Re-matching, Experimentation, and Cross-subsidization*

Daniel Fershtman Northwestern University

Alessandro Pavan Northwestern University

PRELIMINARY AND INCOMPLETE

Abstract

We study mediated many-to-many matching in dynamic two-sided markets in which agents' private valuations for potential partners evolve stochastically over time, either as the result of exogenous shocks, or as the result of experimentation. In many environments, the platform's profit-maximizing matching rule is either myopic or takes the form of a "virtual index rule" capturing the current and future profitability of each link between agents, accounting for the endogenous changes in the partners' matching values. We show how the optimal matching rules can be implemented via a sequence of scoring auctions. We contrast matching dynamics under profit maximization with their counterparts under welfare maximization. When all agents benefit from interacting with all other agents from the opposite side, profit maximization involves fewer and shorter interactions. This conclusion, however, need not hold when certain agents dislike certain interactions.

JEL Classification Numbers: D82, C73, L1.

Keywords: matching, experimentation, bandit problems, asymmetric information, stochastic processes, incentives, mechanism design

1 Introduction

Most matching markets are intrinsically dynamic, due to the gradual resolution of uncertainty about match values as well as to shocks that alter the desirability of the existing matching allocations. The quality of the relationship between workers and employers matched by an employment agency, between the advertisers and consumers matched by a media platform, or between a pair of firms linked by a business-to-business platform are all likely to change over time either as the result of exogenous shocks, or the gradual learning about the attractiveness of the individual interactions.

*For helpful comments and suggestions, we thank Benjamin Hermalin and seminar participants at Berkeley, UNC-Duke, the Toulouse School of Economics, and the University of Western Ontario.

Despite the intrinsic dynamic nature of many matching markets, the theoretical literature has devoted relatively little attention to the study of such dynamics, taking a mostly static approach. This paper takes a first step in the direction of extending the analysis to dynamic markets by studying centralized many-to-many matching in an environment in which agents' values for their potential partners evolve (stochastically) over time. A key difficulty is that, while changes in agents' values are often anticipated, they are not necessarily observed by the matching intermediary. As a result, the intermediary must provide agents on each side of the market with the incentive to reveal their information as it evolves over time.

We consider a general model, but with specific examples in mind. Consider a project-finance consulting firm matching consortiums of firms seeking finance for a project with commercial banks. Alternatively, consider a commercial lobbying firm, mediating interactions between policy-makers and interest groups. Finally, consider a private medical tourism intermediary matching patients from abroad seeking specialized treatments with local physicians providing such treatments. In all of these examples, matching is typically many-to-many, with interactions mediated by a profit-maximizing platform collecting payments from both sides of the market. In addition, matching is dynamic, with agents' valuations changing over time, either because of exogenous shocks (including the arrival of exogenous information) or because of the information generated by previous experiences. In response to these changes, agents may be matched with a different set of partners over time. Finally, the matching services provided by the platform may be costly. In addition to time-varying match-specific costs, the platform may face a capacity constraint restricting the number of interactions that it may accommodate at each point in time. For instance, in the case of a private medical tourism intermediary, a capacity is imposed by the medical facility the agency contracts with.

The analysis characterizes optimal mechanisms, maximizing either social welfare or the profits of a private matching intermediary. It allows us to address the following questions: What is the effect of private experimentation on matching dynamics? How do the dynamics under profit maximization compare with their welfare-maximizing counterparts? How do the dynamics when the evolution of match values are private compare to the case in which they are public? What are the effects on matching dynamics of a binding capacity constraint on the number of interactions that the platform can accommodate? What type of auctions permit the intermediary to sustain the desired matching dynamics?

The basic ingredients of the model are the following. The payoff that each agent derives from each other agent is governed by two components, a time-invariant vertical characteristic that is responsible for the overall importance that the agent assigns to interacting with agents from the opposite side of the market, and a vector of time-varying relation-specific values capturing the evolution of the agent's valuations for each potential partner. The latter values evolve stochastically over time and may turn negative, reflecting the idea that agents may dislike interacting with certain agents from the opposite side of the market.

We study both the case in which these relation-specific values evolve exogenously, as well as the case in which they evolve endogenously as a function of previous interactions. The latter case is meant to capture an experimentation environment in which agents gradually learn the attractiveness of their partners via individual interactions.

Importantly, the agents' interactions are mediated by a platform that may face a constraint on the number of interactions it can accommodate. Depending on the market under consideration, this may reflect time, resource, or facility constraints. We study the effect of such constraints on matching dynamics.

The platform's problem consists in designing a many-to-many matching mechanism specifying how individual links and payments evolve over time as a function of the evolution of the agents' match values. When the platform chooses to link a pair of agents, it internalizes both the current and future payoffs that the pair of agents expect from the relationship, as well as all the externalities that a specific match exerts on all other agents due to the capacity constraint. In particular, the platform may find it optimal to cross-subsidize some of the interactions, for example by matching an agent with a partner she dislikes if the latter partner's current and/or expected future willingness to pay for such a match is large enough.

In the first part of the paper, we consider an environment in which the evolution of the relation-specific match values is observable, in which case the agents' vertical characteristics are the only source of informational asymmetry (recall that an agent's vertical "type" is meant to capture the value that the agent assigns to interacting with a generic agent from the opposite side, i.e., prior to conditioning on the specific profiles of the agents from the opposite side that join the platform). Such an environment is realistic in many circumstances. For example, an employment agency may be able to observe the evolution of the quality of the match between a firm and a worker. Importantly, even if the quality is not directly observed by the platform, the latter can always elicit such information at no cost from the partners when such information is jointly observed by the worker and the firm (for example using Cremer and Mc Lean-type of mechanisms). This environment also serves as a useful benchmark for the more complex case in which the value that each agent derives from each potential partner is the agent's private information.

When the values evolve exogenously over time, the optimal matching rule is *myopic*: in each period, it matches all agents whose *joint virtual surplus* is the highest, up to capacity, and excluding those matches for which the joint virtual surplus is negative. A pair's joint virtual surplus coincides with the pair's joint true surplus, adjusted by "handicaps" that control for the cost to the platform of linking the agents, in terms of informational rents that must be left to the agents.

When, instead, the match values evolve endogenously over time as a result of the experimentation in previous interactions, the optimal matching rule takes the form of a *virtual index rule*. In each period, the platform matches all pairs of agents for which the joint virtual index is the highest, up to capacity, and excluding those pairs for which the index is negative. As in other experimentation environments, such indexes summarize both current and future expected profitability of each match, taking into account the stochastic evolution of the agents' joint values. For the case of a capacity-constrained platform, in order to maintain tractability of the problem, we start by focusing on the case in which a single pair can be matched at each period and later discuss a few special environments for which the constraint can be extended to any arbitrary number of matches.[1,2]

The second part of the paper extends the analysis of optimal matching mechanisms to an environment in which the evolution of the match values is the agents' private information, that is, each agent privately observes how the value he/she attaches to each partner evolves over time. There are two key difficulties in this environment: (i) first, as is the case in other dynamic mechanism design problems, one needs to control for multi-period deviations; (ii) second, the information that each agent receives in each period is multi-dimensional, corresponding to the vector of values the agent assigns to each potential partner. Notwithstanding these complications, we show that the platform can obtain the same expected profits as in the environment with observable values. However, while, on-path, the matching dynamics are the same regardless of whether the values are privately observed, the mechanism that governs such dynamics is more complex under private information. The mechanism we propose belongs to the family of dynamic "virtual pivot mechanisms" that have been considered in recent years in the dynamic mechanism design literature. In each period, the agents report to the intermediary both their vertical type as well as the collection of relation-specific match values. The mechanism then implements an allocation that maximizes the discounted expected sum of all current and future interactions, but where the joint surplus of each interaction is replaced by a "weighted virtual surplus" that coincides with the virtual surplus in the benchmark with observable shocks, but only on path.

At each period following the initial one, the payments are then determined by a pricing formula similar to the one proposed by Bergemann and Valimaki (2010), but adapted to account for profit-maximization. These prices are then augmented by a correction based on the initial period reports whose role is to make the net present value of the expected payments from each agent coincide with the one in the benchmark with observable values. Importantly, these payments induce truthful reporting at each period both on and off path. In the environment with multidimensional types under examination here, this is accomplished by having the agents report in each period each component of their private information (as in Doepke and Townsend (2006) and, more recently, Kakade, Lobel

[1] When the process governing the evolution of the match values is exogenous, the results hold for any level of the capacity constraint.

[2] The case where a single match is feasible in each period is the one typically considered in the literature on dynamic assignment problems with stochastic arrivals. See, for example, Gershkov and Moldovau (2012), and Bloch and Cantala (2014).

and Nazerzadeh (2013)), and then making their payoff coincide with their flow marginal contribution to "weighted virtual surplus".

The properties of the profit-maximizing mechanism highlight the advantages of long-term contractual relationships in the provision of matching services over time. Specifically, agents from both sides of the market may be treated differently based on their initial private information, even when their current and expected future preferences coincide. To illustrate this, we show that the optimal matching rule can be implemented by means of a sequence of scoring auctions. In each period, agents first select a membership level and pay the corresponding membership fee. They then bid for all potential matches. The allocations and the corresponding payments are then determined by a scoring rule where the score of each match depends only on (a) current joint bids, (b) current membership status of each partner, (c) period-0 membership status, and (d) number of past interactions (in the experimentation model). In each period, high membership status is more costly, but buys a higher score and hence more favorable treatment.

The results also allow us to contrast matching dynamics under public (welfare-maximizing) and private (profit-maximizing) provision of matching services. When agents assign a nonnegative value to all interactions at every period, profit maximization results in fewer and shorter interactions. Specifically, absent a capacity constraint, each pair of agents is matched for an inefficiently shorter period of time. When the platform's capacity constraint is binding, certain pairs of agents may interact more under profit maximization than under welfare maximization. However, the aggregate number of interactions in each period under profit maximization is always lower than under welfare maximization. Interestingly, the above conclusion need not extend to markets in which certain agents derive a negative payoff from interacting with other agents (equivalently, a payoff lower than their outside option). In this case, profit maximization may result in an inefficiently large number of matches, for any number of periods.

1.1 Related Literature (incomplete)

Centralized dynamic matching markets. Much of the recent literature on centralized dynamic matching primarily focuses on markets without transfers in which agents are matched once, with dynamics stemming from the arrival and possibly departure of agents to and from the market. In the context of kidney exchange, Ünver (2010) studies optimal mechanisms for two-way and multi-way exchanges, minimizing total waiting costs, in a market with stochastic arrivals of donors and recipients. Optimal dynamic matching in markets in which agents gradually arrive over time and may be matched at most once is also the focus of Akbarpour, Li, and Oveis Gharan (2014), Anderson, Ashlagi, Gamarnik, and Kanoria (2014), Baccara, Lee, and Yariv (2015) and Herbst and Schickner (2015). A key tradeoff in such environements is between avoiding waiting costs and the benefit of waiting for the market to thicken. A related strand of literature studies the assignment of objects through waiting lists (see Bloch and Cantala 2014, Leshno 2014, and Schummer 2015 for recent developments).[3]

The key difference with respect to this literature is that in the current paper agents may be matched with different partners over time in response to changes in their valuations for one another. In particular, dynamics stem from gradual changes in preferences rather than population dynamics.

Matching with transfers. The paper is also related to the literature on profit-maximization in matching markets with private information and transfers. Damiano and Li (2007) and Johnson (2013) consider a *one-to-one* matching intermediary that faces asymmetric information about the agents' vertical characteristics that determine match values. 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). These papers derive conditions on primitives for a profit-maximizing intermediary to induce positive assortative matching. In contrast to these papers, Gomes and Pavan (2015) study *many-to-many* matching in a flexible setting where agents may differ in their consumer value (willingness-to-pay) and input value (salience). The key difference with respect to this body of research is that the present paper considers a dynamic environment in which match values evolve either exogenously or endogenously over time.

Position and Scoring Auctions. In the context of procurement settings, scoring auctions are often used to combine the offers of sellers for each of multiple dimensions (price and various attributes). See, for example, Che (1993) and Asker and Cantillon (2008) for treatment of such scoring auctions. Our implementation of the optimal match allocation using sequential scoring auctions has a similar flavor in that multiple aspects are combined to determine the score assigned to each potential match. Another related literature studies auctions for selling sponsored links. For example, Varian (2007) and Edelman, Ostrovsky and Schwarz (2007) provide theoretical analysis of the generalized second-price auction, variations of which are commonly used by search engines.[4] Relative to this literature, we consider sequential auctions that take into account the evolution of agents' preferences and incentives over time, and on both sides of the market.

Experimentation and Bandits. The paper is also related to the literature on experimentation in screening settings. In particular, the mechanisms proposed here can be seen as the analogs, in a matching environments, of the bandit auctions of Pavan et al. (2014) and Kakade et al. (2013) for the sale of an indivisible item. See also Bergemann and Valimaki (2006) for a survey on bandit

[3] While these papers focus on optimal matching in dynamic environments, a related literature explores appropriate stability notions for dynamic environments. See, for example, Damiano and Lam (2005) and Kurino (2009), who focus on markets with fixed population and matches changing over time. The main difference from the current work is the focus on stability, rather than optimality. Doval (2015) develops a notion of stability for an environment in which agents arrive over time and derives conditions for the existence of stable allocations.

[4] Other related works include Athey and Ellison (2011), Börgers, Cox, Pesendorfer and Petricek (2013), and Gomes (2014).

problems in economics.

Dynamic Mechanism Design. From a methodological standpoint, we draw from recent results in the dynamic mechanism design literature. In particular, the necessary and sufficient conditions for incentive compatibility in the present paper adapt to the environment under examination results in Theorems 1 and 3 in Pavan, Segal, and Toikka (2014). That paper provides a general treatment of incentive compatibility in dynamic settings. It extends previous work by Baron and Besanko (1984), Besanko (1985), Courty and Li (2000), Battaglini (2005), Board (2007), Eso and Szentes (2007, 2013), and Kapicka (2013), among others, by allowing for more general payoffs and stochastic processes and by identifying the role of impulse responses as the key driving force for the dynamics of optimal contracts. See Bergemann and Strack (2015) for a recent extension of the Myersonian approach in Pavan et al. to continuous time and Borges (2015) and Bergemann and Pavan (2015) for a discussion of the recent developments of the dynamic mechanism design literature. The agents' private information in most of this work is exogenous. Instead, the evolution of the agents' private information is endogenous in the experimentation setting considered in this paper, as well as in the bandit auction of Pavan et al. (2014), in the dynamic virtual pivot mechanism of Kakade et al. (2013), and in the taxation model of Makris and Pavan (2015).

Particularly related is also the strand of the dynamic mechanism design literature that investigates how to implement dynamically efficient allocations in settings in which the agents' types change over time, thus extending the Vickrey–Clarke–Groves (VCG) and d'Aspremont–Gérard-Varet (AGV) results from static to dynamic settings (see, for example, Bergemann and Valimaki, 2010, Athey and Segal, 2013, and the references therein).[5]

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 Weyl (2010), 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. In contrast, the present paper allows for general matching rules and for more flexible payoff structures. In particular, it does not restrict agents' willingness to pay to coincide with their attractiveness. Most importantly, it focuses on a dynamic environment in which match values change over time. Cabral (2011) also considers a dynamic model with network effects but in which values are constant over time.

[5] Another stream of the dynamic mechanism design literature considers both efficient and profit-maximizing mechanisms in settings where the agents' private information is static, but where interesting dynamics originate by agents or objects arriving stochastically over time (see the recent monograph by Gershkov and Moldovanu, 2014, as well as the overview by Bergemann and Said, 2012).

2 The Model

Agents, matches, and preferences

A platform mediates the interactions among agents from two sides of a market, A and B. There are $n_A \in \mathbb{N}$ agents on side A and $n_B \in \mathbb{N}$ agents on side B, with $N_A = \{1, ..., n_A\}$ and $N_B = \{1, ..., n_B\}$ denoting the corresponding sets of agents on the two sides. Time is discrete, indexed by $t = 0, 1, ..., \infty$. Agents live for infinitely many periods and can change partners infinitely many times.

Below, we describe various features of the environment from the perspective of a generic agent from side A. A similar description applies to side B.

The flow period-t utility that agent $i \in N_A$ from side A derives from being matched to agent $j \in N_B$ from side B is given by

$$u_{ijt}^A(\theta_i^A, \varepsilon_{ijt}^A) = \theta_i^A \cdot \varepsilon_{ijt}^A.$$

The parameter θ_i^A is time- and match-invariant and controls for the overall importance that agent i assigns to interacting with agents from the opposite side of the market (that is, θ_i^A parametrizes the value that agent i assigns to a generic agent from the opposite side, prior to conditioning on the specific profile of the latter agent). The parameter ε_{ijt}^A, instead, is match-specific and controls for the attractiveness of agent j from side B in the eyes of agent i. These match-specific values evolve over time, reflecting the change in the agents' true (or perceived) attractiveness. They can either represent the evolution of the agents' beliefs about fixed, but unknown, match qualities, or variations in attractiveness triggered by stochastic changes in the environment. For expositional convenience, hereafter we refer to θ_i^A as to the agent's (vertical) type and to ε_{ijt}^A as to the agent's period-t match value for agent j. We maintain that types are the agents' own private information. As for the match values ε_{ijt}^A, we will consider both the case in which they are observed by the platform and the case in which they are agent i's private information (as anticipated in the Introduction, the results for the case in which the sequence of ε_{ijt}^A is observed by the platform in turn coincide with those for an environment in which each match-specific pair of values $(\varepsilon_{ijt}^A, \varepsilon_{ijt}^B)$ is observed jointly by agents i and j, but not by the platform; the reason is that, in this latter case, $(\varepsilon_{ijt}^A, \varepsilon_{ijt}^B)$ can be elicited from agents i and j at no cost for the platform).

Agent i's type θ_i^A is drawn from an absolutely continuous cumulative distribution function (cdf) F_i^A with density f_i^A strictly positive over $\Theta_i^A = [\underline{\theta}_i^A, \overline{\theta}_i^A]$, with $\underline{\theta}_i^A \geq 0$. Agents' types are drawn independently across agents and from the match-specific values, ε. We denote by $\eta_i^A(\theta_i^A) = (1 - F_i^A(\theta_i^A))/f_i^A(\theta_i^A)$ the Mill's ratio (i.e., the multiplicative inverse of the hazard rate) of agent i's type distribution. As is standard in mechanism design, we assume that the functions $\eta_i^A(\cdot)$ are nonincreasing and refer to this property as "regularity". Furthermore, we assume that all *virtual types*, defined as $\phi_i^A(\theta_i^k) \equiv \theta_i^A - \eta_i^A(\theta_i^A)$, are nonnegative. Importantly, note that while we restrict the agents' virtual types to be nonnegative, we allow the agents' match-specific values ε_{ijt}^{k} to be negative, $k = A, B$, reflecting the possibility that an agent may derive a negative utility from interacting with certain agents from the opposite side.[6]

For any $t \geq 1$ and any pair of agents $(i, j) \in N_A \times N_B$, let $X_{ijt} = \{0, 1\}$, and interpret $x_{ijt} = 1$ as the decision to match agent i from side A to agent j from side B, in period t, and $x_{ijt} = 0$ as the decision to leave the two agents unmatched.[7]

Agents may be matched to more than a single agent from the opposite side. In each period $t \geq 1$, the platform has the capacity to match any number $m \leq M$ of pairs of agents from opposite sides, independently of past matching allocations. Note that M is a constraint on the stock of existing matches. In each period, the platform can delete some of the previously formed matches and create new ones. We only impose that the total number of existing matches be smaller or equal to M in all periods. We will consider both the case in which M is sufficiently large such that this capacity constraint never binds (i.e., $M \geq n_A \cdot n_B$), as well as the case $M < n_A \cdot n_B$ in which this constraint may be binding.[8] The set of feasible period-t matches is therefore given by

$$X_t = \left\{ x_t \in \prod_{(i,j) \in N_A \times N_B} X_{ijt} : \sum_{i \in N_A} \sum_{j \in N_B} x_{ijt} \leq M \right\}$$

whereas the set of sequences of feasible matching allocations is denoted by $X = \prod_{t=1}^{\infty} X_t$.

All agents are expected-utility maximizers and maximize the expected discounted sum of their flow payoffs using the common discount factor $\delta \in (0, 1]$. Let $p_t = (p_{it}^A, p_{jt}^B)_{i \in N_A, j \in N_B}$ denote the payments collected by the platform from the two sides of the market in period-t and $p = (p_t)_{t=0}^{\infty}$ an entire sequence of payments. Note that, while the matching starts in period one, we allow the platform to start collecting payments from the agents in period $t = 0$, after the agents have observed their types θ, but before they observe their values for specific partners. This assumption is motivated by the idea that, in many applications of interest, agents learn the identities (or the profiles) of the agents from the opposite side of the market who joined the platform only after "getting on board". Also note that payments are allowed to be negative, reflecting the idea that the platform may want to cross-subsidize certain interactions.

