informs
http://pubsonline.informs.org/journal/msom

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 24, No. 1, January–February 2022, pp. 125–142
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Dynamic Type Matching

Ming Hu,[a] Yun Zhou[b]

[a] Operations Management and Statistics, Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada;
[b] DeGroote School of Business, McMaster University, Hamilton, Ontario L8S 4L8, Canada

Contact: ming.hu@rotman.utoronto.ca, https://orcid.org/0000-0003-0900-7631 (MH); zhouy185@mcmaster.ca,
https://orcid.org/0000-0001-5905-5995 (YZ)

Received: April 9, 2019
Revised: May 9, 2020; August 26, 2020; September 9, 2020
Accepted: September 10, 2020
Published Online in Articles in Advance: March 18, 2021

https://doi.org/10.1287/msom.2020.0952

Abstract. ***Problem definition***: We consider an intermediary's problem of dynamically matching demand and supply of heterogeneous types in a periodic-review fashion. Specifically, there are two disjoint sets of demand and supply types, and a reward for each possible matching of a demand type and a supply type. In each period, demand and supply of various types arrive in random quantities. The platform decides on the optimal matching policy to maximize the expected total discounted rewards, given that unmatched demand and supply may incur waiting or holding costs, and will be fully or partially carried over to the next period. ***Academic/practical relevance*****:** The problem is crucial to many intermediaries who manage matchings centrally in a sharing economy. ***Methodology*****:** We formulate the problem as a dynamic program. We explore the structural properties of the optimal policy and propose heuristic policies. ***Results*****:** We provide sufficient conditions on matching rewards such that the optimal matching policy follows a priority hierarchy among possible matching pairs. We show that those conditions are satisfied by vertically and unidirectionally horizontally differentiated types, for which quality and distance determine priority, respectively. ***Managerial implications*****:** The priority property simplifies the matching decision within a period, and the trade-off reduces to a choice between matching in the current period and that in the future. Then the optimal matching policy has a match-down-to structure when considering a specific pair of demand and supply types in the priority hierarchy.

Funding: M. Hu was supported by the Natural Sciences and Engineering Research Council of Canada [Grant RGPIN-2015-06757]. Y. Zhou was supported by the Natural Sciences and Engineering Research Council of Canada [Grant RGPIN-2019-07050] and the Social Sciences and Humanities Research Council of Canada [Grant 430-2018-00790].
Supplemental Material: The online supplement and appendices are available at https://doi.org/ 10.1287/msom.2020.0952.

Keywords: dynamic matching • sharing economy • match-down-to policy • horizontal differentiation • vertical differentiation • Monge sequence • assignment problem • priority structure

1. Introduction

We consider a firm that manages the matching of demand and supply in a centralized manner. In each period, demand and supply of various types arrive in random quantities. Each type represents a distinct set of characteristics of demand or supply, and the matching between demand and supply generates type- and time-dependent rewards. The firm determines which types of demand and supply to match, and to what extent, so as to maximize the total expected rewards.

This problem is important to many firms and organizations that make dynamic resource-allocation decisions with randomly arriving supply. For example, it relates to a class of problems where a firm allocates substitutable types of supply (e.g., goods or services) to fulfill customer demand. For a more specific example, let us consider a firm that sells used goods (e.g., used books, cars, or toys). Items of the same product could be heterogeneous with respect to their conditions (which can be considered as heterogeneous supply types). Customers differ in the condition of the item that they prefer or require. In addition, to satisfy customers with an item in their requested condition, the firm may also offer an upgrade (to a superior condition, free, or at a cost) or a downgrade (with a monetary incentive). A similar example is the service upgrading problem for a ride-hailing platform (e.g., Uber, which offers the economical service UberX and premium services such as UberBlack and UberSelect), which uses crowd-sourced, independent drivers (and thus has random supply). In this example, a better vehicle can be used to serve a customer who requested a no-frills service.

Another class of related problems concerns homogeneous goods or services supplied by different providers.

For example, let us consider the transportation of donated food by charities such as food banks. Trucks for Change, a nonprofit logistics broker, connects Food Banks Canada to shippers with spare capacity at a discounted price or for free. In this example, the shippers may differ in their price. Orders to be shipped, on the other hand, may differ in their urgency. Another example is the assignment of workers to customer requests in online labor marketplaces, where both workers and hirers are differentiated on the basis of their ratings. Workers with higher ratings generate higher rewards for themselves (e.g., the online labor market platform Handy offers the cleaning workers four salary rates based on their rating) and for the hirer on average (assuming that a higher rating leads to greater hirer satisfaction on average). Similarly, a hirer with a higher rating may generate better experiences for the worker on average.

In those examples, the firms or organizations need to make the matching and allocation decisions dynamically, with any unused supply and unmet demand either lost or carried over partially or fully to the future. Moreover, the examples share the following features.

First, as mentioned, there are *heterogeneous types* of demand and supply. Matching each pair of demand type and supply type yields a distinct unit matching reward, which can be social surplus or profit. Second, the matching is *centrally controlled* by the firm. Lastly, both demand and supply arrive *randomly* over time and independently of the matching decisions. For example, in the case of used goods, supply is uncertain since it is often procured through donations. Supply for ride-hailing platforms and online labor marketplaces is random because of their reliance on crowdsourced drivers or workers. In the transportation problem for donated food, the availability of participating shippers' spare capacity is random over time.

To model the dynamic matching problem with these features, we will consider a multiperiod stochastic dynamic programming framework. With heterogeneous types of demand and supply arriving randomly in every period, one needs to track the available quantity of demand and supply of each type as a multidimensional state. The optimal decision in each period is in general state-dependent and extremely complex. The problem is therefore challenging. The following provides an overview of the main results and contributions of the paper.

Our paper establishes the modified and strong modified Monge conditions. Those conditions define a partial order over the set of demand-supply pairs and shed light on the order in which the firm should match different pairs in each period. For two special classes of the problem, we show that the conditions lead to simplified optimal matching policies. In particular, in the first class of problems we mentioned earlier (e.g., the upgrading and substitution problems in used-goods inventory management and ride hailing), types are horizontally differentiated, in the sense that the unit reward is higher if one matches demand types with supply types that are closer to the customers' preferences or requests, because upgrading or substitution may incur a cost. In the dynamic capacity management problem studied by Shumsky and Zhang (2009) and Yu et al. (2015), products are horizontally differentiated in a similar way, and our paper further considers random demand and supply arrivals. In the second class of problems mentioned earlier (e.g., the donated-food transportation problem), the unit reward is higher if we use a lower-cost supplier, regardless of the demand types served. A similar example is the vertically differentiated product line in Keskin and Birge (2019) (the authors study a dynamic selling mechanism to decide the line of products to offer in each period). For both classes of problems, we show that the optimal matching policy has a match-down-to structure.

Motivated by the insights and properties of the optimal policy obtained from the modified and strong modified Monge conditions, we propose heuristic policies to compute the optimal matching decisions for the aforementioned two classes of problems. The heuristics follow the match-down-to structure. When matching each pair of demand type and supply type, they determine how much demand and supply to reserve (for potentially more rewarding matching in later periods) based on a threshold level dependent on the imbalance between demand and supply. The simple structure of the heuristic policies connects back to the threshold-type policies in inventory management (e.g., base-stock policies) and quantity-based revenue management (e.g., admission policies based on protection levels).

2. Literature Review

The proposed dynamic matching framework can be viewed as a generalization of two foundations of operations management, that is, inventory management where the firm orders the supply centrally (Zipkin 2000), and revenue management where the firm regulates the demand side with a fixed supply side (Talluri and van Ryzin 2006); and of a combination of the two, that is, joint pricing and inventory control (Chen and Simchi-Levi 2012). Our work is partially motivated by the sharing economy. See, for example, Benjaafar and Hu (2020) and Bernstein et al. (2020) for the literature and latest development on the sharing economy. Unlike the existing work in inventory and revenue management, the supply in the sharing economy is crowdsourced and hence may have uncertainty.

Driven by real-life applications, economists, computer scientists, and operations researchers have studied a variety of two-sided matching problems (see, e.g., Roth and Sotomayor 1990, Abdulkadiroğlu and Sönmez 2013 for a survey), which include the college admissions problem (with the marriage problem as a special case), kidney exchange, and the online bipartite matching problem.

The college admissions problem and the marriage problem are preference-based, and they focus on finding stable matchings in a static and deterministic setting. In those problems, parties on both demand and supply sides submit preferences over options to the matching agency (see, e.g., Ashlagi and Shi 2016). Soliciting preferences may not, however, be practical for day-to-day real-time operations. To handle such situations, we assign a monetary contribution to the matching between a pair of demand and supply types instead of adopting preferences by demand and supply.

In a typical kidney exchange situation, patients and donors arrive in pairs of incompatible patients and donors. Subject to compatibility constraints, researchers have designed efficient matching mechanisms based on cycles or chains of patient-donor pairs to maximize the number of matchings (see, e.g., Roth et al. 2004, 2007 for static problems and Ünver 2010 for a dynamic problem). Our model differs by allowing arbitrary unbalanced arrivals of demand and supply, with the objective of maximizing total reward (e.g., social surplus or profit).

Baccara et al. (2020) consider two demand types and two supply types of vertical differentiation. With a supermodular reward structure, one arrival on both sides of the market in each period and fully backlogged types if unmatched, they show that the optimal matching between the congruent pair of demand and supply is greedy and that the matching between the incongruent pair has a threshold-type structure. In contrast, we consider a more general setting with (i) any number of demand and supply types, (ii) any number of arrivals of any type in each period, (iii) a general reward structure including horizontally and vertically differentiated types as special cases, and (iv) arbitrary carry-over rates for unmatched types. In particular, Proposition 1 in this paper shows an analogous result for two demand types and two supply types of horizontal differentiation.

Online bipartite matching problems have many applications, such as allocation of display advertisements. Initiated by Karp et al. (1990), the classical version considers a bipartite graph $G = (U, V, E)$ and assumes that the vertices in U arrive in an online fashion; that is, only when a vertex $u \in U$ (e.g., a web viewer) arrives are its adjacent edges (e.g., the viewer's interests) revealed. The problem has many variants, all with the focus on maximizing the number of matchings and analyzing the competitive ratios of the algorithms (see Manshadi et al. 2012 for a more recent literature review). The main difference from our model is the online feature. Besides that, there is no explicit notation of inventory, with one side (e.g., advertisers) always there and the other (e.g., impressions) getting lost if not matched. Instead of worst-case analysis, we maximize the expected reward.

Operations researchers have studied two-sided matching by analyzing a stochastic system (e.g., a queueing system) and its mean-field/fluid counterpart. Through a mean-field approach, Arnosti et al. (2021) study a decentralized two-sided matching market and show that limiting the visibility of applicants can significantly improve the social surplus. With a fluid approach, Zenios et al. (2000) and Su and Zenios (2006) study the efficiency-equity trade-off in kidney allocations and Akan et al. (2012) the efficiency-urgency trade-off in liver allocations. Using double-sided queues, Zenios (1999) studies the transplant waiting list and Afèche et al. (2014) investigate trading systems of crossing networks. Su and Zenios (2004) analyze a queueing model with service discipline first come first served (FCFS) or last come first served (LCFS) to examine the role of patient choices in the kidney transplant waiting system. Adan and Weiss (2012) show that the stationary distribution of FCFS matching rates for two infinite multitype sequences is of product form. Gurvich and Ward (2014) study the dynamic control of matching queues with the objective of minimizing holding costs. Focusing on the fluid approximation and its asymptotic optimality, the authors observe that, in principle, the controller may choose to wait until some inventory of items builds up to facilitate more rewardable matches in the future. Kanoria and Saban (2017) study a dynamic fluid matching model in which agents on one side receive proposals from those on the other side and decide whether to pay screening costs to discover the value of the proposing agent. In contrast to these papers, we focus on a stochastic model (versus the fluid counterpart) and optimal decision making (versus performance evaluation).

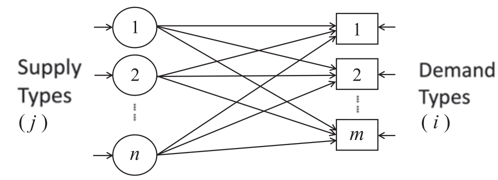
Our dynamic matching framework generalizes the classical transportation problem. Monge (1781) observes the key idea underlying the Monge properties that ensure the optimality of a greedy solution. Hoffman (1963) formalizes the idea and provides a rigorous proof. Monge's work has led to extensive studies on the optimal transport problem by mathematicians, economists, and computer scientists, and is also generalized to multi-index transportation problems (see, e.g., Queyranne et al. 1998). Burkard (2007) provides a detailed review on the development and applications of the Monge condition to optimization problems. Recently, Estes and Ball (2020) consider a dynamic transportation problem in which new quantities join the supply and demand nodes over time. The paper provides conditions for the greedy solution to be oracle (i.e., distribution-free) optimal and conditions under which an existing policy can be improved. Unlike Estes and Ball (2020), our paper aims to maximize the expected discounted reward, and the modified Monge condition we define leads to the optimality of the match-down-to-threshold-type policies under certain reward structures (e.g., the horizontal model and the vertical model), for which we further propose methods to compute the threshold levels approximately. Our model is also more general in allowing unmatched demand or supply to be lost or partially carried over to the future.

