cdot)$ is weakly convex). Together with the monotonicity requirements of incentive compatibility, conditions (2a) and (2b) imply that the *cross-subsidization benefits* (i.e., the welfare gains from matching a positive-valuation agent to a negative-valuation one) are *maximized* under a threshold structure. This is because, under the above conditions, among all agents with positive valuations, those with the highest valuations are (a) the least salient ones (and hence the cheapest to serve), and (b) the ones with the highest marginal utility for matching intensity (and hence the most valuable to cross-subsidize).

Importantly, these conditions are "sharp": When either of these conditions is violated or, alternatively, when the vertical characteristics are observable, one can construct natural examples where threshold rules fail to be optimal.

Our second main result provides a precise characterization of the thresholds. We use variational techniques to obtain an Euler equation that equalizes (i) the marginal gains from expanding the matching sets on one side to (ii) the marginal losses that, by reciprocity, arise on the opposite side of the market. Intuitively, this equation endogenously partitions agents from each side into two groups. The first group consists of agents playing 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 playing 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 (cross-subsidization).

Our characterization results enable us to compare the matching allocations that result from the public provision of matching services (which we assumed is motivated by welfare maximization) to those that result from the private provision of such services (which we assume is motivated by profit maximization). Interestingly, profit-maximization in vertical matching markets may result in inefficiently small matching sets for *all* agents, including those "at the top" of the distribution (i.e., with the highest valuations for matching intensity). The reason is that the costs of cross-subsidizing such agents is higher under profit maximization, due to the informational rents that must be given to the agents-inputs.

Our analysis also delivers testable predictions about the effects of shocks that alter the salience of the agents. In particular, we show that a shock that increases the salience of all agents from a given side (albeit not necessarily uniformly across agents) induces a profit-maximizing platform to offer matching sets of *higher* intensity to those agents with *low* valuations and matching sets of *lower* intensity to those agents with *high* valuations. In terms of surplus, these shocks make low-valuations agents better off at the expense of high-valuation ones. Perhaps surprisingly, this result implies that a B2B platform may respond to an increase in the attractiveness of firms from one side of the market by offering fewer connections to those firms with the highest valuations for matching intensity.

Although formulated in a two-sided matching environment, all our results have implications also for *one-sided* vertical matching markets. Indeed, the one-sided environment is mathematically equivalent to a two-sided matching market where both sides have symmetric primitives and matching rules are constrained to be symmetric across sides. As it turns out, in two-sided matching markets with symmetric primitives, the optimal matching rules are naturally symmetric, in which case the latter constraint is non binding. All our results thus have implications also for such problems in organization and personnel economics that pertain to the optimal design of teams in the presence of peer effects.

The rest of the paper is organized as follows. Below, we close the introduction by briefly reviewing the most pertinent literature. Section 2 describes the model. Section 3 contains all results. Section 4 discusses a few extensions, while Section 5 concludes. All proofs are in the Appendix at the end of the document.

Related Literature

The paper is primarily related to the following literatures.

Vertical Matching Markets. Following the pioneering work of Becker (1973), Legros and Newman (2002) derive necessary and sufficient conditions for positive assortative (one-to-one) matching in settings with market imperfections (such as transaction costs or moral hazard). In turn, Shimer and Smith (2000) and Eeckhout and Kircher (2010) extend 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 or market imperfections, and consider many-to-many matching rules.

In a vertical market with one-to-one matches, 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 a one-to-one positive assortative way. Johnson (2013) 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 study *many-to-many* matching in a flexible setting where agents may differ in their consumer value (valuation) and input value (salience).

Price Discrimination. The availability of transfers and the presence of asymmetric information relates this paper to the literature on second-degree price discrimination (e.g., Mussa and Rosen (1978), Maskin and Riley (1983), Wilson (1997)). 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 implies that the cost of procuring an input is endogenous and depends in a nontrivial way on the entire matching rule.

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 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).[2]

Two-Sided Markets. Markets where agents purchase access to other agents are the focus of the literature on two-sided markets (see Rysman (2009) for a survey, and Bedre-Defolie and Calvano (2013) and Lee (2013) for recent developments). This literature, however, restricts attention to a single network or to mutually exclusive networks. Our contribution is in allowing for general matching rules, in distinguishing the agents' valuations from their salience, and in accommodating for nonlinear preferences over matching intensity.

2 Model and Preliminaries

2.1 Environment

A platform matches agents from two sides of a market. Each side $k \in \{A, B\}$ is populated by a unit-mass continuum of agents. Each agent from each side $k \in \{A, B\}$ has a type $v_k \in V_k \equiv [\underline{v}_k, \overline{v}_k] \subseteq \mathbb{R}$ that parametrizes both the agent's valuation for matching intensity, that is, the value that the agent assigns to interacting with agents from the opposite side, and the agent's "salience", which we denote by $\sigma_k(v_k) \in \mathbb{R}_+$.

Each v_k is drawn from an absolutely continuous distribution F_k (with density f_k), independently across agents. As is standard in the mechanism design literature, we assume that F_k is *regular* in the sense of Myerson (1981), meaning that the virtual valuations for matching $v_k - [1 - F_k(v_k)]/f_k(v_k)$ are continuous and nondecreasing.

[2] 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.

Given any (Borel measurable) set $\mathbf{s}$ of types from side $l \neq k$, the payoff that an agent from side $k \in \{A, B\}$ with type v_k obtains from being matched, at a price p, to the set $\mathbf{s}$ is given by

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

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

$$|\mathbf{s}|_l \equiv \int_{v_l \in \mathbf{s}} \sigma_l(v_l) dF_l(v_l) \tag{2}$$

is the *matching intensity* of the set $\mathbf{s}$.

The case where an agent from side k dislikes interacting with agents from the other side is thus captured by a negative valuation $v_k < 0$. 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 $k \in \{A, B\}$. The functions $g_k(\cdot)$, $k = A, B$, in turn capture increasing (or, alternatively, decreasing) marginal utility (or, alternatively, disutility) for matching intensity.

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

Example 1 ***(advertising exchange)*** *The platform is an advertising exchange matching advertisers from side A to online publishers from side B.*[3] *All trades and monetary transfers are mediated by the platform. The matching sets designed by the platform determine which advertisers appear in each publisher's website. The payoff of an advertiser whose ad is displayed by a set $\mathbf{s}$ of publishers is given by*

$$\pi_A(\mathbf{s}, p; v_A) = v_A \cdot \left(\int_{v_B \in \mathbf{s}} \sigma_B(v_B) dF\left(v_B\right) \right) - p$$

where $v_A \in V_A \subseteq \mathbb{R}_+$ captures the advertiser's (gross) value for having its ad displayed and where $\sigma_B(v_B)$ captures the salience (equivalently, the prominence) of a publisher with opportunity cost $v_B \in V_B \subseteq \mathbb{R}_-$.[4] *In turn, each publisher's payoff from selling capacity (equivalently, space) to a set $\mathbf{s}$ of advertisers is given by*

$$\pi_B(\mathbf{s}, p; v_B) = v_B \cdot \int_{v_A \in \mathbf{s}} dF\left(v_A\right) - p,$$

where the first term is the total opportunity cost from all impressions sold and where $p \in \mathbb{R}_-$ is the total amount of money (net of commissions) that the publisher collects through the platform. $\diamondsuit$

Example 2 ***(business-to-business platform)****. The platform is an intermediary matching firms from two sides of a market. The setting is the same as in Example 1, but with $\sigma_k(v_k) \equiv 1$ for all*

[3] Advertising exchanges are large online matching platforms where advertisers purchase advertising space offered by online publishers. The major ad exchanges include Right Media, owned by Yahoo!, and DoubleClick AdEx, owned by Google. Second generation ad exchanges include OpenX and AppNexus.

[4] The publisher's opportunity cost is the value of its best alternative use of advertising space, *outside* the matching platform (e.g., the value of promoting its own products, or selling the space through side contracts with advertisers).

$v_k \in V_k$, $k = A, B$. *In other words, businesses have preferences which are linear in the size of the matching sets, in which case*

$$\pi_k(\mathbf{s}, p; v_k) = v_k \cdot \int_{v_l \in \mathbf{s}} dF_l\left(v_l\right) - p$$

for $k = A, B$, *with* v_k *denoting the gross value that each firm assigns to reaching more firms from the other side of the market. 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 Lucking-Reyley and Spulber (2001) and Jullien (2012)).* $\diamondsuit$

The next example considers a market in which the matching values are supermodular, as in the literature on positive assortative one-to-one matching.

Example 3 ***(supermodular matching values)***. *The platform is an online intermediary matching musicians from side* A *to music venues from side* B. *The match between a musician of type* v_A *and a venue of type* v_B *yields a payoff* $\alpha \cdot (v_A \cdot v_B)$ *to the musician and a payoff* $(1-\alpha) \cdot (v_A \cdot v_B)$ *to the venue (e.g., the proceeds of each concert are split between the musician and the venue according to a generalized Nash bargaining protocol). In this specification, the salience of each agent coincides with his valuation (i.e.,* $\sigma_k(v_k) = v_k$ *for all* $v_k \in V_k$, *with* $V_k \subset \mathbb{R}_+$ *and* $g_k(x) = x$, $k = A, B$*). Each musician's payoff is then equal to*

$$\pi_A(\mathbf{s}, p; v_A) = \alpha \cdot v_A \cdot \int_{v_B \in \mathbf{s}} v_B dF_B\left(v_B\right) - p$$

whereas each venue's payoff is equal to

$$\pi_B(\mathbf{s}, p; v_A) = (1-\alpha) \cdot v_B \cdot \int_{v_A \in \mathbf{s}} v_A dF_A\left(v_A\right) - p$$

where the prices here denote the commissions paid to the platform. $\diamondsuit$

In the three examples above the marginal utility (or, alternatively, disutility) for matching intensity is constant (as preferences are linear). In contrast, the next two examples describe environments where the marginal (dis)utility naturally depends on the matching intensity experienced by the agent.