[6]Under the assumed multiplicative structure $u_{ijt}^A = \theta_i^A \cdot \varepsilon_{ijt}^A$, allowing the virtual values $\phi_i^A(\theta_i^k) = \theta_i^A - \eta_i^A(\theta_i^A)$ to also take on negative values is not desirable, for it faciliates confusion, given that the match-specific shocks ε_{ijt}^A are already allowed to take on negative values.

[7]Note that implicit in this formalization is a *reciprocity condition* imposing that whenever agent $i \in N_A$ is matched to agent $j \in N_B$, then agent j is matched to agent i (see also Gomes and Pavan (2015) for the role played by a similar condition in a static environment).

[8]It should also be clear from the analysis below that all our results extend to the case in which M changes over time.

The Bernoulli payoff function for each agent $i \in N_A$ from side A is given by

$$U_i^A(\theta, \varepsilon, x, p) = \sum_{t=0}^{\infty} \delta^t \sum_{j \in N_B} x_{ijt} u_{ijt}^A - \sum_{t=0}^{\infty} \delta^t p_{it}^A$$

where $\theta = (\theta_{it}^A, \theta_{jt}^B)_{i \in N_A, j \in N_B}$ and $\varepsilon = (\varepsilon_{ijt}^k)_{i \in N_A, j \in N_B, k \in \{A,B\}}^{t=1,\ldots,\infty}$. The platform's Bernoulli payoff function is the discounted sum of the payments collected from all agents,

$$U_0(\theta, \varepsilon, x, p) = \sum_{t=0}^{\infty} \delta^t \left(\sum_{i \in N_A} p_{it}^A + \sum_{j \in N_B} p_{jt}^B - \sum_{i \in N_A} \sum_{j \in N_B} x_{ijt} c_{ijt} \right),$$

where $c_{ijt} \geq 0$ is the flow cost of matching the pair (i, j).

Evolution of match values

Agents' flow utilities u_{ijt}^A from interacting with other agents are correlated over time, both through the fully persistent vertical component θ_i^A and through the partially persistent match-specific value ε_{ijt}^k. In particular, we assume that these values evolve over time according to the following process. For each pair $(i, j) \in N_A \times N_B$, and each period $t \geq 1$, let $\mathcal{E}_{ijt}^A \subseteq \mathbb{R}$ denote the set of possible values that agent i from side A may derive from interacting with agent j from side B, in period t. The evolution of these values is governed by the kernels $G = (G_{ijt}^k)_{(i,j) \in N_A \times N_B, k \in \{A,B\}}^{t=1,\ldots,\infty}$. In particular, the period-1 values ε_{ij1}^A are drawn from $\mathcal{E}_{ij1}^A$ according to the cdf G_{ij1}^A. In each subsequent period $t > 1$, the period-t value ε_{ijt}^A is drawn from $\mathcal{E}_{ijt}^A$ according to the cdf $G_{ijt}^A(\varepsilon_{ijt}^A \mid \varepsilon_{ijt-1}^A, x^{t-1})$ where $\varepsilon_{ijt-1}^A \in \mathcal{E}_{ijt-1}^A$ is the match value in the preceding period and where $x^{t-1} = (x_s)_{s=1}^{t-1}$ denotes the entire sequence of previous interactions.

We will consider two environments, capturing different features of dynamic matching markets with private information. The first environment is one in which the evolution of the match values is exogenous to the platform's decisions. In the second environment, the evolution of the match values depends on past interactions, with properties reflecting (private) experimentation.

Exogenous processes. For all $(i, j) \in N_A \times N_B$, $t > 1$ and $\varepsilon_{ijt-1}^A \in \mathcal{E}_{ijt-1}^A$, $G_{ijt}^A(\cdot \mid \varepsilon_{ijt-1}^A, x^{t-1})$ does not depend on x^{t-1}.

Experimentation model. For any $(i, j) \in N_A \times N_B$, the following properties hold: (i) whenever $x_{ijt-1} = 1$, the dependence of $G_{ijt}^A(\varepsilon_{ijt}^A \mid \varepsilon_{ijt-1}^A, x^{t-1})$ on x^{t-1} is only through $\sum_{s=1}^{t-1} x_{ijs}$ (ii) whenever $x_{ijt-1} = 0$, $G_{ijt}^A(\cdot \mid \varepsilon_{ijt-1}^A, x^{t-1})$ is a Dirac measure at $\varepsilon_{ijt}^A = \varepsilon_{ijt-1}^A$, i.e.,

$$G_{ijt}^A(\varepsilon_{ijt}^A \mid \varepsilon_{ijt-1}^A, x^{t-1}) = \mathbf{1}_{\{\varepsilon_{ijt}^A \geq \varepsilon_{ijt-1}^A\}}.$$

(iii) there exists a sequence $(\omega_{ijs}^A)_{s=1}^{\infty}$ drawn from an exogenous distribution, such that, for any number S_{ij} of past interactions between agent i from side A and agent j from side B, the period-t match value ε_{ijt}^A is given by a deterministic function of $(\omega_{ijs}^A)_{s=1}^{S_{ij}}$, uniformly over t.

The latter formulation, henceforth referred to as the *experimentation model,* captures the following key properties: (1) agents' values for potential partners change only as a result of experimentation. That is, if $x_{ijt-1} = 0$, then $\varepsilon^A_{ijt} = \varepsilon^A_{ijt-1}$ almost surely; (2) The processes governing the agents' match values are Markov time-homogeneous: if agents $(i, j) \in N_A \times N_B$ are matched in period $t-1$, then the distribution of agent i's period-t value for agent j, ε^A_{ijt}, depends only on agent i's period-$(t-1)$ value ε^A_{ijt-1} and the number of times the two agents have been matched prior to period t, $\sum_{s=1}^{t-1} x_{ijs}$.

The above formulation can capture, for example, an environment in which agents gradually learn about their (unknown) true values for interacting with agents from the other side.

Example 1 (Gaussian learning) Suppose that every agent $i \in N_A$ from side A has a constant match values ω^A_{ij} for being matched with agent $j \in N_B$ from side B, which is unknown to the agent. Agent i starts with a prior belief that $\omega^A_{ij} \sim N(\varepsilon^A_{ij1}, \tau^A_{ij})$, where the variance τ^A_{ij} is common knowledge but where the initial prior mean ε^A_{ij1} is the agent's private information. Each ε^A_{ij1} is drawn from a distribution G^A_{ij1}. Each time agent i is matched to agent j, agent i receives a conditionally i.i.d. private signal[9] $\xi^A_{ij} \sim N(\omega^A_{ij}, \sigma^A_{ij})$ and updates his expectation of ω^A_{ij} using standard projection formulae. Take ε^A_{ijt} to be agent i's posterior about ω^A_{ij} in period t. Given ε^A_{ijt-1}, if $x_{ijt-1} = 1$ then $G^A_{ijt}(\cdot \mid \varepsilon^A_{ijt-1}, x^{t-1}_{ij})$ is the c.d.f. for the posterior expectation as a result of observing $\sum_{s=1}^{t-1} x_{ijs}$ signals. $\square$

As mentioned above, fixing $x = (x_t)_{t=1}^{\infty}$, the value that each agent i from side A derives from being matched to any agent j from side B is independent of the value the same agent assigns to any other agent from side B. This assumption facilitates the characterization of the optimal mechanism by favoring an index representation of the optimal policy. The assumption that values (as well as types) are independent across agents in turn guarantees that the platform cannot extract the entire surplus from the agents using payments similar to those in Cremer and McLean (1988). Notwithstanding this clarification, note that an agent's match values may still depend on other agents' private information through the observable past matches x^{t-1}.

Mechanisms, efficiency, and profit maximization

The evolution of the matches between the two sides is governed by the platform through a *dynamic matching mechanism,* $\mathcal{M} \equiv (\chi, \psi)$. The latter consists of a *matching rule* $\chi \equiv (\chi_t)_{t=1}^{\infty}$ and a *payment rule* $\psi \equiv (\psi_t)_{t=0}^{\infty}$, where the functions $\chi_t : \Theta \times \mathcal{E}^t \to X_t$ and $\psi_t : \Theta \times \mathcal{E}^t \to \mathbb{R}^{N_A + N_B}$ map histories of reports about types, θ, and match values $\varepsilon^t = (\varepsilon^k_{ijs})^{s=1,\ldots,t}_{i \in N_A, j \in N_B, k \in \{A,B\}} \in \mathcal{E}^t$ into period-t allocations and payments, respectively (in environments in which the match values are observable, the mechanism naturally conditions on the observed values rather than the reported ones). We denote by $\mathcal{X}$ the set

[9] By this, we mean that each signal ξ^A_{ij} can be written as $\xi^A_{ij} = \omega^A_{ij} + \varsigma^A_{ij}$ with the innovations ς^A_{ij} drawn from a Normal distribution with mean 0 and variance σ^A_{ij}, independently from all other random variables.

of feasible matching rules. Note that, while reporting and payments start in period zero, the actual matching begins in period 1.

A matching rule χ, combined with the distributions F and the kernels G, $k = A, B$, defines a stochastic process over $\Theta \times \mathcal{E}$. We denote this process by $\lambda[\chi]$ and then denote by $\lambda[\chi]|\theta_i^A$ the process from the perspective of agent i from side A after the latter observes θ_i^A. To guarantee that the expected payoff of each agent is well defined and satisfies a certain envelope formula (more below), we assume that for all $i \in N_A$, there exists a constant $K_i^A > 0$ such that, for any $\chi \in \mathcal{X}$ $\mathbb{E}^{\lambda[\chi]|\theta_i^A}\left[\sum_{t=1}^{\infty} \delta^t \sum_{j \in N_B} |\varepsilon_{ijt}^A| \cdot \chi_{ijt}(\theta, \varepsilon^t)\right] \leq K_i^A$.

Following Myerson (1986), we restrict attention to mechanisms that are *on-path incentive compatible*: in these mechanisms agents are asked to report their private information in each period and find it optimal to report truthfully on the equilibrium path. That is, at any period $t \geq 0$, conditional on having reported truthfully in the past, each agent finds it optimal to continue reporting truthfully in the continuation game that starts with period t, when all other agents are expected to report truthfully in all periods.

A mechanism is said to be *individually rational* if each agent finds it optimal to participate in the mechanism at period zero, again when all agents are expected to report truthfully in all periods.

A dynamic matching mechanism is profit-maximizing if it maximizes

$$\Omega_{\mathcal{M}}^P = \mathbb{E}^{\lambda[\chi]}\left[\sum_{t=0}^{\infty} \delta^t \left(\sum_{i \in A} \psi_{it}^A(\theta, \varepsilon^t) + \sum_{j \in N_B} \psi_{jt}^B(\theta, \varepsilon^t) - \sum_{i \in N_A} \sum_{j \in N_B} \chi_{ijt}(\theta, \varepsilon^t) c_{ijt}\right)\right]$$

and is efficient if it maximizes

$$\Omega_{\mathcal{M}}^W \equiv \mathbb{E}^{\lambda[\chi]}\left[\sum_{t=1}^{\infty} \delta^t \sum_{i \in N_A} \sum_{j \in N_B} \chi_{ijt}(\theta, \varepsilon^t) \cdot \left(u_{ijt}^A(\theta_i^A, \varepsilon_{ijt}^A) + u_{ijt}^B(\theta_j^B, \varepsilon_{ijt}^B) - c_{ijt}\right)\right],$$

among all mechanisms that are on-path incentive compatible and individually rational.

Remark. The above definitions of incentive compatibility and individual rationality require that each agent (a) finds it optimal to report truthfully on-path, and (b) participate in period zero. Below we show that the platform can, at no additional cost, guarantee stronger forms of incentive compatibility and individual rationality, by which each agent finds it optimal to remain in the mechanism and report truthfully at all histories, irrespective of the agent's previous reports (that is, even if she has lied in the past) and irrespective of the agent's beliefs about other agents' types and past and current values (but provided the agent expects these latter agents to report truthfully in the continuation games). Such stronger requirements are often referred to as *ex-post periodic* in the dynamic mechanism design literature.

3 Optimal mechanisms – observable match values

We start by considering an environment in which the platform observes the evolution of the match-specific values. Alternatively, as mentioned above, the results in this section also apply to an environment in which each pair of match-specific values $(\varepsilon^A_{ijt}, \varepsilon^B_{ijt})$ is jointly observed by the pair $(i,j) \in N_A \times N_B$, but not by the platforms; in fact, in this latter case, the platform can always induce the pair of agents to report truthfully $(\varepsilon^A_{ijt}, \varepsilon^B_{ijt})$ at no cost to the platform. Throughout the analysis, we maintain the assumption that the vertical types θ are the agents' private information.

3.1 Incentive compatibility and average monotonicity

In this environment, the expected discounted payoff that type $\theta^A_i \in \Theta^A_i$ of agent i from side A obtains from reporting $\hat{\theta}^A_i \in \Theta^A_i$, when all other agents report truthfully, is given by (here $\theta^A_{-i} = ((\theta^A_l)_{l\in N_A, l\neq i}, (\theta^B_l)_{l\in N_B})$):

$$U^A_i(\theta^A_i; \hat{\theta}^A_i) \equiv \mathbb{E}^{\lambda[x]|\hat{\theta}^A_i}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B} \theta^A_i \varepsilon^A_{ijt}\chi_{ijt}(\hat{\theta}^A_i, \theta^A_{-i}, \varepsilon^t) - \sum_{t=0}^{\infty}\delta^t \psi^A_{it}(\hat{\theta}^A_i, \theta^A_{-i}, \varepsilon^t)\right].$$

Incentive compatibility then requires that, for all $i \in N_A$, all $\theta^A_i, \hat{\theta}^A_i \in \Theta^A_i$, $U^A_i(\theta^A_i) \geq U^A_i(\hat{\theta}^A_i; \theta^A_i)$, where $U^A_i(\theta^A_i) = U^A_i(\theta^A_i; \theta^A_i)$, while individually rationality requires that, for all $i \in N_A$, all $\theta^A_i \in \Theta^A_i$, $U^A_i(\theta^A_i) \geq 0$. Analogous conditions obviously apply to all agents from side B.

In static mechanism design, it is well known that incentive compatibility requires the allocation rule to satisfy appropriate monotonicity conditions. Dynamic analogs of such conditions have been discussed in the recent dynamic mechanism design literature (see, e.g., Pavan, Segal, and Toikka, 2014, for an overview). In the simple dynamic environment under consideration here, incentive compatibility requires that, for each agent, the expected net present value of her future allocations, discounted by the match-specific values, be nondecreasing in the period-0 type report. Formally, for any matching rule $\chi \in \mathcal{X}$, any $l \in N_k$, $k = A, B$, let[10]

$$D^k_l(\theta^k_l; \chi) = \begin{cases} \mathbb{E}^{\lambda[x]|\theta^k_l}\left[\sum_{t=1}^{\infty}\delta^t \sum_{h\in N_{-k}} \varepsilon^k_{lht}\chi_{lht}(\theta, \varepsilon^t)\right] & \text{if } k = A \\ \mathbb{E}^{\lambda[x]|\theta^k_l}\left[\sum_{t=1}^{\infty}\delta^t \sum_{h\in N_{-k}} \varepsilon^k_{hlt}\chi_{hlt}(\theta, \varepsilon^t)\right] & \text{if } k = B \end{cases} \tag{1}$$

denote the expected net present value of the agent's future allocations, discounted by the agent's match values.

Definition 1 (average monotonicity) *The matching rule $\chi \in X$ satisfies average monotonicity if, for all $l \in N_k$, $k = A, B$, the function $D^k_l(\theta^k_l; \chi)$ is nondecreasing in θ^k_l.*

[10] Note that the reason we distinguish between the case in which $k = A$ and the one in which $k = B$ is that the order in the subscripts of the allocations χ_{ijt}, as well as the order in the subscripts in the match-specific values ε^k_{ijt} is not permutable: the first index always refers to side A, while the second to side B.

This condition is the analog of Pavan, Segal and Toikka (2014)'s average monotonicity in the environment under examination here (note that the match-specific values ε_{ijt}^A coincide with Pavan, Segal, and Toikka's "impulse responses" of future types $u_{ijt}^A = \theta_i^A \cdot \varepsilon_{ijt}^A$ to the initial type, θ_i^A; one difference, though, is that here each agent's future type $u_{it}^A = (u_{ijt}^A)_{j \in N^B}$ is multi-dimensional). We then have the following result.

Lemma 1 (necessary and sufficient conditions for IC) *Suppose that match values are observable. (1) Let $\mathcal{M} = (\chi, \psi)$ be an incentive-compatible matching mechanism. Then (a) χ satisfies average monotonicity and (b) ψ satisfies the following formula (aka envelope condition) for all $l \in N_k$, $k = A, B$,*

$$\mathbb{E}^{\lambda[\chi]|\theta_l^k}\left[\sum_{t=0}^{\infty} \delta^t \psi_{lt}^k(\theta, \varepsilon^t)\right] = \theta_l^k D_l^k(\theta_l^k; \chi) - U_l^k(\underline{\theta}_l^k) - \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y; \chi) dy. \tag{2}$$

(2) Let $\chi \in \mathcal{X}$ be a feasible matching rule satisfying average monotonicity and such that $D_l^k(\underline{\theta}_l^k; \chi) \geq 0$ for all $l \in N_k$, $k = A, B$. Then there exists a payment rule ψ such that the mechanism $\mathcal{M} =(\chi, \psi)$ (a) is individually rational and incentive compatible and (b) gives zero surplus to the lowest type of each agent (i.e., $U_l^k(\underline{\theta}_l^k) = 0$, all $l \in N_k$, $k = A, B$).

The arguments in the proof parallel those in static environments. Note that a stronger form of monotonicity would require that the discounted sum of match values be monotone *ex-post*, that is, for any given (θ, ε). However, as will be shown below, such a stronger notion of monotonicity (aka ex-post monotonicity in the dynamic mechanism design literature) is not necessary for IC and fails to hold under the optimal allocation rule, in many environments.

Also note that the payment rule in (2) implies that each agent's period-0 payoff satisfies the following condition:

$$U_l^k(\theta_l^k) = U_l^k(\underline{\theta}_l^k) + \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y; \chi) dy.$$

3.2 Optimal matching rules

The dynamic problem the platform faces thus consists in choosing which pairs of agents to link at each period, as a function of the agents' initial private information (the overall value they attach to reaching the other side), and the evolution of the match-specific values. Below we introduce two rules that will play an important role for the results.

For any pair of potential partners $(i, j) \in N_A \times N_B$, given the reported types θ and the period-t match values ε_t, denote the *joint virtual surplus* from matching i and j in period t by

$$V_{ijt}(\theta, \varepsilon_t) = \phi_i^A(\theta_i^A)\varepsilon_{ijt}^A + \phi_j^B(\theta_j^B)\varepsilon_{ijt}^B - c_{ijt}. \tag{3}$$

Now let $Q_t(\theta, \varepsilon_t)$ denote the set of links for which the joint virtual surplus is non-negative and for which there are at most $M-1$ links with a strictly higher joint virtual surplus. Formally,

$$Q_t(\theta, \varepsilon_t) = \left\{ \begin{array}{l} (i,j) \in N_A \times N_B \text{ s.t. (i) } V_{ijt}(\theta, \varepsilon_t) \geq 0 \text{ and} \\ \text{(ii) } \#\{(l,m) \in N_A \times N_B : V_{lmt}(\theta, \varepsilon_t) > V_{ijt}(\theta, \varepsilon_t)\} < M \end{array} \right\}.$$

Definition 2 (myopic rule) *Let χ^m denote the "myopic" matching rule that links at each period $t \geq 1$ the pairs with the highest nonnegative joint virtual surplus, subject to the platform's capacity constraint. Formally,*

$$\chi^m_{ijt}(\theta, \varepsilon^t) = 1 \;\Rightarrow\; (i,j) \in Q_t(\theta, \varepsilon_t).$$

Furthermore,

$$(i,j) \in Q_t(\theta, \varepsilon_t),\ \#Q_t(\theta, \varepsilon_t) \leq M,\ V_{ijt}(\theta, \varepsilon_t) > 0 \;\Rightarrow\; \chi^m_{ijt}(\theta, \varepsilon^t) = 1.$$

In case $Q_t(\theta, \varepsilon_t)$ contains more than M links, χ^m selects M links from $Q_t(\theta, \varepsilon_t)$ in any arbitrary way.

