Submitted to
manuscript

Multi-Objective Online Ride-Matching

Guodong Lyu
NUS Business School, National University of Singapore

Wang Chi Cheung
Department of Industrial Systems Engineering and Management, National University of Singapore

Chung-Piaw Teo
NUS Business School and Institute of Operations Research and Analytics, National University of Singapore

Hai Wang*
Heinz College of Information Systems and Public Policy, Carnegie Mellon University

We study the following multi-period multi-objective online ride-matching problem. A ride-sourcing platform needs to match passengers and drivers in real time without observing future information, considering multiple objectives such as platform revenue, pick-up distance, and service quality. We develop an efficient online matching policy that adaptively balances the trade-offs among multiple objectives in a dynamic setting, and provide theoretical performance guarantee for the policy. We prove that the proposed adaptive matching policy can achieve a "target-based optimal solution", i.e., a solution that minimizes the Euclidean distance to any pre-determined multi-objective target. Specifically, the outcome under our policy converges to the "compromise solution" if we set the utopia point as the target. Through numerical experiments and industrial testing using real data from a ride-sourcing platform, we demonstrate that our approach indeed obtains solutions that are closest to the pre-determined targets under various settings, in comparison to existing approaches. The policy presents solutions with delicate balance among multiple objectives and brings value to all the stakeholders in the ride-sourcing ecosystem comparing to benchmark policies: (1) drivers with higher service scores are dispatched with more orders and receive higher incomes; (2) passengers are more likely to be served by drivers with higher service scores, and passengers with higher order revenues are served with higher answer rates, at the expense of a small increase in pick-up distance; (3) the platform obtains a higher total revenue.

Key words: Multi-objective Optimization; Ride-Matching; Online Algorithm; Target-based Optimal Solution

1. Introduction

In many cities in the world, dynamic ride-sourcing companies, such as Grab, Lyft, Uber and Didi Chuxing, have been able to leverage on Internet-based platforms to conduct

* Corresponding author: haiw@andrew.cmu.edu

Electronic copy available at: https://ssrn.com/abstract=3356823

on-demand transportation service. These online platforms facilitate the integration of passengers' and drivers' mobility data on smart phones in real time, which enables a convenient matching between passengers and drivers. These online platforms also dispense with the need for drivers to cruise around to search for passengers in the streets. These clear operational advantages have motivated many similar shared service business models in the public transportation arena, and have been a disruptive force to the conventional taxi industry. Didi Chuxing, for instance, has grown from a ride-hailing service provider to a one-stop mobile transportation platform, offering a variety of services from Taxi, Express, Premier, Luxe, Hitch, Bus, Minibus, Designated Driving, Car Rental, Enterprise Solutions and Bike-Sharing etc., to more than 550 million users in China (Data as of July, 2019).

Matching passengers (demand) with drivers (supply) in real-time is a challenging problem for the ride-sourcing platforms. The "closest distance" policy, as a commonly used matching policy, assigns a passenger to the "nearest" available driver, based on the pick-up distance (or pick-up time) estimated from passenger's and driver's location and surrounding traffic conditions (Özkan and Ward 2016). In practice, platforms also consider many other objectives in the matching decisions. One important objective is platform revenue. Platforms usually give priority to passengers with higher order revenues in the matching decisions, and such preference could bring higher total revenue and profit to the platforms. Another important objective is service quality. Patforms also would like to give priority to drivers who provide higher service quality, and such preference could improve the overall service in the ride-sourcing system, encourage drivers to provide better service and improve the long-term reputation of the platforms.

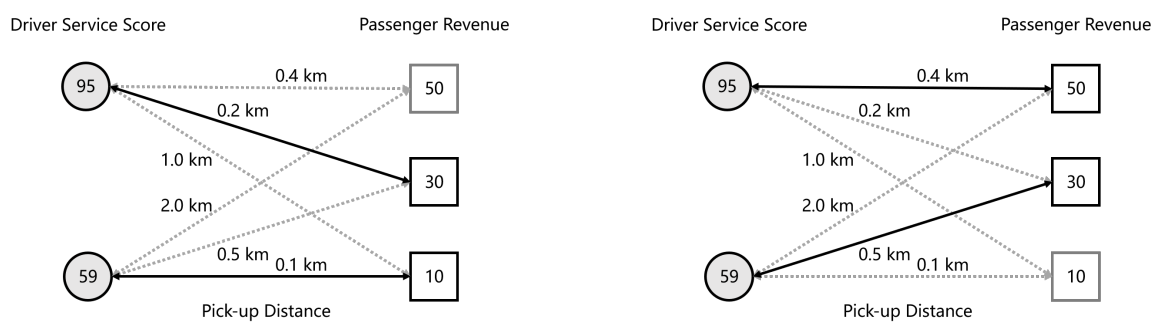
In general, the platforms need to consider multiple objectives when matching passengers and drivers. However, different objectives may generate conflicting matching results. To better illustrate the trade-offs between multiple objectives, we refer to Example 1 for different matching results with the objective of minimizing pick-up distance and the objective of maximizing revenue, when there are more passengers than idle drivers.

Example 1. As shown in Figure 1, assuming three passengers (with order revenue 50, 30, and 10, respectively) are requesting ride-sourcing service and two idle drivers (with service quality measured with score 95 and 59, respectively) are available. Pick-up distances between the first driver and the three passengers are 0.4, 0.2, and 1.0 kilometers, respectively, and between the second driver and the three passengers are 2.0, 0.5, and 0.1

Electronic copy available at: https://ssrn.com/abstract=3356823

kilometers, respectively. (1) If the objective is to minimize pick-up distance, the first driver will be matched to the second passenger and the second driver to the third passenger. The total pick-up distance is 0.3 kilometers and the total revenue is 40. (2) If the objective is to maximize revenue, the first driver will be matched to the first passenger and the second driver to the second passenger. The total pick-up distance is 0.9 kilometers and the total revenue is 80. ■

Figure 1 Matchings with Different Objectives.



(a) Minimize Pick-up Distance (b) Maximize Revenue

Is there a good way to incorporate multiple objectives, e.g., pick-up distance, platform revenue, and service quality, in the matching decisions? A common method is to find an appropriate way to combine the various objectives into a single "merged" objective. However, finding and calibrating the right weights on different objectives is laborious and non-trivial, and there is a lack of theory to guide the selection of such weights in an online setting. To address these challenges, in this paper, we present a novel online matching policy that simultaneously considers multiple objectives in a balanced manner. More precisely, we aim to achieve a solution that has the smallest deviation (in terms of the Euclidean distance) to any pre-determined multi-objective target. We define this solution as the *target-based optimal solution*. This type of target-based solution is widely favored in the multi-objective optimization literature. For example, if we set the "utopia point"[1]—which maximizes the performance of all objectives simultaneously but is in general unattainable—as the target, the target-based optimal solution in such a case is also named as *compromise*

[1] To avoid confusion, we follow the definition in Yu (1973) to define the "utopia point", while we remark that the utopia point is also named as "ideal point" in the multi-objective optimization literature (e.g., Bolanda et al. 2016).

Electronic copy available at: https://ssrn.com/abstract=3356823

solution (Yu 1973, Voorneveld et al. 2011). The formal mathematical definition of the target-based optimal solution will be presented in Section 3.

In the ride-sourcing system, if all the matching scenarios across multiple periods are deterministic and known in advance, it may be feasible to obtain the target-based optimal solution by solving a (deterministic) quadratic programming problem, i.e., by minimizing the Euclidean distance from the attained solution to the target. However, in the online decision problem, the matching scenarios are stochastic and are only revealed to the platforms sequentially over time. Therefore, we are essentially unaware of the specific configuration of the efficient frontier for the multi-objective optimization problem. In this paper, we address the challenges and focus on the multi-objective online ride-matching problem without observing future information. Our methods can be extended to solve other multi-objective optimization problems in an online fashion.

Based on the multi-objective optimization and online convex optimization techniques, our key contributions are summarized as follows:

1. This is the first paper, to the best of our knowledge, to address the online multi-objective ride-matching problem under the framework of online convex optimization. We develop a novel online matching policy to adaptively balance the trade-offs between multiple objectives given any pre-determined target in the ride-sourcing system. We call our policy as *adaptive matching* (AM) policy hereafter.
2. We provide theoretical performance guarantee for the proposed AM policy. We prove that the online solution under the AM policy converges to the target-based optimal solution, i.e., the solution closest to the pre-determined target. The additive error caused by the AM policy increases sub-linearly with the time horizon T, and converges to 0 asymptotically with a sufficiently large T. More precisely, the error term is upper bounded by $O(\log T)$, which is the fastest convergence in literature.
3. We develop a ride-matching simulator to perform extensive numerical experiments and validate that the solutions under the AM policy are closest to the pre-determined targets under various settings. We also implement the AM policy using real ride-sourcing data from our industry partner. Compared to the most natural benchmark "legacy" policy and the widely studied "closest distance" policy, our approach obtains a delicate balance of multiple objectives and brings value to all the stakeholders in the ride-sourcing ecosystem.

Electronic copy available at: https://ssrn.com/abstract=3356823

The rest of this paper is organized as follows. Relevant literature is reviewed in Section 2. We formally formulate the multi-objective online ride-matching problem in Section 3 and describe the online adaptive matching policy in Section 4. Three classes of benchmark policies are introduceds in Section 5. In Section 6, we provide extensive numerical experiments to validate the performance of the proposed matching policy. In Section 7, we implement the proposed policy using real ride-sourcing data, taking into account traffic conditions, passenger patience and cancellation, driver routing and travel times in real road network. Section 8 concludes this paper.

2. Literature Review

We review three streams of literature that are closely relevant to this work, namely, that of ride-matching, multi-objective optimization, and online convex optimization.

Ride-Matching. A surge of efforts have been devoted to designing diverse matching policies and algorithms for the two-sided ride-sourcing market, whereas the majority of these works focused on a single-objective optimization problem. For example, Hu and Zhou (2016) studied the dynamic matching control of a two-sided, discrete-time matching system in which both the demand and supply may leave the platform if the waiting time before getting drivers or passengers are too long, with an objective of maximizing the expected total discounted profit. Özkan and Ward (2016) proposed a linear programming based matching policy that accounts for temporal changing demand and supply and passenger patience, with an objective of maximizing the overall number of passengers being served. Braverman et al. (2017) studied the empty-car routing problem, with an objective of maximizing the availability of empty cars when passengers request ride-sourcing service. In addition, Wang et al. (2017) introduced the concept of stability in the dynamic hitch ride-sharing system and provided mathematical programming approaches to solve stable and nearly stable matching problems, with an objective of minimizing the pick-up detour distance. Akbarpour et al. (2017) studied the dynamic matching problem using some random-graph techniques and they targeted to minimize the total number of perished agents. Xu et al. (2018) studied order dispatch problem in on-demand ride-hailing platforms using reinforcement learning techniques, with an objective of maximizing the immediate reward and future gain.

In the ride-sourcing markets, three parties—passengers, drivers, and the platform—are stakeholders closely related to each other. The matching decisions need to take into account

Electronic copy available at: https://ssrn.com/abstract=3356823

the trade-offs between multiple objectives (or managerial interests in different development stages) . However, as far as we know, few studies shed light on the multi-objective online matching problem in the ride-sourcing markets.

Multi-objective Optimization. A fundamental challenge in multi-objective optimization is to address the trade-offs between multiple criteria. Although it is common to use a set of fixed weights on different objectives to measure their relative importance (e.g., Tsiporkova and Boeva 2006, Marler and Arora 2010), the selection and calibration of such weights are non-trivial. An alternative approach is to characterize a set of Pareto optimal solutions so that the decision maker can choose the most "preferred" one. The space composed of these Pareto optimal solutions is called "efficient frontier" in the literature (Steuer 1986, Yu 1973). For any two Pareto optimal solutions on the efficient frontier, each of them cannot be strictly dominated by the other (Hu and Mehrotra 2012). In general, the Pareto optimal solutions are not unique and different solutions on the efficient frontier can be obtained by varying the weights to these objectives. It would be computationally challenging to enumerate all the solutions on the efficient frontier (Masin and Bukchin 2008). To overcome this difficulty, several filtering techniques have been applied to find a partial set of solutions that can represent all trade-offs on the efficient frontier (e.g., Das and Dennis 1998, Masin and Bukchin 2008). Boland et al. (2017) developed an efficient searching-based algorithm to enumerate the nondominated points for a class of multi-objective integer programming problems. Notably, with different side constraints or utility functions, decision maker would select different Pareto solutions.

Among all the Pareto solutions, Yu (1973) introduced the concept of compromise solution and focused on ℓ_p norms functions to study the distance from the compromise solution to the utopia point. Gearhart (1979) and White (1984) extended this topic by considering different distance functions. When the distance function is given, it is straightforward to derive the compromise solution in the setting where both objective function and feasible domain are well characterized. However, when the feasible domain changes over time, characterizing the compromise solution becomes challenging. In this paper, we develop an efficient online adaptive matching policy to guarantee the compromise solution, and the target-based optimal solution in general, without explicitly characterizing the efficient frontier. To the best of our knowledge, this is the first paper to use online approaches to derive the target-based optimal solution in the multi-objective optimization literature.

Electronic copy available at: https://ssrn.com/abstract=3356823

Online Convex Optimization. The analytical results derived for the proposed online matching policy rely on the online convex optimization (OCO) techniques, which have been widely studied to solve the sequential decision problems over multiple periods. The online gradient descent algorithm, developed by Zinkevich (2003), presents one of the well-known results with sub-linearly increasing regret bounds. In a similar vein, when the loss functions are strongly convex, a logarithmic regret bound can be achieved (e.g., Cesa-Bianchi and Lugosi 2006, Shalev-Shwartz 2011).

Under the OCO framework, Mahdavi et al. (2013) studied the multi-objective optimization problem and reformulated a stochastic constrained problem by forcing different thresholds on the original objective functions. A primal-dual online algorithm was developed to solve this reformulated multi-objective problem. Uziel and El-Yaniv (2017) extended this multi-objective work to the case where the underlying unknown process is stationary and ergodic and proposed a minimax histogram aggregation algorithm to solve this problem. We note that the decision domain is deterministic over the whole time horizon in these two works, while the feasible domain at each matching period changes over time in our problem. This new feature results in different analysis. To this end, we note that our online ride-matching problem is related to the "feasibility" problem studied by Agrawal and Devanur (2015) and Lyu et al. (2018). Agrawal and Devanur (2015) studied a class of online stochastic problem with feasibility constraints which require the average performance of the online solutions lies in (or as close to) a convex set. Lyu et al. (2018) studied the capacity allocation problem in a production network setting, with the Type-II service level targets on demand sides being explicitly specified. The analysis and insights derived in this paper are different from the aforementioned works. Whenever the pre-determined targets are unattainable, we explicitly characterized the performance of the proposed online matching policy, which was not captured by both Agrawal and Devanur (2015) and Lyu et al. (2018). In addition, we note that using OCO to address the ride-matching problem is novel in the literature on shared transportation.