Example 4 ***(nuisance costs)***. *The platform is an online newspaper matching readers from side* A *to advertisers from side* B. *Readers dislike advertising and their intolerance is indexed by the parameter* $v_A \in V_A \subset \mathbb{R}_-$. *The nuisance generated by an advertiser with willingness to pay* $v_B \in V_B \subset \mathbb{R}_+$ *to a reader with intolerance* v_A *is given by*

$$v_A \cdot \sigma_B(v_B) \cdot \left(|\mathbf{s}|_B\right)^{\beta},$$

where $\mathbf{s}$ *is the set of ads displayed to the reader, and where* $\beta \geq 0$ *is the nuisance parameter.*[5] *The reader's total payoff is then given by*

$$\pi_A(\mathbf{s}, p; v_A) = \int_{v_B \in \mathbf{s}} v_A \cdot \sigma_B(v_B) \cdot (|\mathbf{s}|)^{\beta}\, dF\left(v_B\right) - p = v_A \cdot g_A\left(|\mathbf{s}|_B\right) - p,$$

[5] See Kaiser and Wright (2006) and Kaiser and Song (2009) for an empirical assessment of the preferences of readers vis-a-vis advertisers.

where p is the price the reader pays for the online newspaper and where the function $g_A(x) = x^{1+\beta}$ (which is strictly convex for $\beta > 0$), captures the increasing "marginal" nuisance of advertising. An increasing salience function $\sigma_B(\cdot)$ then captures the idea that advertisers with a higher willingness to pay display, on average, ads that are more annoying to the readers, whereas a decreasing $\sigma_B(\cdot)$ captures the opposite case. For simplicity, in this example, advertisers are assumed to have linear preferences as in Example 2. ◇

Example 5 ***(limited attention)***. *The platform is an online media matching viewers from side A to content providers from side B. The utility that a viewer from side A obtains from the content of a provider with willingness to pay v_B is given by*

$$\frac{v_A \cdot \sigma_B(v_B)}{(|\mathbf{s}|_B)^{\alpha}}$$

where $\sigma_B(v_B)$ is the salience of the provider's content (equivalently, its attractiveness), $\mathbf{s}$ is the viewer's matching set, $v_A \in V_A \subset \mathbb{R}_+$ is a parameter indexing the importance the viewer assigns to receiving content, and $\alpha \in [0, 1)$ is a parameter controlling for limited attention. The viewer's total payoff is then equal to

$$\pi_A(\mathbf{s}, p; v_A) = \int_{v_B \in \mathbf{s}} \frac{v_A \cdot \sigma_B(v_B)}{(|\mathbf{s}|_B)^{\alpha}} dF(v_B) - p = v_A \cdot g_A(|\mathbf{s}|_B) - p,$$

with $g_A(x) = x^{1-\alpha}$. Note that when $0 < \alpha < 1$ the function $g(\cdot)$ is strictly concave, capturing the idea that the marginal value of extra content is decreasing in the amount received, possibly due to limited attention. Again, for simplicity, the providers' preferences can be assumed to have the linear structure of Example 2, capturing the idea that providers simply want to reach as many viewers as possible. ◇

Matching Mechanisms

A matching mechanism $M \equiv \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ consists of two pairs (indexed by side) of matching and payment rules. For each type $v_k \in V_k$, the rule $p_k(\cdot)$ specifies the payment asked to an agent from side $k \in \{A, B\}$ with type v_k, while the rule $\mathbf{s}_k(\cdot) \subseteq V_l$ specifies the set of types from side $l \neq k$ included in type v_k's matching set. Note that $p_k(\cdot)$ maps V_k into $\mathbb{R}$ (both positive and negative payments are allowed), while $\mathbf{s}_k(\cdot)$ maps V_k into the Borel sigma algebra over V_l. With some abuse of notation, hereafter we will denote by $|\mathbf{s}_k(v_k)|_l$ the matching intensity of type v_k's matching set.[6]

A *matching rule* $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ is *feasible* if and only if it satisfies the following **reciprocity condition**

$$v_l \in \mathbf{s}_k(v_k) \;\Rightarrow\; v_k \in \mathbf{s}_l(v_l), \tag{3}$$

[6] Restricting attention to deterministic mechanisms is without loss of optimality under the assumptions in the model (The proof is based on arguments similar to those in Strausz (2006)). It is easy to see that restricting attention to anonymous mechanisms is also without loss of optimality given that there is no aggregate uncertainty and that individual identities are irrelevant for payoffs.

which requires that if an agent from side l with type v_l is included in the matching set of an agent from side k with type v_k, then any agent from side k with type v_k is included in the matching set of any agent from side l with type v_l.

Next, denote by $\hat{\Pi}_k(v_k, \hat{v}_k; M) \equiv v_k \cdot g_k(|\mathbf{s}_k(\hat{v}_k)|_l) - p_k(\hat{v}_k)$ the payoff that type v_k obtains when reporting type $\hat{v}_k$, and by $\Pi_k(v_k; M) \equiv \hat{\Pi}_k(v_k, v_k; M)$ the payoff that type v_k obtains by reporting truthfully. A mechanism M is *individually rational* (IR) if $\Pi_k(v_k; M) \geq 0$ for all $v_k \in V_k$, $k = A, B$, and is *incentive compatible* (IC) if $\Pi_k(v_k; M) \geq \hat{\Pi}_k(v_k, \hat{v}_k; M)$ for all $v_k, \hat{v}_k \in V_k$, $k = A, B$.

A matching rule 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.[7]

Efficiency and Profit Maximization

The welfare generated by the mechanism M is given by

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

whereas the profits generated by the mechanism M are given by

$$\Omega^P(M) = \sum_{k=A,B} \int_{V_k} p_k(v_k) dF_k(v_k). \tag{5}$$

A mechanism is efficient (alternatively, profit-maximizing) if it maximizes $\Omega^W(M)$ (alternatively, $\Omega^P(M)$) among all mechanisms that are individually rational, incentive compatible, and satisfy the reciprocity condition (3). Note that the reciprocity condition implies that the matching rule $\{\mathbf{s}_k(\cdot)\}_{k=A,B}$ can be fully described by its side-k correspondence $\mathbf{s}_k(\cdot)$.

It is standard to verify 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 matching intensity of the set $\mathbf{s}_k(v_k)$ is nondecreasing in the valuation v_k;

(ii) the equilibrium payoffs $\Pi_k(\underline{v}_k; M)$ of the agents with the lowest valuation are non-negative;

(iii) the pricing rule satisfies the envelope formula

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

It is also easy to see that in any mechanism that maximizes the platform's profits, the IR constraints of those agents with the lowest valuations bind, i.e., $\Pi_k(\underline{v}_k; M^P) = 0$, $k = A, B$. Using the expression for payments (6), it is then standard practice to rewrite the platform's profit maximization problem in a manner analogous to the welfare maximization problem. One simply needs to replace the true valuations with their virtual analogs (i.e., with the valuations discounted for informational rents).

[7] Implicit in the aforementioned specification is the assumption that the platform must charge the agents before they observe their payoffs. 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).

Formally, for any $k = A, B$, any $v_k \in V_k$, let $\varphi_k^W(v_k) \equiv v_k$ and $\varphi_k^P(v_k) \equiv v_k - [1 - F_k(v_k)]/f_k(v_k)$. 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

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

among all rules that satisfy the monotonicity constraint (i) and the reciprocity condition (3). Bearing these observations in mind, hereafter, we will say that a matching rule $\{\mathbf{s}_k^h(\cdot)\}_{k=A,B}$ is h-optimal if it solves the above h-problem. For future reference, for both $h = W, P$, we also define the *reservation value* $r_k^h \equiv \inf\{v_k \in V_k : \varphi_k^h(v_k) \geq 0\}$ when $\{v_k \in V_k : \varphi_k^h(v_k) \geq 0\} \neq \varnothing$.

3 Optimal Matching Rules

We start by introducing an important class of matching rules and by identifying tight conditions under which restricting attention to such rules entails no loss of optimality. We then proceed by studying properties of optimal rules and conclude with comparative statics.

3.1 Threshold Rules

Consider the following class of matching rules.

Definition 1 ***(threshold rules)***. *A matching rule exhibits a threshold structure (equivalently, is a threshold rule) if for any $v_k \in V_k$, $k = A, B$,*

$$\mathbf{s}_k(v_k) = \begin{cases} [t_k(v_k), \overline{v}_l] & \text{if } v_k \geq \omega_k \\ \emptyset & \text{otherwise,} \end{cases}$$

where the exclusion type $\omega_k \in V_k$ is the valuation below which types are excluded, and where the non-increasing threshold function $t_k(v_k)$ determines the matching sets.

Matching rules with a threshold structure are remarkably simple. Any type below the threshold ω_k is excluded, while a type $v_k > \omega_k$ is matched to any agent from the other side whose type is above the threshold $t_k(v_k)$, with the function $t_k(v_k)$ weakly decreasing. Matching rules with a threshold structure thus exhibit a form of negative assortativeness at the margin: Those agents with a low valuations are matched only to those agents from the opposite side whose valuation is sufficiently high. Furthermore, the matching sets are ordered across types, in the weak (inclusion) set-order sense, i.e., if $v_k < \hat{v}_k$ then $\mathbf{s}_k(v_k) \subseteq \mathbf{s}_k(\hat{v}_k)$.

The next proposition identifies primitive conditions under which threshold rules are optimal. As anticipated in the Introduction, the optimality of threshold rules is not a mere consequence of incentive compatibility. The latter simply requires the matching intensity to be nondecreasing in valuations, but imposes no restrictions on the *composition* of the matching sets. To see this, suppose, for example, that v_k is drawn uniformly from $V_k = [0,1]$ and that $\sigma_k(v_k) = 1$ for all $v_k \in V_k$, $k = A, B$. Then partitional rules of the type

$$\mathbf{s}_k(v_k) = \begin{cases} \left[\frac{1}{2}, 1\right] & \text{if} \quad v_k \in \left[\frac{1}{2}, 1\right] \\ \left[0, \frac{1}{2}\right] & \text{if} \quad v_k \in \left[0, \frac{1}{2}\right], \end{cases}$$

are clearly implementable but do not exhibit a threshold structure. Furthermore, the matching sets induced by incentive compatible rules need not be ordered or connected. For example, continue to assume that valuations are drawn uniformly from $V_k = [0,1]$ but now let $\sigma_k(v_k) = 1 - v_k$, for all $v_k \in V_k$, $k = A, B$. The following matching rule, described by its side-k correspondence, is implementable:

$$\mathbf{s}_k(v_k) = \begin{cases} \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] & \text{if} \quad v_k \in \left[1 - \frac{\sqrt{2}}{2}, 1\right] \\ \left[\frac{1}{3}, \frac{2}{3}\right] & \text{if} \quad v_k \in \left[0, 1 - \frac{\sqrt{2}}{2}\right]. \end{cases}$$

In light of the above observations, it should be clear that the optimality of threshold rules should not be taken for granted. We then proceed by identifying primitive payoff conditions that, along with incentive compatibility, make threshold rules optimal.