Next consider the following forward-looking matching rule, which takes into account not only the current, but also the future expected joint virtual surplus of matching any pair of agents. Specifically, for each pair of agents $(i,j) \in N_A \times N_B$, define the period-t *virtual index* of the pair, given $(\theta, \varepsilon^t, x^{t-1})$ as

$$\gamma_{ijt}\left(\theta, \varepsilon_t, x^{t-1}\right) = \max_{\tau} \mathbb{E}^{\lambda[\chi^o]|\theta, \varepsilon_t, x^{t-1}} \left\{ \frac{\sum_{s=t}^{\tau} \delta^{s-t} V_{ijs}(\theta, \varepsilon_s)}{\sum_{s=t}^{\tau} \delta^{s-t}} \right\},$$

where τ is a stopping time and where $\lambda[\chi^o]|\theta, \varepsilon_t, x^t$ is the probability distribution over $\Theta \times \mathcal{E}^{\geq t}$ that obtains when one starts from $(\theta, \varepsilon_t, x^{t-1})$ and in all periods $s \geq t$ only agents i, j are matched. Importantly, note that, under the assumptions of the experimentation environment described above, the virtual index corresponding to each pair $(i,j) \in N_A \times N_B$ depends on the "state of the system" $(\theta, \varepsilon_t; x^{t-1})$ only through $(\theta^A_i, \theta^B_j, \varepsilon^A_{ij}, \varepsilon^B_{ij})$ and the total number of times $\sum_{s=1}^{t-1} x_{ijs}$ the two agents interacted in the past. It is completely independent of the processes governing the evolution of the other match values.

Now denote by $J_t(\theta, \varepsilon_t, x^{t-1})$ the set of period-t matches for which the virtual index is non-negative and for which there are at most $M-1$ links with a strictly higher virtual index. Formally,

$$J_t(\theta, \varepsilon_t, x^{t-1}) = \left\{ \begin{array}{l} (i,j) \in N_A \times N_B \text{ s.t. (i) } \gamma_{ijt}\left(\theta, \varepsilon_t, x^{t-1}\right) \geq 0 \text{ and} \\ \text{(ii) } \#\left\{(l,m) \in N_A \times N_B : \gamma_{lmt}\left(\theta, \varepsilon_t, x^{t-1}\right) > \gamma_{ijt}\left(\theta, \varepsilon_t, x^{t-1}\right)\right\} < M \end{array} \right\}.$$

Definition 3 (virtual index rule) *Let χ^I denote the "virtual index" matching rule that links at each period $t \geq 1$ the pairs with the highest nonnegative virtual index, subject to the platform's* *capacity constraint. Formally, χ^I is defined inductively as follows:*

$$\chi_{ijt}^I(\theta, \varepsilon^t) = 1 \;\Rightarrow\; (i,j) \in J_t(\theta, \varepsilon_t, \chi^{It-1}(\theta, \varepsilon^{t-1})).$$

Furthermore,

$$\begin{aligned}(i,j) \;&\in\; J_t(\theta, \varepsilon_t, \chi^{It-1}(\theta, \varepsilon^{t-1})),\ \#J_t(\theta, \varepsilon_t, \chi^{It-1}(\theta, \varepsilon^{t-1})) \leq M,\ \gamma_{ijt}(\theta, \varepsilon_t, \chi^{It-1}(\theta, \varepsilon^{t-1})) > 0 \\ &\Rightarrow\; \chi_{ijt}^I(\theta, \varepsilon^t) = 1.\end{aligned}$$

The remainder of this section shows that the myopic matching rule χ^m and virtual index rule χ^I are optimal in the exogenous process environment and the experimentation environment, respectively. The results are summarized in Theorem 1 below.

Theorem 1 (optimal rules) *Consider the environment with observable match values described above.*

(1) Suppose match values evolve exogenously and $D_l^k(\underline{\theta}_l^k; \chi^m) \geq 0$, all $l \in N_k$, $k = A, B$. There exists a payment rule ψ^m such that the mechanism (χ^m, ψ^m) is profit-maximizing.

(2) Suppose the environment satisfies the assumptions of the experimentation model defined above, that $D_l^k(\underline{\theta}_l^k; \chi^I) \geq 0$, all $l \in N_k$, $k = A, B$, and that either $M \geq n_A \cdot n_B$ or $M = 1$. Then there exists a payment rule ψ^I such that the mechanism (χ^I, ψ^I) is profit-maximizing.

The proof of Theorem 1 in the Appendix 1 proceeds in three steps. The first step uses the results in Lemma 1 to express the platform's expected profits in terms of dynamic virtual surplus

$$\mathbb{E}^{\lambda[\chi]}\left[\sum_{t=0}^{\infty} \delta^t \sum_{i \in N_A} \sum_{j \in N_B} V_{ijt}(\theta, \varepsilon_t) \cdot \chi_{ijt}(\theta, \varepsilon^t)\right] \tag{5}$$

where the latter parallels the expression for total surplus, but with the virtual types $\phi_l^k(\theta_l^k)$ replacing the true ones, θ_l^k. More specifically, Lemma 2 in the Appendix establishes that if a matching rule χ maximizes dynamic virtual surplus, gives the lowest type of each agent a sufficiently high gross payoff, and satisfies average monotonicity, then there exists a payment rule ψ such that the pair (χ, ψ) constitutes a profit-maximizing matching mechanism. The second step (Lemma 3 in the Appendix) establishes that the myopic rule χ^m and the virtual index rule χ^I maximize dynamic virtual surplus in the two respective environments of Theorem 1. The the final step (Lemma 4 in the Appendix) establishes that χ^m and χ^I satisfy average monotonicity.

Consider the experimentation environment described above, in which the match values evolve endogenously as a function of past interactions. In this environment, the platform's dynamic programming problem of maximizing (5) cannot be solved pointwise. In particular, the platform's problem in this case defines a *multiarmed bandit problem.* This problem admits an index policy solution when either the capacity constraint is not binding or when the platform can match at most one pair of agents at a time. When, instead, $1 < M < n_A \cdot n_B$, in general, without further restrictions on the environment, it is not guaranteed that an index rule maximizes dynamic virtual surplus. This is because, in general, multiarmed bandit problems in which multiple arms can be activated simultaneously fail to admit a simple index solution. In Section 7 below we present an experimentation environment in which the virtual index rule is optimal for all $M \in \mathbb{N}$. This environment, however, requires discreteness of the indexes.

The following example describes a special case of the experimentation environment in which the surplus generated by a match (randomly) decreases over time. In this case, the virtual index rule χ^I coincides with the myopic rule χ^m.

Example 2 (deteriorating matches) Suppose that for all $t \geq 1$, $k = A, B$, $(i,j) \in N_A \times N_B$, $\varepsilon^k_{ijt+1} \leq \varepsilon^k_{ijt}$ a.s. and $c_{ijt+1} \geq c_{ijt}$ irrespective of the history. In this case, the virtual index γ_{ijt} for each pair of agents coincides with the virtual surplus V_{ijt}, pathwise.[11] Therefore, the virtual index rule χ^I coincides with the myopic rule χ^m. For $M \geq n_A \cdot n_B$ or $M = 1$, the myopic rule is therefore optimal in this experimentation environment.[12] □

More generally, though, the allocations implemented under an index rule differ from those implemented under a myopic rule, due to the value that the platform assigns to experimentation. Interestingly, the platform's value for experimentation need not coincide with the agents' value and this may call for distortions in the selection of the allocations. Notwithstanding this, when the rule is sufficiently monotone, in the sense of the above definition, the platform can induce the agents to report their initial types truthfully. In this respect, note that, when the matches are governed by a myopic rule, match quality is monotone in the reported initial types not only on average, but also ex-post, and period-by-period. That is, for every $t \geq 1$, $\theta \in \Theta$, $\varepsilon \in \mathcal{E}$, for each agent $l \in N_A$,

$$\sum_{h \in N_B} \varepsilon^A_{lht} \cdot \chi^m_{lht}(\theta, \varepsilon^t)$$

is nondecreasing in the agent's reported type θ^A_l (an analogous condition holds for agents on side B).

As anticipated above, this latter property, however, need not hold under the virtual index rule. This is because a higher reported type θ^A_i, by changing the virtual indexes of all of i's potential matches, may lead to reversals in the ordering of these indexes. Because the virtual index rule is forward looking, such reversals, while optimal based on the period-t information, might not be optimal ex-post. As a result, the monotonicity of an agent's discounted sum of current and future match values under χ^I need not hold ex-post. The following example illustrates.

[11] This is because the optimal stopping time satisfies the property $\tau_{ijt}(\theta, \varepsilon_t, x^{t-1}) = \inf\{s > t \mid \gamma_{ijs} \leq \gamma_{ijt}\}$. It is the first time at which the process V_{ijt} reaches a state in which the virtual index drops (weakly) below its initial value.

[12] For $1 < M < n_A \cdot n_B$, however, this property does not guarantee the optimality of χ^I.

Example 3 (failure of ex-post monotonicity) Suppose that $M = 1$, $N_A = \{i\}$, $N_B = \{j, j'\}$, and $c_{ijt} = c_{ij't} = 0$ all t. Consider two types $\hat{\theta}_i^A, \tilde{\theta}_i^A \in \Theta_i^A$ and assume Θ_i^A and the distribution F_i^A are such that the virtual types corresponding to θ_{i1}^A and θ_{i2}^A satisfy $\phi_i^A(\hat{\theta}_i^A) = 0$ and $\phi_i^A(\tilde{\theta}_i^A) = 1$. On side B, suppose that $\Theta_j^B = \{\theta_j^B\}$ and $\Theta_{j'}^B = \{\theta_{j'}^B\}$ and that $\phi_j^B(\theta_j^B) = \theta_j^B = 1$ and $\phi_{j'}^B(\theta_{j'}^B) = \theta_{j'}^B = 0$.

Consider the following processes of the agents' match values. For all $t \geq 1$, $\varepsilon_{ij't}^A = 2$. The value ε_{ijt}^A that i assigns to j is, instead, equal to 0 in period $t = 1$, whereas for all $t \geq 2$, $\varepsilon_{ijt}^A = \varepsilon_{ijt-1}^A$ if $x_{ijt-1} = 0$, while, if $x_{ijt-1} = 1$, ε_{ijt}^A is equal to 100 with probability β and -100 with probability $1-\beta$, for some $\beta \in (0, 1)$. Furthermore, $\varepsilon_{ijt}^B = 1$ all t, whereas the value $\varepsilon_{ij't}^B$ that j' assigns to i does not play any role in this example, so the process for this value can be taken arbitrary.

In this example, the match between i and j' is a "safe" match, yielding a deterministic flow virtual surplus $V_{ij't} = 0$ for each $t \geq 1$ when $\theta_i^A = \hat{\theta}_i^A$ and $V_{ij't} = 2$ when $\theta_i^A = \tilde{\theta}_i^A$. The match between i and j yields a deterministic flow virtual surplus $V_{ijt} = 1$ for all $t \geq 1$ when $\theta_i^A = \hat{\theta}_i^A$. When, instead, $\theta_i^A = \tilde{\theta}_i^A$, this match yields $V_{ij1} = 1$ in period 1, and an expected flow virtual surplus $\mathbb{E}[V_{ijt}] = 1 + 100\beta + (1-\beta)(-100)$ at all $t \geq 2$. When $\theta_i^A = \hat{\theta}_i^A$, the virtual index rule χ^I matches agents i and j in all periods $t \geq 1$. When, instead, $\theta_i^A = \tilde{\theta}_i^A$, for β sufficiently small the virtual index rule χ^I matches agent i with agent j' in all periods $t \geq 1$. An increase in i's report from $\hat{\theta}_i^A$ to $\tilde{\theta}_i^A$ thus implies a change in i's partner from j to j' at all periods $t \geq 1$.

Now consider a sequence of values $\hat{\varepsilon}$ such that $\hat{\varepsilon}_{ijt}^A = 100$ all t, whenever $x_{ijt-1} = 1$. Then, for all $\delta > 1/50$, under ε,

$$\sum_{t=1}^{\infty} \delta^t \sum_{r \in N_B} \hat{\varepsilon}_{irt}^A \chi_{irt}^I(\tilde{\theta}_i^A, \theta_{-i}^A, \hat{\varepsilon}^t) < \sum_{t=1}^{\infty} \delta^t \sum_{r \in N_B} \hat{\varepsilon}_{irt}^A \chi_{irt}^I(\hat{\theta}_i^A, \theta_{-i}^A, \hat{\varepsilon}^t)$$

which, since $\hat{\theta}_i^A < \tilde{\theta}_i^A$, means that the virtual index rule χ^I fails ex-post monotonicity.[13] □

4 Optimal mechanisms – private match values

We now turn to an environment in which only agent i observes the evolution of his match-specific values for agent j and likewise for agent j. That is, the match-specific values are the agents' own private information. As mentioned in the Introduction, since private information evolves over time, the key difficulty is the need to control for multi-period contingent deviations. Nevertheless, we will show that the platform can obtain the same expected profits as in the environment with observable values examined in the previous section, using a matching mechanism that guarantees each agent's participation and truthful reporting even at histories for which she has misreported in the past, and for any belief about other agents' types and values.

The approach here is similar to that in Doepke and Townsend (2006) and in Kakade et al. (2013). It consists in enlarging the message space in all periods $t \geq 1$ so as to give the agents the possibility

[13] While, for simplicity, the example assumes discrete types, the result does not hinge on this property.

to reveal their true types after possible lies in previous periods.[14] Asking the agents to report again their type θ_l^k after period $t = 0$ is a convenient trick which allows us to control for the agents' behavior off path. We use this in turn to establish the optimality of truthful reporting. As implied by the Revelation Principle (see, e.g., Myerson (1986)), such re-reporting never expands the set of implementable policies.

To differentiate the period-0 reported types from the types reported in subsequent periods, we will denote the period-0 reported types by $z \in \Theta$. For any pair of agents $(i, j) \in N_A \times N_B$, any period-0 profile of reported types $z \in \Theta$, any period-t (with $t \geq 1$) profile of reported types $\theta_t \in \Theta$, and any period-t profile of reported match values $\varepsilon_t \in \mathcal{E}_t$, we then let

$$\tilde{V}_{ijt}\left(z, \theta_t, \varepsilon_t\right) \equiv \left(\frac{\phi_i^A(z_i^A)}{z_i^A}\right)\theta_{it}^A\varepsilon_{ijt}^A + \left(\frac{\phi_i^B(z_i^B)}{z_i^B}\right)\theta_{jt}^B\varepsilon_{ijt}^B - c_{ijt}$$

denote the *weighted period-t virtual surplus* from matching agents i and j. Note that such surplus coincides with the virtual surplus defined in the previous section when the agents report truthfully, in which case $z_i^A = \theta_{it}^A$ all t.

Next, consider an *extended* matching mechanism $\mathcal{M} = (\tilde{\chi}, \tilde{\psi})$ in which the matching and payment rules

$$\tilde{\chi}_t : \Theta^2 \times \mathcal{E}_t \times X^{t-1} \to X_t \text{ and } \tilde{\psi}_t : \Theta^2 \times \mathcal{E}_t \times X^{t-1} \to \mathbb{R}^{N_A+N_B}$$

are adjusted so that their arguments now contain the period-0 and period-t reported types $(z, \theta_t) \in \Theta^2$, the period-$t$ match values $\varepsilon_t \in \mathcal{E}_t$, and the history $x^{t-1} \in X^{t-1}$ of past allocations. As it will become clear below, conditioning on past allocations is a convenient way to express the functioning of the mechanism off-path in the case of endogenous processes.

Paralleling the analysis in the previous section, let $\tilde{\chi}^m$ denote the matching rule that links at each period $t \geq 1$ the pairs $(i, j) \in N_A \times N_B$ with highest nonnegative weighted period-t virtual surplus $\tilde{V}_{ijt}$, subject to the platform's capacity constraint. Formally, let

$$\tilde{Q}_t(z, \theta_t, \varepsilon_t) = \left\{ \begin{array}{l} (i, j) \in N_A \times N_B \text{ s.t. (i) } \tilde{V}_{ijt}(z, \theta_t, \varepsilon_t) \geq 0 \text{ and} \\ \text{(ii) } \#\left\{(l, m) \in N_A \times N_B : \tilde{V}_{lmt}(z, \theta_t, \varepsilon_t) > \tilde{V}_{ijt}(z, \theta_t, \varepsilon_t)\right\} < M \end{array} \right\}.$$

Definition 4 (weighted myopic rule) *Let $\tilde{\chi}^m$ denote the "weighted myopic" virtual matching rule linking at each period $t \geq 1$ the pairs with the highest nonnegative weighted joint virtual surplus,*

[14] That is, we let the message space of each agent $l \in N_A$ at each period $t \geq 1$ be equal to $\Theta_l^A \times \prod_{h \in N_B} \mathcal{E}_{lht}^A$, and similarly for each agent $h \in N_B$.

subject to the platform's capacity constraint. That is,[15]

$$\tilde{\chi}^{m}_{ijt}(z,\theta_t,\varepsilon_t,x^{t-1}) = 1 \;\Rightarrow\; (i,j) \in \tilde{Q}_t(z,\theta_t,\varepsilon_t).$$

Furthermore,

$$(i,j) \in \tilde{Q}_t(z,\theta_t,\varepsilon_t),\ \#\tilde{Q}_t(z,\theta_t,\varepsilon_t) \le M,\ \tilde{V}_{ijt}(z,\theta_t,\varepsilon_t) > 0 \Rightarrow\ \tilde{\chi}^{m}_{ijt}(z,\theta_t,\varepsilon_t,x^{t-1}) = 1.$$

Next, for the experimentation environment, consider the following modified version of the virtual index rule defined in the previous section. Let

$$\tilde{\gamma}_{ijt}\left(z,\theta_t,\varepsilon_t,x^{t-1}\right) = \max_{\tau} \mathbb{E}^{\lambda[\tilde{\chi}^{o}]|z,\theta_t,\varepsilon_t,x^{t-1}} \left\{ \frac{\sum_{s=t}^{\tau} \delta^{s-t} \tilde{V}_{ijs}\left(z,\theta_s,\varepsilon_s\right)}{\sum_{s=t}^{\tau} \delta^{s-t}} \right\}$$

denote the period-t weighted virtual index of the pair $(i,j) \in N_A \times N_B$, where τ continues to denote a stopping time, and where $\lambda[\tilde{\chi}^{o}]|z,\theta_t,\varepsilon_t,x^{t-1}$ denotes the distribution over $\Theta \times \mathcal{E}^{\ge t}$ that one obtains starting from $(z,\theta_t,\varepsilon_t,x^{t-1})$ when all future reported types are expected to coincide with those reported in period t and when each agent is expected to be matched to all other agents from period t onwards.[16] As in the previous section, then denote by

$$\tilde{J}_t(z,\theta_t,\varepsilon_t,x^{t-1}) = \left\{ \begin{array}{l} (i,j) \in N_A \times N_B \text{ s.t. (i) } \tilde{\gamma}_{ijt}\left(z,\theta_t,\varepsilon_t,x^{t-1}\right) \ge 0 \text{ and} \\ \text{(ii) } \#\left\{(l,m) \in N_A \times N_B : \tilde{\gamma}_{lmt}\left(z,\theta_t,\varepsilon_t,x^{t-1}\right) > \tilde{\gamma}_{ijt}\left(z,\theta_t,\varepsilon_t,x^{t-1}\right)\right\} < M \end{array} \right\}.$$

the set of pairs with the highest nonnegative weighted virtual index.

Definition 5 (weighted virtual index rule) *Let χ^I denote the "weighted virtual index" rule that links at each period $t \ge 1$ the pairs with the highest nonnegative virtual index, subject to the platform's capacity constraint. That is,*

$$\tilde{\chi}^{I}_{ijt}(z,\theta_t,\varepsilon_t,x^{t-1}) = 1 \;\Rightarrow\; (i,j) \in \tilde{J}_t(z,\theta_t,\varepsilon_t,x^{t-1}).$$

Furthermore, if $\#\tilde{J}_t(z,\theta_t,\varepsilon_t,x^{t-1}) \le M$, then

$$\begin{aligned} (i,j) &\in \tilde{J}_t(z,\theta_t,\varepsilon_t,x^{t-1}),\ \#\tilde{J}_t(z,\theta_t,\varepsilon_t,x^{t-1}) \le M,\ \tilde{\gamma}_{ijt}\left(z,\theta_t,\varepsilon_t,x^{t-1}\right) > 0 \\ &\Rightarrow \tilde{\chi}^{I}_{ijt}(z,\theta_t,\varepsilon_t,x^{t-1}) = 1. \end{aligned}$$

[15] Note that it is only to maintain symmetry with with the weighted virtual index rule $\tilde{\chi}^I$ defined below that we include past allocations x^{t-1} in the arguments of the weighted myopic rule $\tilde{\chi}^m$. Clearly, $\tilde{\chi}^m$ does not condition on past allocations.

[16] Again, this rule need not be feasible when $M < n_A \cdot n_B$. As in the previous section, what matters for the definition of the indexes is only the probability distribution over the future virtual surplus of the pair of agents $(i,j) \in N_A \times N_B$ that one obtains when the two agents are matched in all subsequent periods.

