

# Understanding Delivery Driver Order Acceptance: A Dynamic Discrete Choice Approach with Markov Decision Processes

**Dale Robbennolt**

University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA  
Email: dar4836@utexas.edu

**Jake Robbennolt**

University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA  
Email: jr73453@utexas.edu

**Zhaomiao Guo**

University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA  
Email: zguo@utexas.edu

**Stephen D. Boyles**

University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA  
Email: sboyles@austin.utexas.edu

**Chandra R. Bhat**

University of Texas at Austin  
Department of Civil, Architectural and Environmental Engineering  
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA  
Email: bhat@mail.utexas.edu

March 22, 2026

## Abstract

The rapid growth of peer-to-peer transportation solutions for real-time food delivery and ride-hailing services has transformed urban mobility and local commerce. Despite significant research on platform optimization, limited attention has been given to the behavioral dynamics of delivery drivers. This study employs a dynamic discrete choice modeling approach using a Markov decision process to investigate the decision-making process of delivery drivers when accepting or rejecting order requests. Using a dataset from Meituan, a major delivery platform based in China, we examine how spatial, temporal, and operational factors influence drivers' order acceptance decisions. We demonstrate the importance of considering both the dynamic nature of the decision process and of incorporating individual-specific heterogeneity to account for correlations between the decisions made by the same individual over time. The research reveals critical insights into driver behavior, demonstrating that order acceptance is a sophisticated, forward-looking process. Key findings include a significant preference for pre-booked orders, varying trip length preferences across peak and off-peak periods, and the importance of maintaining an optimal number of concurrent orders. The research contributes to the understanding of driver decision-making in urban transportation systems and provides actionable recommendations for improving platform efficiency and driver satisfaction.

*Keywords:* Dynamic Discrete Choice, Markov Decision Process, Unobserved Heterogeneity, Delivery Driver Dispatch, Food Delivery, Order Acceptance

# 1. Introduction

The rapid growth of peer-to-peer transportation solutions, including delivery services and ride hailing services facilitated by large management platforms, has transformed urban mobility and local commerce over the last few years. In particular, online shopping for groceries and prepared meals has increased dramatically, driven by large investments in real-time e-commerce services and rapidly popularized by changing consumer behaviors and preferences during and after the COVID-19 pandemic (Saphores et al., 2023). However, the long-term benefits to restaurants and stores, consumers, delivery drivers, and the platform owners themselves are still unclear (Sun et al., 2024). Significant research in recent years has focused on how these large platforms can maximize efficiency through their own policies and pricing behaviors as well as on the dynamics of consumer demand for these emerging systems (Yildiz and Savelsbergh, 2019; Ke et al., 2024). But, much less attention has been accorded to the behavioral characteristics of the drivers themselves. The behavioral dynamics of delivery drivers in urban transportation systems represent a complex interplay of spatial, temporal, and economic factors that fundamentally shape the efficiency of modern service and delivery platforms. At the core of these systems lies a fundamental decision point: whether drivers accept or reject offered requests for service and deliveries. As highlighted by Savelsbergh and Ulmer (2022), the acceptance rate is of critical importance for platform efficiency and service quality, as repeated rejections can increase delivery times as unaccepted orders remain idle. As many existing order matching algorithms make strong assumptions about the ways that delivery drivers will make order acceptance decisions (including the common assumption that all orders will be accepted), such rejection decisions made by freelance drivers can lead to instabilities in the matching process and cause downstream delays and operational costs (Hou et al., 2025).

To address the growing importance of these peer-to-peer transportation systems, a large body of research in recent years has focused on how platforms can optimize system-level decisions as well as how these decisions will influence profits and customer demands (Bai et al., 2019; Chen and Levin, 2019; Guda and Subramanian, 2019; Tan et al., 2022). However, while some authors have studied how these decisions impact the total supply of drivers (see, for example, Cachon et al. (2017); Zha et al. (2017)), most assume that drivers will always accept the trips they are assigned or assume some fixed rejection probability for all trips. In practice, the rejection rate for major ride hailing and deal-time delivery services is quite high, with documented cases of rejection rates between 40% and 80% (Gao et al., 2022; Xu et al., 2018). This high rejection rate is detrimental to the profits of the platforms and drivers who lose potential revenue as well as the passengers in ride-hailing services and to businesses and consumers using delivery services. Even when service/delivery requests are eventually fulfilled, high rejection rates can increase delays and cause other inefficiencies. In addition to the short-term losses and delays, high driver rejection rates also erode loyalty of businesses, drivers, and customers (Mao et al., 2019; Xu et al., 2018).