Condition 1 ***[TP] Threshold Primitives:*** *One of the following two sets of conditions holds:*

(a) the functions $g_k(\cdot)$ are weakly concave, and the functions $\sigma_k(\cdot)$ are weakly increasing, for both $k = A$ and $k = B$;

(b) the functions $g_k(\cdot)$ are weakly convex, and the functions $\sigma_k(\cdot)$ are weakly decreasing, for both $k = A$ and $k = B$.

Condition TP covers two alternative scenarios. The first scenario is one where, on both sides, agents have (weakly) concave preferences for matching intensity. In this case, Condition TP also requires that, on both sides, salience increases (weakly) with the valuations. The second scenario covers a symmetrically opposite situation, where agents have (weakly) convex preferences for matching intensity and salience decreases (weakly) with valuations.

To illustrate, note that the preferences in Example 1 (ad exchange) satisfy Condition TP as long as the salience function on side B is monotone (either increasing or decreasing). Likewise, the preferences in Example 2 (business-to-business) clearly satisfy Condition TP, for in this example all agents' preferences are linear and salience is constant on both sides. The preferences in Example 3 (supermodular matching values) also satisfy Condition TP, for in this case preferences are linear and salience is increasing in valuations. The preferences in Example 4 (nuisance costs) satisfy Condition TP as long as salience on side B is non-increasing in valuations, meaning that the ads of those advertisers with the highest willingness to pay are seen, on average, as being the least annoying ones. Finally, the preferences in Example 5 (limited attention) satisfy Condition TP as long as salience on side B is nondecreasing in valuations, meaning that content providers with a high willingness to pay are expected to provide, on average, the most interesting content. More generally, note that when marginal benefits and costs are constant ($g_k(\cdot)$ is linear, $k = A, B$), Condition TP accommodates either a positive or a negative relation between salience and valuations. We then have the following result.

Proposition 1 ***(optimality of threshold rules)*** *Assume Condition TP holds. Then both the profit-maximizing and the welfare-maximizing rules have a threshold structure.*

Below, we first illustrate heuristically the logic behind the arguments that lead to the result in Proposition 1 (the formal proof in the Appendix is significantly more complex and uses results from the theory of stochastic orders to verify the heuristics described below). We then show that the conditions in the proposition are tight, in the sense that, when condition TP is violated, one can construct examples where the optimal rules fail to have a threshold structure.

Sketch of the Proof of Proposition 1. Consider an agent for whom $\varphi_k^h(v_k) \geq 0$. 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 agent a matching set $\mathbf{s}_k(v_k) \supset \{v_l : \varphi_l^h(v_l) \geq 0\}$ that includes all agents from the other side whose φ_l^h-value is non-negative. This is because (i) these latter agents contribute positively to type v_k's payoff and, (ii) these latter agents have non-negative φ_l^h-values, which implies that adding type v_k to these latter agents' matching sets (as required by reciprocity) never reduces the platform's payoff.

Next, consider an agent for whom $\varphi_k^h(v_k) < 0$. It is also easy to see that it is never optimal to assign to this agent a matching set that contains agents from the opposite side whose φ_l^h-values are also negative. The reason is that matching two agents with negative valuations (or, alternatively, virtual valuations) can only decrease the platform's payoff.

These general observations do not hinge on Condition TP. Moreover, they say nothing on how to optimally match agents with a positive (virtual) valuation to agents from the opposite side with a negative (virtual) valuation *(cross-subsidization)*. This is where Condition TP, along with the fact that valuations are private information, plays a role.

Consider first the scenario of Condition TP(a), where $g_k(\cdot)$ is weakly concave and $\sigma_k(\cdot)$ is weakly increasing. Pick an agent from side k with $\varphi_k^h(v_k) > 0$ and suppose that the platform wants to assign to this agent a matching set whose intensity

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

exceeds the matching intensity of those agents from side l with non-negative φ_l^h-values (i.e., for whom $v_l \geq r_l^h$). The combination of the assumptions that (i) salience is weakly increasing in valuations, (ii) g_l are weakly concave, and (iii) valuations are private information, implies that the *least costly* way to deliver intensity q to such an agent is to match him to all agents from side l whose $\varphi_l^h(v_l)$ is the least negative. This is because (a) these latter agents are the most attractive ones, and (b) by virtue of g_l being concave, using the same agents from side l with a negative φ_l^h-valuations *intensively* is less costly than using different agents with negative φ_l^h-valuations. This, in turn, means that type v_k's matching set takes the form $[t_k(v_k), \overline{v}_l]$, where the threshold $t_k(v_k)$ is computed so that

$$\int_{[t_k(v_k),\overline{v}_l]} \sigma_l(v_l) dF_l(v_l) = q. \tag{8}$$

Building on the above ideas, the formal proof in the Appendix uses results from the monotone concave order to verify that, when Condition TP(a) holds, 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-valuations is nonnegative) matching sets of high quality in the most economical way. Note that the threshold rule constructed above is implementable provided that the original matching rule is implementable. In particular, under the new rule, among those agents with negative φ_l^h-valuations, those with a higher valuations may receive larger matching sets.

Next, consider the scenario of Condition TP(b), where $g_k(\cdot)$ is weakly convex on both sides and where $\sigma_k(\cdot)$ is weakly decreasing. Then pick a type v_k from side k with $\varphi_k^h(v_k) < 0$. Recall that using such an agent is costly for the platform. Now imagine that the platform wanted to assign to this type a matching set of strictly positive intensity, $|\mathbf{s}_k(v_k)|_l > 0$. The combination of the assumptions that (i) salience decreases with valuations, (ii) $g_l(\cdot)$ are weakly convex and (iii) types are private information, then implies that the most profitable way of using type v_k as an input is to match him to those agents from side l with the highest positive φ_l^h-valuations. This is because (a) these latter types are the ones that benefit the most from interacting with type v_k (indeed, as required by incentive compatibility, these types have the matching sets with the highest intensity and hence, by the convexity of $g_l(\cdot)$, the highest marginal utility for meeting additional agents) and (b) these latter types are the least salient ones and hence exert the lowest negative externalities on type v_k (recall that $\varphi_k^h(v_k) < 0$). In the scenario covered by Condition TP(b), the reason why a threshold structure is thus optimal is that it *maximizes the benefits (or profits) of cross-subsidization.* Q.E.D.

That agents on both sides have private information about their valuations is crucial for the result in Proposition 1. To see this, consider the following example.

Example 6 ***(role of private information)*** *Agents from sides A and B have valuations drawn uniformly from $V_A = [0, 1]$ and $V_B = [-2, 0]$, respectively. Preferences are linear on both sides, that is, $g_k(x) = x$, $k = A, B$. The salience function on side A is constant, $\sigma_A(v_A) = 1$ all $v_A \in V_B$, whereas the salience function on side B is given by*

$$\sigma_B(v_B) = \begin{cases} 1 & \text{if} \quad v_B \in [-1, 0] \\ 8 & \text{if} \quad v_B \in [-2, -1]. \end{cases}$$

These preferences clearly satisfy Condition TP(b). Now suppose that valuations are publicly observable on both sides and that the platform maximizes welfare. The optimal matching rule is then *given by*

$$t_A(v_A) = \begin{cases} [-2,-1] \cup [-v_A, 0] & \text{if} \quad v_A \geq \frac{1}{4} \\ [-8v_A, -1] \cup [-v_A, 0] & \text{if} \quad \frac{1}{8} \leq v_A < \frac{1}{4} \\ [-v_A, 0] & \text{if} \quad 0 \leq v_A < \frac{1}{8}. \end{cases}$$

Furthermore, no threshold rule yields the same welfare as the above rule. $\diamondsuit$

The key ingredient of the above example is that salience is decreasing in valuations on side B (which is the "input" side, as $v_B \leq 0$). As such, some of the most "expensive" agents from side B are the most attractive ones to the side-A agents. The welfare-maximizing rule (under complete information) then proceeds by evaluating separately each possible match between agents from the two sides (note that this follows from the fact that, in this example, g is linear on both sides, which implies that preferences are separable in the matches). It is then welfare-enhancing to include in the matching sets of side-A agents (whose valuation is positive) a disjoint collection of types from side B. The matching rule in the example, however, fails the monotonicity condition required by incentive compatibility (that is, the total salience of the matching sets is non monotone in valuations). As such, it is not implementable when types are private information.

The example above assumes that agents have linear preferences for matching intensity. Allowing, for example, preferences to be convex only exacerbates the sub-optimality of threshold rules. In this case, "concentrating" matching intensity on those side-B agents with the highest valuations (i.e., lower costs) leads to even higher *marginal* costs of cross-subsidization.

Similar issues arise when salience increases with valuations and preferences for matching intensity are concave (as in Condition TP(a)). In this case, under complete information, non-threshold rules might be optimal, as agents with the highest valuations on a given side might be too salient (relative to agents on the same side with lower valuations) and therefore have large cross-subsidization costs, and might enjoy small benefits from extra matches (as preferences are concave). In such cases, one can construct examples in which the complete-information efficient rule is not a threshold rule (and violates incentive compatibility).

Therefore, considered in isolation, neither Condition TP nor incentive compatibility is itself sufficient for the optimality of threshold rules. It is the *combination* of the cross-subsidization logic outlined in the proof sketch of Proposition 1 with the monotonicity requirements of incentive compatibility that leads to the optimality of threshold rules.

We conclude this subsection by showing that Condition TP is "sharp" in the sense that, under asymmetric information, when Condition TP is relaxed, one can construct primitives under which threshold rules fail to be optimal. The following two examples illustrate.