3. Problem Description and Formulation

We study the multi-period multi-objective online ride-matching problem faced by ride-sourcing platforms in a batch-matching setting. The random matching scenario $\boldsymbol{\omega}_t$ is revealed to the platform sequentially over period $t = 1, 2, \ldots, T$, and the decision maker

Electronic copy available at: https://ssrn.com/abstract=3356823

needs to match passengers $i \in \{1, 2, \ldots, N(\boldsymbol{\omega}_t)\}$ and drivers $j \in \{1, 2, \ldots, M(\boldsymbol{\omega}_t)\}$ in batches at the end of each period. Both $N(\boldsymbol{\omega}_t)$ and $M(\boldsymbol{\omega}_t)$ are random variables that depends on the random scenario $\boldsymbol{\omega}_t$. For ease of exposition, we assume that $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_T$ are independent and identically distributed (i.i.d.) generated according to an *unknown* distribution Ψ, which has a potentially infinite support set $\boldsymbol{\Omega}$.

We denote the bipartite network of this two-sided market at period t as $\mathcal{G}(\boldsymbol{\omega}_t)$. We allow passenger i to be matched to driver j only if the pick-up distance $d_{i,j}$ is smaller than a tolerance threshold (e.g., 3 km); i.e., the pair $(i, j) \in \mathcal{G}(\boldsymbol{\omega}_t)$ only if $d_{i,j} \leq 3$ km. More precisely, the feasible matching region $\mathcal{X}(\boldsymbol{\omega}_t)$ at period t can be formulated as:

$$\mathcal{X}(\boldsymbol{\omega}_t) := \left\{ \boldsymbol{x} \in \mathbb{R}^{|N(\boldsymbol{\omega}_t)| \times |M(\boldsymbol{\omega}_t)|} \;\middle|\; \begin{array}{rl} \sum_{j=1}^{M(\boldsymbol{\omega}_t)} x_{i,j} \leq & 1, \forall i = 1, 2, \ldots, N(\boldsymbol{\omega}_t) \\ \sum_{i=1}^{N(\boldsymbol{\omega}_t)} x_{i,j} \leq & 1, \forall j = 1, 2, \ldots, M(\boldsymbol{\omega}_t) \\ x_{i,j} \in & \{0, 1\}, \forall (i, j) \in \mathcal{G}(\boldsymbol{\omega}_t) \\ x_{i,j} = & 0, \forall (i, j) \notin \mathcal{G}(\boldsymbol{\omega}_t) \end{array} \right\},$$

where the first set of constraints requires that one passenger can be picked up by at most one driver, and the second set requires that one driver can be dispatched to at most one passenger. The decision maker aims to address the following multi-objective ride-matching problem:

$$\begin{aligned} (\mathrm{M}_0)\ \max & \left\{ \frac{1}{T}\left(\sum_{t=1}^{T} h_1(\boldsymbol{x}_t, \boldsymbol{\omega}_t)\right), \frac{1}{T}\left(\sum_{t=1}^{T} h_2(\boldsymbol{x}_t, \boldsymbol{\omega}_t)\right), \ldots, \frac{1}{T}\left(\sum_{t=1}^{T} h_K(\boldsymbol{x}_t, \boldsymbol{\omega}_t)\right) \right\} \\ \text{s.t.}\ & \boldsymbol{x}_t \in \mathcal{X}(\boldsymbol{\omega}_t),\ t = 1, 2, \ldots, T \\ & \boldsymbol{x}_t \text{ non-anticipative},\ t = 1, 2, \ldots, T. \end{aligned}$$

The function $h_k(\cdot, \boldsymbol{\omega}) : \mathcal{X}(\boldsymbol{\omega}) \to \mathbf{R}$ represents the kth objective function under the matching scenario $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, and $h_k(\cdot, \boldsymbol{\omega})$ is assumed to be concave in the decision variable. For example, we can denote the kth objective function $h_k(\boldsymbol{x}, \boldsymbol{\omega}) = \sum_{(i,j) \in \mathcal{G}(\boldsymbol{\omega})} r_i(\boldsymbol{\omega}) x_{i,j}$ as the total revenue using the matching decision $\boldsymbol{x}$ under the scenario $\boldsymbol{\omega}$, where $r_i(\boldsymbol{\omega})$ represents the revenue obtained by serving passenger i. In the *online* problem (M_0), each decision $\boldsymbol{x}_t$ is made at period t only using the historical information $\{\boldsymbol{x}_s, \boldsymbol{\omega}_s\}_{s=1}^{t-1} \cup \{\boldsymbol{\omega}_t\}$ available

Electronic copy available at: https://ssrn.com/abstract=3356823

up to period t, while the future information $\{\boldsymbol{\omega}_{t+1}, \boldsymbol{\omega}_{t+2}, \ldots, \boldsymbol{\omega}_T\}$ is unknown and cannot be used. Equivalently, the decision variable $\boldsymbol{x}_t$ is $\sigma(\{\boldsymbol{x}_s, \boldsymbol{\omega}_s\}_{s=1}^{t-1} \cup \{\boldsymbol{\omega}_t\})$-measurable for all $t = 1, 2, \ldots, T$. For this purpose, we say $\boldsymbol{x}_t$ is *non-anticipative*. We note that online optimization is attractive in practice as it is unnecessary to track the states of the whole system over time, which, in contrast, is required in dynamic programming methods. This allows us to solve the large-scale matching problems efficiently in real time.

To balance the trade-offs between multiple objectives (i.e., different $h_k(\cdot,\cdot)$'s), we formulate a new problem (M_1) with any *pre-determined multi-objective target* $\boldsymbol{U} = (U_1, U_2, \ldots, U_K) \in \mathbf{R}^K$. This problem seeks a sequence of online matching decisions $\{\boldsymbol{x}_t\}_{t=1}^T$ that attains the *target-based optimal solution* with objective performance $\boldsymbol{\tau} = (\tau_1, \tau_2, \ldots, \tau_K) \in \mathbf{R}^K$ closest (in terms of the Euclidean distance $\|\cdot\|_2$) to the pre-determined target $\boldsymbol{U}$. Mathematically, we define the target-based optimal solution as the optimal solution to problem (M_1).

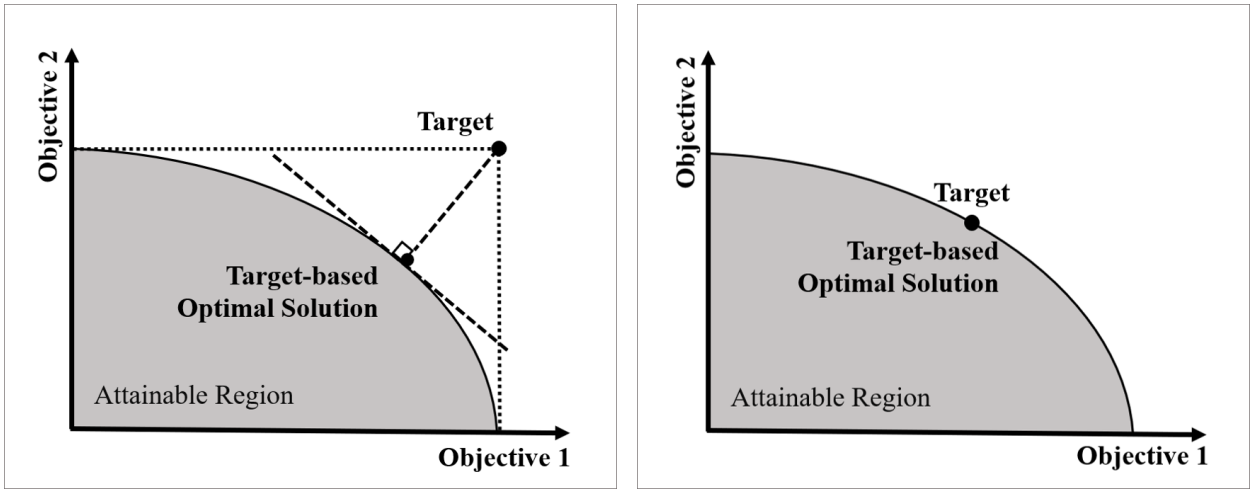
$$
\begin{aligned}
(\mathrm{M}_1) \quad \min \quad & \|(\boldsymbol{U} - \boldsymbol{\tau})^+\|_2 \\
\text{s.t.} \quad & \frac{1}{T}\left(\sum_{t=1}^{T} h_k(\boldsymbol{x}_t, \boldsymbol{\omega}_t)\right) \geq \tau_k, \ k = 1, 2, \ldots, K \\
& \boldsymbol{x}_t \in \mathcal{X}(\boldsymbol{\omega}_t), \ t = 1, 2, \ldots, T \\
& \boldsymbol{x}_t \text{ non-anticipative}, \ t = 1, 2, \ldots, T.
\end{aligned}
$$

The positive function $(\boldsymbol{U} - \boldsymbol{\tau})^+$ returns the component-wise positive part of the vector $(\boldsymbol{U} - \boldsymbol{\tau})$, i.e., $(U_k - \tau_k)^+ = \max(U_k - \tau_k, 0)$. In this way, the Euclidean distance function $\|\cdot\|_2$ imposes a penalization if the attained objective performance τ_k is less than the corresponding target U_k for a component $k \in \{1, 2, \ldots, K\}$.

Figure 2 illustrates the concept of target-based optimal solution. Assuming there are two objectives to be maximized and the shaded area characterizes the attainable region. In (a), if the pre-determined multi-objective target is unattainable, the target-based optimal solution achieves the point that is closest to the unattainable target. In (b), if the pre-determined target is attainable, the target-based optimal solution achieves the target itself. Specifically, if we set target $\boldsymbol{U}$ as the *utopia point*, i.e., U_k is the expected optimal value (the expectation is taken over $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_T \sim \Psi$) in problem ($\mathrm{M}_0$) when $\frac{1}{T}[\sum_{t=1}^T h_k(\boldsymbol{x}_t, \boldsymbol{\omega}_t)]$ is the only objective to maximize, the target-based optimal solution to (M_1) is defined as *compromise solution* in the multi-objective optimization literature (cf. Yu 1973).

Electronic copy available at: https://ssrn.com/abstract=3356823

Figure 2 **Illustration of the target-based optimal solution.**



(a) Unattainable Target

(b) Attainable Target

Note that $\|(\boldsymbol{w})^+\|_2$ is convex in $\boldsymbol{w}$ with $\|(\boldsymbol{w})^+\|_2 = 0$ for all $\boldsymbol{w} \in \mathbf{R}^K_{\leq 0}$. If the scenario realizations $\{\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_T\}$ are known and provided in offline settings, we may obtain a sequence of matching decisions $\{\boldsymbol{x}^*(\boldsymbol{\omega}_t)\}_{t=1}^T$ that minimizes the Euclidean distance to $\boldsymbol{U}$ by solving a large-scale quadratic programming problem. However, this offline optimal decisions $\{\boldsymbol{x}^*(\boldsymbol{\omega}_t)\}_{t=1}^T$ are scenario-dependent and cannot be applied in online decision making since the matching scenarios $\{\boldsymbol{\omega}_t\}_{t=1}^T$ are randomly generated and revealed sequentially over time. Additionally, even we assume that all the matching scenarios are provided in advance, it is still computationally difficult to solve the problem (M_1) in the real-life ride-matching environment due to the mega scale of the optimization problem.

To this end, we aim to provide a tractable framework to solve the multi-objective online matching problem (M_1) for the target-based optimal solution, which will be presented in Section 4. More concretely in the ride-matching problem, collaborating with industry partners, we finalize the following three critical objectives:

- **Platform revenue**: Passengers with higher order revenue should be served with higher priority. We denote $r_i(\boldsymbol{\omega})$ as the platform revenue earned by serving passenger i under scenario $\boldsymbol{\omega}$.
- **Pick-up distance**: Pairs of passenger and driver with shorter pick-up distance should be matched with higher priority. We denote $d_{i,j}(\boldsymbol{\omega})$ as the pick-up distance between passenger i and driver j under scenario $\boldsymbol{\omega}$. Note that an important underlying concern is to maximize the total number of matched pairs, so we introduce a big number

Electronic copy available at: https://ssrn.com/abstract=3356823

$d_m > 3$ km and denote $d_m - d_{i,j}(\boldsymbol{\omega})$ as a so-called "saved pick-up distance". Then we transform the real pick-up distance minimization problem into a saved pick-up distance maximization problem.

- **Service quality**: Drivers providing better service (e.g., measured by higher service score[2]) should be dispatched with higher priority. We denote $s_j(\boldsymbol{\omega})$ as the service score of the dispatched driver j under scenario $\boldsymbol{\omega}$.

Therefore, for any matching scenario $\boldsymbol{\omega}$ and feasible solution $\boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega})$, we let

$$h_1(\boldsymbol{x}, \boldsymbol{\omega}) := \beta_1 \left[\sum_{(i,j) \in \mathcal{G}(\boldsymbol{\omega})} r_i(\boldsymbol{\omega}) x_{i,j} \right],$$

$$h_2(\boldsymbol{x}, \boldsymbol{\omega}) := \beta_2 \left[\sum_{(i,j) \in \mathcal{G}(\boldsymbol{\omega})} \left(d_m - d_{i,j}(\boldsymbol{\omega})\right) x_{i,j} \right], \text{ and}$$

$$h_3(\boldsymbol{x}, \boldsymbol{\omega}) := \beta_3 \left[\sum_{(i,j) \in \mathcal{G}(\boldsymbol{\omega})} s_j(\boldsymbol{\omega}) x_{i,j} \right]$$

denote the total revenue earned, total saved pick-up distance, and total service score, respectively. Note that the magnitudes of the three objective performances are different. Following industry practices suggested by industry partners, we can set the revenue as the baseline and normalize the saved pick-up distance and driver service score to similar magnitudes by introducing certain parameters $\{\beta_1, \beta_2, \beta_3\}$, which do not affect the theoretical properties of the proposed policy and can be adjusted anyway.

4. Online Adaptive Matching Policy

In this section, we develop an online adaptive matching policy to solve problem (M_1). If the pre-determined target is attainable (i.e., can be attained by some non-anticipative policies), the AM policy can converge to achieve such a target; if the pre-determined target is unattainable, the AM policy can converge to the solution that is closest to the target, i.e., converge to the target-based optimal solution.