Thus, a deeper investigation into how drivers make decisions about when to accept or reject service/delivery requests would provide valuable insights into strategies to increase acceptance rates and improve overall system efficiency. In this context, developing an understanding of these decisions is complicated by the fact that the decision-making processes of delivery drivers are inherently dynamic, with each choice affecting not just immediate earnings but also future opportunities by spatiotemporally changing order availability (see Urata et al. (2021)). Specifically, because each decision influences how drivers will move through the transportation network, they must evaluate the benefits of the current delivery offer as well as the potential opportunity costs of accepting or rejecting based on their knowledge of their own future actions and their expectations of future offers (Yu et al., 2019). Therefore, we model this decision process using a dynamic discrete choice modeling (DDCM) approach that evaluates both current period utility and the expected future value of each decision using a Markov decision process (MDP) that accommodates this accumulated value over time. Driven by a spatiotemporal demand imbalance, MDPs have been commonly applied to taxi, ride-sourcing, and delivery drivers' empty driver behavior as they search for new passengers or order requests, although these studies have generally searched for optimal policies given exogenous reward functions (Gao et al., 2018; Lin et al., 2018). However, rather than attempting to determine the profit maximizing actions that delivery drivers should take given these exogenously provided assumptions about the relative valuation of different factors in the decision process, we apply the same underlying principles of the MDPs with the goal

of understanding the current decision processes of delivery drivers. That is, we study the order acceptance decision from the perspective of understanding how delivery drivers make choices and the value they place on different factors in their dynamic decision process. The results from this model provide important insights into the behaviors of delivery drivers, yielding valuable information that can be used to improve models of system performance and better align system-level strategies with driver needs and responses.

The remainder of this paper is organized as follows. Section 2 provides an overview of existing literature on dynamic discrete choice models as well as applications of MDPs in the context of delivery driver behaviors. Section 3 presents the framework used to model this decision process, as well as the characteristics of the dataset used. Section 4 provides the model estimation results and interpretations, along with two methods used to translate the results into appropriate forms for policy recommendations. Finally, Section 5 concludes the paper with a brief summary of the important findings and avenues for future research.

## 2. Literature Review

### 2.1. Delivery Driver Behaviors

Given the growing importance of on-demand transportation solutions, and the reliance of these systems on individual drivers for order fulfillment, a growing stream of literature in the urban-logistics domain has begun to examine the behaviors and choices of delivery and ride-hailing drivers. In particular, there has been growing attention to how delivery drivers make decisions relating to their overall employment (Wang et al., 2025) and timing of their work hours (Guda and Subramanian, 2019; Fielbaum and Tirachini, 2021; Allon et al., 2023; Safira, 2025), helping platforms ensure they have a sufficient fleet of drivers to meet demand. Recent research has also examined the actual driving behaviors of couriers and how they impact safety outcomes (He et al., 2023; Qian et al., 2024; Papantoniou et al., 2025; Barbosa et al., 2025). However, despite the growing attention in this area, few studies have directly addressed how delivery drivers (or other types of on-demand transportation workers) make accept/reject decisions for individual orders. Still, a few recent studies have revealed how these decisions are shaped by a combination of immediate trip attributes as well as platform constraints (Idug et al., 2023). In particular, monetary incentives have been a major focus, as they are both an important motivation for ride acceptance and a key lever by which management platforms can influence the desirability of specific rides (Xu et al., 2018; Ashkrof et al., 2022; Tu et al., 2025). Relatedly, trip distance for orders has been shown to be very closely related to driver productivity (Zuniga-Garcia et al., 2020), and it is used by drivers to assess the benefits of different orders and make acceptance/rejection decisions (Popan, 2024; Tu et al., 2025). Drivers also attempt to avoid any possible delays, rejecting rides that they consider to have higher risk for delays. In the ridehailing context, these decisions are often based on rider ratings (Ashkrof et al., 2020; Tu et al., 2025) and other rider-specific concerns such as the challenges associated with driving individuals with physical impairments (Gebresselassie, 2025). In contrast, in the delivery context, these delays are more likely to be incurred due to waiting at order pickup locations, so drivers may avoid certain businesses that are known to incur longer waiting times (Popan, 2024; Chen et al., 2025). Further, drivers seem more likely to reject requests when they are more confident about getting additional offers (particularly at high-demand times; see Ashkrof et al. (2022); Fatemi Pour et al. (2023); Chen et al. (2025)), as well as when they perceive any particular safety issues at either order pickup or dropoff locations (Popan, 2024). Individual driver characteristics also significantly influence order acceptance decisions. Drivers with stronger problem-focused coping skills handle high workloads and rapid requests more effectively and exhibit higher acceptance rates (Caparas, 2026). Finally, as discussed by Feng and Wang (2023), drivers may improve their overall profit by strategically rejecting certain orders, even if it leads to them having longer idling time, highlighting that drivers must consider the evolution of future orders, not just the current order they are offered.