Example 7 ***(sub-optimality of threshold rules - 1)****.* *Agents from sides A and B have their valuations drawn uniformly from* $V_A = [0, 1]$ *and* $V_B = [-2, 0]$*, respectively. The salience of side-A agents is constant and normalized to one, i.e.,* $\sigma_A(v_A) \equiv 1$ *for all* $v_A \in V_A$*, while the salience of the*

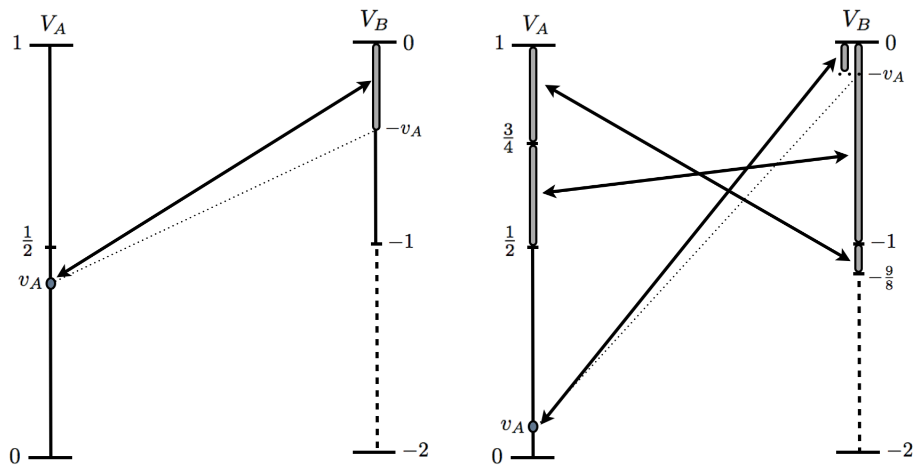

Figure 1: The welfare-maximizing rule among those with a threshold structure (left) and the welfare-improving non-threshold rule (right) from Example 7.

side-B agents is given by

$$\sigma_B(v_B) = \begin{cases} 1 & \textit{if} \quad v_B \in [-1,0] \\ 8 & \textit{if} \quad v_B \in [-2,-1]. \end{cases}$$

Preferences for matching intensity are linear on side B (that is, g_B is the identity function), whereas preferences on side A are given by the concave function[8]

$$g_A(x) = \min\left\{x, \frac{1}{2}\right\}.$$

In this environment, the welfare-maximizing threshold rule is described by threshold function $t_A(v) = t_B(v) = -v$, with exclusion types $\omega_A = 0$ and $\omega_B = -1$, as can be easily verified from Proposition 2 below. Total welfare under the optimal threshold rule is $1/12$. *Now consider the following non-threshold rule, which we describe by its side-A correspondence:*

$$\mathbf{s}_A(v_A) = \begin{cases} \left[-\frac{9}{8}, -1\right] & \textit{if} \quad v_A \in \left[\frac{3}{4}, 1\right] \\ [-1, 0] & \textit{if} \quad v_A \in \left[\frac{1}{2}, \frac{3}{4}\right] \\ [-v_A, 0] & \textit{if} \quad v_A \in \left[0, \frac{1}{2}\right]. \end{cases}$$

It is easy to check that this rule is implementable. Total welfare under this rule equals $3/32 > 1/12$. $\diamondsuit$

The matching rules in this example are illustrated in Figure 1. Intuitively, the reason why threshold rules fail to be optimal in the above example is that they fail to maximize the benefits of cross-subsidization. Agents from side B with valuation $v_B \in [-2,-1]$ are more expensive but significantly

[8] That the function g_A has a kink simplifies the computations but is not important for the result; the sub-optimality of threshold rules clearly extends to an environment identical to the one in the example but where the function g_A is replaced by a sufficiently close smooth concave approximation.

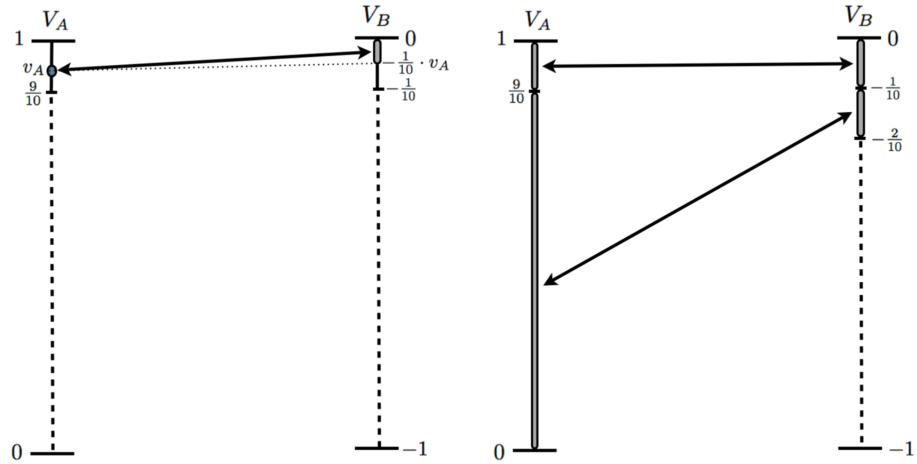

Figure 2: The welfare-maximizing rule among those with a threshold structure (left) and the welfare-improving non-threshold rule (right) from Example 8.

more attractive than agents with valuation $v_B \in [-1, 0]$. That salience is decreasing in valuations (weakly on side A, strictly on side B) per se does not make threshold rules suboptimal. Indeed, as established in Proposition 1 above, were preferences for matching intensity weakly convex on both sides, threshold rules would maximize welfare. Under concavity, however, once the high-valuation agents from side A interact with the high-valuation agents from side B (those with $v_B > -1$), they no longer benefit from interacting with agents from side B whose valuation is low (those with $v_B \leq -1$). This is inefficient, for those side-B agents with a low valuation are in fact the most attractive ones from the eyes of the side-A agents. More efficient cross-subsidization (and hence higher welfare) can be achieved by matching high-valuations agents from side A *only* to low-valuations agents from side B.

The next example considers the opposite case where salience is nondecreasing but preferences are strictly convex on one of the two sides.

Example 8 ***(sub-optimality of threshold rules - 2).*** *Agents from sides A and B have their valuations drawn uniformly from $V_A = [0, 1]$ and $V_B = [-1, 0]$, respectively. The salience of the side-B agents is constant and normalized to one, i.e., $\sigma_B(v_B) \equiv 1$ for all $v_B \in V_B$, while the salience of the side-A agents is given by*

$$\sigma_A(v_A) = \begin{cases} 10 & \textit{if} \quad v_A \in \left[\frac{9}{10}, 1\right] \\ \frac{10}{9} & \textit{if} \quad v_A \in \left[0, \frac{9}{10}\right] . \end{cases}$$

Preferences for matching intensity are linear on side A (that is, g_A is the identity function), whereas *preferences on side B are given by the convex function*[9]

$$g_B(x) = \begin{cases} x & \text{if} \quad x \leq 1 \\ +\infty & \text{if} \quad x > 1. \end{cases}$$

In this environment, the welfare-maximizing threshold rule is described by the threshold function $t_A(v) = -\frac{v}{10}$*, with exclusion types* $\omega_A = \frac{9}{10}$ *and* $\omega_B = -\frac{1}{10}$*, as can be easily verified from Proposition 2 below. Total welfare under such a rule is* $4/10^3$*. Now consider the following non-threshold rule, which we describe by its side-A correspondence:*

$$\mathbf{s}_A(v_A) = \begin{cases} \left[-\frac{1}{10}, 0\right] & \text{if} \quad v_A \in \left[\frac{9}{10}, 1\right] \\ \left[-\frac{2}{10}, -\frac{1}{10}\right] & \text{if} \quad v_A \in \left[0, \frac{9}{10}\right). \end{cases}$$

It is easy to check that this matching rule is implementable. Total welfare under this rule equals $3/10^2 > 4/10^3$. $\diamondsuit$

The matching rules in this example are illustrated in Figure 2. To illustrate the logic of the example, let agents from side A be advertisers and agents from side B be viewers. The advertisers with the highest willingness to pay for viewers, $v_A \in \left[\frac{9}{10}, 1\right]$, are the most salient ones (i.e., their ads are perceived as the most annoying by the viewers). Viewers have convex nuisance costs, as described by the function g_B (in particular, the disutility from advertising becomes arbitrarily large once the salience-adjusted mass of advertising exceeds one). In this example, the optimal threshold rule matches advertisers with a high willingness to pay with those viewers whose tolerance for advertising is sufficiently high, and assigns empty matching sets to all other advertisers and viewers. Because of the convexity of nuisance costs, few advertisers are matched to viewers under such a rule. The alternative rule proposed in the example better distributes advertisers to viewers. Under the proposed rule, advertisers with willingness to pay $v_A \in [0, \frac{9}{10}]$ (whose ads are not particularly annoying) are matched to viewers with moderate tolerance for advertising (i.e., $v_B \in [-\frac{2}{10}, -\frac{1}{10}]$), while advertisers with a high willingness to pay (whose ads are the most annoying ones) are matched with viewers whose tolerance for advertising is the highest. Welfare under the proposed rule is almost ten times higher than under the optimal threshold rule.

3.2 Properties of Optimal Threshold Rules

Assuming throughout the rest of the paper that Condition TP holds, we then proceed by further investigating the properties of optimal threshold rules. To conveniently describe the agents' payoffs, we introduce the function $\hat{g}_k : V_l \to \mathbb{R}_+$ defined by

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

[9] That the function g_B jumps at infinity at $x = 1$ simplifies the exposition but is not important for the result; the sub-optimality of threshold rules clearly extends to an environment identical to the one in the example but where the function g_B is replaced by a sufficiently close smooth convex approximation.

$k,l = A,B$, $l \neq k$. The utility that an agent with type v_k obtains from a matching set $[t_k(v_k), \overline{v}_l]$ can then be written concisely as $v_k \cdot \hat{g}_k\left(t_k(v_k)\right)$. Note that $\hat{g}_k\left(t_k(v_k)\right)$ in decreasing in $t_k(v_k)$, as increasing the threshold $t_k(v_k)$ reduces the intensity of the matching set.