With a multidimensional stochastic dynamic programming formulation, our work is also related to the literature on approximate dynamic programming (ADP). It is well-known that high-dimensional stochastic dynamic programs are difficult to solve, and researchers have developed various ADP techniques to compute the optimal policy. These include rollout algorithms (see, e.g., Bertsekas 2013), and limited lookahead policies and value function approximations (see, e.g., Powell 2016). For recent applications of ADP in operations management, we refer readers to Ke et al. (2019) and the references therein. In contrast, we focus on exploring structural properties of the specific problems to develop matching policies with an intuitive and easy-to-implement structure. We also propose computational methods based on the structural properties of the optimal policy. In the horizontal model, we match demand and supply pairs in the order indicated by the structure of the optimal policy, and match each pair down to a threshold based on the greedy approximation of the expected future reward. In the vertical model, we resort to the one-step-lookahead policy for computation, which is essentially a rollout strategy that improves upon greedy matching, and we show that it retains the top-down structure of the optimal policy.

3. The Model with General Matching Rewards

Consider a planning horizon of T periods. At the beginning of each period, m types of demand (indexed by demand type $i = 1, \ldots, m$) and n types of supply (indexed by supply type $j = 1, \ldots, n$) arrive in random quantities. The pairs of demand and supply are illustrated in Figure 1 as a bipartite graph. For a demand type i and a supply type j, we denote by (i, j) a matchable pair of demand and supply types, which corresponds to an arc in the graph. Without loss of generality, suppose that any demand type is matchable with any supply type, with possibly different rewards. (We may set the unit reward to be an arbitrarily large negative value if a pair is unmatchable.)

Figure 1. Pairs of Demand and Supply



In the beginning of each period t, a random quantity D_i^t of type i demand and a random quantity S_j^t of type j supply arrive. We write the new arrivals of demand and supply in vector form, as $\mathbf{D}^t = (D_1^t, \ldots, D_m^t)$ and $\mathbf{S}^t = (S_1^t, \ldots, S_m^t)$, respectively. The distributions of supply and demand in a period can be exogenously correlated with those in another period, but are independent of any matching decisions. Indeed, for the used-goods inventory management problem with product upgrading and the transportation problem of donated goods that we mentioned earlier, the matching decision does not seem to directly affect future demand or supply. For the ride-hailing problem with service upgrade, the matching decision may affect future supply at the place where the driver drops off the customer. Nevertheless, such endogeneity may be minimal if we focus on a reasonably small region and a short time horizon. Given the high matching volumes in a ride-hailing system, the review period should be very short (e.g., a few seconds). Thus, the matched driver may not be able to complete the current trip until many periods later, and the new supply arrivals correspond to those drivers who just log in or roam into the region. For other problems such as the assignment of workers by online labor market platforms, supply availability could be more affected by long-term factors (e.g., payment mechanisms) than short-term matching decisions. We allow the current matching and future supply to be correlated in an extension in online supplement C.3. (See Hu and Zhou 2020 for an unabridged memo which includes all the online supplements to this paper.)

The state for a given period t comprises the demand and supply levels of various types before matching but after the arrival of demand and supply. We denote, as the system state, the demand vector by $\mathbf{x} = (x_1, \ldots, x_m)$ and the supply vector by $\mathbf{y} = (y_1, \ldots, y_n)$, where x_i and y_j are the quantities of type i demand and type j supply available to be matched. We suppose either that all state variables and realizations of demand and supply are continuous-valued or that all of them are discrete-valued. On observing the state $(\mathbf{x}, \mathbf{y})$, the firm decides on the quantity q_{ij} of type i demand to be matched with type j supply, for any

$i = 1, \ldots, m$ and $j = 1, \ldots, n$. For conciseness, we write the decision variables of matching quantities in a matrix form as $\mathbf{Q} = (q_{ij}) \in \mathbb{R}_+^{m \times n}$. There is a reward r_{ij}^t for matching a unit of type i demand and a unit of type j supply for all i, j. We write the rewards in a matrix form as $\mathbf{R}^t = (r_{ij}^t) \in \mathbb{R}^{m \times n}$. The total matching reward is linear in the matching quantities; that is, $\mathbf{R}^t \circ \mathbf{Q} \equiv \sum_{i=1}^m \sum_{j=1}^n r_{ij}^t q_{ij}$, where $\circ$ is the sum of elements of the entrywise product of two matrices. We allow the unit reward r_{ij}^t to be nonhomogeneous in time.

At the end of period t, the level of type i demand is reduced by the quantity $\sum_{j=1}^n q_{ij}$ (i.e., the total quantity by which type i demand is matched with all supply types), which leads to its postmatching level $u_i := x_i - \sum_{j=1}^n q_{ij}$. Similarly, type j supply will be reduced to $v_j = y_j - \sum_{i=1}^m q_{ij}$. A fraction α of the unmatched demand and a fraction β of the unmatched supply carry over to the next period, where $0 \le \alpha, \beta \le 1$ are exogenously given. In other words, $1 - \alpha$ fraction of demand and $1 - \beta$ fraction of supply leave the system with a zero reward. We allow α and β to take any value between 0 and 1 if the state variables (and demand and supply realizations) are continuous-valued, and we require them to be binary (i.e., 0 or 1) if the state variables are discrete. Therefore, the state for the next period $t+1$ comprises the type i demand level $x_i^{t+1} = \alpha(x_i - \sum_{j=1}^n q_{ij}) + D_i^{t+1}$ (which includes the random demand D_i^{t+1} of type i that joins in the beginning of period $t+1$) and the type j supply level $y_j^{t+1} = \beta(y_j - \sum_{i=1}^m q_{ij}) + S_j^{t+1}$ (which includes the random supply S_j^{t+1}). For ease of notation, we define $\mathbf{1}^k$ as the k-dimension row vector with all its entries equal to one, and may omit the superscript k and infer the length of the vector from the context. We write $\mathbf{u} := (u_1, \ldots, u_m) := \mathbf{x} - \mathbf{1}\mathbf{Q}^{\mathrm{T}}$, $\mathbf{v} := (v_1, \ldots, v_m) := \mathbf{y} - \mathbf{1}\mathbf{Q}$, $\mathbf{x}^{t+1} := \alpha\mathbf{u} + \mathbf{D}^{t+1}$, and $\mathbf{y}^{t+1} := \beta\mathbf{v} + \mathbf{S}^{t+1}$.

The firm's goal is to determine the matching quantities $\mathbf{Q}^* = (q_{ij}^*)$ in each period t to maximize the expected total reward from the current and remaining periods. Let $V_t(\mathbf{x}, \mathbf{y})$ be the optimal expected total reward given that it is in period t and the current state is $(\mathbf{x}, \mathbf{y})$. We formulate the problem as the following stochastic dynamic program. For $t = 1, \ldots, T$,

$$\begin{aligned} V_t(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{Q}} \quad & H_t(\mathbf{Q}, \mathbf{x}, \mathbf{y}) := \mathbf{R}^t \circ \mathbf{Q} \\ & + \gamma E V_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1}), \\ \text{s.t.} \quad & \mathbf{u} = \mathbf{x} - \mathbf{1}\mathbf{Q}^{\mathrm{T}}, \quad \mathbf{v} = \mathbf{y} - \mathbf{1}\mathbf{Q}, \\ & \mathbf{u} \ge \mathbf{0}, \quad \mathbf{v} \ge \mathbf{0}, \end{aligned} \tag{1}$$

where $\gamma \in [0, 1]$ is a discount factor. We assume $V_{T+1}(\mathbf{x}, \mathbf{y}) \equiv 0$ and $r_{ij}^{T+1} \equiv 0$ for all i and j; that is, at the end of the horizon, all unmatched demand and supply leave the system with a zero reward.

A matching policy $\{\mathbf{Q}^t(\mathbf{x}, \mathbf{y})\}_{t=1,\ldots,T}$ specifies the matching quantity matrix $\mathbf{Q}^t$ in each period t for any given state $(\mathbf{x}, \mathbf{y})$. An optimal matching policy for the problem (1) always exists, and it determines the matching quantities between the $m \times n$ pairs of demand and supply types in each period. We remark that there can be multiple (sometimes, even infinitely many) optimal matching decisions in a given period. We will focus on exploring optimal policies that have a match-down-to structure according to a priority hierarchy if such a policy exists.

We refer readers to online Appendix C for the notation used in the paper. We now present a simple numerical example that motivates subsequent analysis.

Example 1. Consider a two-period problem with two demand types and two supply types. In period 1, the available demand is $\mathbf{x} = (x_1, x_2) = (3, 4)$, and the available supply is $\mathbf{y} = (y_1, y_2) = (4, 3)$. Demand and supply of all types to arrive in period 2 are independent of each other, with the following probability distributions: Both type 1 demand and type 2 supply in period 2 follow the two-point distribution that is equal to either 0 or 2 with equal probabilities; both type 2 demand and type 1 supply in period 2 follow the two-point distribution that takes either the value of 0 or 3 with equal probabilities. Both carry-over rates α and β are equal to 1, and the discount factor is $\gamma = 0.9$. The unit matching rewards are given in Table 1, where r_{11}^1 takes one of the five given values. As r_{11}^1 varies, we obtain the optimal matching quantities $\mathbf{Q}^{1*} = (q_{11}^{1*}, q_{12}^{1*}, q_{21}^{1*}, q_{22}^{1*})$ in period 1.

For any of the values of r_{11}^1 we use, $r_{11}^t > r_{12}^t > r_{21}^t > r_{22}^t$ for $t = 1, 2$. Without loss of generality, we can always consider the optimal policy to follow the descending order of unit rewards to match the demand-supply pairs, that is, along the sequence of pairs $(1,1) \to (1,2) \to (2,1) \to (2,2)$. (But one does not need to match a pair to the maximum extent before proceeding to the next pair.)

Table 1. The Unit Matching Rewards and Optimal Matching Quantities for Example 1

Panel A: The unit matching rewards

Period	r_{11}^t	r_{12}^t	r_{21}^t	r_{22}^t
$t = 1$	85, 90, 95, 100, or 105	60	50	5
$t = 2$	85	60	50	5

Panel B: The optimal quantities in period 1 for different values of r_{11}^1

r_{11}^1	q_{11}^{1*}	q_{12}^{1*}	q_{21}^{1*}	q_{22}^{1*}
85	0	3	4	0
90	1	2	3	0
95	2	1	2	0
100	3	0	1	0
105	3	0	1	0

When r_{11}^1 is small enough, that is, $r_{11}^1 = 85, 90, 95$, in period 1, along the sequence the optimal policy may proceed to the next pair without matching the previous pair to the maximum extent. For example, if $r_{11}^1 = 90$, it does not match the pair $(1,1)$ greedily, but just for one unit, after which it will match the pair $(1,2)$ for two units and the pair $(2,1)$ for three units, even though the unit reward r_{11}^1 is higher than the unit reward from any other pair. This is because the combination of the two pairs $(1,2)$ and $(2,1)$ yields a considerably higher total reward (i.e., $r_{12}^1 + r_{21}^1 = 110$) than r_{11}^1, which may lead the optimal policy to match $(1,2)$ and $(2,1)$ simultaneously before fully matching $(1,1)$. On the other hand, although the combination yields a higher total reward, it also consumes more resources, which may reduce the expected reward in period 2. Compared with matching $(1,1)$ for a unit alone, the simultaneous matching of $(1,2)$ and $(2,1)$ (for one unit each) consumes an extra unit of both type 2 demand and type 2 supply. (This reduces the reward in period 2, for example, when there is type 1 demand to be fulfilled in period 2 but insufficient supply to match with them, partly due to the consumption of type 2 supply by the combination in the previous period. In that case, the reward in period 2 could be increased by $r_{12}^2 = 60$, if we had saved a unit of type 2 supply in period 1.) Thus, the optimal policy needs to balance between matching $(1,1)$ and matching the combination of $(1,2)$ and $(2,1)$. In keeping with this intuition, as r_{11}^1 increases, we observe that the matching quantity for $(1,1)$ increases (weakly), while the quantities for $(1,2)$ and $(2,1)$ both decrease.