### 2.2. Dynamic Discrete Choice Models

As mentioned above, delivery drivers make a sequence of decisions, each of which has immediate benefits and influences their alternatives for future decisions. While static discrete choice models assume that in-

dividual decisions are not influenced by the future evolution of their state, dynamic discrete choice models (DDCMs) consider the potential future rewards of forward-looking agents and are, therefore, well suited for the current context. A foundational example of this dynamic approach is Rust (1987), who developed an optimal stopping formulation for bus replacement decisions. Rust considered the actions of a single agent (rather than considering how different individuals make different decisions) and considered that random error components were independently distributed over time according to an extreme value distribution. Rust’s framework demonstrated the feasibility of estimating dynamic programming models for discrete choice problems, establishing a paradigm for analyzing forward-looking agents making sequential decisions under uncertainty. Formulating the model as a nested-fixed-point problem, Rust established a framework to incorporate an MDP into the maximum likelihood estimator to accommodate the dynamic aspect of the decision process. Numerous extensions of Rust’s optimal stopping model have emerged that accommodate dynamic market changes that influence consumer behavior over time, with applications primarily in the context of purchase decisions for durable products under changing market conditions, such as vehicle ownership models (Schiraldi, 2011; Cirillo et al., 2016; Liu and Cirillo, 2018).

In recognition of the importance of forward-looking decision making in the context of routing decisions, as each decision impacts the starting location for subsequent decisions, Fosgerau et al. (2013) adapted the structure of the DDCM to propose a recursive logit model. This study considered the network structure of nodes and links to develop a dynamic, link-based, routing model that directly incorporates the influences of location changes on the decision state and expectations of future choices. Because of this network structure, state transition for the recursive logit model is assumed to be deterministic (Oyama and Hato, 2017). While the use of an extreme value distributed error term enables rapid convergence for dynamic models, this choice of error distribution results in a lack of flexibility when extending to more choice alternatives or correlated decisions (Aguirregabiria and Mira, 2010). While subsequent extensions of the recursive logit model use mixing approaches to generalize the approach and allow for correlations among the choice alternatives (see Leong et al. (2024)), dynamic models do not commonly include random coefficients. Thus, some authors have pointed to the need to consider other error structures, particularly in the context of correlated choice alternatives, and given that iterative nature of DDCMs implies that these models are more susceptible to error misspecification than static models (Larsen et al., 2011).

An additional issue is that while Rust’s model considered only a single agent, much of the development of DDCMs in transportation (including the recursive logit models mentioned above) has ignored individual-specific heterogeneity when modeling dynamic decision processes, even when considering data from many individuals (see Cirillo and Xu (2011)). Several DDCM papers in economics have explored this issue, developing frameworks to consider unobserved individual-specific heterogeneity to allow for correlations in preferences over time. Berry et al. (1995) demonstrated the importance of incorporating individual heterogeneity in dynamic models, and explicit incorporation of unobserved heterogeneity was further developed by Gowrisankaran and Rysman (2012) in the context of digital camcorder purchases. Additionally, while many authors have assumed discrete support for unobserved heterogeneity (see, for example, Kasahara and Shimotsu (2009); Hu and Shum (2012)), there is evidence that models with continuous unobserved heterogeneity can be point identified (Bunting, 2024). Still, the only existing applications in transportation have considered unobserved segments to be discrete (Song et al., 2022). We extend such frameworks by allowing for continuous unobserved heterogeneity, allowing for a flexible representation of behavioral variation across agents in the current delivery context.