We first introduce some key concepts that are important in our online policy. $\alpha_k(s)$ represents the *debt* for objective k at period s, i.e., the difference owe to target U_k by performance $h_k(\cdot, \cdot)$:

$$\alpha_k(s) := U_k - h_k(\boldsymbol{x}_s, \boldsymbol{\omega}_s). \tag{1}$$

[2] Our industry partner uses a so-called "service score" to evaluate the performance and service quality of drivers. The service score is determined by a driver's ratings received from passengers, his historical cancellation rates, and some other detected driving behaviors (e.g., over-speed driving). A higher service score indicates better performance and higher service quality of a driver.

Electronic copy available at: https://ssrn.com/abstract=3356823

At the beginning of period $(t+1)$, we can track the average debt, $w_k(t+1)$, from period 1 to t for each objective $k \in \{1, 2, \ldots, K\}$ as:

$$w_k(t+1) := \frac{1}{t} \sum_{s=1}^{t} \alpha_k(s). \tag{2}$$

Intuitively, if $w_k(t+1) > 0$, it means that the previous decisions over the first t periods have not achieved the required target U_k at the kth objective, and hence the decision at period $(t+1)$ should give a high priority at the kth objective function $h_k(\cdot, \boldsymbol{\omega}_{t+1})$. The larger value of the positive $w_k(t+1)$, the higher priority we should give to the kth objective. Otherwise, if $w_k(t+1) \leq 0$, it means that the previous decisions have already met the required target U_k, and hence the decision at period $(t+1)$ does not need to focus on the kth objective function $h_k(\cdot, \boldsymbol{\omega}_{t+1})$. Altogether, the average debt vector $\boldsymbol{w}(t+1)$ quantifies the relative performance comparing to the target $\boldsymbol{U}$ at multiple objectives, and suggests the priority associated with each of the K objective functions $h_1(\cdot, \boldsymbol{\omega}_{t+1}), h_2(\cdot, \boldsymbol{\omega}_{t+1}), \ldots, h_K(\cdot, \boldsymbol{\omega}_{t+1})$ at period $(t+1)$.

We integrate the intuition described above rigorously in the following online Adaptive Matching policy (AM policy):

<u>Online Adaptive Matching Policy:</u>

For $t = 0, 1, \ldots,$ do:

1. At the beginning of period $(t+1)$, compute the average debt vector $\boldsymbol{w}(t+1) = \{w_k(t+1)\}_{k=1}^{K}$, based on Equation (2).
2. After observing the realized matching scenario $\boldsymbol{\omega}_{t+1}$ at period $(t+1)$, make the matching decision $\boldsymbol{x}_{t+1}^{\text{AM}}$ of period $(t+1)$ as the optimal solution to the following convex optimization problem:

$$\begin{aligned} (\text{AM}(t+1)) \ \max \ & \sum\nolimits_{k=1}^{K} w_k^{+}(t+1) h_k(\boldsymbol{x}_{t+1}, \boldsymbol{\omega}_{t+1}) \\ \text{s.t.} \ \ & \boldsymbol{x}_{t+1} \in \mathcal{X}(\boldsymbol{\omega}_{t+1}) \end{aligned}$$

The kth objective value under the AM policy is expressed as $h_k(\boldsymbol{x}_{t+1}^{\text{AM}}, \boldsymbol{\omega}_{t+1})$.

The online adaptive matching policy can be easily applied to solve large-scale ride-matching problems in real time. It also has good theoretical properties. To present theoretical performance guarantee of the AM policy, we assume that the objective val-

Electronic copy available at: https://ssrn.com/abstract=3356823

ues are bounded under all matching scenarios. Specifically, we define a constant $\Delta_k := \max_{\boldsymbol{\omega}\in\boldsymbol{\Omega},\boldsymbol{x}(\boldsymbol{\omega})\in\mathcal{X}(\boldsymbol{\omega})} h_k(\boldsymbol{x}(\boldsymbol{\omega}),\boldsymbol{\omega})$ for each $k \in \{1,2,\ldots,K\}$, and define constant

$$\Delta := \sqrt{\sum_{k=1}^{K} \Delta_k^2}. \tag{3}$$

Note that Δ only depends on the ranges of objective values, not on the time horizon T.

Next, we provide a Lemma to facilitate the comparison between the AM policy and the offline optimal policy for problem (M_1). More concretely, we denote $y_k^* := \mathbf{E}_{\boldsymbol{\omega}\sim\Psi}[h_k(\boldsymbol{x}^*(\boldsymbol{\omega}),\boldsymbol{\omega})]$ as the performance for each objective function $k \in \{1,2,\ldots,K\}$ with the target-based optimal solution $\{\boldsymbol{x}^*(\boldsymbol{\omega})\}_{\boldsymbol{\omega}\in\boldsymbol{\Omega}}$ to a single-period problem. Let $\boldsymbol{w}^* = (w_1^*, w_2^*, \ldots, w_K^*)$ denote the *optimal debt vector* as $w_k^* = U_k - y_k^*$. In other words, $\boldsymbol{w}^*$ is the debt vector correponding to the target-based optimal solution to a single-priod problem. For a given period t and a realization of average debt vector $\boldsymbol{w}(t)$ from period 1 to $(t-1)$, we denote

$$y_k^{\mathrm{AM}}(t) := \mathbf{E}_{\boldsymbol{\omega}_t\sim\Psi}\left[h_k(\boldsymbol{x}^{\mathrm{AM}}(t),\boldsymbol{\omega}_t)\middle|\boldsymbol{w}(t)\right] \tag{4}$$

as the expected performance under the AM policy for each $k \in \{1,2,\ldots,K\}$. The conditional expectation in (4) is taken over only on the randomness of the period t scenario $\boldsymbol{\omega}_t$, which is distributed as Ψ, while the average debt $\boldsymbol{w}(t)$ is held deterministic.

Lemma 1. *For any given period t and any average debt vector $\boldsymbol{w}(t)$, the following inequality holds with certainty:*

$$\boldsymbol{w}^+(t)^\top \mathbf{E}\left[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)\right] \le \boldsymbol{w}^+(t)^\top \boldsymbol{w}^*. \tag{5}$$

Proof. Accorded by the AM policy, we have

$$\sum_{k=1}^{K} w_k^+(t) h(\boldsymbol{x}_t^{\mathrm{AM}},\boldsymbol{\omega}_t) \ge \sum_{k=1}^{K} w_k^+(t) h(\boldsymbol{x}(\boldsymbol{\omega}_t),\boldsymbol{\omega}_t)$$

for any $\boldsymbol{x}(\boldsymbol{\omega}_t) \in \mathcal{X}(\boldsymbol{\omega}_t)$, conditional on the average debt vector $\boldsymbol{w}(t)$. In particular, given the optimal solution $\{\boldsymbol{x}^*(\boldsymbol{\omega})\}_{\boldsymbol{\omega}\in\boldsymbol{\Omega}}$ for a single-period problem, we also have

$$\sum_{k=1}^{K} w_k^+(t) h(\boldsymbol{x}_t^{\mathrm{AM}},\boldsymbol{\omega}_t) \ge \sum_{k=1}^{K} w_k^+(t) h(\boldsymbol{x}^*(\boldsymbol{\omega}_t),\boldsymbol{\omega}_t).$$

Electronic copy available at: https://ssrn.com/abstract=3356823

Since $\boldsymbol{w}(t)$ is independent of $\boldsymbol{\omega}_t$, taking expectation over $\boldsymbol{\omega}_t \sim \Psi$ yields the following inequality:

$$\sum_{k=1}^{K} w_k^+(t) y_k^{\mathrm{AM}}(t) \geq \sum_{k=1}^{K} w_k^+(t) y_k^*.$$

Furthermore, by the definition of debt $\alpha_k(t) = U_k - h_k(\boldsymbol{x}_t^{\mathrm{AM}}, \boldsymbol{\omega}_t)$, we obtain the inequality (5). ■

Lemma 1 is proved by using the fact that the AM policy provides the optimal solution to problem AM(t) conditional on the average debt vector $\boldsymbol{w}(t)$, which plays a key role in analyzing the convergence from $\boldsymbol{w}^+(T+1)$ to $(\boldsymbol{w}^*)^+$. Next we present the main theorem on the theoretical performance guarantee of the AM policy.

Theorem 1. *Consider the multi-period multi-objective online ride-matching problem* (M_1). *Debt vector* $\boldsymbol{w}(T+1)$ *under the AM policy converges to the optimal debt vector* $\boldsymbol{w}^*$. *Specifically, it satisfies the following non-asymptotic convergence guarantee in expectation:*

$$\mathbf{E}\left[\|\boldsymbol{w}^+(T+1)\|_2^2\right] - \|(\boldsymbol{w}^*)^+\|_2^2 \leq \frac{\Delta^2(1+\log T)}{T}, \tag{6}$$

where the expectation in Equation (6) is taken over $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_T$.

Proof. This theorem is proved by tracking the decreasing rate in $\|\boldsymbol{w}^+(t+1)\|_2^2$. To make the discussion compact, we relegate some useful inequalities (as stated in Claim 1 and 2) in the end of this section, and sketch the key results to support the proof.

Denote $\boldsymbol{w}(0) = 0$ for notational convenience. Then, for any $0 < t \leq T-1$, we have

$$\begin{aligned}
&\|\boldsymbol{w}^+(t+1)\|_2^2 \\
\leq &\|\boldsymbol{w}^+(t)\|_2^2 + 2\boldsymbol{w}^+(t)^\top\left[\boldsymbol{w}(t+1) - \boldsymbol{w}(t)\right] + \|\boldsymbol{w}(t+1) - \boldsymbol{w}(t)\|_2^2 && (7)\\
= &\|\boldsymbol{w}^+(t)\|_2^2 + \frac{2\boldsymbol{w}^+(t)^\top[\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)]}{t} + \frac{\|\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)\|_2^2}{t^2} && (8)\\
\leq &\|\boldsymbol{w}^+(t)\|_2^2 + \frac{2\boldsymbol{w}^+(t)^\top[\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)]}{t} + \frac{\Delta^2}{t^2} && (9)\\
= &\|\boldsymbol{w}^+(t)\|_2^2 + \frac{2\boldsymbol{w}^+(t)^\top\left\{\mathbf{E}\left[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)\right] - \boldsymbol{w}(t)\right\}}{t} + \frac{2\boldsymbol{w}^+(t)^\top\left\{\boldsymbol{\alpha}(t) - \mathbf{E}\left[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)\right]\right\}}{t} + \frac{\Delta^2}{t^2}. && (10)
\end{aligned}$$

Step (7) is due to the second inequality in Claim 2 (with $\boldsymbol{q} \to \boldsymbol{w}(t+1)$, $\boldsymbol{p} \to \boldsymbol{w}(t)$). Step (8) directly follows the definitions of $\boldsymbol{w}(t)$ and $\boldsymbol{\alpha}(t)$, i.e., $\boldsymbol{w}(t+1) - \boldsymbol{w}(t) = \frac{1}{t}\left[\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)\right]$. In step (9), we bound the term $\|\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)\|_2^2/t^2$ by using the assumption $\|\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)\|_2 \leq \Delta$,

Electronic copy available at: https://ssrn.com/abstract=3356823

as specified in Equation (3). Finally, in step (10), we decompose the term $\{\boldsymbol{\alpha}(t) - \boldsymbol{w}(t)\}$ by introducing the conditional expectation $\mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]$.

Next, using Lemma 1, we are able to bound (10):

$$\begin{aligned}\|\boldsymbol{w}^+(t+1)\|_2^2 &\le \|\boldsymbol{w}^+(t)\|_2^2 + \frac{2\boldsymbol{w}^+(t)^\top[\boldsymbol{w}^* - \boldsymbol{w}(t)]}{t} + \frac{2\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\}}{t} + \frac{\Delta^2}{t^2} \\ &\le \|\boldsymbol{w}^+(t)\|_2^2 + \frac{\|(\boldsymbol{w}^*)^+\|_2^2 - \|\boldsymbol{w}^+(t)\|_2^2}{t} + \frac{2\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\}}{t} + \frac{\Delta^2}{t^2},\end{aligned} \tag{11}$$

where step (11) results from the first inequality in Claim 2 (with $\boldsymbol{p} \leftarrow \boldsymbol{w}(t)$ and $\boldsymbol{q} \leftarrow \boldsymbol{w}^*$). It follows that

$$\begin{aligned}\|\boldsymbol{w}^+(t+1)\|_2^2 - \|(\boldsymbol{w}^*)^+\|_2^2 \le \frac{t-1}{t}\left[\|\boldsymbol{w}^+(t)\|_2^2 - \|(\boldsymbol{w}^*)^+\|_2^2\right] \\ + \frac{2\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\}}{t} + \frac{\Delta^2}{t^2}.\end{aligned} \tag{12}$$

Denote $W(t) := t\left(\|\boldsymbol{w}^+(t+1)\|_2^2 - \|(\boldsymbol{w}^*)^+\|_2^2\right)$. We then restate inequality (12) as follows:

$$W(t) \le W(t-1) + 2\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\} + \frac{\Delta^2}{t}. \tag{13}$$

Apply the inequality (13) recursively from $t = T$ to $t = 1$, we have

$$W(T) \le W(0) + 2\sum_{t=1}^{T}\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\} + \sum_{t=1}^{T}\frac{\Delta^2}{t} \tag{14}$$

$$\le 0 + 2\sum_{t=1}^{T}\boldsymbol{w}^+(t)^\top\{\boldsymbol{\alpha}(t) - \mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)]\} + \Delta^2(1 + \log T). \tag{15}$$

Note that $\mathbf{E}[\boldsymbol{\alpha}(t) \mid \boldsymbol{w}(t)] = \mathbf{E}[\boldsymbol{\alpha}(t)]$ under the i.i.d. assumption of the underlying stochastic process. Therefore, taking expectation on both sides of (15) would suffice to provide the desired convergence result in Equation (6) in Theorem 1. ■

With a slight abuse of notation, we use $\{\boldsymbol{x}_t^{\mathrm{AM}}\}_{t=1}^T$ to denote the sequence of matching decisions under the AM policy, which satisfies the second and third sets of constraints in problem (M_1). The non-asymptotic convergence guarantee in Theorem 1 implies that $\mathbf{E}[\|\boldsymbol{w}^+(T+1)\|_2^2] - \|(\boldsymbol{w}^*)^+\|_2^2 \to 0$ as $T \to \infty$. More precisely, let $\boldsymbol{\tau}_T^{\mathrm{AM}} = (\frac{1}{T}\sum_{t=1}^T h_k(\boldsymbol{x}_t^{\mathrm{AM}}, \boldsymbol{\omega}_t))_{k=1}^K$ denote the average objective performance vector under the AM policy and let $\boldsymbol{\tau}^* = (\mathbf{E}[\frac{1}{T}\sum_{t=1}^T h_k(\boldsymbol{x}^*(\boldsymbol{\omega}_t), \boldsymbol{\omega}_t)])_{k=1}^K$ denote the optimal performance that is

Electronic copy available at: https://ssrn.com/abstract=3356823

closest to the pre-determined target $\boldsymbol{U}$ under target-based optimal solutions $\{\boldsymbol{x}^*(\boldsymbol{\omega}_t)\}_{\boldsymbol{\omega}_t \in \boldsymbol{\Omega}}$. Theorem 1 provides the following convergence guarantee for the AM policy:

$$\mathbf{E}_{\boldsymbol{\omega}_1,\boldsymbol{\omega}_2,\ldots,\boldsymbol{\omega}_T \sim \Psi} \left[\|(\boldsymbol{U} - \boldsymbol{\tau}_T^{\mathrm{AM}})^+\|_2^2 \right] - \|(\boldsymbol{U} - \boldsymbol{\tau}^*)^+\|_2^2 \to 0, \; as \; T \to \infty, \tag{16}$$

which indicates that the objective performance vector $\boldsymbol{\tau}_T^{\mathrm{AM}}$ achieved by the AM policy converges to the target-based optimal performance under the pre-determined target $\boldsymbol{U}$.