Now fix the period-0 reported types $z \in \Theta$ and, for any $t \geq 1$, any $(\theta_t, \varepsilon_t, x^{t-1}) \in \Theta \times \mathcal{E}_t \times X^{t-1}$, and any rule $\tilde{\chi} \in \tilde{\mathcal{X}}$, let $\lambda[\tilde{\chi}]|z, \theta_t, \varepsilon_t, x^{t-1}$ denote the distribution over $\Theta \times \mathcal{E}^{\geq t}$ that one obtains starting from $(z, \theta_t, \varepsilon_t, x^{t-1})$ when all future reported types are expected to coincide with those reported in period t (i.e., $\theta_s = \theta_t$, all $s \geq t$) and when allocations in all periods $s \geq t$ are determined by the rule $\tilde{\chi}$.[17] Then let

$$\mathbb{E}^{\lambda[\tilde{\chi}]|z,\theta_t,\varepsilon_t,x^{t-1}} \left[\sum_{s=t}^{\infty} \delta^s \sum_{i \in N_A} \sum_{j \in N_B} \tilde{V}_{ijs}\left(z, \theta_s, \varepsilon_s\right) \tilde{\chi}_{ijs}(z, \theta_s, \varepsilon_s, x^{s-1}) \right] \tag{6}$$

the expected weighted virtual surplus from period t onwards with the weights $\phi_l^k(z_l^k)/z_l^k$ constructed from the reports z.

Suppose that match values evolve exogenously. It is easy to see that the weighted myopic matching rule $\tilde{\chi}^m$ maximizes the objective in (6) for all $M \in \mathbb{N}$. Similarly, the weighted virtual index rule $\tilde{\chi}^I$ maximizes the objective in (6) for $M \geq n_A \times n_B$ or $M = 1$. To see this, note that the objective in (6) differs from that in (5) only in the time-invariant weights assigned to the matches.

Next note that, given any rule $\tilde{\chi}$, when all agents report truthfully in all periods, including period zero (in which case $z_l^k = \theta_{lt}^k = \theta_l^k$, all $l \in N_k$, $\theta_l^k \in \Theta_l^k$, $k = A, B$, $t \geq 0$), then the ex-ante expectation of (6), where the expectation is now also over the period-0 reports, coincides with the ex-ante expectation of dynamic virtual surplus, as defined in (5). Lastly note that, when all agents report truthfully in all periods, the allocations implemented under the weighted myopic rule $\tilde{\chi}^m$ coincide with those implemented under the rule χ^m and, likewise, the allocations under the weighted virtual index rule $\tilde{\chi}^I$ coincide with those under χ^I.

Now suppose the environment satisfies the assumptions of the experimentation model described above. Following an approach similar to that in Bergemann and Valimaki (2010) and in Kakade et al. (2013), for any profile z of period-0 reported types, we will construct payments $\tilde{\psi}^I_{t>0}(z)$ such that, when faced with the mechanism $(\tilde{\chi}^I, \tilde{\psi}^I_{t>0}(z))$, reporting truthfully $(\theta_l^k, (\varepsilon_{lht})_{h \in N_{-k}})$ in each period $t \geq 1$, all $l \in N_k$, $k = A, B$, irrespective of past reports, constitutes a periodic ex-post equilibrium. Furthermore, remaining in the mechanism is also periodic ex-post optimal for all agents. With this result in hand, we will then complete the construction of the payment scheme $\tilde{\psi}^I$ by adding payments $\tilde{\psi}^I_0(z)$ for period $t = 0$ that, together with $\tilde{\psi}^I_{t>0}(z)$, give, on path, to each agent an expected payoff equal to the one in the previous section. We will then argue that, given the complete mechanism $(\tilde{\chi}^I, \tilde{\psi}^I)$, with $\tilde{\psi}^I = (\tilde{\psi}^I_{t=0}(z), \tilde{\psi}^I_{t>0}(z))$, all agents find it optimal to participate and report truthfully also in period zero. Because the mechanism $\tilde{\mathcal{M}} = (\tilde{\chi}^I, \tilde{\psi}^I)$ induces the same matches and gives each agent the same ex-ante expected payoff as the mechanism $\mathcal{M} = (\chi^I, \psi^I)$ in the benchmark environment with observable values, we will then conclude that $\tilde{\mathcal{M}}$ is optimal for the

[17] In order to distinguish between matching rules in this environment with unobservable values and the enlarged message space from matching rules in the previous environment with observable types, we denote matching rules in this environment by $\tilde{\chi}$, and the set of all such feasible matches by $\tilde{\mathcal{X}}$.

platform and yields the same expected revenue as the optimal mechanism in the original environment with observable shocks. Similar conclusions can be established for the environment with exogenous processes.

Construction of the payment scheme $\tilde{\psi}^I$. For any $t \geq 1$, any $(z, \theta_t, \varepsilon_t, x^{t-1})$, any rule $\tilde{\chi} \in \tilde{\mathcal{X}}$, let $\tilde{\lambda}[\tilde{\chi}] \mid z, \theta_t, \varepsilon_t, x^{t-1}$ denote the distribution over $\Theta \times \mathcal{E}^{\geq t} \times X^{\geq t}$ that obtains under the rule $\tilde{\chi}$ when all agents are expected to report truthfully in the continuation game that starts with period t, the period-0 reports were equal to z, the true types are the ones reported in period t, i.e., θ_t, the true period-t values are the one reported in period t, i.e., ε_t, and the history of past allocations is x^{t-1}.[18] Then let

$$\begin{aligned} & W_t(z, \theta_t, \varepsilon_t, x^{t-1}) \\ & \equiv \max_{\tilde{\chi} \in \tilde{\mathcal{X}}} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}] \mid z, \theta_t, \varepsilon_t, x^{t-1}} \left[\sum_{s=t}^{\infty} \delta^{s-t} \sum_{i \in N_A} \sum_{j \in N_B} \tilde{V}_{ijs}\left(z, \theta_s, \varepsilon_s\right) \tilde{\chi}_{ijs}(z, \theta_s, \varepsilon_s, x^{s-1}) \right], \end{aligned} \tag{7}$$

denote the maximal expected weighted virtual surplus in the continuation game that starts in period t with $(z, \theta_t, \varepsilon_t, x^{t-1})$.

Next, for any agent $l \in N_k$ from side $k = A, B$, any $t \geq 1$, let $\tilde{\mathcal{X}}^{-l,k}$ denote the set of feasible matching rules that never assign any partner to agent l. Then, let

$$\begin{aligned} & W_t^{-l,k}(z, \theta_t, \varepsilon_t, x^{t-1}) \\ & \equiv \max_{\tilde{\chi} \in \tilde{\mathcal{X}}^{-l,k}} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}] \mid z, \theta_t, \varepsilon_t, x^{t-1}} \left[\sum_{s=t}^{\infty} \delta^{s-t} \sum_{i \in N_A} \sum_{j \in N_B} \tilde{V}_{ijs}\left(z, \theta_s, \varepsilon_s\right) \tilde{\chi}_{ijs}(z, \theta_s, \varepsilon_s, x^{s-1}) \right] \end{aligned}$$

denote the maximal expected weighted virtual surplus in the continuation game that starts with $(z, \theta_t, \varepsilon_t, x^{t-1})$, in the absence of agent l, i.e., when agent $l \in N_k$ from side k is not included in any of the matches. Finally, let $\tilde{\chi}^{-l,k}$ be any allocation rule that maximizes the weighed virtual surplus in the absence of agent l.

As argued above, in the experimentation environment, when $M \geq n_A \times n_B$ or $M = 1$, $\tilde{\chi}^I$ maximizes (6) over the entire set of feasible matching rules $\tilde{\mathcal{X}}$. Similarly, $\tilde{\chi}^{I-l,k}$ maximizes (6) over the restricted set $\tilde{\mathcal{X}}^{-l,k}$.

Next, define the expected marginal contribution of agent $l \in N_k$ with respect to (6) from period t onwards by

$$R_{lt}^k(z, \theta_t, \varepsilon_t, x^{t-1}) = W_t(z, \theta_t, \varepsilon_t, x^{t-1}) - W_t^{-l,k}(z, \theta_t, \varepsilon_t, x^{t-1}).$$

[18] Note that $\tilde{\lambda}[\tilde{\chi}] \mid z, \theta_t, \varepsilon_t, x^{t-1}$ differs from $\lambda[\tilde{\chi}] \mid z, \theta_t, \varepsilon_t, x^{t-1}$ in that it is a distribution also over current and future matching allocations.

Finally, define the period-t *flow marginal contribution* of agent l from side k recursively by

$$r_{lt}^{k}(z,\theta_t,\varepsilon_t,x^{t-1})=R_{lt}^{k}(z,\theta_t,\varepsilon_t,x^{t-1})-\delta\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_t,\varepsilon_t,x^t}\left[R_{lt+1}^{k}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right],$$

Note that r_{lt}^{k} can be rewritten as

$$\begin{aligned}
&r_{lt}^{k}(z,\theta_t,\varepsilon_t,x^{t-1})=\\
&=\sum_{i\in N_A}\sum_{j\in N_B}\tilde{V}_{ijt}\left(z,\theta_t,\varepsilon_t\right)\cdot\tilde{\chi}_{ijt}^{I}(z,\theta_t,\varepsilon_t,x^{t-1})-\sum_{i\in N_A}\sum_{j\in N_B}\tilde{V}_{ijt}\left(z,\theta_t,\varepsilon_t\right)\tilde{\chi}_{ijt}^{I-l,k}(z,\theta_t,\varepsilon_t,x^{-1})\\
&+\delta\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_t,\varepsilon_t,x^{t-1}}\left[W_{t+1}^{-l,k}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]-\delta\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^{I-l,k}]|z,\theta_t,\varepsilon_t,x^{t-1}}\left[W_{t+1}^{-l,k}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right].
\end{aligned}\tag{8}$$

Given any profile $z\in\Theta$ of period-0 reports, we now construct a sequence of payments $\tilde{\psi}_{t>0}^{I}(z)$ that makes each agent's flow net utility in each period $t>0$ coincide with her flow marginal contribution to the weighted virtual surplus. In particular, for each period $t\geq 1$, the payment for each agent $i\in N_A$ from side A is given by

$$\tilde{\psi}_{it}^{A,I}(z,\theta_t,\varepsilon_t,x^{t-1})=\sum_{j\in N_B}\theta_{it}^{A}\varepsilon_{ijt}^{A}\tilde{\chi}_{ijt}^{I}(z,\theta_t,\varepsilon_t,x^{t-1})-\frac{z_i^A\cdot r_{it}^{A}(z,\theta_t,\varepsilon_t,x^{t-1})}{\phi_i^A(z_i^A)}\tag{9}$$

and similarly for each agent $j\in N_B$ from side B. These payments reflect the externality the agent imposes in the current and future periods on others agents (from both sides of the market). This externality is calculated with respect to the agents' weighted utility. Note that the externality may be positive or negative and therefore the payments may be positive or negative. For example, if an agent is valued highly by the agents she is matched to, her payment may be negative (i.e., the agent may receive a positive transfer from the platform — *cross subsidization*).

Finally, for period $t=0$, the payments $\tilde{\psi}_{l0}^{k,I}(z)$ are constructed so that, in expectation, if all agents report truthfully in all periods, including period $t=0$, the expected net present value of each agent's payments is the same as under the virtual index mechanism for the environment with observable match values. That is, given $\tilde{\psi}_{t>0}^{I}(z)$, the payments $\tilde{\psi}_{l0}^{k,I}(z)$ for each agent $l\in N_k$ from side $k=A,B$ are such that

$$\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z_l}\left[\tilde{\psi}_{l0}^{k,I}(z)+\sum_{t=1}^{\infty}\delta^t\tilde{\psi}_{lt}^{k,I}(z,\theta_t,\varepsilon_t,x^{t-1})\right]=\mathbb{E}^{\lambda[\chi^I]|z_l}\left[\sum_{t=0}^{\infty}\delta^t\psi_{lt}^{k,I}(z,\varepsilon^t)\right],\tag{10}$$

where all expectations are taken under the assumption that all agents are truthful at all periods, including $t=0$. Note that the right hand side of (10) is given by the envelope condition of Lemma 1, computed for the virtual index rule χ^I.

As mentioned above, the payments $\tilde{\psi}^m$ for the case of exogenous match values can be constructed in precisely the same manner as those for the rule χ^I.

By construction, the mechanisms $(\tilde{\chi}^I,\tilde{\psi}^I)$ and $(\tilde{\chi}^m,\tilde{\psi}^m)$ induce the same matching dynamics and yield the same profits as their counterparts (χ^I,ψ^I) and (χ^m,ψ^m) in the benchmark with observable match values. Because the platform is more constrained in the environment under examination than in the benchmark with observable match values, we then have the following result:

Theorem 2 (private match values) *Consider the environment in which match values are the agents' private information.*

(1) Suppose match values evolve exogenously and that $D_l^k(\underline{\theta}_l^k;\chi^m) \geq 0$, *all* $l \in N_k$, $k = A, B$. *Then the matching mechanism* $(\tilde{\chi}^m, \tilde{\psi}^m)$ *is profit-maximizing.*

(2) Suppose the environment satisfies the assumptions of the experimentation model defined above, that $D_l^k(\underline{\theta}_l^k;\chi^I) \geq 0$, *all* $l \in N_k$, $k = A, B$, *and that either* $M \geq n_A \cdot n_B$ *or* $M = 1$. *Then the matching mechanism* $(\tilde{\chi}^I, \tilde{\psi}^I)$ *is profit-maximizing.*

(3) Suppose the environment satisfies the conditions in part (1). Given $(\tilde{\chi}^m, \tilde{\psi}^m)$, *in the continuation game that starts with any period* $t \geq 1$ *and any arbitrary history, participating and reporting truthfully in the current and all subsequent periods is a periodic ex-post equilibrium. The same conclusions hold for the mechanism* $(\tilde{\chi}^I, \tilde{\psi}^I)$ *under the conditions of part (2).*[19]

In the Appendix, we establish the results by first showing that, at any period $t \geq 1$, irrespective of the history, an agent who expects all other agents to report truthfully at present and in future periods finds it optimal to remain in the mechanism and also report truthfully at any period $s \geq t$, irrespective of the agent's beliefs about the other agents' types and values. We then complete the proof by showing that participating and reporting truthfully at period $t = 0$ is also optimal for the agent when he expects all other agents to report truthfully at all periods.

5 Sequential Scoring Auctions

We now discuss how the matching mechanisms described in the previous section can be implemented via a sequence of scoring auctions in which agents repeatedly bid to be matched with potential partners.

The structure of the sequential auctions is the following. In period $t = 0$, agents select a *membership level.* A higher level grants higher status, which translates into more favorable treatment, on average, in each of the subsequent auctions. Accordingly, high membership status comes with a higher membership fee.

In each period $t \geq 1$, agents first renew their membership by choosing their status, and then submit a bid for each of the agents on the opposite side. Each match is then assigned a "score" based on current and past bids, as well as membership status. All pairs receiving the highest non-negative scores are then matched, up to capacity. The computation of the scores and the

[19] At no cost to the platform, the payments $\tilde{\psi}^m$ and $\tilde{\psi}^I$ can also be adjusted to guarantee that participating and reporting truthfully in all periods is a periodic ex-post equilibrium not only from $t = 1$ but also from period $t = 0$.

corresponding payments depend on the environment under examination. To facilitate the comparison to the previous section, we denote the period-0 membership status by z_l^k, and the period-t status by θ_{lt}^k, $l \in N_k$, $t \geq 1$.

Exogenous match values. (0) In period $t = 0$, each agent $l \in N_k$ from each side k is asked to select a membership level. The selection of the period-0 level z_l^k comes with a membership fee equal to $\psi_{l0}^k(z_l^k)$, with the payment functions ψ_{l0}^k defined as in (10).

At any subsequent period $t \geq 1$, the rules of the period-t scoring auction are the following.

(1) Each agent $l \in N_k$ is asked to re-select a period-t membership level, θ_{lt}^k, and to submit a bid for each potential partner on the opposite side, $b_{lt}^k \equiv (b_{ljt}^k)_{j=1}^{n_{-k}}$. Note that agents' bids reflect their utility from interacting with each of the partners, u_{lt}^k, rather then their match values, ε_{lt}^k.

(2) Each pair of agents $(i, j) \in N_A \times N_B$ is then assigned a score $B_{ijt} \equiv \lambda_i^A(z_i^A) b_{ijt}^A + \lambda_j^B(z_j^B) b_{ijt}^B - c_{ijt}$, where $\lambda_l^k(z_l^k) \equiv \phi_l^k(z_l^k)/z_l^k$, and the functions $\phi_l^k(z_l^k)$ are defined as in the previous sections. The pairs of agents $(i, j) \in N_A \times N_B$ with the highest nonnegative score are then matched, subject to the platform's capacity constraint. We denote the resulting matching rule by χ.

(3) Payments are as follows. If i is not matched to any agent $j \in N_B$ from the opposite side in period t, he does not make any payments in this period. If, instead, i is matched to some agents from side B, his payment depends on the total number of matches he secures. Specifically, if i is matched to $K \geq 1$ agents from side B, he pays

$$(1/\lambda_i^A(z_i^A)) \left(\mathbf{B}_{it}^A(K) + \sum_{j \in N_B} (c_{ijt} - \lambda_j^B(z_j^B) b_{ijt}^B) \chi_{ijt} \right),$$

where $\mathbf{B}_{it}^A(K)$ is defined as the sum of the K highest non-negative scores among the pairs that are unmatched in period t and that do not include i. A similar structure applies to each side-B agent.

Note that an agent's payments may be negative, reflecting the possibility of cross-subsidization. Also note that, in this environment, the period-t scores and payments are not forward-looking, and depend only on the period-0 membership choices and the period-t bids. Finally, note that, contrary to standard Vickrey-type auctions, multiple agents may be charged for the same externality they impose on other agents.

Experimentation model ($M \geq n_A \cdot n_B$). Contrary to the case of exogenous match values, with experimentation, the matching rule is forward-looking. As a result, in each period $t > 0$, the rules of the auction in parts (2) and (3) above are amended as follows.

(2) For each pair $(i, j) \in N_A \times N_B$, the function B_{ijt} is computed as in the case of exogenous match values. Each pair (i, j) 's score is now defined as the Gittins index computed using B_{ijt} as in Section 4.[20] All pairs with nonnegative score are then matched.

[20] Note that calculating the indexes also requires knowledge of the agents' match-specific values ε_{ijt}^A, as well as the number of times i and j were matched in the past. The former can be recovered by dividing each bid b_{ijt}^A by the period-t membership choice θ_{it}^A. The same procedure permits to recover the side-B's values.

(3) Any agent $i \in N_A$ who is matched to any agent $j \in N_B$ from side B is asked to pay $(1/\lambda_i^A(z_i^A)) \left(c_{ijt} - \lambda_j^B(z_j^B)b_{ijt}^B\right)$ for this match. Overall, each agent i's period-t payment is thus equal to

$$(1/\lambda_i^A(z_i^A)) \left(\sum_{j \in N_B} (c_{ijt} - \lambda_j^B(z_j^B)b_{ijt}^B)\chi_{ijt} \right).$$

Note that, in this environment, the period-t scores depend on the period-0 and period-t membership choices, and the period-t bids (as well as information about the number of previous interactions). The period-t payments, however, depend only on the period-0 membership choices and period-t bids. More importantly, note that while the period-t scores are forward-looking in this environment, the period-t payments are not.

Experimentation model ($M = 1$). In this case, both the matching and the payment rule are forward looking. To reflect this, stages (2) and (3) of the period-t scoring auction are now amended as follows.

(2) The score of each pair $(i, j) \in N_A \times N_B$ is defined as their Gittins index, computed exactly as in the experimentation model with $M \geq n_A \cdot n_B$. The pair of agents $(i, j) \in N_A \times N_B$ with the highest nonnegative score is then matched.

(3) If agent $i \in N_A$ is matched to an agent $j \in N_B$ from side B, i is asked to pay an amount

$$(1/\lambda_i^A(z_i^A)) \left(c_{ijt} - \lambda_j^B(z_j^B)b_{ijt}^B + (1-\delta)W_t^{-i,A}\right),$$

where $W_t^{-i,A}$ is as defined in the previous section. If agent i is not matched, he does not make any payment. The payments of side-B agents are defined analogously.

Note that the period-t payments and scores in this environment depend on the period-0 and period-t membership choices and the period-t bids (and the number of previous interactions). Furthermore, in contrast to the case in which $M \geq n_A \cdot n_B$, payments are now also forward-looking, reflecting the idea that an agent who is matched to another agent in period t must compensate all other agents (including those with whom he is matched) for the dynamic externality generated by his period-t match.[21]

We then have the following result.