When r_{11}^1 is sufficiently large, that is, $r_{11}^1 = 100, 105$, along the sequence in the descending order of unit rewards, in period 1 the optimal matching decision has the following match-down-to structure: it either proceeds to the next pair after greedily matching the current pair or stops after partially matching the current pair. It first matches $(1,1)$ greedily for three units (i.e., until type 1 demand runs out), and then $(1,2)$ for zero units (since type 1 demand is no longer available), and finally stops the matching procedure after matching $(2,1)$ for one unit without matching $(2,2)$. We note that although the combination of $(1,2)$ and $(2,1)$ still yields a higher total unit reward than $(1,1)$, the optimal policy chooses to prioritize $(1,1)$ over the combination for $r_{11}^1 = 100, 105$.

Example 1 demonstrates the complexity of trading off among various pairs of demand and supply types. We observe from the example that, to determine the extent to which we should match each pair, one needs to compare not just individual pairs of demand and supply by their unit rewards, but also an individual pair with a combination of two other pairs intertemporally. For a problem with more types of demand and supply than in Example 1, it becomes even more challenging to characterize the optimal policy, as one may need to compare combinations of multiple pairs. In Example 1, the optimal policy has a match-down-to structure when r_{11}^1 is sufficiently large. To generalize this observation, we will explore conditions under which the optimal policy has a similar structure (i.e., following a certain order of pairs to match demand with supply).

3.1. The Monge Sequence and the Single-Period Problem

Consider the single-period problem with available demand $\mathbf{x} = (x_1, \ldots, x_m)$ and supply $\mathbf{y} = (y_1, \ldots, y_n)$. If demand and supply are balanced (i.e., $\sum_{i=1}^m x_i = \sum_{j=1}^n y_j$), Hoffman (1963) shows that it is optimal to greedily match demand with supply along the Monge sequence if such a sequence exists. The original definition of the Monge sequence is based on the unit transportation costs from supply locations to demand locations. We restate its definition by using the unit reward r_{ij} for matching type i demand with type j supply in the single period ($i = 1, \ldots, m, j = 1, \ldots, n$).

Definition 1 (Hoffman 1963)**.** A Monge sequence is a sequence of all demand-supply pairs in the set $\{(i,j) \mid i = 1, \ldots, m, j = 1, \ldots, n\}$ such that, if a pair (i,j) is ordered before both (i',j) and (i,j') in the sequence, then the unit matching rewards satisfy the condition $r_{ij} + r_{i'j'} \geq r_{i'j} + r_{ij'}$.

3.2. The Modified Monge Condition

For the dynamic problem, we define the modified Monge condition to rank demand-supply pairs.

Definition 2 (Modified Monge Condition)**.** For two neighboring pairs (i,j) and (i',j) of demand and supply types, we say that (i,j) weakly precedes a neighboring pair (i',j) if $r_{ij}^t - r_{i'j}^t \geq \alpha\gamma \max_{j''=1,\ldots,n}(r_{ij''}^{t+1} - r_{i'j''}^{t+1})^+$ for any period $t = 1, \ldots, T$. Likewise, we say that (i,j) weakly precedes (i,j') if for any period t, $r_{ij}^t - r_{ij'}^t \geq \beta\gamma \max_{i''=1,\ldots,m}(r_{i''j}^{t+1} - r_{i''j'}^{t+1})^+$.

For a pair (i,j) to weakly precede its neighboring pair (i',j), the previous condition requires the former to generate a sufficiently higher unit reward than the latter in every period t. More specifically, the unit reward r_{ij}^t should exceed $r_{i'j}^t$ by a minimum margin, which depends on the difference in the unit rewards involving demand types i and i' in the next period.

The Monge sequence (Definition 1) specifies a complete order of all pairs of demand and supply. In contrast, the modified Monge condition defines a partial order over the demand-supply pairs (see online supplement C.2 for a formal proof) and compares two pairs at a time. However, the condition does not require that any two pairs of demand and supply types be comparable. In the following theorem, we show that there exists an optimal policy that respects this partial order.

Theorem 1. *There exists an optimal policy $\pi^* = \{\mathbf{Q}^{t*}\}_{t=1,\ldots,T}$ such that if any pair (i,j) weakly precedes a neighboring pair (i',j) (respectively, (i,j')), in each period t either the matching quantity $q^{t*}_{i'j} = 0$ (respectively, $q^{t*}_{ij'} = 0$) or the postmatching level $u^{t*}_i = 0$ (respectively, $v^{t*}_j = 0$).*

Theorem 1 shows that if (i,j) weakly precedes (i',j), matching of (i',j) is possible only if there is no remaining type i demand to compete for type j supply. To see the underlying intuition, we compare two options. Option 1 is to match (i,j) for one unit in period t, and option 2 is to match (i',j) for a unit instead. In period t, we receive a higher reward under option 1 by an amount of $r^t_{ij} - r^t_{i'j}$. However, option 2 leads to α more units of type i demand and α fewer units of type i' demand in period $t+1$. If those α units of type i demand are matched with some supply type j'' in period $t+1$, under option 1 we may instead match α units of type i' demand with type j'' supply in period $t+1$. In this way, the difference in the rewards received in period 2 between options 1 and 2 is $-\alpha r^{t+1}_{ij''} + \alpha r^{t+1}_{i'j''}$, and the difference in the total discounted rewards over both periods is $r^t_{ij} - r^t_{i'j} - \gamma\alpha(r^{t+1}_{ij''} - r^{t+1}_{i'j''}) \geq 0$. This implies that (i,j) should be assigned a higher matching priority than (i,j'), in the sense that we will not regret having used option 1 instead of option 2.

Remark 1. In online supplement C.1, we extend Definition 2 to allow a pair to weakly precede a nonneighboring pair, and we show that an optimal policy does not match any pair (i',j') weakly preceded by another pair (i,j) if both type i demand and type j supply are still available.

Interestingly, although the weak modified Monge condition implies higher matching priority for one pair over another pair it weakly precedes, the combination of two weakly preceded pairs can possibly preempt the matching of a preceding pair, as we demonstrate in the following example.

Example 2. Consider a two-period problem with two demand types and two supply types, and the following unit rewards: in period 1, $r^1_{11} = 8$, $r^1_{12} = 4$, $r^1_{21} = 5$, and $r^1_{22} = 1$; in period 2, $r^2_{11} = 8$, $r^2_{12} = 5$, $r^2_{21} = 6$, and $r^2_{22} = 2.5$. The carry-over rates and the discount factors are $\alpha = \beta = \gamma = 1$. We can verify that the pair $(1,1)$ weakly precedes both $(1,2)$ and $(2,1)$. However, in period 2, the optimal policy may try to match $(1,2)$ and $(2,1)$ simultaneously before matching the preceding pair $(1,1)$. For example, if in period 2 the available demand and supply are $x^2_1 = x^2_2 = y^2_1 = y^2_2 = 1$, the optimal policy matches both pairs $(1,2)$ and $(2,1)$ for one unit but does not match $(1,1)$ or $(2,2)$, since $r^2_{12} + r^2_{21} = 11 > r^2_{11} + r^2_{22}$.

Next, suppose that in period 1 the available demand and supply are $x^1_1 = 1$, $x^1_2 = 0$, $y^1_1 = 1$, and $y^1_1 = 0$, and the firm anticipates demand and supply arrival in period 2 as $D^2_1 = 0$, $D^2_2 = 1$, $S^2_1 = 0$, and $S^2_2 = 1$ with probability 1. The optimal policy does not match any pair in period 1 in order to save all type 1 demand and type 1 supply for matching the pairs $(1,2)$ and $(2,1)$ both for one unit in period 2. This results in a total reward $r^2_{12} + r^2_{21} = 11$ over the two periods. Alternatively, if we match the pair $(1,1)$ for a unit in period 1, in period 2 we can only match the pair $(2,2)$ for one unit, resulting in a lower total reward $r^1_{11} + r^2_{22} = 10.5$.

Example 2 shows that the optimal policy may let two weakly preceded pairs preempt a preceding pair in a period, and it may also reserve a pair of demand and supply types to match two pairs it weakly precedes in a later period. Assumption 1 rules out the former possibility and will ensure the optimality of the match-down-to policy for two special reward structures.

Assumption 1 (Strong Modified Monge Condition)**.** *In period t, for any pair (i,j), if it weakly precedes two neighboring pairs (i',j) and (i,j'), then $r^t_{ij} + r^t_{i'j'} \geq r^t_{ij'} + r^t_{i'j}$.*

The strong modified Monge condition resembles the Monge sequence. The inequality condition in Definition 1 is based on the given complete order (i.e., the given sequence), whereas the inequality condition in Assumption 1 is based on the partial order defined by the modified Monge condition. In that sense, the strong modified Monge condition is a weaker version of the Monge sequence.

Theorem 2. *Suppose Assumption 1 holds in a subset of periods $\tau \in \mathcal{T} \subseteq \{1,\ldots,T\}$. There is an optimal policy $\pi^* = \{\mathbf{Q}^{t*}\}_{t=1,\ldots,T}$ such that, in addition to the properties in Theorem 1, it also satisfies the following property: If a pair (i,j) weakly precedes both (i',j) and (i,j'), either $q^{\tau*}_{ij'} = 0$ or $q^{\tau*}_{i'j} = 0$ in any period $\tau \in \mathcal{T}$.*

Theorem 2 is intuitive. If a pair (i,j) weakly precedes any two neighboring pairs (i',j) and (i,j'), and both $q^{\tau*}_{ij'}$ and $q^{\tau*}_{i'j}$ are positive, then according to Assumption 1, we can simply reduce $q^{\tau*}_{ij'}$ and $q^{\tau*}_{i'j}$ by a small amount and increase the matching quantities for (i,j) and (i',j') by the same amount, to increase the reward in period τ without affecting the postmatching levels.

Without loss of generality, we can consider any policy to follow the partial order defined by the modified Monge condition (without needing to match a pair to the maximum extent before moving on to subsequent pairs), and simultaneously match a pair and those not comparable with it (referred to as its parallel pairs). Theorems 1 and 2 further imply that a pair (i,j) must be matched to the maximum extent before the optimal policy can match any (neighboring) pairs it weakly precedes. In fact, if the optimal policy does not match (i, j) to the maximum extent but proceeds to match subsequent pairs it weakly precedes, then either there is remaining type i demand or type j supply at the end of the period, or both i and j are further consumed by pairs weakly preceded by (i, j). However, those possibilities are at odds with the properties in Theorem 1 or Theorem 2.

As long as Assumption 1 holds in a certain period τ, the optimal policy has the priority structure mentioned previously in that period. Following the priority structure, it may not match demand and supply greedily. Instead, it may match a pair (i, j) only partially in period τ (in which case any pair weakly preceded by (i, j) is not matched in period τ) due to at least one of the following reasons:

a. Reserve some type i demand (respectively, type j supply) to match a pair (i, j') (respectively, (i', j)) that weakly precedes (i, j) in a later period.

b. Reserve both type i demand and type j supply to match two pairs (i, j') and (i', j) that are weakly preceded by (i, j) but jointly yield a higher reward in a later period.

c. Strategically delay the matching between i and j and match the same pair in a later period, if the unit reward between i and j is higher in that later period.

In Example 2, Assumption 1 holds in period 1 only, and we see that the optimal policy may withhold the matching of the pair $(1, 1)$ in period 1 due to (b). However, if Assumption 1 holds in all periods, the firm will not have the incentive (b) to withhold demand or supply. Indeed, by Theorem 2, if Assumption 1 also holds in period $t + 1$, any type i demand and type j supply reserved in period t should still be matched in period $t + 1$ first before the optimal policy simultaneously matches any two pairs (i, j') and (i', j) that are weakly preceded by (i, j). In Theorem 3, greedy matching of a pair becomes optimal when we further eliminate incentives (a) and (c).

Theorem 3. *Suppose that Assumption 1 holds in all periods. There exists an optimal policy $\pi^* = \{\mathbf{Q}^{t*}\}$ such that it satisfies the properties in Theorems 1 and 2 and also the following property: For any pair (i, j), if it weakly precedes all its neighboring pairs and the unit reward r_{ij}^t satisfies the condition that $r_{ij}^t \geq \gamma \max\{\alpha, \beta\} r_{ij}^{t+1}$ (for $t = 1, \ldots, T-1$), then the optimal policy matches type i demand with type j supply as much as possible; that is, $q_{ij}^{t*} = \min\{x_i, y_j\}$ for all $t = 1, \ldots, T$.*

The condition of (i, j) weakly preceding all neighboring pairs eliminates the incentive in (a), Assumption 1 holding in all periods eliminates (b), and the condition $r_{ij}^t \geq \gamma \max\{\alpha, \beta\} r_{ij}^{t+1}$ eliminates (c). We say that (i, j) is a perfect pair if it satisfies all the conditions in Theorem 3.

Moreover, if Assumption 1 is satisfied in a period t, then, for two special cases of the problem, Theorems 1 and 2 imply the optimality of the match-down-to structure in period t.