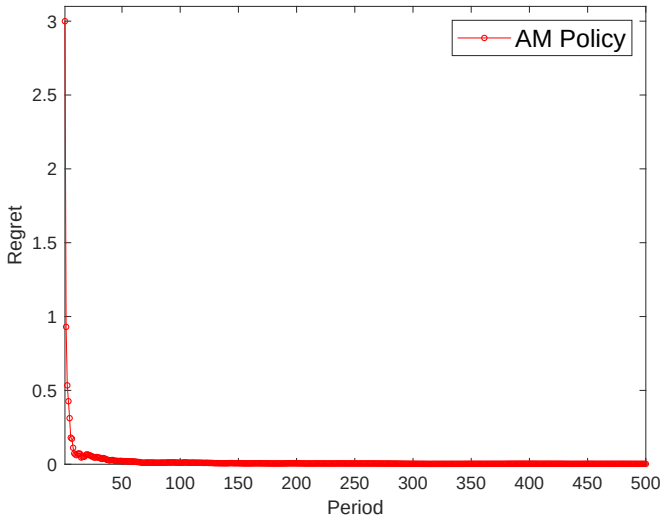
The AM policy has a logarithm convergence rate as shown in Equation (6). Next, we use Example 2 to illustrate the convergence rate.

Example 2. Consider a ride-matching problem over T periods. At each time period $t = 1, 2, \ldots, T$, the random matching scenario $\boldsymbol{\omega}_t$ is i.i.d. generated. For ease of exposition, we assume that the number of drivers $M(\boldsymbol{\omega}_t)$ and the number of passengers $N(\boldsymbol{\omega}_t)$ are fixed to be 10 across periods. The order revenue $r_i(\boldsymbol{\omega}_t)$ for passenger $i \in \{1, 2, \ldots, N(\boldsymbol{\omega}_t)\}$, and the service score $s_j(\boldsymbol{\omega}_t)$ for driver $j \in \{1, 2, \ldots, M(\boldsymbol{\omega}_t)\}$ are uniformly generated between 0 and 1. Similarly, the distance $d_{i,j}(\boldsymbol{\omega}_t)$ between each pair of passenger and driver also follows a uniform distribution between 0 and 1. First, we solve three single-objective optimization problems and derive the Utopia target $\boldsymbol{U}$. Second, we use the Sampling Average Approximation to solve the quadratic programming problem (M_1) to derive the "real" compromise solution $\boldsymbol{\tau}^*$. Note that this approach can only solve small-scale problems. Third, we implement the AM policy to get the performance $\boldsymbol{\tau}_T^{\mathrm{AM}}$. Since both $\boldsymbol{\tau}^*$ and $\boldsymbol{\tau}_T^{\mathrm{AM}}$ are sample dependent, we repeat this experiment for 100 times to compute the expected value of the gap $\|(\boldsymbol{U} - \boldsymbol{\tau}_T^{\mathrm{AM}})^+\|_2^2 - \|(\boldsymbol{U} - \boldsymbol{\tau}^*)^+\|_2^2$, which is defined as *Regret* in the online optimization literature. As shown in Figure 3, we vary period t from 1 to 500 and plot the expected performance gap of the AM policy. The performance gap converges to 0 quickly, e.g., when T exceeds 100. Notably, the convergence rate follows a logarithm style, which is consistent with the result in Theorem 1. ■

We make the following three remarks about the AM policy. First, the AM policy is indeed non-anticipatory, since the average debt vector $\boldsymbol{w}(t+1)$ at period $(t+1)$ is calculated based on the debt from period 1 to t without observing future information. Second, the AM policy does not require to track other status of the ride-sourcing system and can be easily applied to solve large-scale ride-matching problems in real time. Third, the logarithm convergence rate is the fastest in the OCO literature.

Electronic copy available at: https://ssrn.com/abstract=3356823

Figure 3 Convergence Guarantee of the AM Policy.



Note that if the multi-objective target $\boldsymbol{U}$ is attainable by some non-anticipative policies, we must have $\|(\boldsymbol{w}^*)^+\|_2^2 = 0$ and hence $\mathbf{E}\left[\|\boldsymbol{w}^+(T+1)\|_2^2\right] \to 0$ as $T \to \infty$, which implies that the objective performance under the AM policy can achieve the target.

Corollary 1. *Consider the multi-period multi-objective online ride-matching problem* (M_1). *If the multi-objective target is attainable under some non-anticipative policies, the average debt* $w_k(T+1)$ *for each objective* k *under the AM policy converges to 0. Specifically, it satisfies the following non-asymptotic convergence guarantee in expectation:*

$$\mathbf{E}\left[\|w_k^+(T+1)\|_2^2\right] \leq \frac{\Delta^2(1+\log T)}{T}, \ \forall k = 1, 2, \ldots, K, \tag{17}$$

where the expectation in Equation (17) is taken over $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_T$.

Proof. The results directly follow by Theorem 1, and the fact that $\|(\boldsymbol{w}^*)^+\|_2^2 = 0$ whenever the multi-objective target is attainable under some non-anticipative policies, i.e.,

$$\mathbf{E}\left[\|w_k^+(T+1)\|_2^2\right] \leq \mathbf{E}\left[\|\boldsymbol{w}^+(T+1)\|_2^2\right] \leq \frac{\Delta^2(1+\log T)}{T}, \ \forall k = 1, 2, \ldots, K.$$

■

By Corollary 1, the AM policy can be used to examine the attainability of any predetermined target. If the Euclidean norm of the debt vector $\boldsymbol{w}(T+1)$ would not converge to 0 as T increases by using the AM policy, then the corresponding target cannot be attained by any other non-anticipative policies.

Electronic copy available at: https://ssrn.com/abstract=3356823

In the end, we present the inequalities in Claim 1 and 2 that are used to support the proof of Theorem 1.

Claim 1. *Given any two vectors $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^K$, the following two inequalities hold:*

$$\begin{cases} (\boldsymbol{p}^+)^\top [\boldsymbol{q}^+ - \boldsymbol{p}^+] & \geq (\boldsymbol{p}^+)^\top [\boldsymbol{q} - \boldsymbol{p}] \\ 2\,(\boldsymbol{p}^+)^\top [\boldsymbol{q}^+ - \boldsymbol{p}^+] + \|\boldsymbol{q}^+ - \boldsymbol{p}^+\|_2^2 & \leq 2\,(\boldsymbol{p}^+)^\top [\boldsymbol{q} - \boldsymbol{p}] + \|\boldsymbol{q} - \boldsymbol{p}\|_2^2 \end{cases} \tag{18}$$

Proof. It is sufficient to prove the two inequalities in component-wise, i.e., for any $k = 1, 2, \ldots, K$, we claim:

$$p_k^+ \times (q_k^+ - p_k^+) \geq p_k^+ \times (q_k - p_k), \text{ and} \tag{19}$$

$$2 \times p_k^+ \times (q_k^+ - p_k^+) + (q_k^+ - p_k^+)^2 \leq 2 \times p_k^+ \times (q_k - p_k) + (q_k - p_k)^2. \tag{20}$$

Inequality (19) can be easily validated by considering the following four cases:

- If $p_k \geq 0, q_k \geq 0$, then the LHS of inequality (19) is equivalent to the RHS;
- If $p_k \geq 0, q_k < 0$, then inequality (19) reduces to $-p_k^2 \geq -p_k^2 + p_k q_k$, which is true;
- If $p_k < 0, q_k \geq 0$, then both sides of inequality (19) are equal to 0;
- If $p_k < 0, q_k < 0$, then both sides of inequality (19) are also equal to 0.

Similarly, inequality (20) can also be checked case by case:

- If $p_k \geq 0, q_k \geq 0$, then the LHS of inequality (20) is equivalent to the RHS;
- If $p_k \geq 0, q_k < 0$, then inequality (20) reduces to $-2p_k^2 + p_k^2 \leq -2p_k^2 + 2p_k q_k + (q_k - p_k)^2 = -p_k^2 + q_k^2$;
- If $p_k < 0, q_k \geq 0$, then inequality (20) reduces to $q_k^2 \leq q_k^2 + p_k^2 + (-2p_k q_k)$;
- If $p_k < 0, q_k < 0$, then inequality (20) reduces to $0 \leq (q_k - p_k)^2$, which is clearly true.

The inequalities (18) are demonstrated by summing (19) and (20) over all $k = 1, 2, \ldots, K$, respectively. ■

Claim 2. *Given any two vectors $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^K$, the following two inequalities hold:*

$$\begin{cases} \|\boldsymbol{q}^+\|_2^2 - \|\boldsymbol{p}^+\|_2^2 \geq 2(\boldsymbol{p}^+)^\top [\boldsymbol{q} - \boldsymbol{p}] \\ \|\boldsymbol{q}^+\|_2^2 - \|\boldsymbol{p}^+\|_2^2 \leq 2(\boldsymbol{p}^+)^\top [\boldsymbol{q} - \boldsymbol{p}] + \|\boldsymbol{q} - \boldsymbol{p}\|_2^2. \end{cases} \tag{21}$$

Proof. Note that the function $\psi(\boldsymbol{x}) := \|\boldsymbol{x}\|_2^2$ is both differentiable and convex, it is straightforward to see $\|\boldsymbol{q}^+\|_2^2 - \|\boldsymbol{p}^+\|_2^2 \geq 2(\boldsymbol{p}^+)^\top [\boldsymbol{q}^+ - \boldsymbol{p}^+]$ for any $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^K$. In addition, we can re-express $\|\boldsymbol{q}^+\|_2^2 - \|\boldsymbol{p}^+\|_2^2 = 2(\boldsymbol{p}^+)^\top [\boldsymbol{q}^+ - \boldsymbol{p}^+] + \|\boldsymbol{q}^+ - \boldsymbol{p}^+\|_2^2$. Combining the inequalities (18) in Claim 1, we obtain the inequalities (21) in this Claim. ■

Electronic copy available at: https://ssrn.com/abstract=3356823

5. Benchmark Matching Policies

Multi-objective optimization is a common challenge in matching problems for ride-sourcing markets. To validate and evaluate the performance of the proposed AM policy, we introduce three classes of benchmark matching policies to address the multiple objective challenge, including (1) weighted-sum matching policy, (2) multistage matching policy, and (3) stable matching policy (Wang et al. 2017), which are detailed as follows.

5.1. The Weighted-Sum Matching (WM) Policy

This is one of the most popular policies to address multi-objective optimization problems, due to its computational simplicity. The weight on each objective is pre-selected and fixed across different matching scenarios and multiple periods. For the optimization problem with three objectives, i.e., platform revenue, pick-up distance, and service quality, given fixed weights (w_1, w_2, w_3), the weighted-sum matching model under any matching scenario $\boldsymbol{\omega}$ can be represented as:

$$(WM)\ \max\ w_1 h_1(\boldsymbol{x}, \boldsymbol{\omega}) + w_2 h_2(\boldsymbol{x}, \boldsymbol{\omega}) + w_3 h_3(\boldsymbol{x}, \boldsymbol{\omega})$$

$$s.t.\ \ \boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega}).$$

Denoting the optimal decision under the WM policy as $\boldsymbol{x}^{WM}$, we can describe the expected performance of this policy for the kth objective as:

$$\tau_k^{WM} := \mathbf{E}\left[h_k(\boldsymbol{x}^{WM}, \boldsymbol{\omega})\right],\ k = 1, 2, 3. \tag{22}$$

In practice, (w_1, w_2, w_3) are usually pre-selected based on relative importance of the objectives according to managerial judgment. Although the weights can be adjusted when the manager observes poor performance of an objective, it is more an adhoc adjustment without scientific guidance. We evaluate diffierent sets of weights under the WM policy and report the results with $w_1 = w_2 = w_3 = 1$ in Section 6 for illustrative purpose. In Section 7, we implement a more practical WM policy, i.e., the legacy policy, in which the weights on different objectives were obtained from brute-force simulations and back testing to maximize the total revenue.

Electronic copy available at: https://ssrn.com/abstract=3356823

5.2. The Multistage Matching (MM) Policy

This is a priority-based policy that addresses the multi-objective optimization problems according to a pre-determined priority sequence of the multiple objectives. More formally, this policy splits a multi-objective problem into a hierarchy of single-objective subproblems. The objective with higher priority is placed at a higher stage to solve.

For illustration, we consider three priority sequences: (1) Revenue $\succ$ Distance $\succ$ Service, (2) Distance $\succ$ Revenue $\succ$ Service, and (3) Service $\succ$ Distance $\succ$ Revenue.