Equipped with this notation, we can then recast the platform's problem as choosing a pair of non-increasing threshold functions $(t_k^h(\cdot))_{k\in\{A,B\}}$ along with two scalars (ω_A, ω_B) so as to maximize the objective

$$\Omega^h(M) = \sum_{k=A,B} \int_{\omega_k^h}^{\overline{v}_k} \hat{g}_k(t_k^h(v_k)) \cdot \varphi_k^h(v_k) dF_k(v_k) \tag{9}$$

subject to the reciprocity constraint

$$t_k^h(v_k) = \inf\{v_l : t_l^h(v_l) \leq v_k\} \tag{10}$$

for all $v_k \in [\omega_k^h, \overline{v}_k]$, $k,l = A,B$, $l \neq k$. Note that the reciprocity constraint (10) renders the platform's problem a nonstandard control problem (as each of the two controls $t_k(\cdot)$, $k \in \{A,B\}$, is required to coincide with the generalized inverse of the other).

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 2 ***(separation)*** *The h-optimal matching rule entails*

1. *separation if there exists a (positive measure) set $\hat{V}_k \subset V_k$ such that, for any $v_k, v_k' \in \hat{V}_k$, $t_k^h(v_k) \neq t_k^h(v_k')$,*
2. *exclusion at the bottom on side k if $\omega_k^h > \underline{v}_k$,*
3. *bunching at the top on side k if $t_l^h(\omega_l^h) < \bar{v}_k$.*

The rule is maximally separating if $t_k^h(\cdot)$ is strictly decreasing over the interval $[\omega_k^h, t_l^h(\omega_l^h)]$, which, hereafter, we refer to as the "separating range".

Accordingly, separation occurs when *some* agents on the same side receive different matching sets. Exclusion at the bottom occurs when all agents in a neighborhood of $\underline{v}_k$ are assigned empty matching sets. Bunching at the top occurs when all agents in a neighborhood of $\bar{v}_k$ receive identical matching sets. In turn, maximal separation requires that, as valuations increase, matching sets strictly expand whenever they are "interior" (in the sense that $t_k^h(v_k) \in (\omega_l^h, t_k^h(\omega_k^h))$).

The following regularity condition guarantees that the optimal rules are maximally separating.

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

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

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

As will be clear shortly, the optimal matching rules entail maximal separation if and only if Condition MR holds for every valuation in the separating range. Accordingly, this condition is the analog of Myerson's standard regularity condition in two-sided matching problems.

To understand the condition, 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 valuation v_k as a *consumer* (as his virtual valuation $\varphi_k^h(v_k)$ is proportional to the marginal revenue produced by the 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 valuation above v_k). Therefore, the above regularity condition requires that, under a threshold rule, the contribution of an agent as a consumer (as captured by his virtual valuation) increases faster than his contribution as an input.

Remark 1 *Note that Condition TP(b) implies Condition MR. Under Condition TP(a), instead, Condition MR imposes additional restrictions on the distributions F_k, the salience functions $\sigma_k(\cdot)$, and the preference functions $g_k(\cdot)$. When $g_k(\cdot)$ is linear (as in Examples 1, 2 and 3), Condition MR is equivalent to*

$$\frac{\varphi_k^h(v_k)}{\sigma_k(v_k)}$$

being strictly increasing. In this case, MR holds provided that salience does not increase "too fast" with an agent's valuation relative to his h-valuation, $\varphi_k^h(v_k)$. A similar requirement applies when $g_k(\cdot)$ is strictly concave (see Example 10 below).

To better appreciate the platform's trade-offs at the optimum, it is convenient to define the *marginal surplus function* $\triangle_k^h : V_k \times V_l \to \mathbb{R}$ according to

$$\triangle_k^h(v_k, v_l) \equiv -\hat{g}_k'(v_l) \cdot \varphi_k^h(v_k) \cdot f_k(v_k) - \hat{g}_l'(v_k) \cdot \varphi_l^h(v_l) \cdot f_l(v_l), \tag{11}$$

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.

Proposition 2 ***(optimal rules)*** *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 $v_k \in V_k$, $k = A, B$ if $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$.*[10]

When, instead, $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$, the h-optimal matching rule is maximally separating and entails

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

2. *exclusion at the bottom on side l and no bunching at the top on of side k if $\triangle_k^h(\bar{v}_k, \underline{v}_l) < 0$.*[11]

[10] The statement above holds true even when Condition MR is violated, provided that either $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, or $\varphi_k^h(\underline{v}_k) > 0$ for $k = A, B$. Condition MR is only needed in the case where $\varphi_k^h(\underline{v}_k) > 0 > \varphi_l^h(\underline{v}_l)$ for $k, l \in \{A, B\}$.

[11] 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.

Finally, the threshold function $t_k^h(\cdot)$ is implicitly defined by the Euler equation

$$\triangle_k^h(v_k, t_k^h(v_k)) = 0 \tag{12}$$

for any v_k in the separating range $[\omega_k^h, t_l^h(\omega_l^h)]$.

The optimal matching rule thus entails separation whenever the marginal surplus function evaluated at the lowest valuations on both sides of the market is negative: $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. When, instead, this condition fails, each agent from each side is matched to any other agent from the opposite side: $\mathbf{s}_k^h(v_k) = \Theta_l$ for all $v_k \in V_k$, $k = A, B$.

When separation occurs, Proposition 2 sheds light on the optimal cross-subsidization strategy employed by the platform. To illustrate, consider the case of profit-maximization (the arguments for welfare maximization are analogous), and let $\underline{v}_k < 0$, for $k = A, B$. An important feature of the profit-maximizing rule is that $t_k^P(v_k) \leq r_l^P$ if and only if $v_k \geq r_k^P$, where the reservation type r_k^P is the lowest type for whom $\varphi_l^P(v_k) \geq 0$. This implies that agents from each side of the market are endogenously partitioned in two groups. Those agents with positive virtual valuations (equivalently, with valuations $v_k \geq r_k^P$) 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 valuations (equivalently, with valuation $v_k < r_k^P$) 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). At the optimum, the platform recovers the "costs" of procuring agents-inputs from the gains obtained by agents-consumers.

The Euler equation (12) in the proposition then describes the optimal level of cross-subsidization for each type. In particular this equation can be rewritten as

$$\underbrace{-\hat{g}_k'(t_k^P(v_k)) \cdot \varphi_k^P(v_k) \cdot f_k(v_k)}_{\textit{marginal gains}} \quad = \quad \underbrace{\hat{g}_l'(v_k) \cdot \varphi_l^P(t_k^P(v_k)) \cdot f_l(t_k^P(v_k))}_{\textit{marginal losses}}. \tag{13}$$

At the optimum, the platform equalizes the marginal gains and the marginal losses of expanding the matching set of each agent in the separating range. When v_k corresponds to an agent-consumer (i.e., when $\varphi_k^P(v_k) > 0$), the left-hand side of (13) is the marginal revenue of expanding the agent's matching set, starting from a situation in which the agent is matched already to all agents from the other side whose valuation is above $t_k^P(v_k)$. In turn, the right-hand-side of (13) is the marginal cost associated with procuring extra agents-inputs from the opposite side; under a threshold rule, this cost is the loss that the platform incurs by expanding the matching set of an agent from side l whose valuation is $v_l = t_k^P(v_k)$, starting from a situation where such an agent is already matched to all agents from side k whose valuation exceeds $v_k = t_l^P(v_l)$, as required by reciprocity (recall that $t_l^P(t_k^P(v_k)) = v_k$). The terms $\hat{g}_k'(t_k^P(v_k))$ and $\hat{g}_l'(v_k)$ adjust the marginal utilities to account for the effect of the new matches on the supra-marginal matches (i.e., those matches above the profit-maximizing thresholds).

Note that optimality also 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.

Remark 2 *Condition MR is necessary and sufficient for the marginal surplus function $\triangle_k^h(v_k, v_l)$ to satisfy the following single-crossing property: whenever $\triangle_k^h(v_k, v_l) \geq 0$, then $\triangle_k^h(v_k, \hat{v}_l) > 0$ 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 can be seen from the Euler equation (12), this single-crossing property is equivalent to the threshold function $t_k^h(\cdot)$ being strictly decreasing over the separating range. Therefore, Condition MR is the "weakest" regularity condition that rules out non-monotonicities (or bunching) in the matching rule.*

The next two examples illustrate the characterization of Proposition 2.

Example 9 ***(ad exchange)*** *Consider a profit-maximizing ad exchange, as in Example 1. Assume that advertisers and publishers have valuations drawn from a uniform distribution over $V_A = [0, 1]$ and $V_B = [-1, 0]$, respectively. The profit-maximizing matching rule is then described by the threshold function*

$$t_B^P(v_B) = \frac{1}{2} - \frac{v_B}{\sigma_B(v_B)},$$

where the "cost-prominence ratio" $\frac{-v_B}{\sigma_B(v_B)}$ is strictly decreasing (as required by Condition MR). This rule is broadly consistent with the practice of most exchanges of eliciting information from each publisher about both its opportunity cost and a measure of the "prominence" of its website, and then determine which advertisers are displayed on the publisher's website based on scoring rules that weigh the two dimensions. $\diamondsuit$

Example 10 ***(nuisance costs)*** *Consider an advertising platform (say an online newspaper) matching advertisers to readers with convex nuisance costs, as in Example 4. Assume that readers and advertisers have valuations drawn from a uniform distribution over $V_A = [-1, 0]$ and $V_B = [0, 1]$, respectively, and that all ads are equally annoying to the readers, i.e. $\sigma_B(v_B) = 1$ for all v_B. It is easy to check that Conditions TP(b) and MR are then satisfied. From Proposition 2, the welfare-maximizing rule is then described by the threshold function*

$$t_B^W(v_B) = -\frac{v_B}{(1 - v_B)^\beta} \cdot \frac{1}{1 + \beta},$$

where $\beta \geq 0$ parametrizes the nuisance costs. As the nuisance cost increases, i.e. as β goes up, all advertisers obtain smaller matching sets. $\diamondsuit$

3.3 Welfare-maximizing vs Profit-Maximizing Rules

We now turn to the distortions brought in by profit-maximization relative to the welfare-maximizing matching rule. Consider the following example.