Special Case 1. *Any* two neighboring pairs are comparable by the modified Monge condition, that is, one weakly precedes the other (without loss of generality, assume no mutual precedence; see online supplement C.2). We then organize the pairs into levels. Level 1 pairs are not weakly preceded by any other pair, and inductively, level ℓ pairs are those preceded only by pairs of levels $1, \ldots, \ell - 1$. Then, the optimal policy follows the descending order of priority (i.e., the ascending order of level indices) and matches the parallel pairs of the same level simultaneously. In any period, unless it fully matches a pair (i, j) of level ℓ, no lower-priority pair weakly preceded by (i, j) is matched.

Special Case 2. Suppose that there is only one level 1 pair in special case 1, and we assume the following properties: (i) If we remove either the demand type or the supply type of the level 1 pair as a node from the bipartite graph, in the remaining graph there is just one pair of the highest (priority) level; and (ii) successively, each time we remove either the demand or supply type of the highest-level pair from the remaining graph, there is just one pair of the highest level left, regardless of the types previously removed. Then, the optimal policy has a match-down-to structure (see the proof in online supplement E). When the optimal policy matches the highest-level pair, it either matches it greedily, or matches it halfway to some level and stops the matching procedure. In the former case, either demand or supply of that pair is exhausted; we remove the exhausted type, and move on to the new highest-level pair in the remaining graph.

More general problems than the two special cases need not satisfy the (strong) modified condition, and the match-down-to policy may not be optimal. Nevertheless, one can enforce such a policy along a certain priority sequence or structure as a heuristic, as in Example 3.

Example 3. We revisit Example 1 and focus on the matching in period 1. We can readily verify that both pairs $(1, 2)$ and $(2, 1)$ weakly precede $(2, 2)$ for all values of r_{11}^1 we adopt (i.e., $r_{11}^1 = 85, 90, 95, 100, 105$). The pair $(1, 1)$ weakly precedes $(1, 2)$ only for $r_{11}^1 = 105$, and it weakly precedes $(2, 1)$ only for $r_{11}^1 = 100, 105$. Moreover, the strong modified Monge condition (i.e., Assumption 1) is satisfied in period 1 when $r_{11}^1 = 105$. Theorems 1 and 2 ensure that the match-down-to structure along the sequence $(1, 1) \to (1, 2) \to (2, 1) \to (2, 2)$ is optimal when $r_{11}^1 = 105$. The computational results in Example 1 show that this structure is also optimal for $r_{11}^1 = 100$, but not for $r_{11}^1 = 85, 90, 95$. Nevertheless, we enforce that structure along the sequence $(1, 1) \to (1, 2) \to (2, 1) \to (2, 2)$ for all $r_{11}^1 = 85, 90, 95, 100, 105$, and compare the corresponding expected discounted reward over the two periods

Table 2. The Reward Under the Best Match-Down-To Policy vs. the Optimal Expected Discounted Reward in Example 3

r_{11}^1	85	90	95	100	105
Reward under the best match-down-to policy	438.75	450.5	465.5	480.5	495.5
Optimal reward	457.63	459.56	468.69	480.5	495.5

with the optimal expected total discounted reward, as in Table 2.

As r_{11}^1 increases, the modified Monge condition is violated to a lesser extent (e.g., for $(1,1)$ to weakly precede $(1,2)$, we need $r_{11}^1 - r_{12}^1 \geq \gamma \max_{i=1,2}(r_{i1}^2 - r_{i2}^2)^+$; since the right side remains constant, the condition is violated to a lesser extent as r_{11}^1 increases), and so too is a condition similar to the strong modified Monge condition, $r_{11}^1 + r_{22}^1 \geq r_{12}^1 + r_{21}^1$. (Note that we enforce the match-down-to structure along the sequence $(1,1) \to (1,2) \to (2,1) \to (2,2)$; the latter condition ensures that the combination of $(1,2)$ and $(2,1)$ does not preempt the matching of the pair $(1,1)$, which we assign the highest priority.) We observe that the gap between the optimal reward and the reward achieved by the match-down-to structure shrinks as r_{11}^1 increases.

In Example 3, the performance of the match-down-to policy (relative to the optimal policy) improves as the modified Monge condition (and a condition similar to the strong modified Monge condition) is violated to a lesser extent along the sequence. This seems to suggest that the match-down-to structure along a priority sequence or hierarchy may work well if the conditions hold approximately and that the performance loss may be quantified by the degree to which the conditions are violated. We leave further development of those observations for future research and conclude this section with a remark on incorporating waiting and holding costs.

Remark 2. Suppose any unmatched demand incurs a unit waiting cost w, and any unmatched supply incurs a unit holding cost h. If a demand (respectively, supply) unit is never matched from period t to period T, the total discounted waiting (respectively, holding) cost is $w\sum_{\tau=t}^{T}(\gamma\alpha)^{\tau-t}$ (respectively, $h\sum_{\tau=t}^{T}(\gamma\beta)^{\tau-t}$). We define the adjusted unit reward $r_{ij}^{t,\mathrm{adj}}$ as the sum of the unit matching reward r_{ij}^t and the waiting costs saved for i and holding costs saved for j by matching them in the current period t, that is, $r_{ij}^{t,\mathrm{adj}} = r_{ij}^t + w\sum_{\tau=t}^{T}(\gamma\alpha)^{\tau-t} + h\sum_{\tau=t}^{T}(\gamma\beta)^{\tau-t}$. One can verify that if any demand-supply pairs satisfy the (strong) modified condition and/or the inequality condition in Theorem 3 for the unit rewards $\{r_{ij}^t\}_{\forall i,j,t}$, they also satisfy the same conditions for the adjusted unit rewards $r_{ij}^{t,\mathrm{adj}}$. Thus, all the results in this section remain true.

4. Horizontally Differentiated Types

We consider horizontally differentiated demand and supply types that have idiosyncratic preferences over the types on the other side. From the firm's perspective, matching a demand (supply) type with a supply (demand) type closer to its preference yields a higher unit reward in all periods.

4.1. Two Demand Types and Two Supply Types

We first study the model with two demand types and two supply types under Assumption 2. Type 1 (type 2) supply is closer to the preference of type 1 (type 2) demand, and vice versa.

Assumption 2. *Suppose the following conditions hold.*

i. *For any period* t, $r_{12}^t \leq \min\{r_{11}^t, r_{22}^t\}$ *and* $r_{21}^t \leq \min\{r_{11}^t, r_{22}^t\}$.

ii. *The differences in the unit rewards satisfy the conditions* $r_{11}^t - r_{12}^t \geq \beta\gamma(r_{11}^{t+1} - r_{12}^{t+1})$, $r_{11}^t - r_{21}^t \geq \alpha\gamma(r_{11}^{t+1} - r_{21}^{t+1})$, $r_{22}^t - r_{21}^t \geq \beta\gamma(r_{22}^{t+1} - r_{21}^{t+1})$, *and* $r_{22}^t - r_{12}^t \geq \alpha\gamma(r_{22}^{t+1} - r_{12}^{t+1})$, *for* $1 \leq t \leq T-1$.

iii. *For* $i,j = 1,2, 1 \leq t \leq T-1$, $r_{ij}^t \geq \max\gamma\{\alpha,\beta\}r_{ij}^{t+1}$.

Condition (i) of the assumption implies that the demand type and the supply type of the same index are closer to each other's preference and their matching yields a higher unit reward than the matching between a pair of types of different indices. Condition (ii) requires that the reward difference (adjusted by the discount factor and the carry-over rates) between the matching with the closer type and with the farther type decreases in time. It is satisfied, for example, if the unit rewards are stationary over time. Condition (iii) ensures that the firm has no incentive to strategically delay the matching between a pair (in order to match the same pair in a future period). Next, we present an example that satisfies Assumption 2.

Example 4 (Used-Goods Inventory Management with Substitution). Consider the used-goods inventory management problem with items in one of two conditions, say, conditions 1 and 2, with condition 1 being superior to condition 2 (e.g., condition 1 could be the latest edition of a book, and condition 2 an older edition). Let the price and cost for an item in condition i be f_i and c_i, respectively ($i = 1,2$). The two demand types, type 1 and type 2, seek to buy items in condition 1 and condition 2, respectively. Other than assigning an item to its intended demand type, the firm can also offer type 2 demand an upgrade to an item in condition 1 with a price discount c_d, and type 1 demand a downward substitution to an item in condition 2 with compensation c_b. (Customers may differ in the amount of compensation they would demand to accept the downward substitution; in that case, c_b represents the average compensation that a customer would demand.) Let us define the stationary unit matching reward r_{ij} as the profit margin by fulfilling one unit of type i demand with an item in condition j. Thus, $r_{ii} = f_i - c_i$ for $i = 1, 2$, $r_{12} = f_2 - c_2 - c_b$ and $r_{21} = f_1 - c_1 - c_d$. We can verify that Assumption 2 is satisfied if $f_2 - c_2 - c_b \leq f_1 - c_1 \leq f_2 - c_2 + c_d$, which requires (i) the profit margin of an item in the superior condition to be sufficiently high compared with the margin of an item in the inferior condition (provided that both are assigned to the intended demand type), and (ii) the price discount is steep enough for type 2 demand (which seeks inferior but cheaper items) to be willing to buy an item in the superior condition. In particular, the two inequalities are satisfied if the superior condition item is more profitable than the inferior condition item (if assigned to intended customers) and all upgrades are free. Whereas we do not explicitly consider the waiting or holding costs of unmatched demand and supply, they can be easily incorporated into the unit rewards. (See Remark 2.)

We can now characterize the optimal matching policy under Assumption 2. It can be readily verified that the pairs $(1, 1)$ and $(2, 2)$ both weakly precede $(1, 2)$ and $(2, 1)$ and that the strong modified Monge condition is satisfied in all periods. Moreover, $(1, 1)$ and $(2, 2)$ are perfect pairs as defined by the conditions in Theorem 3. We refer to $(1, 2)$ and $(2, 1)$ as imperfect pairs. The optimal policy will have two rounds of matching, as in the following proposition. First, it greedily matches the perfect pairs. Then, it matches an imperfect pair to a threshold level, saving some demand and supply for the matching of perfect pairs in later periods.

Proposition 1. *Given the state* $(\mathbf{x}, \mathbf{y}) = (x_1, x_2, y_1, y_2)$ *in period* t, *the optimal policy first matches the perfect pairs greedily, that is,* $q^*_{11} = \min\{x_1, y_1\}$ *and* $q^*_{22} = \min\{x_2, y_2\}$. *Then, it may or may not match an imperfect pair, contingent on the remaining available demand and supply after the greedy matching of the perfect pairs. In particular, there exist univariate functions* $p^t_{s_1}(IB)$ *and* $p^t_{s_2}(IB)$ *of the total imbalance* $IB := x_1 + x_2 - y_1 - y_2$ *between demand and supply, such that*

i. *If there is no remaining demand (when* $x_1 \leq y_1$ *and* $x_2 \leq y_2$*) or no remaining supply (when* $x_1 > y_1$ *and* $x_2 > y_2$*), the optimal policy can match neither imperfect pair.*

ii. *If there is remaining type* 1 *demand and type* 2 *supply (when* $x_1 > y_1$ *and* $x_2 < y_2$*), the optimal policy matches type* 1 *demand with type* 2 *supply for the quantity* $q^{t*}_{12} = [y_2 - x_2 - p^t_{s_2}(IB)]^+$.

iii. *If there is remaining type* 2 *demand and type* 1 *supply (when* $x_2 > y_2$ *and* $x_1 < y_1$*), the optimal policy matches type* 2 *demand with type* 1 *supply for the quantity* $q^{t*}_{21} = [y_1 - x_1 - p^t_{s_1}(IB)]^+$.

According to Proposition 1, if $x_1 > y_1$ and $x_2 < y_2$, after the first-round matching the quantity of available type 1 demand is $z_1 := x_1 - y_1$ and that of type 2 supply is $z_2 := y_2 - x_2$ (and there is no remaining type 2 demand nor type 1 supply). In the second round, the optimal policy chooses a matching quantity q_{12} to reduce type 2 supply to the threshold $p^t_{s_2}(IB)$ if the available type 2 supply $y_2 - x_2$ after round 1 is above that threshold, and otherwise does not match the pair. Similarly, if $x_1 < y_1$ and $x_2 > y_2$, then there is a remaining quantity $-z_1 := -(x_1 - y_1)$ of type 1 supply and $-z_2 := -(y_2 - x_2)$ of type 2 demand after the first-round matching. In the second round, the optimal policy matches the pair $(2, 1)$ to reduce type 1 supply to the threshold $p^t_{s_1}(IB)$ if the available type 1 supply is above that threshold after round 1, and otherwise does nothing. We have defined thresholds on the supply side to characterize the optimal matching of imperfect pairs. Equivalently, one can also define thresholds on the demand side.

We further explore properties of the thresholds $p^t_{s_1}(IB)$ and $p^t_{s_2}(IB)$ for the case with equal demand and supply carry-over rates and for the case with perishable demand or supply.