For priority sequence (1), Revenue $\succ$ Distance $\succ$ Service, the MM policy solves a single-objective optimization problem to maximize revenue at the first stage:

$$g_{1,1}^{MM}(\boldsymbol{\omega}) := \max_{\boldsymbol{x}} \left\{ h_1(\boldsymbol{x}, \boldsymbol{\omega}) : \boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega}) \right\}.$$

After obtaining the optimal revenue $g_{1,1}^{MM}(\boldsymbol{\omega})$ under scenario $\boldsymbol{\omega}$, the MM policy sets this value as a constraint in a single-objective optimization problem to maximize saved pick-up distance (i.e., minimize real pick-up distance) at the second stage:

$$g_{1,2}^{MM}(\boldsymbol{\omega}) := \max_{\boldsymbol{x}} \left\{ h_2(\boldsymbol{x}, \boldsymbol{\omega}) : h_1(\boldsymbol{x}, \boldsymbol{\omega}) \geq g_{1,1}^{MM}(\boldsymbol{\omega}),\ \boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega}) \right\}.$$

For the third stage, the MM policy enforces $g_{1,1}^{MM}(\boldsymbol{\omega})$ and $g_{1,2}^{MM}(\boldsymbol{\omega})$ in the constraints and solves a single-objective optimization problem to maximize service quality:

$$g_{1,3}^{MM}(\boldsymbol{\omega}) := \max_{\boldsymbol{x}} \left\{ h_3(\boldsymbol{x}, \boldsymbol{\omega}) : h_1(\boldsymbol{x}, \boldsymbol{\omega}) \geq g_{1,1}^{MM}(\boldsymbol{\omega}),\ h_2(\boldsymbol{x}, \boldsymbol{\omega}) \geq g_{1,2}^{MM}(\boldsymbol{\omega}),\ \boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega}) \right\}.$$

Denoting $\boldsymbol{x}^{MM-1}$ as the optimal solution to the third-stage problem under the first priority sequence, Revenue $\succ$ Distance $\succ$ Service, we can calculate the corresponding expected performance on the kth objective as:

$$\tau_k^{MM-1} := \mathbf{E}\left[h_k(\boldsymbol{x}^{MM-1}, \boldsymbol{\omega}) \right],\ k = 1, 2, 3. \tag{23}$$

Similarly, for priority sequence (2), Distance $\succ$ Revenue $\succ$ Service, and priority sequence (3), Service $\succ$ Distance $\succ$ Revenue, we can obtain the corresponding expected performance on the kth objective as:

$$\tau_k^{MM-2} := \mathbf{E}\left[h_k(\boldsymbol{x}^{MM-2}, \boldsymbol{\omega}) \right],\ k = 1, 2, 3, \tag{24}$$

and

$$\tau_k^{MM-3} := \mathbf{E}\left[h_k(\boldsymbol{x}^{MM-3}, \boldsymbol{\omega}) \right],\ k = 1, 2, 3, \tag{25}$$

where $\boldsymbol{x}^{MM-2}$ and $\boldsymbol{x}^{MM-3}$ denote the optimal solution under priority sequences (2) and (3), respectively.

Electronic copy available at: https://ssrn.com/abstract=3356823

5.3. The Stable Matching (SM) Policy

This policy defines stable matching between passengers and drivers. The matching is stable if no passenger and driver that have not been matched as a pair (e.g., they are matched to others or unmatched) would prefer to be matched to each other. Such a potential preferred pair is called a "blocking pair" in the stable matching literature. If no blocking pair is present in a matching solution, the matching solution is stable. Wang et al. (2017) introduced the stable matching concept in a dynamic hitch ride-sharing problem, and showed that this policy can increase the stability of matching solutions. Although the stable matching policy is not designed for multi-objective matching problems to shorten the distance to target, it can be appropriately adopted to solve the matching problem with three objectives. Note that in the ride-sourcing system, passengers always prefer drivers with shorter pick-up distance and higher service quality, and drivers always prefer passengers with shorter pick-up distance and higher order revenue. We can formulate the multi-objective ride-matching problem in a stable matching model: one objective is put into the objective function and the remaining two are considered in the stability constraints.

For illustration, we also introduce three stable matching formulations: (1) Maximize total platform revenue with stability constraints that force passengers to prefer drivers with higher service quality and drivers to prefer passengers with shorter pick-up distance. (2) Maximize total saved pick-up distance with stability constraints that force passengers to prefer drivers with higher service quality and drivers to prefer passengers with higher order revenue. (3) Maximize total service quality with stability constraints that force passengers to prefer drivers with shorter pick-up distance and drivers to prefer passengers with higher order revenue.

Stable matching formulation (1) can be written as:

$$\begin{aligned}
&\max_{\boldsymbol{x}} \; h_1(\boldsymbol{x}, \boldsymbol{\omega}) \\
&s.t. \; \sum_{j' \succeq_i j} x_{i,j'} + \sum_{i' \succeq_j i} x_{i',j} + x_{i,j} \geq 1, \; \forall (i,j) \in \mathcal{G}(\boldsymbol{\omega}), \\
&\qquad \boldsymbol{x} \in \mathcal{X}(\boldsymbol{\omega}).
\end{aligned}$$

where $j' \succeq_i j$ denotes that passenger i prefers driver j' to driver j, i.e., $s_{j'}(\boldsymbol{\omega}) \geq s_j(\boldsymbol{\omega})$. Similarly, $i' \succeq_j i$ denotes that driver j prefers passenger i' to passenger i, i.e., $d_{i',j}(\boldsymbol{\omega}) \leq d_{i,j}(\boldsymbol{\omega})$. The set of stability constraints prevents blocking pair (i,j). Namely, either passenger i is

Electronic copy available at: https://ssrn.com/abstract=3356823

matched to a driver j', who is at least as good as driver j (from the perspective of passenger i), or driver j is assigned to passenger i', who is at least as good as passenger i (from the perspective of driver j). Such stability constraints ensure weakly stable matching.

Denoting $\boldsymbol{x}^{SM-1}$ as the optimal solution to stable matching formulation (1), i.e., revenue-maximization stable matching, we can calculate the corresponding expected performance on the kth objective as:

$$\tau_k^{SM-1} := \mathbf{E}\left[h_k(\boldsymbol{x}^{SM-1}, \boldsymbol{\omega})\right],\ k = 1, 2, 3. \tag{26}$$

Similarly, for the other two stable matching formulations, we can obtain the corresponding expected performance on the kth objective as:

$$\tau_k^{SM-2} := \mathbf{E}\left[h_k(\boldsymbol{x}^{SM-2}, \boldsymbol{\omega})\right],\ k = 1, 2, 3, \tag{27}$$

and

$$\tau_k^{SM-3} := \mathbf{E}\left[h_k(\boldsymbol{x}^{SM-3}, \boldsymbol{\omega})\right],\ k = 1, 2, 3, \tag{28}$$

where $\boldsymbol{x}^{SM-2}$ and $\boldsymbol{x}^{SM-3}$ denote the optimal solution to stable matching formulations (2) and (3), respectively.

6. Numerical Experiments

We provide two streams of numerical experiments to illustrate the trade-offs between multiple objectives and evaluate the performance of different matching policies in the ride-sourcing markets. In this Section, we evaluate the theoretical performance guarantee of the AM policy comparing to the three aforementioned benchmark policies: the WM Policy, MM Policy, and SM Policy. In Section 7, we test the practical performance of the AM policy comparing to the legacy policy and "closest distance" policy using real data from the largest ride-sourcing platform in China. We simulate the ride-matching environment using Java programming language and solve the optimization problem by Gurobi (8.0.1) solver. All the experiments are performed on a 2.70 GHz i7-6820HQ CPU Windows PC with 16GB RAM.

6.1. Ride-Matching Dataset Description

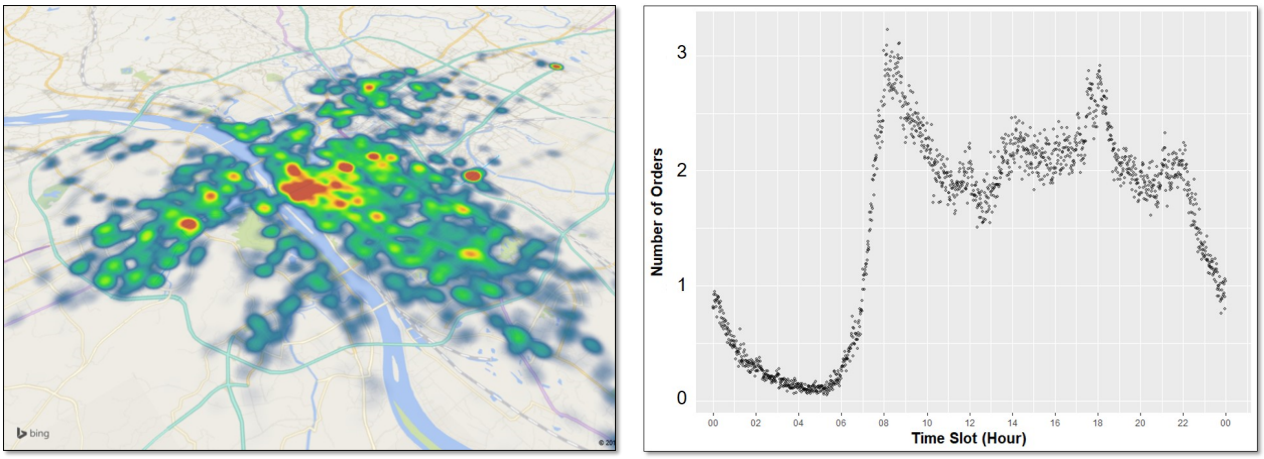
We collect the ride-sourcing records of a Tier-2 city (abbreviated as City A) in China. On the demand side, the dataset for orders contains the detailed information for each travel request (by the passenger), including order booking time, latitude and longitude of

Electronic copy available at: https://ssrn.com/abstract=3356823

the trip's origin and destination, order revenue, etc. On the supply side, the dataset on driver-trace contains minute-level tracking information of all active drivers on the platform, including latitude and longitude of the driver's location, and his/her service status, etc. The dataset records three service statuses of active drivers: (1) Busy, for drivers who are delivering passengers; (2) Occupied, for drivers who are on their way to pick up passengers; and (3) Idle, for drivers who are not assigned to serve any passenger. If a driver logs out from the platform, his/her status is recorded as inactive, and no passenger requests will be assigned to him/her. Based on the service status, we can obtain the shift time of each driver.

We choose one typical weekday records of City A to perform the matching policy comparison. Figure 4 depicts the spatial and temporal distribution of orders on the demand side. We observe that the central regions (highlight in red) of the city are crowded with ride-sourcing requests at the daily aggregated level. In terms of temporal distribution, we find that the morning peak hour, from 8 a.m. to 9 a.m., has highest demand. Demand then fluctuates between 10 a.m. and 17 p.m. During the evening peak hour, demand again climbs until 18 p.m., then declines in the late evening. The volume of orders hits bottom between 4 a.m. and 5 a.m. Typically, during off-peak hours (e.g., 15 p.m.-17 p.m.), the number of orders is relatively stationary. However, during peak hours (e.g., 17 p.m.-19 p.m.), the number of orders fluctuates significantly.

Figure 4 Spatial and Temporal Distribution of Orders. The absolute number of orders is normalized.

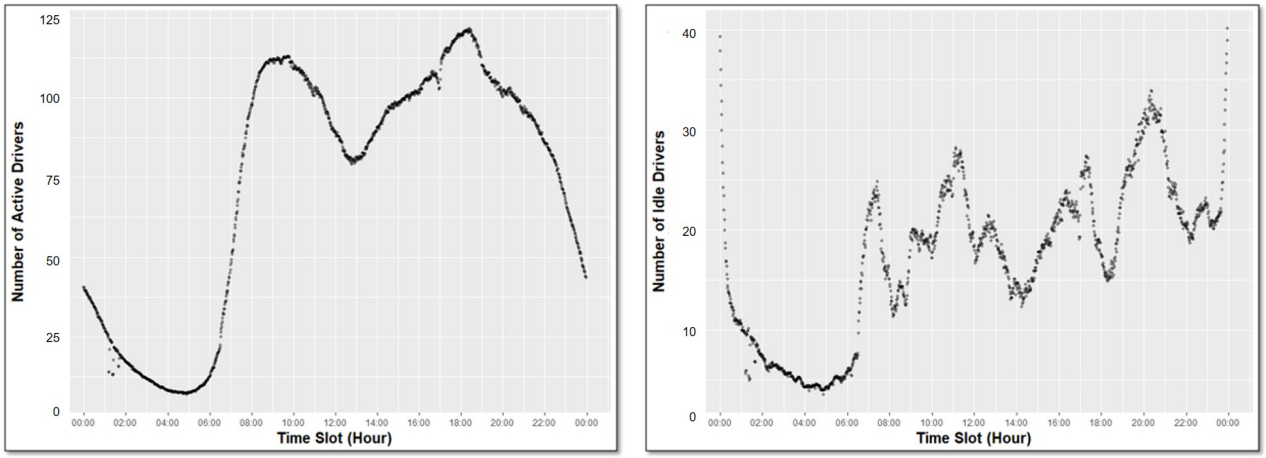


(a) Daily Aggregated Spatial Distribution

(b) Daily Temporal Distribution

Electronic copy available at: https://ssrn.com/abstract=3356823

Figure 5 Temporal Distributions of Active Drivers and Idle Drivers. The absolute number of drivers is normalized.



(a) Number of Active Driver

(b) Number of Idle Drivers

On the supply side, Figure 5 characterizes the temporal distribution of active drivers and idle drivers at minute level. The number of active drivers working on the platform paints a similar temporal trajectory to that of orders: Drivers are more active between 8 a.m. and 11 a.m. and between 16 p.m. and 20 p.m., when there are more ride-sourcing requests. After 17 p.m., part-time drivers join the platform after they have finished their full-time jobs, and hence the evening peak is higher than the morning peak. After midnight and in the early morning, many drivers log out from the platform. Since the number of idle drivers is jointly determined by the number of active drivers and the number of orders, this fluctuates drastically over the whole day.

6.2. Exogenous Ride-Matching Environment

In this section, we compare the performance of the AM Policy with the aforementioned three benchmark policies using real data during both off-peak hours with stationary demand (15:00 p.m.-16:59 p.m.) and peak hours with fluctuating demand (17:00 p.m.-18:59 p.m.). Following common industrial practice, we output the attained performance for platform revenue, saved pick-up distance, and service quality. We also compare the answer rate ([answered order]/[total order]) and the Euclidean distance to the pre-determined target. To understand the trade-offs between the three objectives with independent and identical distribution of matching scenarios across periods as assumed in Theorem 1, we consider a simplified ride-matching environment, in which the drivers on the supply side are exogenously generated and independent of the matching decisions made in previous periods. The

Electronic copy available at: https://ssrn.com/abstract=3356823

following simplifications, which are released in a more realistic environment in subsection 6.3, are made to control matching scenarios in this subsection.

- Orders at each scenario are revealed sequentially according to the recorded time stamps.
- Idle drivers are bootstrapped from the driver-trace set given the number of idle drivers at each scenario. We can adjust the number of idle drivers to simulate different supply scenarios.
- Matched passengers and drivers leave the market, and we do not record the following driver traces.
- Unmatched passengers and drivers also leave the market and will not be considered at the next period.
- The maximum allowed pick-up distance is set as 3.00 km.
- The batch-matching interval is fixed as 6 seconds, with 1,200 matching periods over a two-hour simulation.