Proposition 1 (sequence of scoring auction) *The sequence of scoring auctions described above admits an equilibrium in which all agents make truthful membership choices (i.e., $z_{lt}^k = \theta_{lt}^k = \theta_l^k$, $l \in N_k$, $t \geq 1$) and bid truthfully (i.e., $b_{lt}^k = u_{lt}^k$, $l \in N_k$, $t \geq 1$) in all periods. Such an equilibrium*

[21] One might be tempted to consider an auction similar to the one for exogenous match vales, in which payments are computed by simply replacing each static score with its index. However, such payments are larger than the ones above, and consequently agents may have an incentive to shade their bids. Suppose, for example, that agent $i \in N_A$ is matched to some other agent in period t and that there is only one alternative match in which agent i is not part of, whose index is $\gamma > 0$. Then $\gamma/(1-\delta) \geq W_t^{-i,A}$.

implements the same match allocations and yields the same expected revenue as the optimal direct mechanism in the previous section.[22]

To appreciate the role of the period-0 membership status in determining how focal an agent is in the market, consider for example the case of exogenous match values. Note that higher membership status z_i^A implies a higher weight $\lambda_i^A(z_i^A)$ is placed on agent i's bids (due to the regularity assumption). If i favors agent $j \in N_B$ in period $t \geq 1$ ($\varepsilon_{ijt}^A > 0$), the period-t score corresponding to the match (i, j) is increasing in z_i^A, while if i does not favor agent j ($\varepsilon_{ijt}^A < 0$), this score is decreasing in z_i^A. A high z_i^A therefore favorably affects i's allocation of interactions, both in terms of his competition with other agents for matches, and within each of his matches, shifting the length of the interaction to his benefit.

6 Welfare- vs profit-maximizing matching mechanisms

We now turn to the distortions created by the profit-maximizing provision of matching services relative to its welfare-maximizing counterpart.

As in static mechanism design problems, the platform introduces distortions to the matching allocation in order to reduce the agents' expected information rents. It follows from the results above that the welfare-maximizing matching rule has the same structure as the profit-maximizing one, but with the true types, θ_l^k, replacing the virtual ones, $\phi_l^k(\theta_l^k)$. As a result, the profit maximizing matching rule is distorted from the welfare maximizing one not only in the initial period, but in all subsequent periods as well.

We then have the following result:[23]

Theorem 3 (distortions – non-negative values) *Suppose that, for all $(i, j) \in N_A \times N_B$, $k = A, B$, and $t \geq 1$, $\mathcal{E}_{ijt}^k \subseteq \mathbb{R}_+$. The relationship between the profit-maximizing matching rule χ^P and its welfare-maximizing counterpart χ^W in each of the environments considered in Theorem 1 above satisfies the following properties:*[24]

(1) Suppose $M \geq n_A \cdot n_B$. For any $t \geq 1$, $(\theta, \varepsilon^t) \in \Theta \times \mathcal{E}^t$, $(i, j) \in N_A \times N_B$, $\chi_{ijt}^P(\theta, \varepsilon^t) = 1$ implies $\chi_{ijt}^W(\theta, \varepsilon^t) = 1$.

[22] When match values are exogenous, bidding truthfully is in fact a weakly dominant strategy at all periods $t \geq 1$ (at $t = 0$, the mechanism is IC, but bidding truthfully is not necessarily a weakly dominant strategy).

[23] For this result, we assume that, in each period, the profit maximizing rule uses all M matching slots, unless there are strictly fewer than M pairs for which virtual surplus (in case of exogenous processes) or virtual index (in the experimentation model) is nonnegative. The same qualification applies to the welfare-maximizing rule (replacing virtual surplus with true surplus and virtual index with true index).

[24] Note that the same results apply to the comparison between the rules $\tilde{\chi}^m$ and $\tilde{\chi}^I$ for the environment in which match values are private information and the efficient rule χ^W. In fact, as shown in the previous section, $\tilde{\chi}^m$ and $\tilde{\chi}^I$ induce the same on-path allocations as χ^m and χ^I in the benchmark with observable match values.

(2) For any $t \geq 1$, $(\theta, \varepsilon^t) \in \Theta \times \mathcal{E}^t$, $\sum_{(i,j)\in N_A \times N_B} \chi^W_{ijt}(\theta, \varepsilon^t) \geq \sum_{(i,j)\in N_A \times N_B} \chi^P_{ijt}(\theta, \varepsilon^t)$.

When match values are nonnegative, in any of the environments considered above, the profit-maximizing matching rule thus exhibits a form of downward distortions - the number of matches in each period is inefficiently low relative to what is efficient. Furthermore, when the platform does not face a binding capacity constraint, in each period $t \geq 1$, the set of matches created under profit maximization is a subset of those under welfare maximization. Furthermore, in the experimentation model, when $M \geq n_A \cdot n_B$, once a link between two agents is severed, it is never reactivated. Therefore, matches are gradually broken over time both under profit and welfare maximization.

When, instead, $M = 1$, even if for each pair $(i, j) \in N_A \times N_B$, and any $(\theta, \varepsilon_t, x^{t-1})$ the virtual index is lower than its counterpart under profit maximization (which coincides with the standard Gittins index), the ranking of the virtual indexes under profit maximization need not coincide with the ranking under welfare maximization. As a result, in the experimentation model, certain partners may interact for a longer period of time under profit maximization than under welfare maximization. What remains true is that, if at a given point in time a welfare-maximizing platform finds it optimal to shut down by leaving all agents unmatched, then a profit-maximizing platform that has not shut matching down yet will also do it in the same period.

A natural question to ask is whether the results in Theorem 3 extend to the case of $1 < M < n_A \cdot n_B$ for the experimentation model.[25] That is, if a planner (platform) were to follow the (virtual) index rule in this case, would part (2) of Theorem 3 still hold? As the following example shows, the answer to this question is no. That is, welfare-maximization does not necessarily involve more activity than profit-maximization in each period. In fact, this is not even guaranteed intertemporally.

Example 4 (experimentation model with $1 < M < n_A \cdot n_B$) Suppose $N_A = \{1, 2, 3\}$, $N_B = \{1\}$ and $M = 2$. Consider the following deterministic (given θ) processes of the surpluses and virtual surpluses generated by each match, as a function of the number of previous interactions k:[26]

[25] Theorem 3 addresses the case of $1 < M < n_A \cdot n_B$ for exogenous values.

[26] Suppose that $\Theta^A_2 = \Theta^A_3 = \Theta^B_1 = \{1\}$ and $\theta^A_1 \sim U[2, 3]$, and consider a realization $\theta^A_1 = 2$. The above processes can then be generated by setting $c_{11t} = c_{31t} = 1$, $c_{21t} = 0$, $\varepsilon^B_{i1t} = 0$, all $i = 1, 2, 3$ all $t \geq 1$, and match values for side-A given by the table below (with k denoting the number of previous interactions with agent 1 from side B):

	$k = 0$	1	2	$k \geq 3$
$\varepsilon^A_{11}(k)$	4	0	0	0
$\varepsilon^A_{21}(k)$	4	2	2	2
$\varepsilon^A_{31}(k)$	6	2	0	0

Surplus	$k=0$	1	2	$k\geq 3$	Virtual surplus	$k=0$	1	2	$k\geq 3$
(1,1)	7	-1	-1	-1	(1,1)	3	-1	-1	-1
(2,1)	4	2	2	2	(2,1)	4	2	2	2
(3,1)	5	1	-1	-1	(3,1)	5	1	-1	-1

As in Example 2, since matches are deteriorating, the (virtual) indexes coincide with the (virtual) surpluses, and hence the (virtual) index rule simply matches the two pairs of agents generating the highest nonnegative (virtual) surplus.[27] That is, a planner matches in period 1 the pairs $\{(1,1),(3,1)\}$, in period 2 the pairs $\{(2,1),(3,1)\}$, and $\{(2,1)\}$ thereafter. Similarly, a profit-maximizing platform matches in period 1 the pairs $\{(2,1),(3,1)\}$, in period 2 the pairs $\{(1,1),(2,1)\}$, in period 3 the pairs $\{(2,1),(3,1)\}$ again, and $\{(2,1)\}$ thereafter.

Thus, at $t=3$, the number of interactions under profit-maximization is strictly greater than that under welfare-maximization. In fact, for all $T\geq 1$, the total number of interactions up to period T is always greater under profit-maximization.[28] □

Interestingly, when values are allowed to be negative, the conclusions in Theorem 3 need not hold. The platform may find it optimal to distort in the opposite direction. This is because, when values are negative, a pair's virtual surplus may be larger than its true surplus. As a result, there may exist periods, or even entire continuations, in which more matches are created under profit-maximization than under welfare-maximization. The following example illustrates.

Example 5 (negative values and upward distortions) Let $N_A = N_B = \{1\}$, $M = 1$, and $c_{11t} = 0$ all $t\geq 1$. Assume that $\theta_1^A \sim U[1,2]$, $\Theta_1^B = \{1\}$, and that at each period $t\geq 1$, regardless of previous matches or realizations, $\varepsilon_{i11}^B = 1$ and ε_{11t}^A is equal to either 2 or -2, with equal probability. Consider the realizations of true types $\theta_1^A = \theta_1^B = 1$, for which $\psi_1^A(\theta_1^A) = 0$ and $\psi_1^B(\theta_j) = 1$, and the sequence of realizations $\varepsilon_{11t}^A = -2$ for all $t\geq 1$. Then, for all $t\geq 1$, $\chi_{11t}^W(\theta,\varepsilon^t) = 0$ but $\chi_{11t}^P(\theta,\varepsilon^t) = 1$. □

The above results are consistent with what is discussed in Pavan, Segal and Toikka (2014). That paper shows that, when the agents' private information in each period is unidimensional, the key force responsible for the dynamics of distortions under profit maximization is the impulse response of an agent's future types to her initial type. In the model under consideration here, agents' private information in each period is multidimensional. The period-t impulse response of the value u_{ijt}^A that agent $i\in N_A$ from side A assigns to agent $j\in N_B$ from side B to the agent's initial type

[27] For a sufficiently low δ, these matching rules will also be optimal. See Section 7 for a discussion on the optimality of the index rule for arbitrary capacity constraints.

[28] What remains true is that, if under profit-maximization matching stops after a finite number of periods, then the number of interactions overall (i.e., summing over all periods $t\geq 1$) is weakly greater under welfare-maximization.

θ_i^A is here the match value ε_{ijt}^A. Interestingly, notwithstanding the complexity brought in by the multidimensionality of the agents' flow private information, the key insight in Pavan, Segal and Toikka (2014) that the dynamics of distortions are driven by the dynamics of the impulse responses extends to the matching environment under consideration here.

Finally, note that a familiar result in the mechanism design literature is the absence of distortions "at the top of the distribution." In our environment, this result does not hold. In fact, distortions may affect even those agents for whom the virtual values coincide with the true ones (i.e., $\phi_l^k(\theta_l^k) = \theta_l^k$). The reason is similar to the one discussed in Gomes and Pavan (2015). The allocations of these types depend not only on their own virtual value, but also on the virtual values of their potential partners.

7 Arbitrary capacity constraints

Consider again the benchmark environment with observable match values. As discussed in Section 3, under the assumptions of the experimentation model, when $1 < M < n_A \cdot n_B$, in general there is no guarantee that, without further assumptions, χ^I maximizes dynamic virtual surplus. This is because, in general, multiarmed bandit problems in which multiple arms can be activated simultaneously fail to admit a simple index solution.

In this section, we introduce a condition for the experimentation environment for which χ^I is optimal for all $M \in \mathbb{N}$. This environment requires discreteness of agents' types and values. While the framework above assumed that types were continuous[29], it can easily be modified in the following manner to allow for discrete types (again, for simplicity, we describe the adjustment on side A, with the understanding that the same adjustment applies to side B).

Suppose that each agent i's type from side A is drawn from a commonly known distribution F_i^A over a finite set $\Theta_i^A \equiv \{\theta_{i1}^A, ..., \theta_L^A\}$, with $\theta_{il}^A \geq 0$ for all $l \in \{1, ..., L\}$ and $\theta_{il}^A < \theta_{il+1}^A$. Denote $f_{il}^A \equiv \Pr(\tilde{\theta}_i^A = \theta_{il}^A)$ and $F_{il}^A \equiv \Pr(\tilde{\theta}_i^A \leq \theta_{il}^A) = \sum_{l=1}^{l} f_{il}^A$. For each $\theta_{il}^A \in \Theta_i^A$, $l \in \{1, ..., L-1\}$, denote by $\eta_i^A(\theta_{il}^A) \equiv (1 - F_{il}^A)/f_{il}^A$ the inverse hazard rate of the distribution of agent i's type. In this discrete environment, the definition of virtual types must be adjusted so that

$$\phi_i^A(\theta_{il}^A) \equiv \theta_{il}^A - \eta_i^A(\theta_{il}^A) \cdot \left(\theta_{il+1}^A - \theta_{il}^A\right),$$

and regularity amounts to requiring that the functions $\eta_i^A(\theta_{il}^A) \cdot \left(\theta_{il+1}^A - \theta_{il}^A\right)$ be strictly decreasing. Results analogous to those in Lemma 1 and Lemma 2 then apply to the above discrete-type environment.[30]

The approach below adapts arguments in Pandelis and Teneketzis (1999) to our dynamic matching environment. For each pair of agents $(i, j) \in N_A \times N_B$, consider the following *auxiliary process* for

[29] Note that no such restriction was imposed on the match values.

[30] Certain adjustments are required for Lemma 1 and Lemma 2 when moving to a discrete type space. These adjustments are similar to those that apply in the static case.

the virtual surplus generated from matching this pair of agents, defined pathwise using the virtual indexes according to[31]

$$\underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) \equiv \inf_{s \leq t} \left\{ \gamma_{ijs}(\theta, \varepsilon_s, x^{s-1}) \right\}$$

where the virtual indexes γ_{ijt} continue to be defined as in the previous sections. Note that each stochastic process $\underline{V}_{ijt}$ is nonincreasing in the sense that, with probability one, $\underline{V}_{ijt} \leq \underline{V}_{ijt-1}$.

Next, suppose that the environment satisfies the following separability condition.

Definition 6 (separability) *The environment is separable if, for any $t \geq 1$, any two pairs of agents, $(i,j),(i',j') \in N_A \times N_B$, any $(\theta, \varepsilon^t, x^{t-1}) \in \Theta \times \mathcal{E}^t \times X^{t-1}$, if $\underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) > \underline{V}_{i'j't}(\theta, \varepsilon^t, x^{t-1})$ then $\underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) \cdot (1-\delta) \geq \underline{V}_{i'j't}(\theta, \varepsilon^t, x^{t-1})$.*

We then have the following result:

Proposition 2 (separable environment) *Suppose the environment is separable and that match values are observable. For any $M \in \mathbb{N}$, there exists a payment rule ψ^I such that the matching mechanism (χ^I, ψ^I) in which the matching rule is a virtual index rule is profit maximizing.*

As we show in the Appendix, the role of the separability condition is to guarantee that, in a fictitious environment in which the reward processes are auxiliary ones, the optimal rule is myopic. The proof then proceeds by observing that, given any rule χ, dynamic virtual surplus when the rewards are given by the auxiliary processes (that is, by $\underline{V}_{ijt}$) is greater than, or equal to, dynamic virtual surplus, under the same rule, when the rewards are the ones in the primitive model (that is, they are given by V_{ijt}). Furthermore, in the special case in which the rule is the virtual index one, χ^I, as defined in the previous sections, dynamic virtual surplus in the two environments is the same. The optimality of a virtual index rule for the primitive environment then follows from the above properties along with the fact that the myopic rule for the fictitious environment implements the same allocations as the virtual index rule for the primitive environment.

The following example, in which virtual surpluses follow deterministic decreasing processes, illustrates how the virtual index rule may be suboptimal when separability is not satisfied.

Example 6 (separability) Suppose $N_A = \{1, 2, 3\}$, $N_B = \{1\}$ and $M = 2$. Consider the following processes of the virtual surpluses generated by each of the potential matches, as a function of the

[31] Such auxiliary process is also known as the *lower envelope process* (see Mandelbaum (1986)).

number of previous interactions k:[32]

$k =$	0	1	2	≥ 3
$V_{11}(k)$	3	3	3	-1
$V_{21}(k)$	4	2	2	-1
$V_{31}(k)$	5	1	1	-1

The virtual index rule (which is in this case myopic) first matches $\{(2,1),(3,1)\}$, followed by $\{(1,1),(2,1)\}$, $\{(1,1),(2,1)\}$, $\{(1,1),(3,1)\}$ and $\{(3,1)\}$. Consider an alternative matching rule, which matches pairs $\{(1,1),(3,1)\}$, followed by $\{(1,1),(2,1)\}$, $\{(1,1),(2,1)\}$, $\{(2,1),(3,1)\}$ and $\{(3,1)\}$. The difference in the corresponding profits is then $1 - 2\delta + \delta^3$, which for $\delta > \frac{\sqrt{5}}{2} - \frac{1}{2}$ is negative. Thus, for all $\delta > \frac{\sqrt{5}}{2} - \frac{1}{2}$, the virtual index rule is suboptimal. Clearly, separability is violated for $\delta > \frac{\sqrt{5}}{2} - \frac{1}{2}$. For sufficiently low δ, separability is satisfied, and the virtual index rule becomes optimal. □

We conclude this section by noting that while, to our knowledge, no optimal solutions are known for the general multiarmed bandit problem in which multiple arms can be activated simultaneously, there are several results pertaining to more restricted environments, as well as asymptotic results, of potential interest to the matching problem studied in the present paper. For example, Bergemann and Valimaki (2001) show that an index rule is optimal for the case of stationary multiarmed bandits, in which there are countable infinitely many ex-ante identical arms (and approximately optimal in the limit as the number of arms goes to infinity).

For the (more general) case of restless bandits, i.e., arms that evolve and yield rewards even when they are not activated, under the infinite horizon average reward criterion, Weber and Weiss (1990) provide conditions guaranteeing that an index policy that always pulls the M arms with highest Whittle index is asymptotically optimal (see also Whittle (1988)). In the special case in which passive arms are static and yield no reward, this index reduces to the Gittins index.

The above asymptotic results can be applied to the analysis of certain matching markets in which the number of agents on each side is large.

8 Conclusions

This paper examines the dynamics of matching allocations in a mediated, many-to-many, two-sided market in which agents' preferences for potential partners evolve over time, either exogenously, or as a function of previous interactions.

[32] As in Example 4, one can easily construct an environment consistent with the assumptions of the experimentation model that generates these processes.

It first characterizes matching dynamics under profit-maximizing contracts when match values are observable by the platform (equivalently, when each couple of agents jointly observes the evolution of the payoffs of each of the two agents in the couple). It considers both the case in which the platform may match any number of pairs, as well as the case in which it has limited matching capacity. It then extends the characterization to an environment in which agents privately observe variations in their valuations.

The properties of the profit-maximizing matching mechanism highlight the advantages of long-term contractual relationships in the provision of matching services over time. The results are then used to shed light on the inefficiencies associated with the private provision of matching services.

The framework is flexible enough to admit as special cases such environments in which learning occurs immediately upon matching (that is, in which agents learn their valuation for a partner immediately after the first interaction), as well as one-time interactions, in which agents enjoy interacting at most once with each partner.

Many extensions seem interesting. For example, while certain results can be adapted to accommodate for the possibility that agents' match values depend on the entire history of past interactions (e.g., agents may care about their partners' previous partners), extending the analysis to allow for more general forms of correlation in agents' preferences is challenging but worth exploring. In future work, we also intend to extend the analysis to markets such as those for online sponsored search, in which platforms may be unable to collect payments from one side of the market.

We conclude by noting that while matching dynamics in our model reflect changes in agents' preferences for potential partners, another line of recent research explores matching dynamics generated by the stochastic arrival or departure of agents into and from the market (see, for example, Akbarpour et al. 2014, Anderson et al. 2014, and Baccara et al. 2015). Combining the two lines of research is expected to generate further insights pertaining to the dynamics of matching allocations in various markets of interest.

Appendix

Proof of Lemma 1. *Part (1).* We want to show that if $\mathcal{M} = (\chi, \psi)$ is incentive compatible, then χ satisfies average monotonicity (i.e., $D_l^k(\theta_l^k; \chi)$ is nondecreasing in θ_l^k, all $l \in N_k$, $k = A, B$). Take any pair of types $\theta_i^A, \tilde{\theta}_i^A \in \Theta_i^A$ of agent $i \in N_A$ from side A (the arguments for all agents from side

B is analogous). Incentive compatibility requires that

$$U_i^A(\theta_i^A) = \mathbb{E}^{\lambda[\chi]|\theta_i^A}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j\in N_B}\theta_i^A\varepsilon_{ijt}^A\chi_{ijt}(\theta_i^A,\theta_{-i}^A,\varepsilon^t) - \sum_{t=0}^{\infty}\delta^t\psi_{it}^A(\theta_i^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$\geq \mathbb{E}^{\lambda[\chi]|\tilde{\theta}_i^A}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j\in N_B}\theta_i^A\varepsilon_{ijt}^A\chi_{ijt}(\tilde{\theta}_i^A,\theta_{-i}^A,\varepsilon^t) - \sum_{t=0}^{\infty}\delta^t\psi_{it}^A(\tilde{\theta}_i^A,\theta_{-i}^A,\varepsilon^t)\right]$$

and

$$U_i^A(\tilde{\theta}_i^A) = \mathbb{E}^{\lambda[\chi]|\tilde{\theta}_i^A}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j\in N_B}\tilde{\theta}_i^A\varepsilon_{ijt}^A\chi_{ijt}(\tilde{\theta}_i^A,\theta_{-i}^A,\varepsilon^t) - \sum_{t=0}^{\infty}\delta^t\psi_{it}^A(\tilde{\theta}_i^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$\geq \mathbb{E}^{\lambda[\chi]|\theta_i^A}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j\in N_B}\tilde{\theta}_i^A\varepsilon_{ijt}^A\chi_{ijt}(\theta_i^A,\theta_{-i}^A,\varepsilon^t) - \sum_{t=0}^{\infty}\delta^t\psi_{it}^A(\theta_i^A,\theta_{-i}^A,\varepsilon^t)\right].$$

Combining the two inequalities, we have that

$$(\theta_i^A - \tilde{\theta}_i^A)D_i^A(\tilde{\theta}_i^A;\chi) \leq U_i^A(\theta_i^A) - U_i^A(\tilde{\theta}_i^A) \leq (\theta_i^A - \tilde{\theta}_i^A)D_i^A(\theta_i^A;\chi),$$

which can be satisfied only if $D_i^A(\theta_i^A;\chi)$ is nondecreasing in θ_i^A.