Proposition 2. *The thresholds have the following properties.*

i. *If demand and supply have the same carry-over rate, that is,* $\alpha = \beta$, *the thresholds* $p^t_{s_1}(IB)$ *and* $p^t_{s_2}(IB)$ *decrease in* IB, *with the rate of decrease no greater than* 1.

ii. *If demand is perishable (i.e.,* $\alpha = 0$ *such that it does not carry to the future), then in each period* t *there exists state-independent values* $\bar{p}^t_{s_1}$ *and* $\bar{p}^t_{s_2}$ *such that the thresholds* $p^t_{s_1}(IB)$ *and* $p^t_{s_2}(IB)$ *are given by* $p^t_{s_1}(IB) = \max\{IB^-, \bar{p}^t_{s_1}\}$ *and* $p^t_{s_2}(IB) = \max\{IB^-, \bar{p}^t_{s_2}\}$, *respectively.*

Intuitively, a greater IB implies a greater imbalance of demand relative to supply. Under the conditions in part (i) (i.e., equal carry-over rates), this implies that the chance that unmatched demand will be fulfilled by its perfect-match supply type in future periods becomes smaller. Therefore, it becomes more imperative to allow more consumption of demand by lowering the threshold on the supply side in the current period. Part (i) also shows that the decrease in thresholds cannot exceed 1 for per unit of increment in IB. This result helps reduce the complexity of computing the thresholds, especially for the model with discrete-valued states and decisions. For example, suppose that we have obtained the value of $p^t_{s_1}(IB)$ and need to compute $p^t_{s_1}(IB + 1)$. Then the latter is equal to

$p^t_{s_1}(IB)$ or $p^t_{s_1}(IB) - 1$, whichever yields a higher expected total reward.

If demand is perishable, part (ii) of Proposition 2 shows that the second-round matching in period t is characterized by the state-independent values $\bar{p}^t_{s_1}$ and $\bar{p}^t_{s_2}$. We note that the second-round matching cannot reduce the available supply below $IB^- = (x_1 + x_2 - y_1 - y_2)^-$ (in fact, the available supply reduces to IB^- exactly in the second-round matching if greedy matching is used). In the second round, the optimal policy aims at reducing available supply to the state-independent value (i.e., $\bar{p}^t_{s_1}$ or $\bar{p}^t_{s_2}$), or as close to it as possible (i.e., if the state-independent value is below IB^-, the optimal policy will just reduce the available supply to the lowest possible level that is IB^-). Intuitively, this is because demand in the current period does not carry over and thus a fixed amount of supply should be reserved for the perfect-match demand type to arrive in the future.

4.2. Multiple Demand and Supply Types

We now consider multiple demand and supply types. To model horizontal differentiation, we consider a line segment $\mathcal{L}$ endowed with a direction (without loss of generality, from right to left), on which the types are distributed. The location of a (demand or supply) type on $\mathcal{L}$ represents its characteristics or preferences. We consider unidirectional matching, that is, a demand type i and a supply type j are matchable if and only if the supply type can reach the demand type by traveling along the designated direction (represented by the notion $i \leftarrow j$). Suppose that the unit reward for matching type i demand with type j supply is a linearly decreasing function of their distance, that is, $r^t_{ij} = R^t_i - \text{dist}_{i\leftarrow j}$, where R^t_i is the baseline unit reward received from matching type i demand and a supply type with the same location as demand type i, and $\text{dist}_{i\leftarrow j}$ is the Euclidean distance from supply type j to demand type i on $\mathcal{L}$. Without loss of generality, we assume that the demand-type index i and the supply-type index j are both increasing along the direction of $\mathcal{L}$.

The given setting applies to problems where each supply type represents a specific grade of supply and each demand type requires a certain minimum grade. Along the designated direction, the grade of supply and the minimum grade required by the demand types decrease, implying that only upward substitution is allowed and the unit reward is higher when the firm assigns a demand type to a supply grade closer to its original requirement.

We consider an assumption on the baseline unit reward R^t_i.

Assumption 3. i. *The baseline reward R^t_i decreases along the direction of $\mathcal{L}$, that is, $R^t_i \ge R^t_{i+1}$ for $i = 1, \ldots, m-1$.*

ii. *$R^t_i - R^t_{i+1} \ge \gamma\alpha(R^{t+1}_i - R^{t+1}_{i+1})$ for $i = 1, \ldots, m-1$, and $t = 1, \ldots, T-1$.*

Part (i) of Assumption 3 assumes that a demand type of a smaller index (which represents, e.g., a higher requirement on supply grade) yields a higher baseline reward. Part (ii) assumes that the difference in the baseline rewards (adjusted by the discounted factor and demand carry-over rate) decreases over time. We can verify that the following example satisfies Assumption 3.

Example 5. Consider the service upgrading problem with crowdsourced suppliers (e.g., a ride-hailing platform's economic versus premium services). There are n supply types that represent a line of n different services, with type 1 being the most premium service and type n the most economic one. There are also $m = n$ demand types, and type i demand seeks type i supply, for $i = 1, \ldots, n$. Let c_i be the marginal cost of type i supply, and f_i the price for type i demand of purchasing type i supply, for $i = 1, 2, \ldots, n$. Naturally, both c_i and f_i are decreasing in i. The firm has the option to offer type i demand a free upgrade to a supply type $j < i$ (which is more premium) but does not offer a downward substitution. We define the stationary unit matching reward $r^t_{ij} = r_{ij}$ as the profit margin for fulfilling type i demand with type j supply, that is, $r_{ij} := f_i - c_j = f_i - c_i - (c_j - c_i)$.

We can think of demand type i and supply type j sharing the same location on the line segment $\mathcal{L}$, and the indices of the types increase (i.e., the service becomes less superior) along the direction of $\mathcal{L}$. We define $R_i := f_i - c_i$ as the marginal profit for fulfilling type i demand without upgrading, and $\text{dist}_{i\leftarrow j} := c_j - c_i$ as the distance from supply type j to demand type i, for $j \le i$. Then, $r_{ij} = R_i - \text{dist}_{i\leftarrow j}$. As in Remark 2, we can incorporate the waiting or holding costs into the rewards.

With the unit reward given by $r^t_{ij} = R^t_i - \text{dist}_{i\leftarrow j}$, we have the following lemma.

Lemma 1. *Suppose Assumption 3 holds. For two matchable pairs (i, j) and (i', j) of demand and supply, (i, j) weakly precedes (i', j) if and only if j is closer to i than to i' along the direction of $\mathcal{L}$. Likewise, a matchable pair (i, j) weakly precedes (i, j') if and only if j is closer to i along the direction of $\mathcal{L}$. Also, the strong modified Monge condition (Assumption 1) is satisfied in all periods.*

Lemma 1 implies that any two matchable neighboring pairs are comparable by the modified Monge condition, and the pair with a shorter distance (between its demand and supply types) should be assigned a higher priority. As in special case 1, we can classify the pairs into priority levels. Suppose there are L levels in total, and Proposition 3 follows from special case 1.

Proposition 3. *There exists an optimal policy such that it matches the demand-supply pairs in the descending order of the priority levels. If a level ℓ $(1 \le \ell \le L-1)$ pair (i,j) is not matched to the maximum extent (i.e., remaining type i demand and type j supply are available after the matching), the optimal matching quantity of its neighboring pairs of any level $\ell' > \ell$ will be 0.*

Proposition 3 specifies the order in which we should match demand with supply. A level 1 pair may be matched greedily, such as in the following corollary.

Corollary 1. *If (i,j) is a level 1 pair and $R_i^t - \mathrm{dist}_{i\leftarrow j} \ge \gamma \max\{\alpha,\beta\}(R_i^{t+1} - \mathrm{dist}_{i\leftarrow j})$ for $t = 1,\dots,T-1$, the optimal policy matches type i demand with type j supply greedily in all periods.*

The condition $R_i^t - \mathrm{dist}_{i\leftarrow j} \ge \gamma \max\{\alpha,\beta\}(R_i^{t+1} - \mathrm{dist}_{i\leftarrow j})$ in Corollary 1 holds, for example, if the baseline unit reward R_i^t is decreasing in time, which is satisfied in Example 5. In that example, intuitively the firm should greedily match customers with their intended products or service since it leads to higher profit than any upgrades in any future periods. In general, when we match a pair (i,j) of level $\ell \ge 2$, we may want to withhold some type i demand and/or some type j supply, for them to meet better matches in a future period. Next, we consider a heuristic method for determining the extent to which each pair should be matched along the priority structure.

4.2.1. The Match-Down-to Heuristic.

In each period t, we will match demand types with supply types in the order specified in Proposition 3. Suppose that immediately before we match demand type i with supply type j in period t, their available quantities are $\check{x}_i$ and $\check{y}_j$, respectively. If we were to match i with j greedily, at the end of the period the remaining quantities would be $(\check{x}_i - \check{y}_j)^+$ and $(\check{x}_i - \check{y}_j)^-$, respectively. If we withhold a quantity p from matching i and j greedily, the reward in the current period t will decrease by $r_{ij}^t p$, while the withheld quantities of type i demand and type j supply may generate (potentially greater) rewards in the future periods. To find the quantity to withhold, we estimate the marginal future benefits of reserving type i demand and type j supply.

Let $V_{t+1}^g(\mathbf{x},\mathbf{y})$ be the expected discounted reward from period $t+1$ onward under greedy matching in the order specified by Proposition 3, given the state $(\mathbf{x},\mathbf{y})$ at the beginning of period $t+1$. To approximate the future benefit of reserving p units of type i demand (from greedily matching with type j supply) in period t, we ignore the carry over of any demand and supply type other than i and j in period t into period $t+1$. Then, the future benefit of reserving p units of type i demand is approximated by $FB_i^d(p, IB_{ij}) := \gamma EV_{t+1}^g(\alpha(IB_{ij}^+ + p)\mathbf{e}_i^m + \mathbf{D}^{t+1}, \mathbf{S}^{t+1}) - \gamma EV_{t+1}^g(\alpha IB_{ij}^+ \mathbf{e}_i^m + \mathbf{D}^{t+1}, \mathbf{S}^{t+1})$, where $IB_{ij} := \check{x}_i - \check{y}_j$, and $\mathbf{e}_\ell^k$ is the k-dimension vector with the ℓth entry equal to 1 and all other entries being 0 (we may omit the superscript k if the vector size can be inferred from the context). Likewise, we approximate the future benefit of reserving p units of type j supply by $FB_j^s(p, IB_{ij}) := \gamma EV_{t+1}^g(\mathbf{D}^{t+1}, \beta(IB_{ij}^- + p)\mathbf{e}_j^n + \mathbf{S}^{t+1}) - \gamma EV_{t+1}^g(\mathbf{D}^{t+1}, \beta IB_{ij}^- \mathbf{e}_j^n + \mathbf{S}^{t+1})$. We then solve the problem $\max_{p\ge 0} -r_{ij}^t p + FB_i^d(p, \check{x}_i - \check{y}_j) + FB_j^s(p, \check{x}_i - \check{y}_j)$ to obtain the protection level p for the matching between i and j, which is equivalent to

$$\max_{p\ge 0} -r_{ij}^t p + \gamma EV_{t+1}^g\left(\alpha\left(IB_{ij}^+ + p\right)\mathbf{e}_i^m + \mathbf{D}^{t+1}, \mathbf{S}^{t+1}\right) + \gamma EV_{t+1}^g\left(\mathbf{D}^{t+1}, \beta\left[IB_{ij}^- + p\right]\mathbf{e}_j^n + \mathbf{S}^{t+1}\right). \tag{2}$$

To solve (2), we can evaluate the expected values $EV_{t+1}^g(\alpha[IB_{ij}^+ + p]\mathbf{e}_i^m + \mathbf{D}^{t+1}, \mathbf{S}^{t+1})$ and $EV_{t+1}^g(\mathbf{D}^{t+1}, \beta[IB_{ij}^- + p]\mathbf{e}_i^n + \mathbf{S}^{t+1})$ by Monte Carlo simulation (see online supplement D for details). Clearly, the solution to (2), denoted by $\hat{p}_{ij}^t(IB_{ij})$, depends on the demand and supply levels only through the difference $IB_{ij} = \check{x}_i - \check{y}_j$. We match i with j until type i demand reduces to $IB_{ij}^+ + \hat{p}_{ij}^t(IB_{ij})$, if the demand level $\check{x}_i$ is higher than $IB_{ij}^+ + \hat{p}_{ij}^t(IB_{ij})$. Otherwise, we do not match i with j. (Equivalently, this means that we will match i with j until type j supply reduces to $IB_{ij}^- + \hat{p}_{ij}^t(IB_{ij})$, if it is possible to do so.) We summarize the heuristic in Algorithm 1.