Under these simplifications, the information on sampled idle drivers and orders are controlled to follow identical distribution across matching policies at every matching period, and hence the matching policy at a specific matching period only affects the outcome at that period. To simulate diverse supply scenarios with sufficient, balanced, and insufficient drivers, we first estimate the real number of idle drivers, denoted as n_0, based on the driver-trace set. Then we control supply size by adjusting a coefficient θ to generate θn_0 idle drivers in each scenario, i.e., a large θ implies a sufficient number of drivers in the system while a small θ reveals a driver shortage problem. In the numerical experiments, we set $\theta = 0.2, 0.05$, and 0.025 to construct supply scenarios with sufficient, balanced, and insufficient drivers, respectively.

We choose the utopia point as the pre-determined target (i.e., utopia target) and characterize the utopia point approximately by solving the multistage matching models (i.e., Equations (23), (24), and (25)). Policy comparisons during off-peak hours (e.g., 15:00 p.m.-16:59 p.m.) are summarized in Tables 1, 2, and 3, and during peak hours (e.g., 17:00 p.m.-18:59 p.m.) are summarized in Tables 4, 5, and 6. We report the attained performance for platform revenue, saved pick-up distance, service quality, answer rate, and the Euclidean distance to the utopia point. When there is no confusion, the scripts 1, 2, and 3 under the MM policy represent three priority sequences (i.e., platform revenue first, pick-up

Electronic copy available at: https://ssrn.com/abstract=3356823

distance first, and service quality first), and the scripts under the SM policy represent three objectives (i.e., platform revenue maximization, saved pick-up distance maximization, and service quality maximization).

When supply is sufficient as shown in Tables 1 and 4, around 95% of passengers can be served under different policies, and hence different matching policies yield similar revenue. Since there are always enough drivers to provide ride-sourcing service for most passengers, both MM policies (1) and (2) tend to dispatch the closest driver to serve a nearby passenger. MM policy (3) favors drivers with higher service quality and provides significantly higher expected service scores, with sacrifice on pick-up distance. For the SM policy, we observe that the answer rate is less than other policies, since stability constraints prevent some feasible matching pairs (e.g., blocking pairs). Surprisingly, the SM policy with the service quality maximization provides the worst service quality, but attains satisfactory revenue and pick-up distance—i.e., stability constraints play a greater role than the objective function in this multi-objective stable matching optimization setting. For the WM policy, we test different sets of weights and report the results with weights $w_1 = w_2 = w_3 = 1$ for illustrative purpose. The performance under the WM policy is prosaic for all objectives. Although the WM policy could be improved if the "right" weights are implemented, selection of such weights is unknown. Finally, the proposed AM policy achieves satisfactory platform revenue and service quality at the cost of sacrificing a little bit of pick-up distance. In terms of the Euclidean distance to the utopia target, the AM policy provides the shortest distance, compared with all other policies.

When supply and demand are balanced as shown in Tables 2 and 5, in general, drivers can select passengers with higher order revenue, and passengers can be matched to those drivers with higher service quality. Consequently, the platform revenue obtained under different MM policies are no longer similar. The SM policy provides stable matching solutions and also slightly decreases the system-wide performance on the three objectives. The AM policy achieves satisfactory outcomes for all objectives and has the shortest Euclidean distance to the utopia target.

When supply is insufficient as shown in Tables 3 and 6, most drivers can be assigned to serve passengers. Different policies provide similar service quality, but their performances on platform revenue differ widely. The AM policy still outperforms other policies in terms of the shortest Euclidean distance to the utopia target.

Electronic copy available at: https://ssrn.com/abstract=3356823

Table 1 Comparison of Matching Policies: sufficient supply during off-peak hours

Objective	WM Policy	MM Policy 1	MM Policy 2	MM Policy 3	SM Policy 1	SM Policy 2	SM Policy 3	Utopia Target	AM Policy
Platform Revenue	777.76	**778.83**	777.52	777.52	774.15	777.36	776.86	778.83	778.72
Saved Pick-up Dis.	773.40	779.03	**779.13**	740.64	743.66	730.87	775.29	779.13	754.33
Service Quality	444.93	394.57	394.57	**655.45**	639.79	643.18	393.09	655.45	644.51
Answer Rate	0.95	0.95	0.95	0.95	0.94	0.95	0.94		0.95
Euc. Distance	210.60	260.88	260.89	38.51	39.06	49.82	262.4		**27.11**

Table 2 Comparison of Matching Policies: balanced supply during off-peak hours

Objective	WM Policy	MM Policy 1	MM Policy 2	MM Policy 3	SM Policy 1	SM Policy 2	SM Policy 3	Utopia Target	AM Policy
Platform Revenue	622.51	**646.04**	621.07	621.08	581.43	624.15	618.69	646.04	645.32
Saved Pick-up Dist.	603.2	600.97	**603.35**	595.73	561.99	556.63	564.08	603.35	597.18
Service Quality	314.98	314.04	314.04	**358.28**	338.61	342.03	293.46	358.28	355.48
Answer Rate	0.76	0.76	0.76	0.76	0.71	0.72	0.71		0.76
Euc. Distance	49.27	44.3	50.8	26.09	79.19	54.09	80.57		**6.81**

Table 3 Comparison of Matching Policies: insufficient supply during off-peak hours

Objective	WM Policy	MM Policy 1	MM Policy 2	MM Policy 3	SM Policy 1	SM Policy 2	SM Policy 3	Utopia Target	AM Policy
Platform Revenue	390.79	**455.96**	385.89	385.89	367.18	440.36	437.23	455.96	454.67
Saved Pick-up Dis.	374.75	368.2	**374.81**	374.03	356.85	351.12	353.38	374.81	369.2
Service Quality	195.34	195.32	195.32	**200.90**	194.65	196.02	186.28	200.90	200.09
Answer Rate	0.47	0.47	0.47	0.47	0.45	0.45	0.45		0.47
Euc. Distance	65.4	8.65	70.29	70.08	90.79	28.78	32		**5.81**

Table 4 Comparison of Matching Policies: sufficient supply during peak hours

Objective	WM Policy	MM Policy 1	MM Policy 2	MM Policy 3	SM Policy 1	SM Policy 2	SM Policy 3	Utopia Target	AM Policy
Platform Revenue	874.48	**875.07**	874.09	874.09	871.34	873.97	873.5	875.07	874.61
Saved Pick-up Dis.	888.15	897.88	**897.99**	851.21	855.43	840.05	894.63	897.99	869.11
Service Quality	546.3	458.32	458.31	**777.47**	760.65	764.03	457.11	777.47	756.87
Answer Rate	0.96	0.96	0.96	0.96	0.95	0.95	0.95		0.96
Euc. Distance	231.38	319.16	319.17	46.79	45.91	59.49	320.39		**35.47**

Table 5 Comparison of Matching Policies: balanced supply during peak hours

Objective	WM Policy	MM Policy 1	MM Policy 2	MM Policy 3	SM Policy 1	SM Policy 2	SM Policy 3	Utopia Target	AM Policy
Platform Revenue	733.68	**754.61**	732.54	732.58	682.97	729.98	722.53	754.61	751.95
Saved Pick-up Dis.	724.26	722.29	**724.45**	713.66	672.93	667.37	677.84	724.45	715.8
Service Quality	378.61	377.32	377.32	**440.17**	414.09	418.64	352.38	440.17	433.39
Answer Rate	0.8	0.8	0.8	0.8	0.75	0.76	0.75		0.8
Euc. Distance	65.02	62.89	66.61	24.54	92.02	65.79	104.45		**11.31**

6.3. Endogenous Ride-Matching Environment

In this section, we elaborate the performances of different policies in the endogenous matching environment, where the matching scenario at each period depends on the decisions in

Electronic copy available at: https://ssrn.com/abstract=3356823

Table 6 Comparison of Matching Policies: insufficient supply during peak hour

Objective	WM Policy	MM Policy			SM Policy			Utopia Target	AM Policy
		1	2	3	1	2	3		
Platform Revenue	470.27	**544.36**	464.78	464.77	438.18	522.42	519.09	544.36	543.65
Saved Pick-up Dis.	461.43	453.66	**461.49**	460.44	435.71	427.91	432.22	461.49	454.25
Service Quality	240.42	240.37	240.37	**247.47**	238.89	240.53	227.69	247.47	247.38
Answer Rate	0.51	0.51	0.51	0.51	0.48	0.49	0.48		0.51
Euc. Distance	74.43	10.57	79.89	79.6	109.59	40.71	43.43		**7.27**

previous periods. The same datasets during both off-peak hours (15:00 p.m.-16:59 p.m.) and peak hours (17:00 p.m.-18:59 p.m.) are used for simulation. More concretely,

- Orders at each scenario are revealed sequentially according to the recorded time stamps. Passengers' patience levels (tolerance for waiting time before leaving the market without being assigned to a driver) are uniformly (randomly) generated from 6 seconds to 60 seconds.
- Number of drivers and their status (inactive, idle active, and busy active) are calibrated from the historical data. For idle active drivers, we perform a random walk to simulate their cruising routes. For busy active drivers, we simulate their traveling paths over a grid map. We assume a constant travel speed of 30 kilometers per hour for all drivers.
- Unmatched passengers and drivers are carried over to the next period.
- The maximum allowed pick-up distance is set as 3.00 km.
- The batch-matching interval is fixed as 6 seconds, with 1,200 matching periods over a two-hour simulation.

We compare the performance of different policies in Section 6.3.1. We then investigate the impact of pre-determined targets in Section 6.3.2 and discuss the impact of matching interval with market thickness in Section 6.3.3.

6.3.1. Policy Comparison Due to the endogeneity of the matching scenarios, the MM policy does not always achieve the best performance at each objective. In this subsection, we set the pre-determined target on each objective as the best performance obatained from other policies. More precisely, let τ_k^0 denote the largest value of the kth objective achieved by the WM, MM, and SM polices, and we set the kth target $U_k := \tau_k^0$. We summarize the comparison of matching policies during off-peak and peak hours in Tables 7 and 8, respectively. The majority of observations in the exogenous setting (cf. Section 6.2) are preserved in this endogenous environment, except that the answer rate is heavily affected by the endogeneity issue. This in turn undermines the performance of MM policies. For

Electronic copy available at: https://ssrn.com/abstract=3356823

instance, the MM-1 policy that aims to maximize platform revenue does not achieve the highest revenue even though it prioritizes the revenue maximization. In fact, the availability of idle drivers decreases under the MM-1 policy since more drivers are assigned to serve passengers with higher order revenue and longer O-D distance. MM-2 and MM-3 policies provide satisfactory performances on pick-up distance and service quality respectively, but with poor performances on revenue. Similar to the discussion in Section 6.2, the SM policies rule out some feasible matching pairs and in turn have lower answer rates. In particular, if revenue is put into the stability constraint as in SM-2 and SM-3, the stability constraint forces the drivers to pick-up the high-revenue passengers with longer O-D distances and hence the answer rates decrease. Notably, the AM policy still performs better in terms of the shortest Euclidean distance to the pre-determined target under endogenous matching scenarios.

Table 7 Comparison of Matching Policies during off-peak hours

Objective	WM Policy	MM Policy			SM Policy			Utopia Target	AM Policy
		1	2	3	1	2	3		
Platform Revenue	**3880.30**	3876.05	3770.89	3763.36	3755.24	3861.06	3847.22	3880.30	3871.14
Saved Pick-up Dis.	3850.57	4241.24	**4512.91**	3603.14	3646.08	3318.24	4233.02	4512.91	3911.94
Service Quality	4849.52	3686.04	3833.00	5028.89	**5050.76**	4698.89	3680.51	5050.76	4828.94
Answer Rate	0.66	0.66	0.68	0.68	0.68	0.65	0.65		0.66
Euc. Distance	692.39	1391.76	1222.96	917.70	875.99	1245.72	1399.20		**640.84**

Table 8 Comparison of Matching Policies during peak hours

Objective	WM Policy	MM Policy			SM Policy			Utopia Target	AM Policy
		1	2	3	1	2	3		
Platform Revenue	**4184.17**	4175.24	4062.76	4069.25	4037.79	4166.83	4169.40	4184.17	4186.56
Saved Pick-up Dis.	4190.53	4525.60	**4914.88**	4014.18	4048.64	3604.71	4525.03	4914.88	4268.70
Service Quality	5234.85	4026.79	4250.89	**5438.75**	5418.75	5057.07	4056.61	5438.75	5216.20
Answer Rate	0.63	0.63	0.66	0.66	0.66	0.62	0.62		0.64
Euc. Distance	752.69	1464.94	1194.39	908.22	879.01	1364.94	1436.43		**683.63**

6.3.2. Impact of Targets We next examine the performances of the AM policy with different pre-determined targets. More precisely, we set the kth target $U_k := \gamma\tau_k^0$, where $\gamma \geq 1$ represents an expansion coefficient. Note that we set $\gamma = 1.0$ in Section 6.3.1. We enumerate $\gamma = \{1.0, 1.5, 2.0, 2.5, 5.0, 10.0, 100.0\}$ and compare the performances under different pre-determined targets over 4 hours from 15:00 p.m. to 18:59 p.m. Intuitively, if γ is sufficiently large, debt $w_k(t+1) = \gamma\tau_k^0 - \frac{1}{t}\sum_{s=1}^{t} h_k(\boldsymbol{x}_s^{\text{AM}}, \boldsymbol{\omega}_s)$ is more stable since $h(\cdot,\cdot)$

Electronic copy available at: https://ssrn.com/abstract=3356823

has smaller impacts compared to γ, and then the AM policy degenerates to a weighted-sum matching policy with fixed weights. Additionally, instead of a static target over the matching horizon of four hours, we also evaluate a hour-dependent target, i.e., the multi-objective target changes hour by hour. For each hour, we take the best performance from other policies in that hour as the target.

Table 9 summarizes the impact of pre-determined targets on the performances, including the Euclidean distance to the same original target $\boldsymbol{\tau}^0$. As expected, among all the static target cases, the solution with $\gamma = 1.0$ achieves the shortest distance to the target $\boldsymbol{\tau}^0$, which is consistent to our theoretical results. For a large value of γ, we observe that the performances of the AM policy become similar to the WM policy. Notably, the hour-dependent target induces the shortest Euclidean distance to $\boldsymbol{\tau}^0$.