Next note that, for any $\hat{\theta}_i^A \in \Theta_i^A$, $U_i^A(\theta_i^A;\hat{\theta}_i^A)$ is differentiable and Lipschitz continuous in θ_i^A with derivative

$$\frac{\partial U_i^A(\theta_i^A;\hat{\theta}_i^A)}{\partial\theta_i^A} = D_i^A(\tilde{\theta}_i^A;\chi).$$

Standard envelope theorems (e.g., Milgrom and Segal, 2002), then imply that, if $\mathcal{M}=(\chi,\psi)$ is incentive compatible, then $U_i^A(\theta_i^A)$ is Lipschitz continuous with derivative a.e. equal to

$$\frac{dU_i^A(\theta_i^A)}{d\theta_i^A} = D_i^A(\theta_i^A;\chi).$$

This in turn implies that

$$U_i^A(\theta_i^A) = U_i^A(\underline{\theta}_i^A) + \int_{\underline{\theta}_i^A}^{\theta_i^A} D_i^A(y;\chi)dy,$$

Equivalently, the payment rule ψ must satisfy

$$\mathbb{E}^{\lambda[\chi]|\theta_i^A}\left[\sum_{t=0}^{\infty}\delta^t\psi_{it}^A(\theta,\varepsilon^t)\right] = \theta_i^A D_i^A(\theta_i^A;\chi) - U_i^A(\underline{\theta}_i^A) - \int_{\underline{\theta}_i^A}^{\theta_i^A} D_i^A(y;\chi)dy.$$

Part (2). Let χ be a feasible matching rule satisfying average monotonicity and such that $D_l^k(\underline{\theta}_l^k;\chi) \geq 0$ for all $l \in N_k$, $k = A, B$. Then take any payment rule ψ satisfying

$$\mathbb{E}^{\lambda[\chi]|\theta_l^k}\left[\sum_{t=0}^{\infty}\delta^t\psi_{lt}^k(\theta,\varepsilon^t)\right] = \theta_l^k D_l^k(\theta_l^k;\chi) - \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y;\chi)dy$$

for any $\theta_l^k \in \Theta_l^k$, any $l \in N_k$, $k \in \{A, B\}$. Clearly, under the proposed payments, the expected payoff of each agent $l \in N_k$ from side $k = A, B$ is given by

$$U_l^k(\theta_l^k) = \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y;\chi)dy$$

which is non-negative given that $D_l^k(\underline{\theta}_l^k;\chi) \geq 0$ and that χ satisfies average monotonicity. Furthermore, under the proposed mechanism $\mathcal{M}$ =(χ, ψ),

$$\begin{aligned} U_l^k(\theta_l^k) &= \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y;\chi)dy = \int_{\underline{\theta}_l^k}^{\tilde{\theta}_l^k} D_l^k(y;\chi)dy + \int_{\tilde{\theta}_l^k}^{\theta_l^k} D_l^k(y;\chi)dy \\ &\geq U_l^k(\tilde{\theta}_l^k) + (\theta_l^k - \tilde{\theta}_l^k) D_l^k(\tilde{\theta}_l^k;\chi) = U_i^k(\theta_i^k, \tilde{\theta}_i^k) \end{aligned}$$

where the inequality follows again from the fact that χ satisfies average monotonicity. We conclude that the mechanism $\mathcal{M}$ =(χ, ψ) (a) is individually rational and incentive compatible and (b) gives zero surplus to the lowest type of each agent (i.e., $U_l^k(\underline{\theta}_l^k) = 0$, all $l \in N_k$, $k = A, B$). ■

Proof of Theorem 1. As mentioned in the main text, the proof proceeds in three steps. The first step uses the results in Lemma 1 to express the platform's expected profits in terms of dynamic virtual surplus, net of the sum of all lowest types' expected payoff. More specifically, Lemma 2 below establishes that if a matching rule χ maximizes dynamic virtual surplus, gives the lowest type of each agent a sufficiently high gross payoff, and satisfies average monotonicity, then there exists a payment rule ψ such that the pair (χ, ψ) constitutes a profit-maximizing matching mechanism. The second step (Lemma 3 below) establishes that the myopic rule χ^m and the virtual index rule χ^I maximize dynamic virtual surplus in the two respective environments of Theorem 1. The the final step (Lemma 4 below) establishes that χ^m and χ^I satisfy average monotonicity.

Step 1. We start with the following result.

Lemma 2 (dynamic virtual surplus) *Consider the environment with observable match values described above. Let $\chi^P \in \mathcal{X}$ be any matching rule that maximizes dynamic virtual surplus, as defined in (5). Suppose that χ^P satisfies average monotonicity and that $D_l^k(\underline{\theta}_l^k, \chi^P) \geq 0$, for all $l \in N_k$, $k = A, B$. Then there exists a payment rule ψ^P such that the matching mechanism $\mathcal{M}$ =(χ^P, ψ^P) constitutes a profit-maximizing matching mechanism. Furthermore, ψ^P is such that the participation constraints of the lowest types hold with equality, i.e., $U_l^k(\underline{\theta}_l^k) = 0$, for all $l \in N_k$, $k \in \{A, B\}$.*

Proof of Lemma 2. Using Lemma 1, we have that, in any mechanism $\mathcal{M} = (\chi, \psi)$ that is incentive compatible and individually rational for the agents, the NPV of the payments the platform receives from each type θ_l^k of agent $l \in N_k$, $k = A, B$, is given by

$$\mathbb{E}^{\lambda[\chi]|\theta_l^k}\left[\sum_{t=0}^{\infty} \delta^t \psi_{lt}^k(\theta, \varepsilon^t)\right] = \theta_l^k D_l^k(\theta_l^k;\chi) - U_l^k(\underline{\theta}_l^k) - \int_{\underline{\theta}_l^k}^{\theta_l^k} D_l^k(y;\chi)dy.$$

The law of iterated expectations along with integration by parts then implies that the platform's expected revenue is given by

$$\mathbb{E}^{\lambda[\chi]}\left[\sum_{t=0}^{\infty}\delta^t\left(\sum_{i\in N_A}\sum_{j\in N_B}\left(\phi_i^A(\theta_i^A)\varepsilon_{ijt}^A+\phi_j^A(\theta_j^A)\varepsilon_{ijt}^B-c_{ijt}\right)\chi_{ijt}(\theta,\varepsilon^t)\right)\right]-\sum_{k=A,B}\sum_{l\in N_k}U_l^k(\underline{\theta}_l^k). \quad (11)$$

Now observe that the first term in (11) is dynamic virtual surplus, as defined in (5). Now let $\chi^P\in\mathcal{X}$ be any feasible matching rule that maximizes 5. Suppose that χ^P satisfies average monotonicity and that $D_l^k(\underline{\theta}_l^k,\chi^P)\geq 0$, for all $l\in N_k$, $k=A,B$. Then Lemma 1 implies that there exists a payment rule ψ^P such that the matching mechanism $\mathcal{M}=(\chi^P,\psi^P)$ (a) is individually rational and incentive compatible and (b) gives zero surplus to the lowest type. That the mechanism $\mathcal{M}=(\chi^P,\psi^P)$ constitutes a profit-maximizing mechanism then follows from the fact that the platform's profits under $\mathcal{M}=(\chi^P,\psi^P)$ is given by

$$\mathbb{E}^{\lambda[\chi^P]}\left[\sum_{t=0}^{\infty}\delta^t\sum_{i\in N_A}\sum_{j\in N_B}\left(\phi_i^A(\theta_i^A)\varepsilon_{ijt}^A+\phi_j^A(\theta_j^A)\varepsilon_{ijt}^B-c_{ijt}\right)\chi_{ijt}^P(\theta,\varepsilon^t)\right] \quad (12)$$

whereas its profits under any other mechanism $\mathcal{M}=(\chi,\psi)$ that is incentive compatible and individually rational for the agents is given by (11). That the expression in (12) is larger than the one in (11) follows from the fact that χ^P maximizes the first term in (11) along with the fact that individual rationality requires that $U_l^k(\underline{\theta}_l^k)\geq 0$ all $l\in N_k$, $k=A,B$. ■

Step 2. The next lemma establishes that χ^m and χ^I maximize (5) in the two respective environments of Theorem 1.

Lemma 3 (maximization of DVS) *(1) Suppose match values evolve exogenously. Then the myopic rule χ^m maximizes (5). (2) Suppose the environment satisfies the assumptions of the experimentation model defined above and that either $M\geq n_A\cdot n_B$ or $M=1$. Then the virtual index rule χ^I maximizes (5).*

Proof of Lemma 3. Suppose match values evolve exogenously. That the myopic rule χ^m maximizes (5) follows from the fact that it maximizes the objective in (5) period by period and state by state.

Next, consider the experimentation environment defined in the model setup. The assumptions on the evolution of the match values imply that the problem of maximizing (5) can be viewed as a multiarmed bandit problem, with each arm corresponding to a potential match, and with the reward obtained from activating each arm (i,j) when its state is $(\theta_i^A,\theta_j^B,\varepsilon_{ijt}^A,\varepsilon_{ijt}^B,\sum_{s=1}^{t-1}x_{ijs})$ given by $V_{ijt}(\theta_i^A,\theta_j^B,\varepsilon_{ijt}^A,\varepsilon_{ijt}^B)$. In order to endow the platform with the possibility of not matching any pairs, define an additional arm that yields a reward $V_t=0$ in every period. For $M=1$, that the virtual index rule χ^I maximizes dynamic virtual surplus is then immediate (see, for example, Whittle (1982)). When, instead, $M\geq n_A\cdot n_B$, since the platform's capacity constraint never binds, that χ^I

maximizes dynamic virtual surplus follows from the fact that we can treat the problem as one of solving $n_A \cdot n_B$ separate two-armed bandit problems, one for each potential pair, with the rewards of one arm corresponding to those from matching that pair and the ones for the other (safe) arm equal to zero. ■

Step 3. The final step in the proof of Theorem 1 consists in establishing that the myopic and the virtual index rules satisfy average monotonicity.

Lemma 4 (consistency with average monotonicity) *(1) Suppose match values evolve exogenously. Then χ^m satisfies average monotonicity. (2) Suppose the environment satisfies the assumptions of the experimentation model defined above and either $M \geq n_A \cdot n_B$ or $M = 1$. Then χ^I satisfies average monotonicity.*

Proof of Lemma 4. Suppose that the environment satisfies the assumptions of the experimentation model and that either $M \geq n_A \cdot n_B$ or $M = 1$. Take an arbitrary agent $i \in N_A$ (the arguments for any agent $j \in N_B$ are analogous). Fix a profile of types for the other agents, θ^A_{-i}, and consider an increase in the report of agent i from θ^A_i to θ'^A_i, where $\theta'^A_i > \theta^A_i$. Denote by $\lambda[\chi^I]|\theta^A_i, \theta^A_{-i}$ the process over $\Theta \times \mathcal{E}$ generated by combining χ^I with the distributions F over Θ and the kernels G, given the reports $(\theta^A_i, \theta^A_{-i})$ (similarly, for $\lambda[\chi^I]|\theta'^A_i, \theta^A_{-i}$). Then,

$$\begin{aligned}
&\mathbb{E}^{\lambda[\chi^I]|\theta'^A_i,\theta^A_{-i}}\left[\sum_{t=1}^{\infty}\delta^t \sum_{r\in N_A\backslash\{i\}}\sum_{j\in N_B}\left(\phi^A_r(\theta^A_r)\varepsilon^A_{rjt} + \phi^B_j(\theta^B_j)\varepsilon^B_{rjt} - c_{rjt}\right)\chi^I_{rjt}(\theta'^A_i, \theta^A_{-i}, \varepsilon^t)\right] \\
&+ \mathbb{E}^{\lambda[\chi^I]|\theta'^A_i,\theta^A_{-i}}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi^A_i(\theta'^A_i)\varepsilon^A_{ijt} + \phi^B_j(\theta^B_j)\varepsilon^B_{ijt} - c_{ijt}\right)\chi^I_{ijt}(\theta'^A_i, \theta^A_{-i}, \varepsilon^t)\right] \\
&\geq \mathbb{E}^{\lambda[\chi^I]|\theta^A_i,\theta^A_{-i}}\left[\sum_{t=1}^{\infty}\delta^t \sum_{r\in N_A\backslash\{i\}}\sum_{j\in N_B}\left(\phi^A_r(\theta^A_r)\varepsilon^A_{rjt} + \phi^B_r(\theta^B_r)\varepsilon^B_{rjt} - c_{rjt}\right)\chi^I_{rjt}(\theta^A_i, \theta^A_{-i}, \varepsilon^t)\right] \\
&+ \mathbb{E}^{\lambda[\chi^I]|\theta^A_i,\theta^A_{-i}}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi^A_i(\theta'^A_i)\varepsilon^A_{ijt} + \phi^B_j(\theta^B_j)\varepsilon^B_{ijt} - c_{ijt}\right)\chi^I_{ijt}(\theta^A_i, \theta^A_{-i}, \varepsilon^t)\right]
\end{aligned}$$

and similarly

$$\mathbb{E}^{\lambda[\chi^I]|\theta_i^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{r\in N_A\setminus\{i\}}\sum_{j\in N_B}\left(\phi_r^A(\theta_r^A)\varepsilon_{rjt}^A+\phi_j^B(\theta_j^B)\varepsilon_{rjt}^B-c_{rjt}\right)\chi_{rjt}^I(\theta_i^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$+\,\mathbb{E}^{\lambda[\chi^I]|\theta_i^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi_i^A(\theta_i^A)\varepsilon_{ijt}^A+\phi_j^B(\theta_j^B)\varepsilon_{ijt}^B-c_{ijt}\right)\chi_{ijt}^I(\theta_i^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$\geq \mathbb{E}^{\lambda[\chi^I]|\theta_i'^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{r\in N_A\setminus\{i\}}\sum_{j\in N_B}\left(\phi_r^A(\theta_r^A)\varepsilon_{rjt}^A+\phi_j^B(\theta_j^B)\varepsilon_{rjt}^B-c_{rjt}\right)\chi_{rjt}^I(\theta_i'^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$+\,\mathbb{E}^{\lambda[\chi^I]|\theta_i'^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi_i^A(\theta_i^A)\varepsilon_{ijt}^A+\phi_j^B(\theta_j^B)\varepsilon_{ijt}^B-c_{ijt}\right)\chi_{rjt}^I(\theta_i'^A,\theta_{-i}^A,\varepsilon^t)\right]$$

where the inequalities follow from the fact that the rule χ^I maximizes dynamic virtual surplus, not just in expectation but for any vector of reported types $\theta \in \Theta$.

Combining the two inequalities, we have that

$$\mathbb{E}^{\lambda[\chi^I]|\theta_i'^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi_i^A(\theta_i'^A)-\phi_i^A(\theta_i^A)\right)\varepsilon_{ijt}^A\chi_{ijt}^I(\theta_i'^A,\theta_{-i}^A,\varepsilon^t)\right]$$
$$\geq \mathbb{E}^{\lambda[\chi^I]|\theta_i^A,\theta_{-i}^A}\left[\sum_{t=1}^{\infty}\delta^t \sum_{j\in N_B}\left(\phi_i^A(\theta_i'^A)-\phi_i^A(\theta_i^A)\right)\varepsilon_{ijt}^A\chi_{ijt}^I(\theta_i^A,\theta_{-i}^A,\varepsilon^t)\right].$$

Equivalently, since the above holds for all θ_{-i}^k, using the fact that the match values are independent of the initial types and that the initial types are independent across agents, we have that

$$\left(\phi_i^A(\theta_i'^A)-\phi_i^A(\theta_i^A)\right)\cdot\left(D_i^A(\theta_i'^A;\chi^I)-D_i^A(\theta_i^A;\chi^I)\right)\geq 0.$$

Therefore, $\phi_i^A(\theta_i'^A) > \phi_i^A(\theta_i^A)$ implies $D_i^A(\theta_i'^A,\chi^I) \geq D_i^A(\theta_i^A,\chi^I)$. The assumption that virtual values $\phi_i^A(\cdot)$ are strictly increasing then gives the result.

The same argument shows that, when match values are exogenous, the rule χ^m satisfies average monotonicity. In fact, since the process of match values is exogenous and the matching rule χ^m is myopic, a similar argument applies for each period and given each state. That is, the average monotonicity of χ^m holds ex-post, and period-by-period. ■

Combining the results in the above three lemma establishes the results in the theorem. Q.E.D.

Proof of Example 3. To see that, when $\theta_i^A = \tilde{\theta}_i^A$, χ^I matches agents i and j' at all $t \geq 1$, note first that χ^I is optimal (see Theorem 1). Since the virtual surplus from matching i and j' is constant, if χ^I matches i and j at any period, we can assume that it does so in period $t = 1$. For sufficiently small β, the dynamic virtual surplus in case i and j are matched in period 1 becomes strictly smaller than $2/(1-\delta)$, which is the dynamic virtual surplus obtained from matching i and j' for all $t \geq 1$.

Next, to see that ex-post monotonicity is violated, note that given a sequence of values $\hat{\varepsilon}$ such that $\hat{\varepsilon}^A_{ijt} = 100$ for all t whenever $x_{ijt-1} = 1$, $\sum_{t=1}^{\infty} \delta^t \sum_{r \in N_B} \hat{\varepsilon}^A_{irt} \chi^I_{irt}(\tilde{\theta}^A_i, \theta^A_{-i}, \hat{\varepsilon}^t) = 2/(1-\delta)$, and $\sum_{t=1}^{\infty} \delta^t \sum_{r \in N_B} \hat{\varepsilon}^A_{irt} \chi^I_{irt}(\hat{\theta}^A_i, \theta^A_{-i}, \hat{\varepsilon}^t) = 100 \cdot \delta/(1-\delta)$. The latter is greater for all $\delta > 1/50$. ■

Proof of Theorem 2. The proof is in two steps. Step 1 shows that, at any period $t \geq 1$, irrespective of past reports, an agent who expects all other agents to report truthfully at present and in future periods finds it optimal to remain in the mechanism and also report truthfully at any period $s \geq t$, irrespective of the agent's beliefs about the other agents' types and values. Step 2 then completes the proof by showing that participating and reporting truthfully also at period $t = 0$ is optimal for each agent who expects all other agents to report truthfully at all periods. The proof below focuses on the experimentation model. The proof for the environment with exogenous processes follows from the same arguments and hence is omitted.

Step 1. Given the nature of the mechanism $(\tilde{\chi}^I, \tilde{\psi}^I)$, it suffices to show that, at any $t \geq 1$, and for any $(z, \theta_t, \varepsilon_t, x^{t-1})$, any agent $l \in N_A$ who expects all other agents to report truthfully at all $s \geq t$, finds it optimal to stay in the mechanism and report $(\theta^A_l, (\varepsilon^A_{ls}))$ truthfully in all periods $s \geq t$ (as usual, the arguments for any agent $l \in N_B$ are analogous and hence omitted).

To see this, suppose that agent l's true type and period-t match values are $(\theta^A_l, \varepsilon^A_{lt})$, and denote by

$$\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta^A_l,\theta^A_{-l},\varepsilon^A_{lt},\varepsilon^A_{-lt},x^{t-1};(\hat{\theta}^A_{lt},\hat{\varepsilon}^A_{lt})}\left[R^A_{lt+1}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]$$

the expected marginal contribution of agent l from period $t+1$ onwards when, in period t, l reports $(\hat{\theta}^A_{lt}, \hat{\varepsilon}^A_{lt})$ and then reports truthfully in all future periods (assuming all other agents report truthfully in period t and in all subsequent periods). The distribution $\tilde{\lambda}[\tilde{\chi}^I]|z, \theta^A_l, \theta^A_{-l}, \varepsilon^A_{lt}, \varepsilon^A_{-lt}, x^{t-1}; (\hat{\theta}^A_{lt}, \hat{\varepsilon}^A_{lt})$ is here over $\Theta \times \mathcal{E}^{\geq t} \times X$ and coincides with the distribution $\tilde{\lambda}[\tilde{\chi}^I]|z, \theta^A_l, \theta^A_{-l}, \varepsilon^A_{lt}, \varepsilon^A_{-lt}, x^{t-1}$ when $(\hat{\theta}^A_{lt}, \hat{\varepsilon}^A_{lt}) = (\theta^A_{lt}, \varepsilon^A_{lt})$.