Algorithm 1 (Match-Down-to Heuristic for the Horizontal Model in Any Period t, Given the Demand and Supply Levels $(\mathbf{x},\mathbf{y})$ in the Beginning of the Period)

1: $\check{\mathbf{x}} \leftarrow \mathbf{x}$ and $\check{\mathbf{y}} \leftarrow \mathbf{y}$
2: **for** each pair of demand and supply (i,j) along the priority structure **do**
3: $\quad IB_{ij} \leftarrow \check{x}_i - \check{y}_j$, and compute the protection level $\hat{p}_{ij}^t(IB_{ij})$ by solving (2)
4: $\quad$ Match i with j until type i demand reduces to $IB_{ij}^+ + \hat{p}_{ij}^t(IB_{ij})$ (and simultaneously, type j supply reduces to $IB_{ij}^- + \hat{p}_{ij}^t(IB_{ij})$) or as close to it as possible
5: $\quad \check{x}_i \leftarrow \min\{\check{x}_i, IB_{ij}^+ + \hat{p}_{ij}^t(IB_{ij})\}$ and $\check{y}_j \leftarrow \min\{\check{y}_j, IB_{ij}^- + \hat{p}_{ij}^t(IB_{ij})\}$
6: **end for**

Although Algorithm 1 is motivated by Proposition 3, this heuristic matching policy may not satisfy the structure in the proposition (i.e., not to match a lower-level pair unless the matching between all the weakly preceding pairs is no longer possible). In particular, even if the matching between a pair (i,j) has not been exhausted, under Algorithm 1, the heuristic may proceed to match a neighboring pair (i,j') or (i',j) that is weakly preceded by (i,j). This is because we approximate the expected future rewards by the expected values under greedy matching from period $t+1$ onward, which may not possess some properties of the optimal value function that are needed to ensure the structure in Proposition 3. Nonetheless, one can modify the matching policy so that it satisfies the structure in Proposition 3 (if this improves the expected total discounted matching reward) by redirecting a matching quantity from a lower-level pair to a neighboring higher-level pair. Next, we present a numerical example to illustrate the heuristic.

Example 6. We revisit Example 5 with $m = n = 3$ and $T = 5$ periods. Suppose that the marginal costs for the three supply types are $c_1 = 45$, $c_2 = 38$, and $c_3 = 14$, and the prices paid by the three types of customers are $f_1 = 100$, $f_2 = 60$, and $f_3 = 46$. The unit rewards are stationary over time and are given by $r_{ij} = f_i - c_j$ for $j \le i = 1,2,3$. We use the discount factor $\gamma = 0.9$ and assume that all unmatched demand and supply carry to the next period, that is, $\alpha = \beta = 1$. Following the matching order specified in Proposition 3, we will first match the level 1 pairs, $(1,1)$, $(2,2)$, and $(3,3)$, followed by the level 2 pairs $(2,1)$ and $(3,2)$, and finally the level 3 pair $(3,1)$. We require that arrivals of demand and supply of all types be independent and identically distributed and follow the discrete uniform distribution over the set $\{0,1,2,3,4\}$ throughout the horizon, and that the matching quantities take integer values.

Suppose that in the beginning of period 1, the initial demand levels are given by $\mathbf{x}^1 = (x_1^1, x_2^1, x_3^1) = (2,8,9)$, and the initial supply levels are given by $\mathbf{y}^1 = (y_1^1, y_2^1, y_3^1) = (8,3,2)$. We apply the procedure in Algorithm 1 to compute the protection levels. When we solve (2) to compute the protection levels, we generate $N = 1,000$ sample paths to evaluate the two expected values in (2) approximately.

In period 1, the pairs $(1,1)$, $(2,2)$, and $(3,3)$ are matched greedily (by Corollary 1). Thus, the matching quantities between those pairs are $q_{11}^1 = 2$, $q_{22}^1 = 3$, and $q_{33}^1 = 2$. The remaining demand levels become $\check{\mathbf{x}} = (0,5,7)$, and the remaining supply levels become $\check{\mathbf{y}} = (6,0,0)$.

Then, we proceed to match type 2 demand with type 1 supply. The difference between type 2 demand and type 1 supply is $\check{x}_2 - \check{y}_1 = -1$, and our computation yields the protection level as $\hat{p}_{21}^1(-1) = 3$. Therefore, the matching quantity between type 2 demand and type 1 supply is $q_{21}^1 = 2$. We update the remaining demand and supply levels as $\check{\mathbf{x}} = (0,3,7)$ and $\check{\mathbf{y}} = (4,0,0)$.

The pair $(3,2)$ cannot be matched because type 2 supply is not available. It remains to match $(3,1)$. The difference between type 3 demand and type 1 supply is 3, and we obtain the protection level $\hat{p}_{31}^1(3) = 15$. Thus, the matching quantity between type 3 demand and type 1 supply is $q_{31}^1 = 0$.

To compute the expected total discounted reward under the heuristic, we randomly generate $M = 1,000$ sample paths of demand and supply realizations (denoted by $\varphi_1, \ldots, \varphi_M$). Starting from a given initial state $(\mathbf{x}^1, \mathbf{y}^1)$ in period 1, we apply Algorithm 1 along each sample path and calculate the corresponding total discounted reward. We then calculate the average total discounted reward over the 1,000 sample paths as the expected total reward under the heuristic. For $(\mathbf{x}^1, \mathbf{y}^1) = (2,8,9,8,3,2)$, our heuristic leads to an expected total discounted reward equal to 1,088.77.

We further derive an upper bound on the optimal expected total discounted reward by assuming that in period 1 the firm has perfect information about demand and supply realizations in all periods. In other words, for each sample path φ_k $(k = 1, \ldots, M)$, the firm can perfectly predict all demand and supply realizations along the sample path, and therefore solve a linear program to obtain the optimal matching quantities in all periods and the corresponding expected total discounted reward (for the given initial state $(\mathbf{x}^1, \mathbf{y}^1)$), which we denote by $TR^{\text{perfect\_info}}(\mathbf{x}^1, \mathbf{y}^1, \varphi_k)$. By averaging over all the M sample paths, we obtain an approximate upper bound $U(\mathbf{x}^1, \mathbf{y}^1) := \sum_{k=1}^{M} TR^{\text{perfect\_info}}(\mathbf{x}^1, \mathbf{y}^1, \varphi_k)/M$ on the optimal expected total discounted reward. For $(\mathbf{x}^1, \mathbf{y}^1) = (2,8,9,8,3,2)$, we obtain an approximate upper bound equal to $U(\mathbf{x}^1, \mathbf{y}^1) = 1,111.76$. Thus, our heuristic achieves at least $\frac{1,088.77}{1,111.76} \times 100\% \approx 97.93\%$ of the optimal expected total discounted reward.

Finally, for 200 randomly selected initial states $(\mathbf{x}^1, \mathbf{y}^1)$ such that $\max\{x_1^1, x_2^1, x_3^1, y_1^1, y_2^1, y_3^1\} \le 20$, our heuristic achieves over 96.62% of the upper bound from the perfect information case.

5. Vertically Differentiated Types

We now consider vertically differentiated demand and supply types. Each demand or supply type is associated with a quality level, and it generates a higher reward if matched with a supply or demand type of higher quality. In other words, the unit matching reward between type i demand and type j supply is an increasing function $r_{ij}^t = f^t(a_i, b_j)$ of a_i and b_j, which represent the quality of demand type i and supply type j, respectively. For simplicity, we consider an additive reward function $r_{ij}^t = f_d^t(a_i) + f_s^t(b_j)$, where f_d^t and f_s^t are increasing in a_i and b_j, respectively. We write $r_{id}^t := f_d^t(a_i)$ and $r_{js}^t := f_s^t(b_j)$. Without loss of generality, suppose that a demand or supply type with a smaller index has a higher quality and therefore leads to a higher unit reward, that is, $r_{1d}^t > r_{2d}^t > \cdots > r_{md}^t$ and $r_{1s}^t > r_{2s}^t > \cdots > r_{ns}^t$. In addition, we assume that the unit rewards (after adjustment by the discount factor and the carry-over rates) are decreasing in time as in the following assumption.

Assumption 4. *For any $t = 1, \ldots, T-1$, $i = 1, \ldots, m$, and $j = 1, \ldots, n$, we assume $r_{id}^t - r_{i+1,d}^t \ge \gamma\alpha(r_{id}^{t+1} - r_{i+1,d}^{t+1})$ and*

$r_{js}^t - r_{j+1,s}^t \geq \gamma\beta(r_{js}^{t+1} - r_{j+1,s}^{t+1})$ *(where* $r_{m+1,d}^t = r_{n+1,s}^t := 0$ *for any period t).*

We present two examples that satisfy Assumption 4.

Example 7. We consider the following dynamic transportation problem of donated goods, and focus on the transportation from a given origin and destination. There are n logistics providers offering to ship the goods at regular or at discounted or free rates, using their spare capacity. Let c_j be the rate offered by provider j, for $j = 1, \ldots, n$. Suppose without loss of generality that $c_1 \leq \cdots \leq c_n$. In each period, the spare capacity of each provider is uncertain (since it depends on how much capacity is consumed by its regular business volume). Any unused spare capacity in a period is lost (i.e., $\beta = 0$). Donated goods to be shipped also arrive in random quantities. Suppose that the donated goods can be classified into m types according to their urgency. Unshipped goods in a period partially carry over to the next period with rate $\alpha \geq 0$. Let b be the social benefit for each unit of good shipped, and w_i be the per-unit waiting cost for any unshipped type i goods in a period. We assume that $w_1 \geq \cdots \geq w_m$. If we assign a unit of type i goods to any provider j in period t, a total amount of discounted waiting cost $w_i \sum_{\tau=t}^T (\gamma\alpha)^{\tau-t}$ is saved. Thus, the reward for assigning a unit of type i goods to provider j in period t is $r_{ij}^t = b - c_j + w_i \sum_{\tau=t}^T (\gamma\alpha)^{\tau-t}$. Let $r_{id}^t := w_i \sum_{\tau=t}^T (\gamma\alpha)^{\tau-t}$ and $r_{js}^t := b - c_j$, and we can readily verify that Assumption 4 is satisfied.

Example 8. Consider an online labor market platform that dynamically assigns workers to customers' requests. Customers and workers are divided into m and n types, respectively, according to their ratings. Without loss of generality, suppose that type i customers have higher ratings than type $i+1$ customers for $i = 1, \ldots, m-1$ and that type j workers have higher ratings than type $j+1$ workers for $j = 1, \ldots, n-1$. Workers with higher ratings may receive a higher wage (e.g., Handy offers its cleaning workers four wage rates according to their rating), and on average they also generate higher surplus for their customers (e.g., through better service quality). Customers with higher ratings generate higher surplus for the worker (e.g., through a better experience or more generous tipping, etc.). Let s_j be the compensation for type j workers, f_j the surplus generated by a type j worker for any customer the worker serves, and g_i the surplus generated by a type i customer for the customer's worker (in addition to the payment to the worker). Suppose that s_j and f_j are decreasing in j, and g_i is decreasing in i, and the platform aims to maximize the total welfare of customers and workers. The unit reward for matching a type i customer and a type j worker is $r_{ij}^t = g_i + f_j + s_j$. With $r_{id}^t := g_i$ and $r_{js}^t := f_j + s_j$, Assumption 4 is satisfied.

5.1. The Optimal Policy

Under Assumption 4, we can verify that any two neighboring pairs are comparable by the modified Monge condition (the pair with a higher-quality demand or supply type weakly precedes the other), and the strong modified Monge condition holds in all periods (see the proof of Proposition 4 in Hu and Zhou 2020). The model reduces to special case 2 (if some demand or supply types are removed, the highest-quality demand type and supply type among the remaining types form the only pair of the highest level). Thus, the optimal policy has a match-down-to structure, which we call the top-down matching structure, as described in Algorithm 2 (since it matches higher-quality types first).

Algorithm 2 (Top-Down Matching Procedure Up to a Total Matching Quantity $\bar{Q}$ in a Period)

1: $i \leftarrow 1$, $j \leftarrow 1$, total matching quantity $Q \leftarrow 0$
2: **while** $Q < \bar{Q}$ **do**
3: Match type i demand with type j supply for a quantity q_{ij} such that either one of them runs out or the total matching quantity Q reaches $\bar{Q}$
4: $Q \leftarrow Q + q_{ij}$
5: $i \leftarrow i + 1$ if type i demand runs out, and $j \leftarrow j + 1$ if type j supply runs out
6: **end while**

Proposition 4. *Suppose that Assumption 4 holds. There exists an optimal matching policy that follows the top-down matching procedure up to some total matching quantity* $\bar{Q}^t$ *in each period t.*

Algorithm 2 shows that in each period the top-down matching policy is fully determined by the total matching quantity. In the next lemma, we reduce the optimal matching problem in any period to a one-dimensional convex optimization problem with respect to the total matching quantity.