Example 11 ***(supermodular matching values)*** *Let the environment be as in Example 3 and assume that all v_k are drawn from a uniform distribution over $[\underline{v}, \overline{v}]$, with $\underline{v} > 0$ and $2\underline{v} < \bar{v}$, $k = A, B$. Because matching any two agents generates positive surplus, the welfare-maximizing rule matches each agent to any other agent from the opposite side. Next, consider the profit-maximizing rule. It is easy*

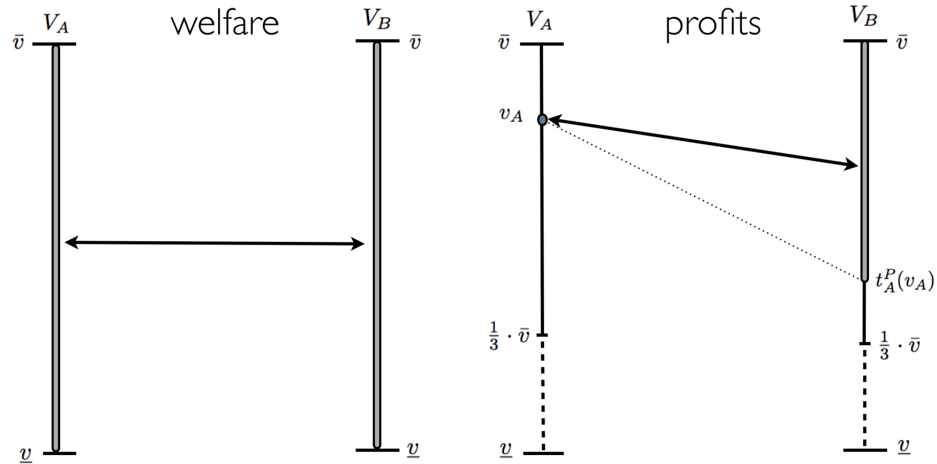

Figure 3: The welfare-maximizing matching rule (left) and the profit-maximizing matching rule (right) from Example 11 (supermodular matching values).

to check that Conditions TP(a) and MR are satisfied. Because $\triangle_k^P(\underline{v},\underline{v}) = 2\underline{v}(2\underline{v}-\bar{v}) < 0$*, it follows from Proposition 2 that 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 $(\omega_k^P, \bar{v}) = (\frac{\bar{v}}{3}, \bar{v})$*. 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.* $\diamondsuit$

The matching rules in this example are illustrated in Figure 3. As indicated in the next Proposition, the distortions in this example are general properties of profit-maximizing rules (the proof follows directly from Proposition 2).

Proposition 3 ***(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 who is not excluded 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* $v_k \geq \omega_k^P$*,* $k = A, B$*.*

The intuition for both effects can be seen from Condition (12): under profit-maximization, the platform only internalizes the cross-effects on marginal revenues (which are proportional to the virtual valuations), rather than the cross-effects on welfare (which are proportional to the true valuations). Contrary to other mechanism design problems, profit-maximization in a matching market may result in inefficiencies for *all* types, including those "at the top" of the distribution (i.e., with the highest valuations for matching intensity). The reason is that, although the virtual valuations of these types coincide with the true valuations, the cost of cross-subsidizing these types is higher under profit maximization than under welfare maximization, due to the inframarginal losses implied by reciprocity on the opposite side.

3.4 Comparative Statics: The Detrimental Effects of Becoming More Attractive

Shocks that alter the cross-side effects of matches are common in vertical matching markets. Changes in market conditions, for example, affect the pricing strategies of B2B platforms, for they affect the attractiveness of business connections for the same population of firms.

The next definition formalizes the notion of a change in attractiveness. We restrict the attention here to a platform maximizing profits in a market where all agents from each side value positively interacting with agents from the opposite side (i.e., $\underline{v}_k \geq 0$ for $k \in \{A, B\}$). For simplicity, we also restrict attention to markets in which preferences for matching intensity are linear (i.e., $g_A(x) = g_B(x) = x$, all $x \in \mathbb{R}_+$).

Definition 3 ***(higher attractiveness)*** *Consider a market in which all agents value positively interacting with agents from the opposite side, i.e., $\underline{v}_k \geq 0$ for $k = A, B$. Side k is more attractive under $\hat{\sigma}_k(\cdot)$ than under $\sigma_k(\cdot)$ if $\hat{\sigma}_k(v_k) \geq \sigma_k(v_k)$ for all $v_k \in V_k$, with the inequality strict for a positive-measure subset of V_k.*

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

Proposition 4 ***(increase in attractiveness)*** *Consider a market in which (a) conditions TP and MR hold, (b) all agents value positively interacting with agents from the opposite side (i.e., $\underline{v}_k \geq 0$ for $k \in \{A, B\}$), and (b) preferences for matching intensity are linear (i.e., $g_A(x) = g_B(x) = x$, all $x \in \mathbb{R}_+$). Suppose side k becomes more attractive. Then a profit-maximizing 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 valuation and decrease for those agents with a high valuation, i.e., $\hat{\mathbf{s}}_k^P(v_k) \supseteq \mathbf{s}_k^P(v_k)$ if and only if $v_k \leq r_k^P$;*

2. *low-valuation agents from side k are better off, whereas the opposite is true for high-valuation ones, i.e., there exists $\hat{\nu}_k \in (r_k^P, \bar{v}_k]$ such that $\Pi_k(v_k; \hat{M}^P) \geq \Pi_k(v_k; M^P)$ 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 valuation 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 valuation $v_k \leq r_k^P$. 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 $v_k \leq r_k^P$

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

meaning that all agents from side k with valuation $v_k \leq r_k^h$ 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 $v_k \geq r_k^h$, then either payoffs increase for all agents from side k, or there exists a threshold $\hat{\nu}_k > r_k^h$ 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 $v_k \leq \hat{v}_k$.

Finally, consider the effect of an increase in the attractiveness of side k on the price that agents have to pay for matching intensity. For any matching intensity q_k, let $\rho_k^P(q_k)$ denote the total price that each agent from side k has to pay for any matching set of intensity q_k under the profit-maximizing mechanism M^P. By optimality, the tariff $\rho_k^P(\cdot)$ has to satisfy

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

We then have the following result.

Corollary 1 ***(effect of an increase in attractiveness on prices)*** *Under the assumptions of Proposition 4, if the attractiveness of side k increases (in the sense of Definition 3), 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 intensity $q_k \leq \hat{q}_k$, 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 thus triggers an increase in the price that the platform charges on side k for matching sets of high intensity and a decrease in the price it charges for matching sets of low intensity.

4 Extensions

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

Imperfect Correlation between Salience and Valuation. To simplify the exposition, the baseline model assumes that salience is a deterministic function of the valuations. All our results extend to environments where the two dimensions are imperfectly correlated and where agents have private information about both dimensions. We formally establish this result in the Appendix by first relaxing Condition TP to require that salience and valuation be positively (or, alternatively, negatively) affiliated. This is the natural generalization of the assumption that σ_k be increasing (or, alternatively, decreasing) in v_k, as required by Condition TP. Under this condition, we then show that the optimal matching rules have a threshold structure, with the thresholds depending on valuations but not on salience. Note that the result is not a mere consequence of the fact that individual preferences are invariant in the agents' own salience. Combined with incentive compatibility, the latter property only implies that the matching intensity is invariant in the agent's own salience, thus permitting the *composition* of the matching sets to depend on salience. Once this result is established, it is then immediate that all other results in the paper extend to this richer environment.

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., $\mathbf{s}_A(v) = \mathbf{s}_B(v)$ for all $v \in V_A = V_B$.

Under the new constraint that matching rules be symmetric across the two sides, maximizing (9) 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 $\psi_l^h(\cdot) = \psi_k^h(\cdot)$). Indeed, 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.

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.[12] It is easy to see that the analysis developed above 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 rule identified in the paper.[13] 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.

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 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. Let c_k denote the quasi-fixed cost that the platform must incur for each agent from side k whose matching set is nonempty. The h-optimal mechanism can then be obtained through the following two-step procedure:

1. Step 1: Ignore quasi-fixed costs and maximize (9) among all weakly decreasing threshold func-

[12] See also Wilson (1989).

[13] 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.

tions $t_k^h(\cdot)$.

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:

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

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.

Robust Implementation. In the direct revelation version of the matching game, each agent from each side is asked to submit a report v_k which leads to a payment $p_k^h(v_k)$, as defined in (6), and grants access to all agents from the other side of the market who reported a valuation 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.[14]

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 $k = A, B$).[15] 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

$$\varrho_k^h(v_k, (v_l^j)^{j\in[0,1]}) = v_k \cdot g_k \left(\left| \{j \in [0,1] : v_l^j \geq t_k(v_k)\} \right|_l \right) - \int_{\underline{v}_k}^{v_k} g_k \left(\left| \{j \in [0,1] : v_l^j \geq t_k(x)\} \right|_l \right) dx, \tag{15}$$

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 Lebesgue measure of agents from side $l \neq k$ reporting a valuation above $t_k(v_k)$. Given the above payment rule, it is *weakly dominant* for each agent to report truthfully. This follows from the fact that, *given any* profile of reports $(v_l^j)^{j\in[0,1]}$ by agents from the opposite side, the intensity 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.[16]

[14] 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).

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

[16] 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 $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_l$ in (15) with $\left|\{j \in [0,1] : v_l^j \geq t_k(v_k)\}\right|_l \equiv \int_{\{j: v_l^j \geq t_k(v_k)\}} \underline{\sigma}_l d\lambda(j)$, where $\underline{\sigma}_l \equiv \min\{\sigma_l(v_l) : v_l \in V_l\}$. 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.

5 Concluding Remarks

The analysis reveals how matching patterns reflect optimal cross-subsidization between sides in centralized markets. We deliver two main results. First, we identify primitive conditions under which the optimal matching rules have a simple threshold structure, according to which agents with a low valuation for matching are included only in the matching sets of those agents from the opposite side whose valuation is sufficiently high. In other words, matching sets are assortative in the set-inclusion order. Second, the optimal matching rules are determined by a simple formula that equalizes the marginal gains in welfare (or, alternatively, in profits) with the cross-subsidization losses that the platform must incur on the opposite side of the market. We show that the optimal rules *endogenously* separate agents into consumers and inputs. At the margin, the "costs" of procuring agents-inputs is recovered from the gains from agents-consumers (cross-subsidization).