Using the one-stage-deviation principle, the fact that agents are asked to report their type in each period $t \geq 1$, and the fact that, under the proposed mechanism, after any history, each agent's continuation payoff (under truthtelling by all agents) coincides with the weighed contribution to virtual social welfare, we thus only need to show that, from any period-t report $(\hat{\theta}^A_{lt}, \hat{\varepsilon}^A_{lt}) \in \Theta^A_l \times \mathcal{E}^A_{lt}$

by agent l,

$$\sum_{j\in N_B}\theta_l^A\varepsilon_{ljt}^A\tilde{\chi}_{ljt}^I(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})-\tilde{\psi}_{lt}^{A,I}(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$+\frac{\delta z_l^A}{\phi_l^A(z_l^A)}\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_l^A,\theta_{-l}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1}}\left[R_{lt+1}^A(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]$$
$$\geq\sum_{j\in N_B}\theta_l^A\varepsilon_{ljt}^A\tilde{\chi}_{ljt}^I(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})-\tilde{\psi}_{lt}^{A,I}(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$+\frac{\delta z_l^A}{\phi_l^A(z_l^A)}\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_l^A,\theta_{-l}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1};(\hat{\theta}_{lt}^A,\hat{\varepsilon}_{lt}^A)}\left[R_{lt+1}^A(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]$$

That is, when agent i is guaranteed a continuation payoff (from period $t+1$ onward, after any possibly history) equal to her expected marginal contribution to the weighted virtual surplus, then it must be optimal for him to report truthfully in period t when she intends to report truthfully in all future periods. Note that the reason why the future payoff is scaled by $z_l^A/\phi_l^A(z_l^A)$ is that, by construction, the mechanism $(\tilde{\chi}^I,\tilde{\psi}^I)$ provides an agent with true type $\phi_l^A(z_l^A)\theta_l^A/z_l^A$ a continuation payoff equal to $R_{lt+1}^A(z,\theta_{t+1},\varepsilon_{t+1},x^t)$. Given that the agent's true type is θ_l^A, the agent needs to discount the future payoff by $z_l^A/\phi_l^A(z_l^A)$.

Now, from the construction in the main text, observe that the left hand side of the above inequality can be rewritten in terms of the functions W_t and $W_t^{-l,A}$ as follows

$$\frac{z_l^A}{\phi_l^A(z_l^A)}\left[W_t(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})-W_t^{-l,A}(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})\right].$$

That is, agent l's expected continuation payoff under truthtelling from period t onward is equal to her expected contribution to the total weighted virtual surplus, scaled by the weight constructed based on his period-0 report. Thus we need to show that

$$\frac{z_l^A}{\phi_l^A(z_l^A)}\left[W_t(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})-W_t^{-l,A}(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})\right]$$
$$\geq\sum_{j\in N_B}\theta_l^A\varepsilon_{ljt}^A\tilde{\chi}_{ljt}^I(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})-\tilde{\psi}_{lt}^{A,I}(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$+\frac{\delta z_l^A}{\phi_l^A(z_l^A)}\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_l^A,\theta_{-l}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1};(\hat{\theta}_{lt}^A,\hat{\varepsilon}_{lt}^A)}\left[R_{lt+1}^A(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right].\qquad(13)$$

Using (8), we can rewrite the period-t payment given the report $(\hat{\theta}_{lt}^A, \hat{\varepsilon}_{lt}^A)$ as follows

$$
\begin{aligned}
&\tilde{\psi}_{lt}^{A,I}(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&= \sum_{j \in N_B} \hat{\theta}_l^A \hat{\varepsilon}_{ljt}^A \tilde{\chi}_{ljt}^I(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) - \frac{z_l^A}{\phi_l^A(z_l^A)} r_{lt}^A(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&= \sum_{j \in N_B} \hat{\theta}_l^A \hat{\varepsilon}_{ljt}^A \tilde{\chi}_{ljt}^I(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&- \frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i \in N_A} \sum_{j \in N_B} \tilde{V}_{ijt}\left(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A\right) \cdot \tilde{\chi}_{ijt}^I(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&+ \frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i \in N_A} \sum_{j \in N_B} \tilde{V}_{ijt}\left(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A\right) \cdot \tilde{\chi}_{ijt}^{I,-l,A}(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{-1}) \\
&- \frac{z_l^A}{\phi_l^A(z_l^A)} \delta \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}} \left[W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \\
&+ \frac{z_l^A}{\phi_l^A(z_l^A)} \delta \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^{I,-l,A}]|z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}} \left[W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \qquad (14)
\end{aligned}
$$

After rearranging, we have that

$$
\begin{aligned}
&\tilde{\psi}_{lt}^{A,I}(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&= -\frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i \in N_A \setminus \{l\}} \sum_{j \in N_B} \tilde{V}_{ijt}\left(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A\right) \cdot \tilde{\chi}_{ijt}^I(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&- \frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{j \in N_B} \left(\frac{\phi_j^B(z_j^B)}{z_j^B} \theta_j^B \varepsilon_{ljt}^B - c_{ljt}\right) \cdot \tilde{\chi}_{ljt}^I(z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}) \\
&- \frac{z_l^A}{\phi_l^A(z_l^A)} \delta \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \hat{\theta}_{lt}^A, \theta_{-lt}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}} \left[W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \\
&+ \frac{z_l^A}{\phi_l^A(z_l^A)} W_t^{-l,A}(z, \theta_{lt}^A, \theta_{-lt}^A, \varepsilon_{lt}^A, \varepsilon_{-lt}^A, x^{t-1})
\end{aligned}
$$

Also note that

$$
\begin{aligned}
&\frac{\delta z_l^A}{\phi_l^A(z_l^A)} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \theta_l^A, \theta_{-l}^A, \varepsilon_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}; (\hat{\theta}_{lt}^A, \hat{\varepsilon}_{lt}^A)} \left[R_{lt+1}^A(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \\
&= \frac{\delta z_l^A}{\phi_l^A(z_l^A)} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \theta_l^A, \theta_{-l}^A, \varepsilon_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}; (\hat{\theta}_{lt}^A, \hat{\varepsilon}_{lt}^A)} \left[W_{t+1}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t) - W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \\
&= \frac{\delta z_l^A}{\phi_l^A(z_l^A)} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \theta_l^A, \theta_{-l}^A, \varepsilon_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}; (\hat{\theta}_{lt}^A, \hat{\varepsilon}_{lt}^A)} \left[W_{t+1}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right] \\
&- \frac{\delta z_l^A}{\phi_l^A(z_l^A)} \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z, \hat{\theta}_l^A, \theta_{-l}^A, \hat{\varepsilon}_{lt}^A, \varepsilon_{-lt}^A, x^{t-1}} \left[W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)\right]
\end{aligned}
$$

where the last equality uses the fact that $W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t)$ is independent of agent l's type and values and the fact that the period-t decisions are invariant in agent l's *true* type and values.

Therefore, the right hand side of the inequality (13) is equivalent to

$$\frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i\in N_A}\sum_{j\in N_B} \tilde{V}_{ijt}\left(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A\right)\cdot \tilde{\chi}_{ijt}^I(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$+\frac{\delta z_l^A}{\phi_l^A(z_l^A)}\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_l^A,\theta_{-l}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1};(\hat{\theta}_{lt}^A,\hat{\varepsilon}_{lt}^A)}\left[W_{t+1}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]$$
$$-\frac{z_l^A}{\phi_l^A(z_l^A)}W_t^{-l,A}(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$

It therefore suffices to show that

$$\frac{z_l^A}{\phi_l^A(z_l^A)}W_t(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$\geq \frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i\in N_A}\sum_{j\in N_B} \tilde{V}_{ijt}\left(z,\theta_{lt}^A,\theta_{-lt}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A\right)\cdot \tilde{\chi}_{ijt}^I(z,\hat{\theta}_{lt}^A,\theta_{-lt}^A,\hat{\varepsilon}_{lt}^A,\varepsilon_{-lt}^A,x^{t-1})$$
$$+\frac{z_l^A}{\phi_l^A(z_l^A)}\delta\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_l^A,\theta_{-l}^A,\varepsilon_{lt}^A,\varepsilon_{-lt}^A,x^{t-1};(\hat{\theta}_{lt}^A,\hat{\varepsilon}_{lt}^A)}\left[W_{t+1}(z,\theta_{t+1},\varepsilon_{t+1},x^t)\right]$$

This follows directly from the fact that the rule $\tilde{\chi}^I$ maximizes (6).

The arguments above establish that the mechanism $(\tilde{\chi}^I,\tilde{\psi}^I)$ is periodic ex-post IC for all periods $t\geq 1$. That it is also periodic IR follows from the fact that each agent's continuation payoff coincides with her expected marginal contribution to aggregate weighted virtual surplus (which is always non-negative) scaled by the positive weight $z_i^k/\phi_i^k(z_i^k)$.

Step 2. It remains to show that, when agent $l\in N_k$ from side $k\in\{A,B\}$ expects all other agents to report truthfully at all periods, it is optimal for her to participate and report truthfully at $t=0$.

We first introduce the following notation. Denote by $\tilde{U}_l^k(z_l^k,\theta_l^k)$ the expected payoff for agent $l\in N_k$ from side k under the mechanism $(\tilde{\chi}^I,\tilde{\psi}^I)$ when her true type is $\theta_l^k\in\Theta_l^k$, she reports $z_l^k\in\Theta_l^k$ in period 0, she plans to report truthfully at all periods $t\geq 1$, and expects all other agents to report truthfully at all periods. For example, when $k=A$, this is equal to

$$\begin{aligned}\tilde{U}_l^A(z_l^A,\theta_l^A) \;&\equiv\; \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z_l^k,\theta_l^k}\left[\sum_{t=1}^{\infty}\delta^t\sum_{j\in N_B}\theta_l^A\varepsilon_{ljt}^A\tilde{\chi}_{ljt}^I(z,\theta_t,\varepsilon_t,x^{t-1})\right]\\ &\quad -\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z_l^k,\theta_l^k}\left[\sum_{t=0}^{\infty}\delta^t\tilde{\psi}_{lt}^{k,I}(z,\theta_t,\varepsilon_t,x^{t-1})\right],\end{aligned}$$

Then, for any agent $l\in N_k$ from each side $k\in\{A,B\}$, let

$$\tilde{D}_l^k(z_l^k,\theta_l^k;\tilde{\chi}^I)=\begin{cases}\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}]|z_l^k,\theta_l^k}\left[\sum_{t=1}^{\infty}\delta^t\sum_{h\in N_{-k}}\varepsilon_{lht}^k\tilde{\chi}_{lht}^I(z,\theta_t,\varepsilon_t,x^{t-1})\right] & \text{if } k=A\\ \mathbb{E}^{\tilde{\lambda}[\tilde{\chi}]|z_l^k,\theta_l^k}\left[\sum_{t=1}^{\infty}\delta^t\sum_{h\in N_{-k}}\varepsilon_{hlt}^k\tilde{\chi}_{hlt}^I(z,\theta_t,\varepsilon_t,x^{t-1})\right] & \text{if } k=B\end{cases}$$

We first show that $\tilde{\chi}^I$ satisfies a monotonicity condition with respect to z_l^k analogous to the average monotonicity property in the environment with observable match values.

Lemma 5 *For all $l \in N_k$, $\theta_l^k \in \Theta_l^k$, $k = A, B$, the function $\tilde{D}_l^k(z_l^k, \theta_l^k; \tilde{\chi}^I)$ is nondecreasing in z_l^k.*

Proof of Lemma 5. Arguments similar to those in the proof of Lemma 4 imply that, for any $l \in N_k$, $\theta_l^k \in \Theta_l^k$, $\tilde{z}_l^k, \hat{z}_l^k \in \Theta_l^k$, with $\tilde{z}_l^k > \hat{z}_l^k$, $k = A, B$,

$$\left(\frac{\phi_l^k(\tilde{z}_l^k)}{\tilde{z}_l^k} - \frac{\phi_l^k(\hat{z}_l^k)}{\hat{z}_l^k}\right) \cdot \theta_l^k \cdot \left(D_l^k(\tilde{z}_l^k, \theta_l^k; \tilde{\chi}^I) - D_l^k(\hat{z}_l^k, \theta_l^k; \tilde{\chi}^I)\right) \geq 0.$$

Regularity implies that $\phi_l^k(\tilde{z}_l^k)/\tilde{z}_l^k > \phi_l^k(\hat{z}_l^k)/\hat{z}_l^k$, which, along with the fact that $\theta_l^k > 0$, in turn implies that $D_l^k(\tilde{z}_l^k, \theta_l^k; \tilde{\chi}^I) \geq D_i^k(\hat{z}_l^k, \theta_l^k; \tilde{\chi}^I)$. ■

Now fix the period-0 report $\hat{z}_l^k$. Because the mechanism $(\tilde{\chi}^I, \tilde{\psi}^I)$ is periodic ex-post incentive compatible from period $t = 1$ onward, we have that the the payoff $\tilde{U}_l^A(z_l^k, \theta_l^k)$ that agent l obtains by reporting $\hat{z}_l^k$ in period $t = 0$ and truthfully in all subsequent periods coincides with the value function for the problem consisting in choosing a reporting strategy from period $t = 1$ onwards, given the period-0 report $\hat{z}_l^k$. Using arguments similar to those that establish Theorem 1 in Pavan, Segal, and Toikka (2014), then permit us to establish that $\tilde{U}_l^A(z_l^k, \theta_l^k)$ must satisfy the following envelope condition

$$\tilde{U}_l^k(z_l^k, \theta_l^k) = \tilde{U}_l^k(z_l^k, \underline{\theta}_l^k) + \int_{\underline{\theta}_l^k}^{\theta_l^k} \tilde{D}_l^k(z_l^k, y; \tilde{\chi}^I)dy. \tag{15}$$

To see that it is optimal for agent l to report truthfully at period $t = 0$, then observe that for any $\theta_l^k, z_l^k \in \Theta_l^k$,

$$\begin{aligned}
\tilde{U}_l^k(z_l^k, \theta_l^k) &= \tilde{U}_l^k(z_l^k, z_l^k) + \int_{z_l^k}^{\theta_l^k} \tilde{D}_l^k(z_l^k, y; \tilde{\chi}^I)dy \\
&\leq \tilde{U}_l^k(z_l^k, z_l^k) + \int_{z_l^k}^{\theta_l^k} \tilde{D}_l^k(y, y; \tilde{\chi}^I)dy \\
&= U_l^k(z_l^k) + \int_{z_l^k}^{\theta_l^k} D_l^k(y; \chi^I)dy = U_l^k(\theta_l^k) = \tilde{U}_l^k(\theta_l^k, \theta_l^k)
\end{aligned}$$

where the first equality follows from (15), the inequality follows from Lemma 4, and all remaining equalities from the fact that, when the agent reports truthfully in all periods her payoff under $(\tilde{\chi}^I, \tilde{\psi}^I)$ coincides with her payoff under (χ^I, ψ^I) in the environment with observable values.

We conclude that reporting truthfully is optimal also in period zero. That participating is also optimal follows from the fact that the agent's expected payoff is the same as in the environment with observable values.

The conclusions for the mechanism $(\tilde{\chi}^m, \tilde{\psi}^m)$ follow from the same arguments as those used above for the mechanism $(\tilde{\chi}^I, \tilde{\psi}^I)$. Q.E.D.

Proof of Proposition 1. First, observe that if agents bid truthfully in all periods then the resulting allocation in the corresponding sequence of matching auctions in each of the environments coincides with the allocation of the respective optimal direct mechanism, as the score of each pair is equal to the virtual surplus their match generates. Furthermore, note that any deviation that is feasible in the sequence of scoring auctions is also feasible in the respective direct mechanism.[33] Therefore, by Theorem 2, it remains only to show that the period-$t \geq 1$ payments in the sequence of scoring auctions coincide with those in the optimal direct mechanism (9) for each of the environments (period-0 payments are the same by construction).

For the case of exogenous match values, the payment of agent $l \in N_A$ given in (9) can be written as

$$\frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{j=1}^{n_B} \left(c_{ljt} - \frac{\phi_j^B(z_j^B)}{z_j^B} \theta_j^B \varepsilon_{iljt}^B \right) \cdot \tilde{\chi}_{ljt}^I(z, \theta_t, \varepsilon_t, x^{t-1})$$
$$+ \frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{i \in N_A \setminus \{l\}} \sum_{j=1}^{n_B} \tilde{V}_{ijt}(z, \theta_t, \varepsilon_t) \cdot \left[\tilde{\chi}_{ijt}^{I,-l,A}(z, \theta_t, \varepsilon_t, x^{t-1}) - \tilde{\chi}_{ijt}^I(z, \theta_t, \varepsilon_t, x^{t-1}) \right],$$

as the agent has no dynamic externality. Similarly, in the experimentation model with $M \geq n_A \cdot n_B$, l's payment is

$$\frac{z_l^A}{\phi_l^A(z_l^A)} \sum_{j=1}^{n_B} \left(c_{ljt} - \frac{\phi_j^B(z_j^B)}{z_j^B} \theta_j^B \varepsilon_{ljt}^B \right) \cdot \tilde{\chi}_{ljt}^I(z, \theta_t, \varepsilon_t, x^{t-1}),$$

as no other pairs are excluded as a result of l's presence. In the experimentation model with $M = 1$, if l is matched to an agent on side B, then his payment (9) reduces to,

$$\frac{z_l^A}{\phi_l^A(z_l^A)} \left(\sum_{j=1}^{n_B} \left(c_{ljt} - \frac{\phi_j^B(z_j^B)}{z_j^B} \theta_j^B \varepsilon_{ljt}^B \right) \cdot \tilde{\chi}_{ljt}^I(z, \theta_t, \varepsilon_t, x^{t-1}) + (1-\delta) W_t^{-l,A}(z, \theta_t, \varepsilon_t, x^{t-1}) \right),$$

applying the observation that $\mathbb{E}^{\tilde{\lambda}[\tilde{\chi}^I]|z,\theta_t,\varepsilon_t,x^{t-1}} \left[W_{t+1}^{-l,A}(z, \theta_{t+1}, \varepsilon_{t+1}, x^t) \right] = W_t^{-l,A}(z, \theta_t, \varepsilon_t, x^{t-1})$. Otherwise, l's payment is 0. Payments for side-B agents can be expressed analogously.

The above payments clearly coincide with those in the respective sequential auctions. ■

Proof of Theorem 3. For each pair $(i, j) \in N_A \times N_B$ and realizations $(\theta, \varepsilon_t) \in \Theta \times \mathcal{E}_t$, let $V_{ijt}^W(\theta, \varepsilon_t) \equiv \theta_i^A \varepsilon_{ijt}^A + \theta_i^B \varepsilon_{ijt}^B - c_{ijt}$ denote the true social surplus from matching the pair in period t, and define the index γ_{ijt}^W analogously to the virtual index defined in Section 3, with V_{ijt}^W replacing V_{ijt}, i.e.,

$$\gamma_{ijt}^W \left(\theta, \varepsilon_t, x^{t-1}\right) \equiv \max_{\tau} \mathbb{E}^{\lambda[\chi^o]|\theta,\varepsilon_t,x^{t-1}} \left\{ \frac{\sum_{s=t}^{\tau} \delta^{s-t} V_{ijs}^W(\theta, \varepsilon_t)}{\sum_{s=t}^{\tau} \delta^{s-t}} \right\}.$$

[33] This is because, while agents make bids corresponding to their utility, u_{lt}^k, rather than their match values ε_{lt}^k, the platform can recover these values by dividing the bids by the period-t membership choice θ_{lt}^k.

Since match values are nonnegative, for each pair $(i,j) \in N_A \times N_B$, any $t \geq 1$, any $(\theta, \varepsilon_t) \in \Theta \times \mathcal{E}_t$, $V_{ijt}(\theta, \varepsilon_t) \leq V^W_{ijt}(\theta, \varepsilon_t)$. Furthermore, $\gamma_{ijt}\left(\theta, \varepsilon_t, x^{t-1}\right) \leq \gamma^W_{ijt}\left(\tilde{\theta}, \tilde{\varepsilon}_t, \tilde{x}^{t-1}\right)$ for any pair $(\theta, \varepsilon_t, x^{t-1}), (\tilde{\theta}, \tilde{\varepsilon}_t, \tilde{x}^{t-1}) \in \Theta \times \mathcal{E}_t \times X^{t-1}$ such that

$$(\theta^A_i, \theta^B_j, \varepsilon^A_{ijt}, \varepsilon^B_{ijt}, \sum\nolimits_{s=1}^{t-1} x_{ijs}) = (\tilde{\theta}^A_i, \tilde{\theta}^B_j, \tilde{\varepsilon}^A_{ijt}, \tilde{\varepsilon}^B_{ijt}, \sum\nolimits_{s=1}^{t-1} \tilde{x}_{ijs}).$$

Part 1. First, consider the case of exogenous processes. Because the capacity constraint is not binding, in each period the profit-maximizing platform (alternatively, the planner) activates all links for which virtual surplus (alternatively, the true surplus) is nonnegative. The above property, along with the fact that $V_{ijt}(\theta, \varepsilon_t) \leq V^W_{ijt}(\theta, \varepsilon_t)$ all $(i,j) \in N_A \times N_B$, $t \geq 1$, $(\theta, \varepsilon_t) \in \Theta \times \mathcal{E}_t$ then yields the result.