Lemma 2. *The optimal expected discounted reward* $V_t(\mathbf{x}, \mathbf{y})$ *from period t to T is equal to*

$$\begin{aligned}\max_{\bar{Q}} \quad & G_t(\bar{Q}, \mathbf{x}, \mathbf{y}) \\ := & \sum_{i=1}^m r_{id}^t \min\left\{\left(\bar{Q} - \sum_{i'=1}^{i-1} x_{i'}\right)^+, x_i\right\} \\ & + \sum_{j=1}^n r_{js}^t \min\left\{\left(\bar{Q} - \sum_{j'=1}^{j-1} y_{j'}\right)^+, y_j\right\} \\ & + \gamma E V_{t+1}(\alpha\mathbf{u} + \mathbf{D}^{t+1}, \beta\mathbf{v} + \mathbf{S}^{t+1}), \end{aligned} \tag{3}$$

subject to $0 \le \bar{Q} \le \min\{\sum_{i=1}^m x_i, \sum_{j=1}^n y_j\}$, $u_i = [x_i - (\bar{Q} - \sum_{i'=1}^{i-1} x_{i'})^+]^+$ *for* $i = 1, \ldots, m$ *and* $v_j = [y_j - (\bar{Q} - \sum_{j'=1}^{j-1} y_{j'})^+]^+$ *for* $j = 1, \ldots, n$. *Moreover,* $G_t(\bar{Q}, \mathbf{x}, \mathbf{y})$ *is concave in* $\bar{Q}$ *for* $t = 1, \ldots, T$.

Given the state $(\mathbf{x}, \mathbf{y})$ in period t, we denote by $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ the optimal total matching quantity. If $r_{1d}^t \ge \gamma \max\{\alpha, \beta\} r_{1d}^{t+1}$ and $r_{1s}^t \ge \gamma \max\{\alpha, \beta\} r_{1s}^{t+1}$ for $t = 1, \ldots, T-1$ (those conditions are satisfied, e.g., in Examples 7 and 8), it follows from Theorem 3 that the pair $(1,1)$ should be matched greedily, and thus $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) \ge \min\{x_1, y_1\}$. In general, the optimal policy may not match all demand and supply greedily but would reserve some lower-quality supply (respectively, demand) as "safety stock" for high-quality demand (respectively, supply) in later periods.

When the demand carry-over rate α equals the supply carry-over rate β or when demand is perishable, we obtain the following properties of the optimal total matching quantity in a period.

Proposition 5. *Suppose that either* $\alpha = \beta > 0$ *or* $\alpha = 0 < \beta$. *In any period* t:

i. *The optimal total matching quantity* $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ *is increasing in* x_i *and* y_j *(for* $i = 1, \ldots, m$ *and* $j = 1, \ldots, n$*), with the corresponding rates of increase smaller than or equal to one.*

ii. *The optimal total matching quantity* $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ *is more sensitive to changes in the available demand and supply of higher quality, that is, it increases faster in* x_i *than* x_{i+1}*, and in* y_j *than* y_{j+1}*, for* $i = 1, \ldots, m-1, j = 1, \ldots, n-1$.

Like the horizontal model with $m = n = 2$ and equal carry-over rates (see Proposition 2(i)), Proposition 5 helps simplify the computation of the optimal total matching quantity in a given period t. For example, suppose that we have obtained $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$. Then, by Proposition 5(i), we only need to search for the value of $\bar{Q}^{t*}(\mathbf{x} + \mathbf{e}_i, \mathbf{y})$ within the interval $[\bar{Q}^{t*}(\mathbf{x}, \mathbf{y}), \bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) + 1]$. In particular, for the discrete-value system where both the states and matching quantities take integer values, $\bar{Q}^{t*}(\mathbf{x} + \mathbf{e}_i, \mathbf{y})$ must equal either $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y})$ and $\bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) + 1$, whichever yields a higher expected total discounted reward. Further, if the latter yields a higher expected total discounted reward, then for any $k = 1, \ldots, i$ we have $\bar{Q}^{t*}(\mathbf{x} + \mathbf{e}_k, \mathbf{y}) = \bar{Q}^{t*}(\mathbf{x}, \mathbf{y}) + 1$ by Proposition 5(ii).

We remark that, when demand is perishable (i.e., $\alpha = 0$), in any period t the vector of demand levels $\mathbf{x}$ is always equal to the new demand $\mathbf{D}^t$ joining in that period. For the case with perishable supply (i.e., $\beta = 0$), we can obtain analogous properties to Proposition 5 by symmetry.

Lemma 2 and Proposition 5 help us compute the optimal total matching quantity in a given period more efficiently, provided that we can efficiently compute the expected value function EV_{t+1} in (3). However, computation of EV_{t+1} is nontrivial if there are many demand or supply types. Next, we study a heuristic that approximates EV_{t+1} by the reward achieved by the greedy policy.

5.2. The One-Step-Lookahead Heuristic

We consider a heuristic method that approximates the value function $V_{t+1}(\mathbf{x}^{t+1}, \mathbf{y}^{t+1})$ by the expected total matching reward from period $t+1$ to the end of the time horizon under the greedy matching policy (following the top-down order), which we denote by $V_{t+1}^g(\mathbf{x}^{t+1}, \mathbf{y}^{t+1})$. Then, for every period t, we solve (3) for the approximate optimal total matching quantity after replacing the value function V_{t+1} by V_{t+1}^g. To evaluate the function $V_{t+1}^g(\mathbf{x}^{t+1}, \mathbf{y}^{t+1})$ for any given state $(\mathbf{x}^{t+1}, \mathbf{y}^{t+1})$, we resort to Monte Carlo simulation by averaging the total discounted matching reward over randomly generated sample paths of demand and supply realizations. We refer to this heuristic as the one-step-lookahead heuristic because it improves the greedy matching policy by looking one period ahead. We have the following characterization of the one-step-lookahead heuristic.

Proposition 6. *The one-step-lookahead heuristic follows the top-down matching procedure in each period if Assumption 4 holds. Moreover, it performs weakly better than the greedy matching policy (with respect to the expected total discounted reward), and it is optimal for any two-period problem.*

When the demand and supply carry-over rates are equal (i.e., $\alpha = \beta$), we will show that the one-step-lookahead heuristic leads to a matching policy with a simple structure characterized by a set of protection levels $\{p_{s_{ij}}^t(IB)\}_{i=1,\ldots,m,\, j=1,\ldots,n,\, t=1,\ldots,T}$, which depend on the total imbalance between demand and supply $IB := \sum_{i=1}^m x_i - \sum_{j=1}^n y_j$ in a period t (rather than on the full state $(\mathbf{x}, \mathbf{y})$). Specifically, under the top-down matching procedure, immediately before we match i with j, the available supply (of all types) is $\tilde{v}_{n,U}^{ij} := \sum_{l=j}^n y_l - \min\{(\sum_{k=1}^{i-1} x_k - \sum_{l=1}^{j-1} y_l)^+, y_j\}$ (i.e., supply of types $j, \ldots, n$ less the amount of type j supply consumed by type $i-1$ or higher-quality demand). If we were to match i with j to the maximum extent, the available supply would reduce to $\tilde{v}_{n,L}^{ij} := \sum_{l=j}^n y_l - \min\{(\sum_{k=1}^{i} x_k - \sum_{l=1}^{j-1} y_l)^+, y_j\}$. If $\alpha = \beta$, Proposition 7 shows that the one-step-lookahead heuristic will match i with j such that the available supply (of all types) reduces to the protection level $p_{s_{ij}}^t(IB)$ or as close to $p_{s_{ij}}^t(IB)$ as possible within the range $[\tilde{v}_{n,L}^{ij}, \tilde{v}_{n,U}^{ij}]$.

Proposition 7. *Suppose* $\alpha = \beta$. *For each pair* (i, j) *there exists a protection level* $p_{s_{ij}}^t(IB)$ *dependent only on the total imbalance* $IB := \sum_{i'=1}^m x_{i'} - \sum_{j'=1}^n y_{j'}$*, such that when the one-step-lookahead heuristic matches* (i, j)*, it matches the pair* *for a quantity of* $\min\{[\tilde{v}_{n,U}^{ij} - p_{s_{ij}}^t(IB)]^+, \tilde{v}_{n,U}^{ij} - \tilde{v}_{n,L}^{ij}\}$ *to reduce the total supply to* $p_{s_{ij}}^t(IB)$*, or as close to it as possible within the range* $[\tilde{v}_{n,L}^{ij}, \tilde{v}_{n,U}^{ij}]$.

We describe the matching procedure in Proposition 7 with respect to the protection levels $\{p_{s_{ij}}^t(IB)\}_{\forall i,j,t}$ in Algorithm B.3 of online Appendix B. In particular, the one-step-lookahead policy will try to reduce the total supply to the protection level $p_{s_{ij}}^t(IB)$ by matching i with j, and then stop the matching procedure in period t. If type i demand (respectively, type j supply) runs out before the total supply reaches the protection level $p_{s_{ij}}^t(IB)$, it will continue to match type $i+1$ demand with type j supply (respectively, type i demand with type $j+1$ supply). If the total available supply (i.e., $\tilde{v}_{n,U}^{ij}$) is already below the protection level $p_{s_{ij}}^t(IB)$ immediately before we match i and j, the one-step-lookahead policy will stop the matching procedure without matching i and j.

Next, we discuss how to find the protection level $p_{s_{ij}}^t(IB)$ in Proposition 7. Because the protection level $p_{s_{ij}}^t(IB)$ remains the same as long as the total imbalance IB is held constant, we can assume without loss of generality that all the available demand is of type i and all the available supply is of type j when we start to match i with j in period t. If the firm matches i with j until the total supply reduces to p, it receives a reward $r_{ij}^t(\tilde{v}_{n,U}^{ij} - p)$ in period t (recall that $\tilde{v}_{n,U}^{ij}$ is the available supply immediately before matching (i,j)). In the meantime, by the definition of the total imbalance IB, the total demand will reduce to $p+IB$, which results in the postmatching levels $((p+IB)\mathbf{e}_i^m, p\mathbf{e}_j^n)$. Assuming greedy matching from period $t+1$ to T and noting that the total supply cannot be reduced below IB^-, we choose the protection level $p_{s_{ij}}^t(IB)$ as follows:

$$p_{s_{ij}}^t(IB) := \arg\max_{p \geq IB^-}\Big\{-r_{ij}^t p + \gamma EV_{t+1}^g\big(\alpha(IB+p)\mathbf{e}_i + \mathbf{D}^{t+1}, \beta p\mathbf{e}_j + \mathbf{S}^{t+1}\big)\Big\}, \quad (4)$$

to maximize the expected discounted reward from period t to T. With $\alpha = \beta$, the right side of (4) is a convex optimization problem (see Lemma B.11 in online supplement B for a proof). If $\alpha \neq \beta$, Proposition 7 may not hold. Nevertheless, as a heuristic method, we can still use (4) to obtain $p_{s_{ij}}^t(IB)$ and apply Algorithm B.3 to match demand with supply in any period t.

Previously, Proposition 5 shows that the optimal total matching quantity $\bar{Q}^{t*}(\mathbf{x},\mathbf{y})$ is increasing in x_i, with a rate of increase smaller than or equal to 1, for any $i = 1, \ldots, m$. Suppose that the same property also holds for the total matching quantity resulting from the matching procedure in Algorithm B.3 (even though it is not guaranteed, we can nonetheless make this assumption as part of the heuristic method). Then, the protection level $p_{s_{ij}}^t(IB)$ should be decreasing in IB with a rate of decrease smaller than or equal to 1 (i.e., $p_{s_{ij}}^t(IB-1) - 1 \leq p_{s_{ij}}^t(IB) \leq p_{s_{ij}}^t(IB-1)$ for all $IB \geq 1$; this is because the total matching quantity is decreasing in the protection level, and IB is increasing in x_i). This is helpful for reducing the computational burden for the model with discrete-valued states and decisions. Suppose that we have already obtained $p_{s_{ij}}^t(IB-1)$ and want to compute $p_{s_{ij}}^t(IB)$. Instead of solving (4), we can simply compare the objective value of (4) for $p = p_{s_{ij}}^t(IB-1)$ and $p = p_{s_{ij}}^t(IB-1) - 1$, and choose the one that yields the greater objective value.

We now present a numerical example to illustrate the one-step-lookahead policy.