The model is flexible enough to permit interesting comparative statics. For example, we show that, when the attractiveness of one side increases, a profit-maximizing platform responds by reducing the intensity of the matching sets offered to those agents whose valuation is high and by increasing the intensity of the matching sets offered to those agents whose valuation is low.

The above analysis is worth extending in a few important directions. For example, all the results are established assuming 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 is 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 challenging but worth exploring.

Likewise, the analysis focuses on a market with a single platform. Many matching markets are 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 venues for future research (see Damiano and Li (2008), Lee (forthcoming), and Jullien and Pavan (2013) for models of platform competition in settings with a limited degree of price discrimination).

6 Appendix

Proof of Proposition 1. Below we prove a stronger result that supports both the claim in the proposition as well as the claim in Section 4 about the optimality of threshold rules in environments where salience is imperfectly correlated with the valuation and where agents have private information about both dimensions.

To this purpose, we enrich the model as follows. For each $v_k \in V_K$, let $\Psi_k(\cdot|v_k)$ denote the conditional distribution of σ_k, given v_k, $k = A, B$ and denote by $\Lambda_k = F_k \cdot \Psi_k$ the measure defined by the product of F_k and Ψ_k. Now assume that agents observe both v_k and σ_k at the time they interact with the platform. Each agent's type is then given by the bi-dimensional vector $\theta_k \equiv (v_k, \sigma_k) \in \Theta_k \equiv V_k \times \Sigma_k$, with $\Sigma_k \subset \mathbb{R}_+$. In this environment, a matching mechanism $M = \{\mathbf{s}_k(\cdot), p_k(\cdot)\}_{k=A,B}$ continues to be described by a pair of matching rules and a pair of payment rules, with the only difference that $p_k(\cdot)$ now maps Θ_k into $\mathbb{R}$, whereas $\mathbf{s}_k(\cdot)$ maps Θ_k into the Borel sigma algebra over Θ_l, $k, l = A, B$, $l \neq k$. With some abuse of notation, hereafter we will denote by $|\mathbf{s}_k(\theta_k)|_l = \int_{(v_l,\sigma_l)\in\mathbf{s}_k(\theta_k)} \sigma_l d\Lambda_l$ the matching intensity of the set $\mathbf{s}_k(\theta_k)$.

Now consider the following extension of Condition TP in the main text.

Condition TP-extended. *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.[17]

Below we will prove the following claim.

Claim 1 *Assume Condition TP-extended holds. Then both the profit-maximizing ($h = P$) and the welfare-maximizing ($h = W$) rules discriminate only along the valuation dimension (that is, $s_k^h(v_k, \sigma_k) = s_k^h(v_k, \sigma_k')$ for any $k = A, B$, $v_k \in V_k$, $\sigma_k, \sigma_k' \in \Sigma_k$, $h = W, P$) and are threshold rules. That is, there exists a scalar $\omega_k^h \in [\underline{v}_k, \overline{v}_k]$ and a non-increasing function $t_k^h : V_k \to V_l$ such that, for any $\theta_k = (v_k, \sigma_k) \in \Theta_k$, $k = A, B$,*

$$\mathbf{s}_k^h(v_k, \sigma_k) = \begin{cases} [t_k^h(v_k), \overline{v}_l] \times \Sigma_l \text{ if } v_k \in [\omega_k^h, \overline{v}_k] \\ \oslash \text{ otherwise.} \end{cases} \tag{16}$$

The case where salience is a deterministic monotone function of the valuation is clearly a special case of affiliation. It is then immediate that the above claim implies the result in Proposition 1.

To establish the claim, we start by observing that, 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 is obviously a threshold one.

[17] See Milgrom and Weber (1982) for a formal definition of affiliation.

Thus consider the situation where $\varphi_k^h(\underline{v}_k) < 0$ for some $k \in \{A, B\}$. Define $\Theta_k^{h+} \equiv \{\theta_k = (v_k, \sigma_k) : \varphi_k^h(v_k) \geq 0\}$ the set of types θ_k whose φ_k^h-value is non-negative, and $\Theta_k^{h-} \equiv \{\theta_k = (v_k, \sigma_k) : \varphi_k^h(v_k) < 0\}$ 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-extended 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 (16) and that weakly improves upon the original one in terms of the platform's objective.

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-extended.

Lemma 1 *A mechanism M is incentive compatible only if, with the exception of a countable subset of V_k, $|\mathbf{s}_k(v_k, \sigma_k)|_l = |\mathbf{s}_k(v_k, \sigma'_k,)|_l$ 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 $|\mathbf{s}_k(v_k, \sigma_k)|_l \geq |\mathbf{s}_k(v'_k, \sigma'_k)|_l$ for any (v_k, σ_k) and (v'_k, σ'_k) such that $v_k \geq v'_k$. This in turn implies that $\mathbb{E}[|\mathbf{s}_k(v_k, \tilde{\sigma}_k)|_l]$ 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 intensity 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.

Definition 4 *[**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-extended.

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

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

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 $k = A, B$,

$$\int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(v_k, \sigma_k)|_l\right) d\Lambda_k \geq \int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(\sigma_k, v_k)\right|_l\right) d\Lambda_k. \tag{17}$$

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,

$$\int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot g_k\left(\left|\mathbf{s}'_k(v_k, \sigma_k)\right|_l\right) d\Lambda_k \leq \int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot g_k\left(|\hat{\mathbf{s}}_k(v_k, \sigma_k)|_l\right) d\Lambda_k. \tag{18}$$

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 $z_1(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\mathbf{s}'_k(v_k, \tilde{\sigma}_k)|_l \,|v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\hat{\mathbf{s}}_k(v_k, \tilde{\sigma}_k)|_l \,|v_k\right]$, where the distribution over $[\underline{v}_k, r_k^h]$ is given by $F_k(v_k)/F_k(r_k^h)$. 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 valuations and salience, (iii) the fact that the measure $F_k(v_k)$ is absolute continuous with respect to the Lebesgue measure and (iv) Lemma 1, we have that for all $x \in [\underline{v}_k, r_k^h]$,

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

or, equivalently,

$$\int_{\underline{v}_k}^{x} z_1(v_k) dF_k(v_k) \geq \int_{\underline{v}_k}^{x} z_2(v_k) dF_k(v_k). \tag{19}$$

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

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

$$\int_{\underline{v}_k}^{r_k^h} \dot{z}_2(v_k) dF_k(v_k) / F_k(r_k^h) = \int_{\underline{v}_k}^{r_k^h} z_2(v_k) dF_k(v_k) / F_k(r_k^h). \tag{20}$$

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

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

where the first inequality follows from the weak concavity of $g(\cdot)$ along with (20), 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 (18). The results above imply that

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

The first equality follows from changing the order of integration. The second equality follows from the fact that, since $\mathbf{s'}_k(\cdot)$ is implementable, $g_k\left(|\mathbf{s'}_k(v_k,\sigma_k)|_l\right)$ 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 (18), as we wanted to show.

Combining (17) with (18) 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 Claim 1 for the case of markets that satisfy conditions (1a) and (1b) in Condition TP-extended.

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

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 any (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 $|\Theta_l|_l > |\mathbf{s'}_k(\theta_k)|_l \ge \left|\Theta_l^{h+}\right|_l > 0$, 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 $|\mathbf{s'}_k(\theta_k)|_l \ge \left|\Theta_l^{h+}\right|_l = 0$, 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 $0 \le |\mathbf{s'}_k(\theta_k)|_l < \left|\Theta_l^{h+}\right|_l$, then let $\hat{t}_k(v_k)$ be such that

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

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

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

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

$$\int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot |\hat{\mathbf{s}}_k(v_k, \sigma_k)|_l \, d\Lambda_k \ge \int_{\Theta_k^{h-}} \varphi_k^h(v_k) \cdot \left|\mathbf{s}'_k(v_k, \sigma_k)\right|_l d\Lambda_k. \tag{21}$$

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 $\Theta_k^{h+} \ne \varnothing$,

$$\int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot |\hat{\mathbf{s}}_k(v_k, \sigma_k)|_l \, d\Lambda_k \ge \int_{\Theta_k^{h+}} \varphi_k^h(v_k) \cdot \left|\mathbf{s}'_k(v_k, \sigma_k)\right|_l d\Lambda_k \tag{22}$$

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 $z_1(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[|\hat{\mathbf{s}}_k(v_k, \tilde{\sigma}_k)|_l \,|v_k\right]$ and $z_2(v_k) \equiv \mathbb{E}_{\tilde{\sigma}_k}\left[\left|\mathbf{s}'_k(v_k, \tilde{\sigma}_k)\right|_l |v_k\right]$, 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 valuations and salience, (iii) the fact that the measure $F_k(v_k)$ is absolute continuous with respect to the Lebesgue measure and (iv) Lemma 1, we have that for all $x \in [r_k^h, \bar{v}_k]$,

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

or, equivalently,

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

The result in the lemma clearly holds if for all $v_k \in [r_k^h, \bar{v}_k]$, $z_1(v_k) \ge z_2(v_k)$. 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 $\int_{r_k^h}^{\bar{v}_k} z_1(v_k) dF_k(v_k) \ge \int_{r_k^h}^{\bar{v}_k} z_2(v_k) dF_k(v_k)$, 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. $\int_{[r_k^h, \bar{v}_k] \backslash \mathcal{T}} \{\dot{z}_2(v_k) - z_2(v_k)\} \, dF_k(v_k) = \int_{\mathcal{T}} \{z_2(v_k) - z_1(v_k)\} \, dF_k(v_k)$.

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

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

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

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

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 (22). The results above imply that

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

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, $g_k\left(\left|\mathbf{s}'_k(v_k,\sigma_k)\right|_l\right)$ 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 (22), as we wanted to show.