Table 9 Impact of Pre-determined Targets

Objective	WM Policy	AM Policy under Static Target (γ)							Hour-dependent Target
		1.0	1.5	2.0	2.5	5.0	10.0	100.0	
Platform Revenue	4238.78	4194.36	4215.68	4234.11	4244.97	4234.06	4238.95	4248.03	4177.30
Saved Pick-up Dis.	4307.67	4490.67	4330.15	4318.19	4306.39	4289.60	4292.99	4290.53	4559.31
Service Quality	5233.97	5081.61	5242.38	5267.36	5264.57	5252.70	5260.38	5243.36	5205.58
Answer Rate	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.70
Euc. Distance	675.52	585.10	651.96	655.96	668.08	687.43	682.09	689.18	460.63

6.3.3. Impact of Market Thickness We next assess the impact of market thickness on the performance of matching policies. Following the definition in Akbarpour et al. (2017), market thickness refers to the number of potential pairs (between passenger and driver) to be matched in the dynamic matching market. Intuitively, the market becomes thicker for a larger matching interval, conditional on the fact that the passenger/driver arrival rates are larger than their departure rates. In a thick market, both passengers and drivers have better matching choices, e.g., passengers could be served by drivers with higher service quality; drivers may pick-up passengers closer to them. On the other hand, some time-sensitive passengers may quite the market (e.g., cancel the order) without being matched if the matching interval is too large, and hence the answer rate may decrease with the matching interval and in turn hurt the total platform revenue.

Table 10 provides the trade-offs between the market thickness and the answer rate with different matching intervals that range from 6 seconds to 48 seconds. As the matching interval increases, the average service quality provided to passengers increases significantly,

Electronic copy available at: https://ssrn.com/abstract=3356823

and the average saved pick-up distance also increases. However, more passengers churn out of the market without being matched due to a longer waiting time. As a result, the revenue per trip and the answer rate decrease. This observation is consistent for both the pre-determined target (i.e., the best performance from other policies for each objective) and the solution under the AM policy.

Table 10 Effect of Market Thickness: matching intervals range from 6 to 48 seconds

	Pre-determined Target				AM Policy			
Interval (seconds)	6	12	24	48	6	12	24	48
Platform Revenue	150.60	150.22	150.16	148.48	148.95	148.92	148.83	146.62
Saved Pick-up Dis.	169.00	169.51	170.16	171.11	150.52	152.34	154.35	156.15
Service Quality	189.82	190.81	194.89	203.01	185.80	187.56	191.03	199.95
Answer Rate	0.68	0.67	0.65	0.59	0.66	0.66	0.63	0.57

From the view of passengers, a higher waiting time under a larger matching interval is compensated by higher ride-sourcing service quality and shorter pick-up time. That could explain the greedy policy, which matches passengers to drivers as soon as possible, is not commonly used in practice. It is flexible for the platform to adjust the matching interval to create a better matching environment, but we note that it is still an open question to determine the "best" matching interval as it is challenging to explicitly characterize its impact on the whole ride-sourcing system. Besides of the trade-offs between objectives, the computational efficiency is another concern in selecting the matching interval. In the practice of our industry partner, the matching interval is normaly set to be 2 seconds for Tier-1 cities in China, due to the mega-scale of demand and supply.

7. Implementation in Ride-Sourcing Markets

In this section, we demonstrate the potential of implementing the AM policy in the ride-sourcing market. With a much more complex and realistic ride-sourcing simulator, we take many other factors into consideration, including traffic conditions, passenger traveling behavior, driver routing patterns, and travel times in real road network. We collect more ride-sourcing records from two Tier-1 cities (abbreviated as City B and C) in China to test our policy. We compare the performance of the AM policy with the one favored in practice (denote as "Legacy Policy"), as well as the classic "closest distance" policy (CD Policy) that is widely studied in academic literatures (e.g., Özkan and Ward 2016). The legacy policy is essentially a weighted-sum matching policy and the weights on different

Electronic copy available at: https://ssrn.com/abstract=3356823

objectives were obtained from brute-force simulations and back testing to maximize the total revenue. The set of weights in legacy policy seldom changes once it is determined. It is the most natural policy for comparison since it was carefully designed to address the multi-objective ride-matching problems[3]. The details of the ride-matching environments are described as follows:

- Order information at each scenario is sampled from the order set. Passengers' patience levels are calibrated from historical data.
- Driver status, routing patterns, and travel times between each specific pair of origin and destination are simulated using historical data. For idle active drivers, we perform a random walk to simulate their cruising routes. For busy active drivers, we simulate their traveling paths using a virtual GPS routing system considering taffic conditions.
- Unmatched passengers and drivers are carried over to the next period.
- The batch-matching interval is fixed as 2 seconds.

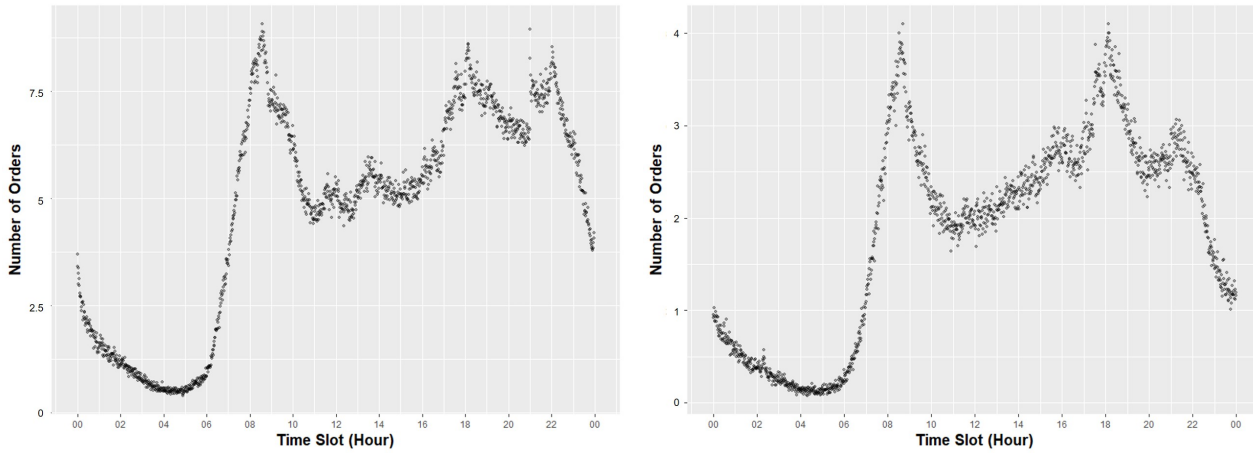
As shown in Figure 6, the number of total orders in City B is more than twice that of City C. However, the number of total registered drivers in City B is only 1.08 times that of City C. In other words, City B suffers from more serious driver shortage and hence drivers from City B have more choices to be dispatched. In addition, City B is larger than City C in terms of geographical size and drivers in City B would take longer time to serve a passenger.

We do not use the rigorous utopia point as defined in Equations (23), (24), and (25) as the target, since we cannot exactly get to know the realized scenario ahead of time in practice. For ease of implementation, we choose the target according to historical performance on the three objectives from real data. Suppose the historical performance on the kth target is τ_k^H, we set $U_k := \gamma\tau_k^H$, where γ denotes an expansion coefficient. Note that the historical performance lies in the feasible region (i.e., on or below the efficient frontier) because it was achieved according to real data, we set $\gamma > 1$ so that the proposed target $\boldsymbol{U}$ lies beyond the efficient frontier and the attained solution under the AM policy is located on the efficient frontier. More concretely, we numerically compare the performance of the AM policy with $\gamma = \{1.5, 2.0, 2.5\}$ and find there is no significant difference among the results, which are consistent to the discussion in Section 6.3.2. We simply set $\gamma = 2.0$ unless stated otherwise.

[3] The legacy policy presented in this section is not exactly identical to the policy implemented in a specific city. For example, we did not use bonus to incentivize drivers to serve passengers from "cold zones" in the numerical experiments.

Electronic copy available at: https://ssrn.com/abstract=3356823

Figure 6 Temporal Order Distributions of City B and C. The absolute number of orders is normalized.



(a) City B (b) City C

Five indicators are introduced to evaluate the performance of different matching policies, including Platform Revenue (per trip), Pick-up Distance (per trip), Service Quality (per trip), Total Answer Count, and Total Revenue. The legacy policy is set as the benchmark, and we report the relative difference, i.e., improvement or deterioration, for each objective for other two policies. For Revenue (per trip), Service Quality (per trip), Total Answer Count, and Total Revenue, we report the difference in terms of percentage. For example:

$$\text{Revenue Diff.} := \frac{\text{Revenue (AM Policy)} - \text{Revenue (Legacy Policy)}}{\text{Revenue (Legacy Policy)}} \times 100\%.$$

For Pick-up Distance (per trip), we report the absolute difference in meters:

$$\text{Pick-up Distance Diff.} := \text{Pick-up Dis. (AM Policy)} - \text{Pick-up Dis. (Legacy Policy)}\ (m).$$

Tables 11 and 12 present the performances of these policies during both off-peak (15:00 p.m.-16:59 p.m.) and peak (17:00 p.m.-18:59 p.m.) hours in City B and City C, respectively. We observe that the performances of legacy policy and CD policy are similar, which means that the pick-up distance is still the top consideration in the legacy policy. Under the AM policy, revenue per trip and service quality are improved at the cost of sacrificing a little bit of pick-up distance. Passengers with higher order revenue are coupled with a longer O-D distance and drivers must take longer time to serve these passengers; therefore, the number of available idle active drivers will decrease under the AM policy. Since we simulated passenger patience level to capture passenger cancellation behavior, the total answer count

Electronic copy available at: https://ssrn.com/abstract=3356823

will decrease because more passengers will leave the market without being matched. Note that revenue per trip increases to a larger extent, and the AM policy generates more total revenue, even though we lose more impatient passengers.

During peak hours, both cities suffer from more serious driver shortage, and more passengers churn out of the market. In addition, drivers are more likely to serve passengers with higher order revenue, and hence revenue per trip increases more significantly. For example, although the answer count in City B decreases by 0.85% (cf. Table 11), the revenue per trip increases by 1.01%. As a result, the total revenue still increases by 0.15%, which implies a significant profit volume when the marginal improvement is translated back to the absolute value. More interestingly, the total revenue increment in City C is as high as 1.38%.

Table 11 Implementation of Three Matching Policies in City B: (+) indicates improvement and (-) for decline, compared with the benchmark policy.

Period	Policy	Revenue	Pick-up Dis. (m)	Service	Answer Count	Total Revenue
Off-Peak	Legacy	9.72	585.33	100.41	1058	10,278.92
	AM	9.78	628.62	101.31	1054	10,315.33
	(%)	0.66% (+)	43.28 (-)	0.90% (+)	-0.30% (-)	0.35% (+)
	CD	9.72	585.00	100.43	1058	10,280.51
	(%)	0.01% (+)	-0.33 (+)	0.03% (+)	0.01% (+)	0.02% (+)
Peak	Legacy	10.43	609.22	100.46	1279	13,332.97
	AM	10.53	664.89	101.43	1268	13,353.56
	(%)	1.01% (+)	55.67 (-)	0.96% (+)	-0.85% (-)	0.15% (+)
	CD	10.43	615.57	100.47	1279	13,332.76
	(%)	0.01% (+)	6.35 (-)	0.01% (+)	-0.01% (-)	0.00% (+)

The revenue, service parameter, and answer count have been normalized according to the baselines [a], [b], and [c], respectively (cf. Table 13).

Next, we implement and compare the performances of matching policies over a whole day (00:00 a.m.-23:59 p.m.) for both cities. For ease of exposition, we use a static target in the AM policy for a whole day. For example, when we implement the AM policy on Friday, we choose the target according to the historical records of previous Fridays with similar weather conditions and then multiply by the expansion parameter $\gamma = 2.0$. Results are summarized in Table 13. As we can see, even when we have a static target for the whole day, the AM policy obtains a delicate balance between multiple objectives and brings value to all the stakeholders in the ride-sourcing ecosystem: passengers, drivers, and the platform.

Electronic copy available at: https://ssrn.com/abstract=3356823

Table 12 Implementation of Three Matching Policies in City C: (+) indicates improvement and (-) for decline, compared with the benchmark policy.

Period	Policy	Revenue	Pick-up Dis. (m)	Service	Answer Count	Total Revenue
Off-Peak	Legacy	7.57	872.86	101.63	446	3,375.77
	AM	7.65	919.82	101.91	443	3,387.89
	(%)	1.03% (+)	46.96 (-)	0.28% (+)	-0.66% (-)	0.36% (+)
	CD	7.56	869.19	101.65	446	3,367.11
	(%)	-0.21% (-)	-3.66 (+)	0.02% (+)	-0.04% (-)	-0.26% (-)
Peak	Legacy	7.15	874.17	101.62	523	3,735.73
	AM	7.38	927.23	101.81	513	3,787.27
	(%)	3.24% (+)	53.05 (-)	0.19% (+)	-1.80% (-)	1.38% (+)
	CD	7.14	886.37	101.55	524	3,737.43
	(%)	-0.16% (-)	12.19 (-)	-0.07% (-)	0.21% (+)	0.05% (+)

The revenue, service parameter, and answer count have been normalized according to the baselines [*a*], [*b*], and [*c*], respectively (cf. Table 13).

Table 13 Implementation of Three Matching Policies in the Ride-sourcing System: (+) indicates improvement while (-) for decline, compared with the benchmark policy.

City	Policy	Revenue	Pick-up Dis. (m)	Service	Answer Count	Total Revenue
City B	Legacy	10.00[a]	645.92	100.00[b]	10000[c]	100,000.00
	AM	10.19	701.12	100.92	9842	100,264.40
	(%)	1.88% (+)	55.20 (m) (-)	0.92% (+)	-1.58% (-)	0.26% (+)
	CD	10.00	649.75	100.00	10002	99,997.56
	(%)	-0.02% (-)	3.83 m (-)	-0.00% (-)	0.02% (+)	-0.00% (-)
City C	Legacy	7.03	801.25	102.59	3750	26,377.90
	AM	7.13	834.76	102.89	3722	26,525.41
	(%)	1.33% (+)	33.51 m (-)	0.29% (+)	-0.76% (-)	0.56% (+)
	CD	7.04	806.70	102.60	3753	26,403.13
	(%)	0.03% (+)	5.45 m (-)	0.00% (+)	-0.06% (-)	0.10% (+)

[*a*] The revenue under legacy policy in City B is normalized to be 10.00. We normalize the remaining parameters associated with revenue in relative to this basis.

[*b*] The service score under legacy policy in City B is normalized to be 100.00. We normalize the remaining parameters associated with service score in relative to this basis.