Next, consider the experimentation model. Again, because the capacity constraint is not binding, in each period the platform (alternatively, the planner) activates all links for which the virtual index (alternatively, the true index) is nonnegative. The result is then proved by induction. To see this, note that the property necessarily holds at $t = 1$ given that $\gamma_{ij1}(\theta, \varepsilon_1, x^0) \leq \gamma^W_{ijt}(\theta, \varepsilon_1, x^0)$ all $(i,j) \in N_A \times N_B$, where x^0 denotes the common null history of past interactions. Suppose now the property holds for all $1 \leq s < t$. Any link that the platform activates at period t has been activated at each preceding period by both the planner and the platform. The result then follows from the fact that for any $(i,j) \in N_A \times N_B$ that has been activated at all periods $s \leq t$ by both the planner and the platform $\gamma_{ijt} \leq \gamma^W_{ijt}$. Because $\chi^P_{ijt}(\theta, \varepsilon^t) = 1$ implies $\gamma_{ijt} \geq 0$, we then have that $\gamma^W_{ijt} \geq 0$. That the welfare-maximizing policy for this environment is the index policy (with indexes γ^W) then gives the result.

Part 2. Consider first the case of exogenous processes. When $M \geq n_A \cdot n_B$ the result follows from Part 1. Thus consider the case in which $M < n_A \cdot n_B$. The result then follows from the following two properties: (a) the set of links with nonnegative virtual surplus is a subset of the set of links with nonnegative true surplus, (b) the cardinality of the set of links selected in each period by the profit-maximizing rule (alternatively, the welfare-maximizing rule) is the minimum between M and the cardinality of the set of links with nonnegative virtual (alternatively, true) surplus.

Next, consider the experimentation model. When $M \geq n_A \cdot n_B$, the result follows again from Part 1. Thus consider the case in which $M = 1$. First observe that, under the profit-maximizing rule, if at some period $t \geq 1$, $\chi^P_{ijt}(\theta, \varepsilon^t) = 0$ all $(i,j) \in N_A \times N_B$, then $\chi^P_{ijs}(\theta, \varepsilon^t) = 0$ all $s > t$, all $(i,j) \in N_A \times N_B$. The same property holds for the welfare-maximizing rule χ^W. Next, observe that if matching stops at period t under profit maximization (alternatively, welfare maximization), then $\gamma_{ijt} < 0$ all $(i,j) \in N_A \times N_B$ (alternatively, $\gamma^W_{ijt} < 0$ all $(i,j) \in N_A \times N_B$).

Now suppose that, under profit maximization, matching is still active in period t (meaning, there exists $(i,j) \in N_A \times N_B$ such that, given the state (θ, ε^t) generated by χ^P in previous periods, $\chi^P_{ijt}(\theta, \varepsilon^t) = 1$). Let $(\theta, \tilde{\varepsilon}^t)$ be the history generated by χ^P in previous periods. Then either

$\sum_{s=1}^{t-1}\chi_{ijs}^{W}(\theta,\varepsilon^{s})=\sum_{s=1}^{t-1}\chi_{ijs}^{P}(\theta,\tilde{\varepsilon}^{s})$ all $(i,j)\in N_A\times N_B$, in which case, for all $(i,j)\in N_A\times N_B$,

$$(\theta_i^A,\theta_j^B,\varepsilon_{ijt}^A,\varepsilon_{ijt}^B,\sum_{s=1}^{t-1}x_{ijs})=(\theta_i^A,\theta_j^B,\tilde{\varepsilon}_{ijt}^A,\tilde{\varepsilon}_{ijt}^B,\sum_{s=1}^{t-1}\tilde{x}_{ijs}).$$

[34] The result then follows from the fact that $\gamma_{ijt}\leq\gamma_{ijt}^{W}$ all $(i,j)\in N_A\times N_B$, along with the index structure of the optimal rules. Or, there exists $(i,j)\in N_A\times N_B$ such that $\sum_{s=1}^{t-1}\chi_{ijs}^{W}(\theta,\tilde{\varepsilon}^{s})<\sum_{s=1}^{t-1}\chi_{ijs}^{P}(\theta,\varepsilon^{s})$. In this case, there must exist $\tau<t$ such that $\chi_{ij\tau}^{P}(\theta,\varepsilon^{\tau})=1$ and

$$(\theta_i^A,\theta_j^B,\tilde{\varepsilon}_{ijt}^A,\tilde{\varepsilon}_{ijt}^B,\sum_{s=1}^{t-1}\chi_{ijs}^{W}(\theta,\tilde{\varepsilon}^{s}))=(\theta_i^A,\theta_j^B,\varepsilon_{ij\tau}^A,\varepsilon_{ij\tau}^B,\sum_{s=1}^{\tau-1}\chi_{ijs}^{P}(\theta,\varepsilon^{s})).$$

That the platform activated such link in period τ in turn implies that $\gamma_{ij\tau}(\theta,\varepsilon_{\tau},x^{t-1})\geq 0$. The index properties described at the beginning of the proof then imply that $\gamma_{ijt}^{W}(\theta,\varepsilon_{t},x^{t-1})\geq 0$. This last property in turn implies that matching is still active in period t also under welfare maximization. Q.E.D.

Proof of Proposition 2. The proof is in two steps. Step 1 shows that, under the separability assumption, the myopic rule selecting in each period the pairs for which the auxiliary reward is the highest among those for which the reward is nonnegative is optimal in the fictitious environment in which the rewards are given by the auxiliary processes. Step 2 in turn uses certain properties relating the auxiliary processes to the original ones to establish the result in the Proposition.

Step 1. Consider a fictitious environment in which, at any period $t\geq 1$, any $(\theta,\varepsilon^{t},x^{t-1})\in\Theta\times\mathcal{E}^{t}\times X^{t-1}$, the period-$t$ reward of matching any pair of agents $(i,j)\in N_A\times N_B$ is given by the auxiliary processes, that is, is equal to $\underline{V}_{ijt}(\theta,\varepsilon^{t},x^{t-1})$, as defined in the main text. Let

$$\underline{Q}_t(\theta,\varepsilon^{t},x^{t-1})=\left\{\begin{array}{l}(i,j)\in N_A\times N_B\text{ s.t. (i) }\underline{V}_{ijt}(\theta,\varepsilon^{t},x^{t-1})\geq 0\text{ and}\\ \text{(ii) }\#\left\{(l,m)\in N_A\times N_B:\underline{V}_{lmt}(\theta,\varepsilon^{t},x^{t-1})>\underline{V}_{ijt}(\theta,\varepsilon^{t},x^{t-1})\right\}<M\end{array}\right\}.$$

Next, let $\underline{\chi}^{m}$ denote the myopic matching rule linking at each period $t\geq 1$ the pairs with the highest nonnegative auxiliary reward, subject to the platform's capacity constraint. That is,

$$\underline{\chi}_{ijt}^{m}(\theta,\varepsilon^{t},x^{t-1})=1\ \Rightarrow\ (i,j)\in\underline{Q}_t(\theta,\varepsilon^{t},x^{t-1}).$$

Furthermore,

$$\begin{aligned}(i,j)\ &\in\ \underline{Q}_t(\theta,\varepsilon^{t},x^{t-1}),\ \#\underline{Q}_t(\theta,\varepsilon^{t},x^{t-1})\leq M,\ \underline{V}_{ijt}(\theta,\varepsilon^{t},x^{t-1})>0\\ &\Rightarrow\ \underline{\chi}_{ijt}^{m}(\theta,\varepsilon^{t},x^{t-1})=1.\end{aligned}$$

Lemma 6 *Consider a fictitious environment in which the rewards are given by the auxiliary processes defined above. Suppose the separability condition of Definition 6 holds. Then the myopic rule* $\underline{\chi}^{m}$ *is*

[34] This last property follows from assumption (iii) of the experimentation model.

optimal for this environment (in the sense that it maximizes the ex-ante expected discounted present value of the rewards).

Proof of lemma 6. Suppose, towards a contradiction, that the claim is not true, meaning that there exists a rule $\chi \neq \underline{\chi}^m$ that dominates $\underline{\chi}^m$ in terms of ex-ante expected discounted present value of the auxiliary rewards. Because $\chi \neq \underline{\chi}^m$, there must exists $t \geq 1$ and a set of "states" $(\theta, \varepsilon^t, x^{t-1}) \in \Theta \times \mathcal{E}^t \times X^{t-1}$ of strictly positive probability under $\lambda[\chi]$ for which one of the following two properties is true, for some $(i,j) \in N_A \times N_B$: (a) $\chi_{ijt}(\theta, \varepsilon^t, x^{t-1}) = 1$ and $(i,j) \notin \underline{Q}_t(\theta, \varepsilon^t, x^{t-1})$, or (b) $(i,j) \in \underline{Q}_t(\theta, \varepsilon^t, x^{t-1}) \wedge \#\underline{Q}_t(\theta, \varepsilon^t, x^{t-1}) \leq M \wedge \underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) > 0$ and $\chi_{ijt}(\theta, \varepsilon^t, x^{t-1}) = 0$.

Consider first the situation corresponding to part (a) above. Because $(i,j) \notin \underline{Q}_t(\theta, \varepsilon^t, x^{t-1})$, either $\underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) < 0$ or

$$\#\left\{(l,m) \in N_A \times N_B : \underline{V}_{lmt}(\theta, \varepsilon^t, x^{t-1}) > \underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1})\right\} \geq M.$$

In the first case, given that the auxiliary rewards are nonincreasing, it is immediate that the myopic rule $\underline{\chi}^m$, by leaving the pair (i,j) unmatched in period t, improves upon χ. Thus consider the second case. There must exist another pair $(i',j') \in N_A \times N_B$ such $\underline{V}_{i'j't}(\theta, \varepsilon^t, x^{t-1}) > \underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) \geq 0$ such that $\chi_{i'j't}(\theta, \varepsilon^t, x^{t-1}) = 0$. Because the environment satisfies the separability condition, this implies $\underline{V}_{i'j't} \geq \underline{V}_{ijt}/(1-\delta)$ (we are dropping here the arguments to ease the notation). Because the auxiliary processes are nonincreasing, the entire stream of rewards that can be obtained by matching (i,j) in period t as well all in some of the subsequent periods is no larger than $\underline{V}_{ijt}/(1-\delta)$ which is smaller than the flow reward that is obtained by matching (i',j') in period t. Hence, by favoring (i',j') over (i,j) in period t, the myopic rule $\underline{\chi}^m$ again improves weakly over the rule χ.

Finally, consider the situation corresponding to part (b) above and denote

$$\chi_t(\theta, \varepsilon^t, x^{t-1}) = \left\{(l,m) \in N_A \times N_B : \chi_{lmt}(\theta, \varepsilon^t, x^{t-1}) = 1\right\}$$

the set of period-t links under the rule χ. If $\chi_t(\theta, \varepsilon^t, x^{t-1}) \subsetneq \underline{Q}_t(\theta, \varepsilon^t, x^{t-1})$, which implies that $\#\chi_t(\theta, \varepsilon^t, x^{t-1}) < M$, then the myopic rule $\underline{\chi}^m$ improves upon χ by adding to the latter a link with a strictly positive reward. If, instead, $\chi_t(\theta, \varepsilon^t, x^{t-1}) \nsubseteq \underline{Q}_t(\theta, \varepsilon^t, x^{t-1})$, then there exists a pair $(i',j') \in \chi_t(\theta, \varepsilon^t, x^{t-1})$ and such that $(i',j') \notin \underline{Q}_t(\theta, \varepsilon^t, x^{t-1})$ such $\underline{V}_{i'j't}(\theta, \varepsilon^t, x^{t-1}) < \underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1})$. Because $\underline{V}_{ijt}(\theta, \varepsilon^t, x^{t-1}) > 0$, the myopic rule, by activating (i,j) and dropping (i',j') necessarily improves upon the rule χ, contradicting the fact that χ dominates $\underline{\chi}^m$. ■

Step 2. Consider the following lemma, whose proof follows from results in Ishikida and Varaiya (1994), expressed in terms of the matching environment under examination.

Lemma 7 *(1) Under any matching policy χ, the ex-ante expected discounted present value of the rewards when the latter are given by the auxiliary processes is weakly higher than ex-ante expected* *discounted present value of the rewards when the latter are given by the original processes (that is, by V).*

(2) For the virtual index rule χ^I, the ex-ante expected discounted present value of the rewards when the latter are given by the auxiliary processes is the same as ex-ante expected discounted present value of the rewards when the latter are given by the original processes.

The result in the proposition follows from the above two lemmas along with the fact that the myopic policy selecting at each period $t \geq 1$ the pairs with the highest nonnegative auxiliary rewards implements the same matches as the virtual index rule χ^I. Q.E.D.

References

[1] Akbarpour, M., Li, S., and Oveis Gharan, S. (2014), "Dynamic Matching Market Design," mimeo.

[2] Anderson, R., Ashlagi, I., Gamarnik, D., and Kanoria, Y. (2014), "A Dynamic Model of Barter Exchange," mimeo.

[3] Asker, J., and Cantillon, E. (2008), "Properties of Scoring Auctions, " *The RAND Journal of Economics,* 39, pp. 69-85.

[4] Athey, S. and Ellison, G. (2011), "Position Auctions with Consumer Search, " *Quarterly Journal of Economics,* Vol. 126, pp. 1213â€“1270.

[5] Athey, S. and I. Segal (2013), "An Efficient Dynamic Mechanism," *Econometrica*, 81(6), 2463–2485.

[6] Baccara, M., S. Lee, and L. Yariv (2015): "Optimal Dynamic Matching," mimeo

[7] Baron, D. and D. Besanko (1984), "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1, 447–470.

[8] Battaglini, M. (2005), "Long-Term Contracting with Markovian Consumers," *American Economic Review*, 95(3), 637–658.

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

[10] Bergemann, D., and P. Strack (2015), "Dynamic Revenue Maximization: A Continuous Time Approach," *Journal of Economic Theory*, forthcoming.

[11] Bergemann, D. and M. Said (2011), "Dynamic Auctions," *Wiley Encyclopedia of Operations Research and Management Science.*

[12] Bergemann, D. and J. Välimäki (2001), "Stationary Multi-Choice Bandit Problems," *Journal of Economic dynamics and Control,* 25(10), 1585-1594.

[13] Bergemann, D. and J. Välimäki (2008), "Bandit Problems," *New Palgrave Dictionary of Economics*, Durlauf and Blume eds.

[14] Bergemann, D. and J. Välimäki (2010), "The Dynamic Pivot Mechanism," *Econometrica*, 78(2), 771–789.

[15] Besanko, D. (1985), "Multiperiod Contracts Between Principal and Agent with Adverse Selection," *Economics Letters*, 17, 33–37.

[16] Bloch, F. and D. Cantala, (2014), "Dynamic Allocation of Objects to Queuing Agents: The Discrete Model,"mimeo.

[17] Board, S. (2007), "Selling Options," *Journal of Economic Theory* 136, 324–340.

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

[19] Borgers, T. (2015), "An Introduction to the Theory of Mechanism Design," Oxford University Press, Oxford.

[20] Borgers, T., Cox, I., Pesendorfer, M., and Petricek, V. (2013), "Equilibrium Bids in Sponsored Search Auctions: Theory and Evidence," *American Economic Journal: Microeconomics,* Vol. 5, pp. 163-187.

[21] Cabral L. (2011): "Dynamic Price Competition with Network Effects," *Review of Economic Studies* 78 (1): 83-111.

[22] Che, Y.K. (1993): "Design Competition through Multi-Dimensional Auctions," *RAND Journal of Economics* Vol. 24, pp. 668-680.

[23] Cremer, J. and R. P. McLean (1988), "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56(6), 1247–1257.

[24] Courty, P. and H. Li (2000), "Sequential Screening," *Review of Economic Studies*, 67, 697–717.

[25] Damiano, E. and H. Li, (2007), "Price Discrimination and Efficient Matchin," *Economic Theory,* Vol. 30, pp. 243-263.

[26] Damiano, E. and R. Lam, (2005), “Stability in dynamic matching markets,“*Games and Economic Behavior,* 52(1):34â€“53

[27] Doepke, M and R.M. Townsend, (2006) “Dynamic Mechanism Design with Hidden Income and Hidden Actions,” *Journal of Economic Theory* 126, 235 – 285

[28] Doval, L. (2015), “A Theory of Stability in Dynamic Matching Markets,” mimeo.

[29] Edelman, B., Ostrovsky, M., and Schwarz, M. (2007), “Internet Advertising and the Generalized Second-Price Auction,” *American Economic Review*, Vol. 97, pp. 242-259.

[30] Eső, P. and B. Szentes (2007), “Optimal Information Disclosure in Auctions and the Handicap Auction,” *Review of Economic Studies*, 74(3), 705–731.

[31] Eső, P. and B. Szentes (2013), “Dynamic Contracting with Adverse Selection: An Irrelevance Result,” Mimeo, LSE and Oxford University.

[32] Gershkov, A. and B. Moldovanu (2012), “Optimal Search, Learning and Implementation,”*Journal of Economic Theory*, 147(3) (2012): 881-909.

[33] Gershkov, A. and B. Moldovanu (2014), “Dynamic Allocation and Pricing: A Mechanism Design Approach,” MIT Press.

[34] Gomes, R. (2014), “Optimal auction design in two-sided markets,” *The RAND Journal of Economics,* 45: 248â€“272.

[35] Gomes, R. and A. Pavan (2015), “Many-to-Many Matching and Price Discrimination,” *Theoretical Economics*, forthcoming.

[36] Herbst, H., and Benjamin S., (2015), “Dynamic Formation of Teams: When Does Waiting for Good Matches Pay Off?,”mimeo.

[37] Ishikida, T., and Varaiya, P. (1994), “Multi-Armed Bandit Problem Revisited,” *Journal of Optimization Theory and Applications,* 83(1), 113-154.

[38] Johnson, T., (2013),“Matching Through Position Auctions,” *Journal of Economic Theory*, Vol. 148, 1700-1713.

[39] Kakade, S., Lobel, I., and H. Nazerzadeh (2011), “Optimal Dynamic Mechanism Design and the Virtual Pivot Mechanism,” *Operations Research*, 61, 837ñ854.

[40] Kapicka, M. (2013), “Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-Order Approach,” *Review of Economic Studies*, 80(3), 1027-1054.

[41] Kurino, M. (2009) "Credibility, efficiency and stability: a theory of dynamic matching markets," *Jena economic research papers, JENA,*

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

[43] Leshno, J. (2015). "Dynamic matching in overloaded systems, "mimeo.

[44] Makris, M., and A. Pavan (2015), "Incentives for Endogeneous Types: Taxation under Learning by Doing," Discussion paper, Northwestern University.

[45] Mandelbaum, A. (1986), "Discrete Multi-Armed Bandits and Multi-Parameter Processes,"*Probability Theory and Related Fields,* 71(1), 129-147.

[46] Milgrom, P. and I. Segal (2002), "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2), 583–601.

[47] Myerson, R. (1986): "Multistage Games With Communication," *Econometrica,* 54 (2), 323–358.

[48] Pandelis, D. G., and Teneketzis, D. (1999), "On the Optimality of the Gittins Index Rule for Multi-Armed Bandits with Multiple Plays," *Mathematical Methods of Operations Research,* 50(3), 449-461.

[49] Pavan, A., I. Segal, and J. Toikka (2014), "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, 82, 601–653.

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

[51] Schummer, J., (2015), "Influencing waiting lists. Technical report,"mimeo, Kellogg School of Management.

[52] Ünver, U. (2010), "Dynamic Kidney Exchange," *Review of Economic Studies,* 77, 372- 414.

[53] Varian, H. (2007), "Position Auctions," *International Journal of Industrial Organization,* Vol. 25, pp. 1163- 1178.

[54] Weber, R. R., and Weiss, G. (1990), "On an Index Policy for Restless Bandits," *Journal of Applied Probability,* 637-648.

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

[56] Whittle, P. (1982), "Optimization Over Time," John Wiley & Sons, Inc..

[57] Whittle, P. (1988), "Restless Bandits: Activity Allocation in a Changing World," *Journal of applied probability,* 287-298.