Combining (21) with (22) 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 Claim 1 under the conditions in part 2 of Condition TP-extended. 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 $t_k^h(v_k) = \underline{v}_l$ for all $v_k \in V_k$ if $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$ and entails separation otherwise.*

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

- First, consider the case where $\varphi_k^h(\underline{v}_k) \geq 0$ for $k = A, B$, implying that $\triangle_k^h(\underline{v}_k, \underline{v}_l) \geq 0$. Because valuations (virtual valuations) are all nonnegative, welfare (profits) is (are) maximized by matching each agent from each side to all agents from the other side, meaning that the optimal matching rule employs a single complete network.

- Next, consider the case where $\varphi_k^h(\underline{v}_k) < 0$ for $k = A, B$, so that $\triangle_k^h(\underline{v}_k, \underline{v}_l) < 0$. 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 $\hat{\omega}_k^h$ denote the threshold type corresponding to the non-separating rule so that agents from side k are excluded if $v_k < \hat{\omega}_k^h$ and are otherwise matched to all agents from side l whose valuation is above $\hat{\omega}_l^h$ otherwise.

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

  Next, suppose that $\hat{\omega}_k^h < r_k^h$ for both $k = A, B$. Starting from this non-separating rule, the platform could then increase its payoff by switching to a separating rule $\mathbf{s}_k^{\diamondsuit}(\cdot)$ such that, for some $k \in \{A, B\}$[18]

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

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

  Finally, suppose that $\hat{\omega}_k^h = r_k^h$ for some $k \in \{A, B\}$ whereas $\hat{\omega}_l^h \leq r_l^h$ for $l \neq k$. The platform could then do better by switching to the following separating rule:

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

  By setting the new exclusion threshold $\hat{\omega}_k^{\#}$ sufficiently close to (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.[19] Note that for this network expansion to be profitable, it is essential

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

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

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 valuation 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

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

  This condition requires that the total effect of a marginal increase of the size of the network on side l (obtained by reducing the threshold $t_k^h(v_k)$ below $\hat{\omega}_l^h$ for all types v_k) be zero. This rewrites as $\int_{\underline{v}_k}^{\bar{v}_k}[\Delta_k^h(v_k, \hat{\omega}_l^h)]dv_k = 0$. 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 $v_k^\# \in (\underline{v}_k, \bar{v}_k)$ such that $\int_{v_k^\#}^{\bar{v}_k} \Delta_k^h(v_k, \hat{\omega}_l^h)dv_k > 0$. 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:

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

  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) :

$$P^F : \qquad \max_{\{\omega_k, t_k(\cdot)\}_{k=A,B}} \sum_{k=A,B} \int_{\omega_k}^{\bar{v}_k} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) \tag{23}$$

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

$$t_k(v_k) = \inf\{v_l : t_l(v_l) \leq v_k\}, \tag{24}$$

$$t_k(\cdot) \text{ weakly decreasing,} \tag{25}$$

$$\text{and} \quad t_k(\cdot) : [\omega_k, \bar{v}_k] \to [\omega_l, \bar{v}_l] \tag{26}$$

with $\omega_k \in [\underline{v}_k, \bar{v}_k]$ and $\omega_l \in [\underline{v}_l, \bar{v}_l]$. Constraint (24) is the reciprocity condition, rewritten using the result in Proposition 1. Constraint (25) is the monotonicity constraint required by incentive compatibility. Finally, constraint (26) 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, $k = A, B$:

$$P_k^F : \qquad \max_{\omega_k, t_k(\cdot), t_l(\cdot)} \int_{\omega_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\bar{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{27}$$

subject to $t_k(\cdot)$ and $t_l(\cdot)$ satisfying the reciprocity and monotonicity constraints (24) and (25), along with the following constraints:

$$t_k(\cdot) : [\omega_k, r_k^h] \to [r_l^h, \bar{v}_l], \qquad t_l(\cdot) : [r_l^h, \bar{v}_l] \to [\omega_k, r_k^h]. \tag{28}$$

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

$$P_k^{Au} : \qquad \max_{t_k(\cdot), t_l(\cdot)} \int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot dF_k^v(v_k) + \int_{r_l^h}^{\bar{v}_l} \hat{g}_l(t_l(v_l)) \cdot \varphi_l^h(v_l) \cdot dF_l^v(v_l) \tag{29}$$

subject to (24),

$$t_k(\cdot), t_l(\cdot) \text{ strictly decreasing,} \tag{30}$$

$$\text{and} \quad t_k(\cdot) : [\underline{v}_k, r_k^h] \to [r_l^h, \overline{v}_l], \quad t_l(\cdot) : [r_l^h, \overline{v}_l] \to [\underline{v}_k, r_k^h] \text{ are } \textit{bijections.} \tag{31}$$

By virtue of (30), (24) can be rewritten as $t_k(v_k) = t_l^{-1}(v_k)$. Plugging this into the objective function (29) yields

$$\int_{\underline{v}_k}^{r_k^h} \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) dv_k + \int_{r_l^h}^{\overline{v}_l} \hat{g}_l(t_k^{-1}(v_l)) \cdot \varphi_l^h(v_l) \cdot f_l^v(v_l) dv_l. \tag{32}$$

Changing the variable of integration in the second integral in (32) 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:

$$P_k^{Au} : \qquad \max_{t_k(\cdot)} \int_{\underline{v}_k}^{r_k^h} \left\{ \hat{g}_k(t_k(v_k)) \cdot \varphi_k^h(v_k) \cdot f_k^v(v_k) - \hat{g}_l(v_k) \cdot \varphi_l^h(t_k(v_k)) \cdot f_l^v(t_k(v_k)) \cdot t_k^{'}(v_k) \right\} dv_k \tag{33}$$

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

$$t_k(\underline{v}_k) = \overline{v}_l \quad \text{and} \quad t_k(r_k^h) = r_l^h. \tag{34}$$

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 sub-differential that satisfies the boundary condition (34).

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*

$$\Delta_k^h(v_k, \eta(v_k)) = 0. \tag{35}$$

Let $\tilde{v}_k \equiv \inf\{v_k \in [\underline{v}_k, r_k^h]$:(35) admits a solution$\}$. The solution to P_k^R is given by

$$\tilde{t}_k(v_k) = \begin{cases} \overline{v}_l & \text{if } \; v_k \in [\underline{v}_k, \tilde{v}_k] \\ \eta(v_k) & \text{if } \; v_k \in (\tilde{v}_k, r_k^h]. \end{cases} \tag{36}$$

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 (35). Condition MR ensures that (i) there exists a $\tilde{v}_k \in [\underline{v}_k, r_k^h)$ such that (35) 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 $\tilde{v}_k > \underline{v}_k$, (35) 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 (33) that $\tilde{t}_k(v_k) = \overline{v}_l$ for all $v_k \in [\underline{v}_k, \tilde{v}_k]$.

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

To see that this is true, fix an arbitrary $(v_k^1, v_k^2) \subseteq [\tilde{v}_k, r_k^h]$ and consider the problem that consists in choosing optimally a step function $\hat{t}_k(\cdot) : (v_k^1, v_k^2) \to \{r_l^h, \overline{v}_l\}$. Because step functions are such that $\hat{t}_k'(v_k) = 0$ at all points of continuity and because $\varphi_k^h(v_k) < 0$ for all $v_k \in (v_k^1, v_k^2)$, it follows that the optimal step function is given by $\hat{t}_k(v_k) = \overline{v}_l$ for all $v_k \in (v_k^1, v_k^2)$. Notice that the value attained by the objective (33) 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 (35) 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 (36). Q.E.D.

Denote by $\max\{P_k^R\}$ the value of program P_k^R (i.e., the value of the objective (33) 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 (24), (25) and (28), 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 (24), (30) and (31). 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 (24). It is clear that the tripe $\omega_k^h, t_k^h(\cdot), t_l^h(\cdot)$ constructed this way satisfies conditions (24), (25) and (28), and is therefore a feasible candidate for program P_k^F. It is also immediate that the value of the objective (27) 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 4. Hereafter, we use the notation "ˆ"for all variables in the mechanism $\hat{M}^P$ corresponding to the new salience function $\hat{\sigma}_k(v_k)$ and continue to denote the variables in the mechanism M^P corresponding to the original function $\sigma_k(v_k)$ 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). 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 with valuation $v_k \leq r_k^P$

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

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 $\Pi_k(v_k; \hat{M}^P) \geq \Pi_k(v_k; M^P)$ if and only if $v_k \leq \hat{\nu}_k$, which establishes part 2 in the proposition. Q.E.D.

Proof of Corollary 1. 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 valuation 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 quality 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.

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

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

[4] Bedre-Defolie, O. and E. Calvano, 2013, "Pricing Payment Cards." *American Economic Journal: Microeconomics*, Vol. 5(3), pp. 206-231.

[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, 2010, "Sorting and Decentralized Price Competition."*Econometrica*, Vol. 78(2), pp. 539-574.

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

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

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

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

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

[18] 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.

[19] Johnson, T., 2013, "Matching Through Position Auctions."*Journal of Economic Theory*, Vol. 148, pp. 1700-1713.

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

[21] Jullien, B. and A. Pavan, 2013, "Platform Pricing Under Dispersed Information," mimeo Toulouse School of Economics and Northwestern University.

[22] 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.

[23] 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.

[24] Koopmans, T. C. and M. Beckmann, 1957, "Assignment problems and the location of economic activities." *Econometrica*, Vol. 25(1), pp. 53-76.

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

[26] Lee, R., 2013, "Vertical Integration and Exclusivity in Platform and Two-Sided Markets."*American Economic Review*, Vol. 103(7), pp. 2960-3000.

[27] Lee, R., "Competing Platforms."forthcoming at the *Journal of Economics and Management Strategy.*

[28] Legros P. and A. Newman, 2002, "Monotone Matching in Perfect and Imperfect Worlds," *Review of Economic Studies*, Vol. 69(4), pp. 925-945.

[29] Legros P. and A. Newman, 2007, "Beauty Is a Beast, Frog Is a Prince: Assortative Matching with Nontransferabilities," *Econometrica*, Vol. 75(4), pp. 1073-1102.

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

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

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

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

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

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

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

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

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

[39] Shimer, R. and L. Smith, 2000, "Assortative Matching and Search."*Econometrica*, Vol. 68(2), pp. 343-369.

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

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

[42] White, A. and G. Weyl, 2010, "Imperfect Platform Competition: A General Framework." Working Paper University of Chicago.

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

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