Example 9. Consider $m = 3$ demand types and $n = 3$ supply types to be matched over $T = 5$ periods. The unit matching rewards are given by $r_{ij}^t = r_{id}^t + r_{js}^t$ ($i = 1,2,3,\ j = 1,2,3$), where $r_{1d}^t = 100$, $r_{2d}^t = 30$, $r_{3d}^t = 5$, $r_{1s}^t = 102$, $r_{2s}^t = 22$, and $r_{3s}^t = 4$ for all $t = 1, \ldots, 5$. Suppose that $\alpha = \beta = 1$ (i.e., all unmatched demand and supply carry over to the future) and the discount factor is $\gamma = 0.9$. In each period t, the new arrivals of type i demand follow the discrete uniform distribution over the values $\{0,1,2,3,4\}$ for all $i = 1,2,3$, and the new arrivals of type j supply follow the discrete uniform distribution over the values $\{0,2,10,12\}$. The initial state in period 1 is $(\mathbf{x}^1, \mathbf{y}^1)$, where $\mathbf{x}^1 = (18,3,10)$ and $\mathbf{y}^1 = (5,8,18)$. For simplicity, we assume that both states and matching quantities take only integer values. We see that the total imbalance in period 1 is $IB = \sum_{i=1}^3 x_i^1 - \sum_{j=1}^3 y_j^1 = 0$.

We now apply the one-step-lookahead heuristic to find the matching quantities in period 1. To that end, we compute the matching thresholds by (4) (to solve the right side of (4), we evaluate the function $EV_{t+1}^g(\alpha(IB+p)\mathbf{e}_i + \mathbf{D}^{t+1}, \beta p\mathbf{e}_j + \mathbf{S}^{t+1})$ with Monte Carlo simulation, by averaging over $N = 1{,}000$ randomly generated sample paths). We obtain the following thresholds: $\hat{p}_{s_{11}}(0) = \hat{p}_{s_{12}}(0) = \hat{p}_{s_{13}}(0) = 0$, $\hat{p}_{s_{21}}(0) = 4$, $\hat{p}_{s_{22}}(0) = 7$, $\hat{p}_{s_{23}}(0) = 8$, $\hat{p}_{s_{31}}(0) = 7$, $\hat{p}_{s_{32}}(0) = 13$, and $\hat{p}_{s_{33}}(0) = 18$.

In period 1, the heuristic first matches the pair $(1,1)$ greedily (since the threshold $\hat{p}_{s_{11}}(0) = 0$), which leads to a remaining quantity 13 of type 1 demand and no remaining type 1 supply. Then, it proceeds to match the pairs $(1,2)$ and $(1,3)$ greedily in sequence. (Note that $\hat{p}_{s_{12}}(0) = \hat{p}_{s_{13}}(0) = 0$, and that there is some remaining type 1 demand but no type 2 supply after the greedy matching of $(1,2)$.) At this moment, the remaining quantities of types 1, 2, and 3 demand are 0, 3, and 10, respectively, and the remaining quantities of types 1, 2, and 3 supply are 0, 0, and 13, respectively.

We then match the pair $(2,3)$. Since $\hat{p}_{s_{23}}(0) = 8$, the heuristic aims to reduce the total remaining supply to eight or as close to it as possible. However, since there are only three units of type 2 demand available, we will match a quantity of three between the pair $(2, 3)$, thereby reducing the total remaining supply to 10 (only type 3 demand and type 3 supply have remaining quantities).

Finally, only 10 units of type 3 demand and 10 units of type 3 supply remain. Given the threshold $\hat{p}_{s_{33}}(0) = 18$, we do not match the pair $(3, 3)$ since the total remaining supply is $10 < 18 = \hat{p}_{s_{33}}(0)$.

To evaluate the performance of the one-step-lookahead heuristic approximately, we randomly generate 1,000 sample paths of demand and supply realizations. With the initial demand and supply levels $\mathbf{x}^1 = (18, 3, 10)$ and $\mathbf{y}^1 = (5, 8, 18)$, we apply the heuristic on each sample path, and compute the average total matching reward, which turns out to be 5,463.58, across all the 1,000 sample paths. We also compute the expected reward under perfect information as 5,544.89, which is an upper bound on the optimal expected reward. We see that the heuristic achieves 98.53% of this upper bound, and thus it achieves at least 98.53% of the optimal expected discounted reward.

We further randomly choose 200 initial states such that each entry of the state is between 0 and 20, and we find the heuristic to achieve 96.03% or more of the optimal expected reward.

6. Concluding Remarks

We consider a stochastic and dynamic matching framework with heterogeneous demand and supply types in discrete time. We propose the modified and strong modified Monge conditions to prioritize demand-supply pairs optimally and study two reward structures that satisfy those conditions for all neighboring pairs. In the unidirectionally horizontal reward structure, distance determines priority, and in the vertical reward structure, quality determines priority. For both reward structures, the optimal policy proceeds along the priority structure, and when it comes to the matching between a specific pair, the optimal policy matches the pair down to a threshold. This structural property of priority and thresholds is a generalization of priority structures seen in the balanced and deterministic transportation problems and of the threshold-type policies seen in the inventory management (e.g., base-stock levels) and quantity-based revenue management (e.g., protection levels).

The proposed framework, which generalizes many classical problems, may lead to interesting future research. For example, we generalize inventory rationing problems in the following sense. For an inventory rationing problem, there is typically one type of good and multiple classes of demand differentiated by their waiting or shortage costs, and the manager decides dynamically the replenishment quantity and allocation of available inventory to each demand class. Our framework allows multiple supply streams with exogenously random arrivals for each stream and interstream substitution. If an inventory rationing problem has exogenously given supply in each period, it becomes a special case of our framework. For example, a firm may have a long-term agreement with its supplier for frequency and quantity of recurring deliveries, but the actual delivery may have random quantities due to supply disruption or random yield. Our framework does not explicitly consider replenishment decisions, and it would be interesting to consider the joint replenishment-matching problem. If the manager can replenish each type of supply and at the same time, decline some of the new supply arrivals and dispose of some current inventory, we conjecture that the optimal replenishment or disposal policy will aim for an optimal target level for each supply type, and the matching decision will have the same structure as in our results.

Our framework also generalizes revenue management problems with multiple types of resources by considering random supply arrivals, which are relevant for many sharing-economy platforms due to their use of independent or crowdsourced suppliers. Our paper focuses on quantity-matching decisions. It would be interesting to study joint pricing and matching decisions (e.g., ride-hailing platforms make both pricing and matching decisions dynamically). One may also consider the possibility of matching multiple demand units with one supply unit (e.g., real-time carpooling) or a demand unit requesting multiple supply units (e.g., in on-demand delivery problems, where the items in an order may be delivered by multiple couriers in parallel), as opposed to our assumption of one-to-one matching. Moreover, it would also be interesting to study our problem from a data-driven perspective. For example, while making matching decisions, a firm may dynamically learn the distribution of demand and supply arrivals or the matching rewards if they are unknown.

Finally, further computational or algorithmic studies can be another direction for future research. We compute the threshold levels heuristically for our horizontal and vertical models by greedy approximation of the future matching rewards. It would be interesting to use the ADP techniques to approximate the value functions and compute the threshold levels from those approximations.

Acknowledgment

The authors gratefully acknowledge Hau Lee (Department Editor), the associate editor, and four anonymous referees for their valuable comments, which significantly improved the paper. We thank Dan Zhang for his careful review and valuable comments on this work as part of the second author's dissertation. We thank Mehmet Begen, Fangruo Chen, Xin Chen, Stan Dimitrov, Dinesh Gauri, Ming-Hui Huang, Benny Martin, Hamid Nazerzadeh, Ozge Sahin, Xuanming Su, Chung-Piaw Teo, and Tunay Tunca for their valuable discussions and comments at various conferences, workshops, and paper competitions.

References

Abdulkadiroğlu, A, Sönmez T (2013) Matching markets: Theory and practice. Acemoglu D, Arellano M, Dekel E, eds. *Advances in Economics and Econometrics: Tenth World Congress* (Cambridge University Press, Cambridge, UK), 3–47.

Adan I, Weiss G (2012) Exact FCFS matching rates for two infinite multitype sequences. *Oper. Res.* 60(2):475–489.

Afèche P, Diamant A, Milner J (2014) Double-sided batch queues with abandonment: Modeling crossing networks. *Oper. Res.* 62(5): 1179–1201.

Akan M, Alagoz O, Ata B, Erenay FS, Said A (2012) A broader view of designing the liver allocation system. *Oper. Res.* 60(4):757–770.

Arnosti N, Johari R, Kanoria Y (2021) Managing congestion in dynamic matching markets. *Manafacturing Service Oper. Management*. Forthcoming.

Ashlagi I, Shi P (2016) Optimal allocation without money: An engineering approach. *Management Sci.* 62(4):1078–1097.

Baccara M, Lee S, Yariv Y (2020) Optimal dynamic matching. *Theoret. Econom.* 15:1221–1278.

Benjaafar S, Hu M (2020) Operations management in the age of the sharing economy: What is old and what is new? *Manufacturing Service Oper. Management* 22(1):93–101.

Bernstein F, DeCroix G, Keskin NB (2020) Competition between two-sided platforms under demand and supply congestion effects. *Manufacturing Service Oper. Management*, ePub ahead of print October 1, https://doi.org/10.1287/msom.2020.0866.

Bertsekas DP (2013) Rollout algorithms for discrete optimization: A survey. Pardalos P, Du DZ, Graham R, eds. *Handbook of Combinatorial Optimization* (Springer, New York), 2989–3013.

Burkard RE (2007) Monge properties, discrete convexity and applications. *Eur. J. Oper. Res.* 176 (1):1–14.

Chen X, Simchi-Levi D (2012) Pricing and inventory management. Özer Ö, Phillips R, eds. *The Oxford Handbook of Pricing Management* (Oxford University Press, Oxford, UK).

Estes A, Ball M (2020) Monge properties, optimal greedy policies, and policy improvement for the dynamic stochastic transportation problem. I*NFORMS J. Comput.*, ePub ahead of print October 13, https://doi.org/10.1287/ijoc.2020.0990.

Gurvich I, Ward A (2014) On the dynamic control of matching queues. *Stochastic Systems* 4(2):479–523.

Hoffman AJ (1963) On simple linear programming problems. Klee V, ed. *Convexity: Proc. Sympos. Pure Math.*, vol. 7 (American Mathematical Society, Providence, RI), 317–327.

Hu M, Zhou Y (2020) Dynamic type matching. Preprint, submitted September 9, https://dx.doi.org/10.2139/ssrn.2592622.

Kanoria Y, Saban D (2017) Facilitating the search for partners on matching platforms. Preprint, submitted July 22, https://dx.doi.org/10.2139/ssrn.3004814.

Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. *Proc. 22nd Annual ACM Sympos. Theory Comput.* (Association for Computing Machinery, Baltimore, Maryland), 352–358.

Ke J, Zhang D, Zheng H (2019) An approximate dynamic programming approach to dynamic pricing for network revenue management. *Production Oper. Management* 28(11):2719–2737.

Keskin NB, Birge JR (2019) Dynamic selling mechanisms for product differentiation and learning. *Oper. Res.* 67(4):1069–1089.

Manshadi VH, Gharan SO, Saberi A (2012) Online stochastic matching: Online actions based on offline statistics. *Math. Oper. Res.* 37(4):559–573.

Monge G (1781) *Délai et remblai* (Mémoire de l'Académie des Sciences, Paris).

Powell WB (2016) Perspectives of approximate dynamic programming. *Ann. Oper. Res.* 241:319–356.

Queyranne M, Spieksma F, Tardella F (1998) A general class of greedily solvable linear programs. *Math. Oper. Res.* 23(4): 892–908.

Roth AE, Sotomayor M (1990) *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* (Cambridge University Press, Cambridge, UK).

Roth AE, Sönmez T, Ünver MU (2004) Kidney exchange. *Quart. J. Econom.* 119(2):457–488.

Roth AE, Sönmez T, Ünver MU (2007) Efficient kidney exchange: Coincident of wants in market with compatibility-based preferences. *Amer. Econom. Rev.* 97(3):828–851.

Shumsky RA, Zhang F (2009) Dynamic capacity management with substitution. *Oper. Res.* 57(3):671–684.

Su X, Zenios SA (2004) Patient choice in kidney allocation: The role of the queueing discipline. *Manufacturing Service Oper. Management* 6(4):280–301.

Su X, Zenios SA (2006) Recipient choice can address the efficiency-equity trade-off in kidney transplantation: A mechanism design model. *Management Sci.* 52(11):1647–1660.

Talluri KT, van Ryzin GJ (2006) *The Theory and Practice of Revenue Management* (Springer, New York).

Ünver MU (2010) Dynamic kidney exchange. *Rev. Econom. Stud.* 77(1):372–414.

Yu Y, Chen X, Zhang F (2015) Dynamic capacity management with general upgrading. *Oper. Res.* 63(6):1372–1389.

Zenios SA (1999) Modeling the transplant waiting list: A queueing model with reneging. *Queueing Systems* 31(3):239–251.

Zenios SA, Chertow GM, Wein LM (2000) Dynamic allocation of kidneys to candidates on the transplant waiting list. *Oper. Res.* 48(4):549–569.

Zipkin PH (2000) *Foundations of Inventory Management* (McGraw-Hill, New York).