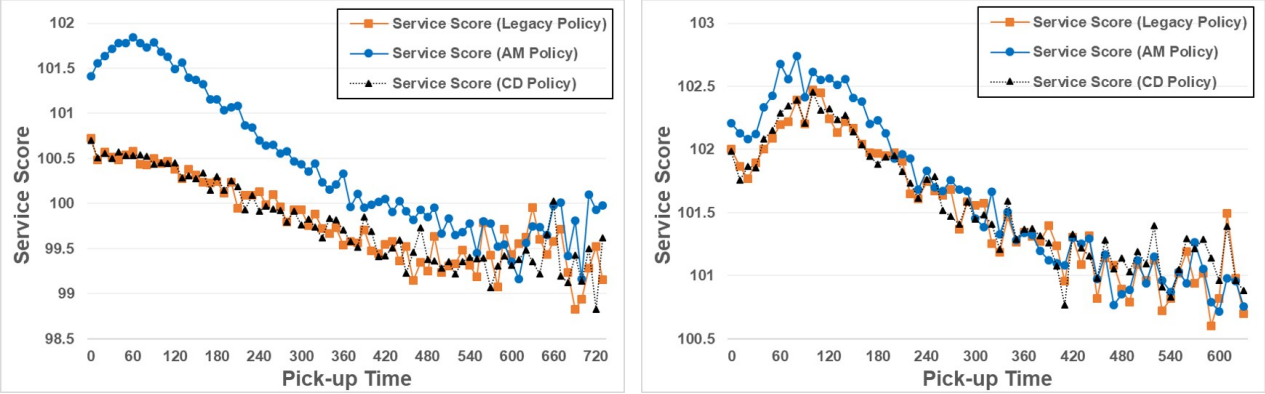
[*c*] The answer count under legacy policy in City B is normalized to be 10000. We also normalize other answer counts and total revenue amounts in relative to this basis.

For Passengers: Compared to the benchmark policy, the AM policy improves service quality per trip by 0.92% and 0.29% for City B and C, respectively. Specifically, as shown in Figure 7, the service quality for passengers with all different pick-up distances have been improved. Figure 8 shows the relations between answer rate and order revenue. Since more short-travel low-revenue orders are requested during peak hours with insufficient drivers, passengers with low order revenue suffer from low answer rates under all the three matching policies. As shown in Table 13, the AM policy increase the revenue per trip by 1.88%

Electronic copy available at: https://ssrn.com/abstract=3356823

and 1.33% in City B and C, respectively, which means more drivers are dispatched to serve these passengers with higher order revenue (i.e., longer travel distance). Figure 8 also demonstrates the higher priority and higher answer rates for passengers with longer travel distance under the AM policy. Note that some ride-sourcing platforms are also investing in the bike-sharing markets, which aim at serving those passengers with shorter travel distance. From this perspective, the AM policy would balance their business strategies in both ride-sourcing and bike-sharing markets, i.e., to provide more comfortable service for passengers with travel request of both long and short distance.

Figure 7 **Service Quality per Trip with Different Pick-up Times: The service quality per trip is the average driver service score attached to these trips belonging to each pick-up time (pick-up distance) interval.**



(a) City B

(b) City C

For Drivers: Compared to the benchmark policy, the AM policy improves service quality per trip, which implies more orders are assigned to drivers with higher service score. Figure 9 demonstrates that the expected total order revenue attached to drivers with higher service scores (e.g., higher than 101) increases under the AM policy. We also find that the order revenue increment for these drivers is indeed due to more orders being dispatched to them. This outcome would motivate drivers to increase their service score and provide better service. We also observe a decreasing trend in total order revenue for the drivers with extreme high service scores under all the three matching policies. One possible explanation is that a large proportion of drivers working on the platform are part-time and their order revenue also depends on their total business hours (i.e., active time as a driver on the platform). The dataset reveals this pattern: these drivers with service scores in the interval

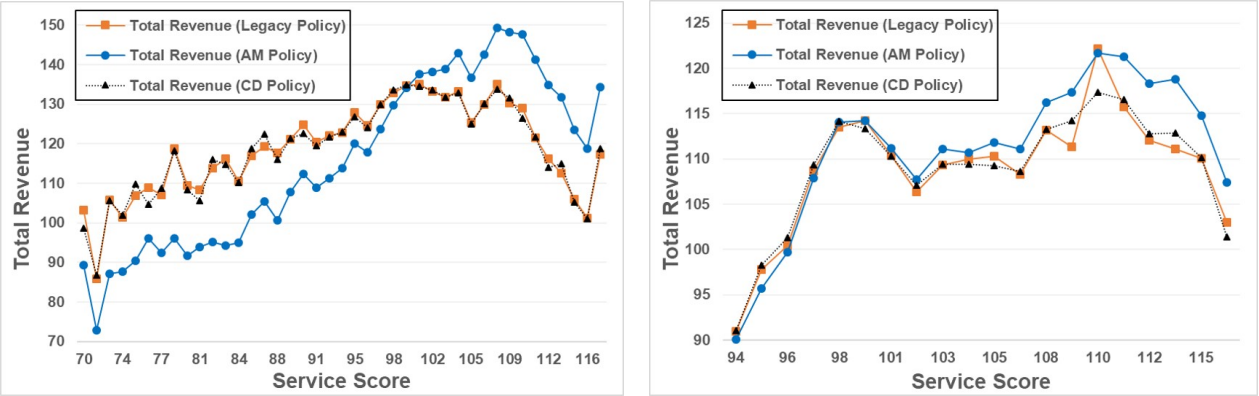
Electronic copy available at: https://ssrn.com/abstract=3356823

Figure 8 **Answer Rates of Passengers with Different Revenues: The answer rate is calculated by (number of orders being matched)/(number of total orders) for passengers belonging each order revenue interval. The answer rate measures the probability of being served.**

(a) City B (b) City C

[98, 108] are more active than the ones with scores in the interval [109,116]. Even so, the AM policy dispatches more orders to these drivers with higher service scores consistently.

Figure 9 **Total Order Revenue per Day attached to Drivers with Different Service Scores: Total revenue for each driver is calculated by aggregating her/his income over the whole day. We plot drivers belonging to each service score interval.**



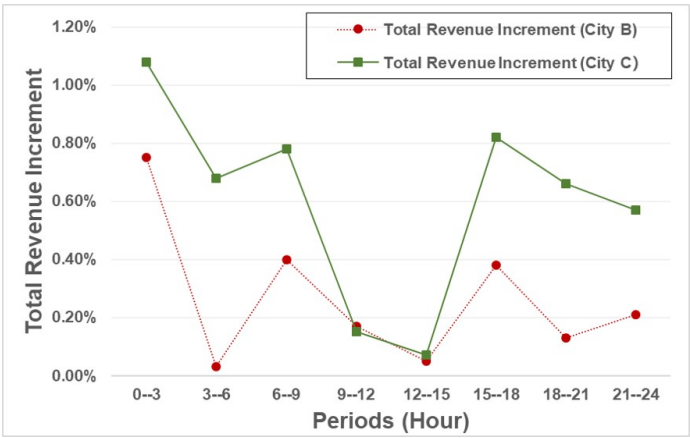
(a) City B (b) City C

For the Platform: Although the pick-up distance per trip increases by 55.20 meters and 33.51 meters, and the answer count reduces by 1.58% and 0.76%, the total platform revenue for the whole day under the AM policy still increases by 0.26% in City B and 0.56% in City C. More precisely, Figure 10 plots the total platform revenue increment in different time periods over the whole day. We observe that the increment is more significant during

Electronic copy available at: https://ssrn.com/abstract=3356823

those periods when drivers start their working shifts, i.e., when drivers switch their inactive status to active. Note that our simulation starts from 0:00 a.m., and the initial active driver pool is identical for all matching policies. Compared to the benchmark policy, the AM policy tends to match idle active drivers to those passengers with higher revenue (with longer travel distance in general), and hence drivers must take longer to serve such orders. As a result, the number of idle active drivers decreases in subsequent periods. The total revenue increment reaches around 1.00% during the first three hours (0 a.m. to 3 a.m.), but decreases between 3 a.m. and 6 a.m. From 6 a.m. to 9 a.m., many drivers start their working shifts, and hence the revenue increment increases again. Since many drivers work part-time in the ride-sourcing market and must work full-time jobs during the daytime, these drivers will leave the platform during the daytime and become active again after 17 p.m. Therefore, the revenue increment becomes much more significant between 15 p.m. and 18 p.m.

Figure 10 **Platform Revenue Increment under the AM Policy, compared with the benchmark policy**



In the long term, if the ride-sourcing platform gives priority to drivers with high service quality, the complaint rate from passengers would decrease and this positive signal would in turn attract more passengers potentially. On the other hand, the increased income for drivers with higher quality may motivate them to work for a longer time on the platform due to positive labor supply elasticity (cf. Chen and Sheldon 2016, Angrist et al. 2017, Sun et al. 2019). Both short- and long-term effects on passengers and drivers can contribute to building a better brand reputation for the ride-sourcing system.

Electronic copy available at: https://ssrn.com/abstract=3356823

8. Conclusion

Motivated by the challenges in matching demand and supply considering multiple objectives in the ride-sourcing markets, we develop a novel online matching policy to solve the multi-objective online ride-matching problem where the platform needs to take actions in real time without observing future information. The proposed online adaptive matching policy adaptively balances the trade-offs between multiple objectives in a dynamic setting. We prove that the AM policy can achieve the target-based optimal solution, i.e., the solution that minimizes the Euclidean distance to any pre-determined multi-objective target, with a non-asymptotic regret bound. The additive error caused by the AM policy increases sub-linearly with the time horizon T, and converges to 0 asymptotically with a sufficiently large T. More precisely, the error term is upper bounded by $O(\log T)$, which is the fastest convergence in literature. We remark that the theoretical property is established in the i.i.d. ride-matching environment, and we also provide extensive numerical experiments to demonstrate its good performance in a dynamic setting with endogenous matching scenarios. Furthermore, the industrial testing show that this AM policy could lead to a favorable outcome, i.e., our approach obtains a delicate balance of multiple objectives and brings value to all the stakeholders in the ride-sourcing ecosystem: passengers, drivers, and the platform.

In this paper, we incorporate platform revenue, pick-up distance, and service quality in the matching policy. For practical implementation, we can consider other objectives such as the number of passengers being served. In the end, we remark that it would be interesting to implement this adaptive matching policy in the industry and study the long-term effect. We leave these and other issues for future research.

Acknowledgment

The authors are grateful to the Matching Group from Didi Chuxing for their help in our numerical study. Their feedback have helped the authors to clarify many practical issues pertaining to the multi-objective ride-matching model.

References

Agrawal, S., Devanur, N.R. 2015. Fast algorithms for online stochastic convex programming. *In Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, 1405–1424.

Akbarpour, M., Li, S., Oveis Gharan, S. 2017. Thickness and information in dynamic matching markets. *Working Paper.*

Electronic copy available at: https://ssrn.com/abstract=3356823

Angrist, J.D., Caldwell, S., Hall, J.V. 2017. Uber vs. Taxi: A Driver's Eye View. *National Bureau of Economic Research (No. w23891).*

Braverman, A., Dai, J.G., Liu, X., Ying, L. 2016. Empty-car routing in ridesharing systems. *arXiv preprint arXiv:1609.07219.*

Boland, N., Charkhgard, H., Savelsbergh, M. 2016. On the Existence of Ideal Solutions in Multi-objective 0-1 Integer Programs. *Working Paper.*

Boland, N., Charkhgard, H., Savelsbergh, M. 2017. A new method for optimizing a linear function over the efficient set of a multiobjective integer program. *European journal of operational research*, **260(3)**, 904–919.

Cesa-Bianchi, N., G. Lugosi, 2006. Prediction, learning, and games. *Cambridge University Press.*

Chen, M.K., Sheldon, M. 2016. Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform. *In EC*, **p.455**.

Das, I., Dennis, J.E. 1998. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization*, **8(3)**, 631–657.

Gearhart, W.B. 1979. Compromise solutions and estimation of the noninferior set. *Journal of Optimization Theory and Applications*, **28(1)**, 29–47.

Hu, J., Mehrotra, S. 2012. Robust and stochastically weighted multiobjective optimization models and reformulations. *Operations Research*, **60(4)**, 936–953.

Hu, M. and Zhou, Y., 2016. Dynamic type matching. *Working Paper.*

Lyu, G., W. C. Cheung, M. C. Chou, C. P. Teo, Z. Zheng, Y. Zhong. 2017. Capacity allocation in Flexible Production Networks: Theory and applications. *Management Science.*

Mahdavi, M., Yang, T., Jin, R. 2013. Stochastic convex optimization with multiple objectives. *In Advances in Neural Information Processing Systems*, 1115–1123.

Marler, R.T., Arora, J.S. 2010. The weighted sum method for multi-objective optimization: new insights. *Structural and multidisciplinary optimization*, **41(6)**, 853–862.

Masin, M., Bukchin, Y., 2008. Diversity maximization approach for multiobjective optimization. *Operations research*, **56(2)**, 411–424.

Özkan, E. and Ward, A.R., 2016. Dynamic matching for real-time ridesharing. *Working Paper.*

Shalev-Shwartz, S. 2011. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, **4(2)**, 107–194.

Steuer, R.E. 1986. Multiple criteria optimization: theory, computation, and applications *Wiley.*

Sun, H., H. Wang, Z. Wan. 2019. Model and Analysis of Labor Supply for Ride-sharing Platforms in the Presence of Sample Self-selection and Endogeneity. *Transportation Research Part B: Methodological*, **125**, 76–93.

Electronic copy available at: https://ssrn.com/abstract=3356823

Tsiporkova, E. and Boeva, V. 2006. Multi-step ranking of alternatives in a multi-criteria and multi-expert decision making environment. *Information Sciences*, **176(18)**, 2673–2697.

Uziel, G., El-Yaniv, R., 2017. Multi-Objective Non-parametric Sequential Prediction. *In Advances in Neural Information Processing Systems*, 3374–3383.

Xu, Z., Li, Z., Guan, Q., Zhang, D., Li, Q., Nan, J., Liu, C., Bian, W., Ye, J., 2018. Large-Scale Order Dispatch in On-Demand Ride-Hailing Platforms: A Learning and Planning Approach. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, 905–913.

Voorneveld, M., Van den Nouweland, A., McLean, R. 2011. Axiomatizations of the Euclidean compromise solution. *International Journal of Game Theory*, **40(3)**, 427–448.

Wang, X., Agatz, N., Erera, A. 2017. Stable matching for dynamic ride-sharing systems. *Transportation Science*, **52(4)**, 850–867.

White, D.J. 1984. Multiobjective programming and penalty functions. *Journal of Optimization Theory and Applications*, **43(4)**, 583–599.

Yu, P.L. 1973. A class of solutions for group decision problems. *Management Science*, **19(8)**, 936–946.

Zinkevich, M. 2003. Online convex programming and generalized infinitesimal gradient ascent. *In Proceedings of the 20th International Conference on Machine Learning (ICML).*

Electronic copy available at: https://ssrn.com/abstract=3356823