### 2.3. Markov Decisions Processes for Delivery Driver Behaviors

Outside of the discrete choice modeling literature, the forward looking behavior of agents in transportation systems has often been modeled through an MDP and solved using reinforcement learning (RL) techniques (Lai et al., 2024). Recent applications relevant to delivery driver behaviors are attempts to dynamically route fleets of vehicles on shared roads (Lazar et al., 2021), solving variants of the vehicle routing problem (Nazari et al., 2018), and developing dynamic dispatch systems for on-demand couriers (Chen et al., 2019, 2022) and ride-hailing drivers (Kullman et al., 2022; Singh et al., 2019). These approaches have been shown to be effective in developing optimal repositioning strategies for idle vehicles (Chen et al., 2021; Jiao et al.,

2021), as well as dispatching and pricing strategies for autonomous mobility-on-demand systems (Liu et al., 2022; Turan et al., 2020).

Many existing studies focus on the system-level decisions used to optimally control fleets of vehicles, rather than to predict and understand the behavior of individual drivers. When optimizing fleet dispatching, many papers have demonstrated the importance of number of vehicles and passenger requests in each zone (Chen et al., 2021; Liu et al., 2022; Robbennolt and Levin, 2023). Other studies have considered the optimal decisions of individual vehicles (often with the intent of aggregating decisions to the fleet level), demonstrating how the optimal decisions depend on the vehicle location, specific features of the vehicle/driver (such as state of charge of electric vehicles), and location specific characteristics (such as variations in supply/demand, local traffic conditions, and presence of gas stations or electric vehicle charging stations) (Jiao et al., 2021; Robbennolt et al., 2024; Sun et al., 2023; Yu et al., 2019). In addition, there is a large body of literature considering the optimal passenger search process (generally assuming that drivers always accept rides/orders), primarily considering how vehicles should reposition (on physical or virtual networks) to maximize their earnings (Yu et al., 2019; Jiao et al., 2021). Finally, some authors have instead allowed vehicles to choose from many actions, including repositioning, charging, and passenger pickup (if available) (Sun et al., 2023). Thus, papers in this area primarily assume that drivers focus on optimizing their own or collective long-run monetary costs (Jiao et al., 2021; Yu et al., 2019).

The major difference between our proposed approach in the current paper and the RL-based approaches listed above is that we do not propose rewards ahead of time. Instead, we suggest that human drivers may be motivated by many factors other than profit to make decisions, and attempt to understand the relative valuation of these different factors. In fact, depending on the information provided to drivers, an accurate evaluation of potential profits may be difficult (Gao et al., 2022; Rosenblat and Stark, 2016). With this in mind, the underlying principles of the MDPs (defined by states, actions, transition probabilities, and rewards) can still be applied, but with the goal of understanding the current decision process of delivery drivers rather than determining the profit maximizing actions they should take by making a priori assumptions about the relative valuation of different factors in the decision process (Powell, 2011). Once determined, the parameters of the proposed decision process can be used in many of the dispatch policies defined in the studies listed above rather than assuming predefined parameters.

## 2.4. Study in Context

The current study explores how couriers make decisions in the dynamic environment of on-demand delivery systems (in this paper, we will use the terms driver and courier interchangeably to refer to the delivery drivers making the accept/reject decision for offered delivery orders). The paper contributes to the literature in several ways. First, to our knowledge, the current study is the first to systematically model the order acceptance decision of delivery drivers. Although this process has been explored qualitatively (see Ashkrof et al. (2020)), and binary choice models of this decision have been applied with generated utility coefficients to test matching and pricing strategies (see Gao et al. (2022)), we are aware of no existing study that directly models order acceptance decisions in the context of delivery couriers. Doing so provides important insights into the needs of drivers as well as ways that delivery platform operators can improve system-level performance. Second, we model the behaviors of drivers in a dynamic framework using an MDP that accommodates the forward-looking behaviors of drivers. As order acceptance decisions amount to destination choices, involving predictable spatial-temporal movement through the network, this framework captures the fact that drivers must consider the opportunity costs of their evolving state as they make these decisions. This dynamic application of a discrete choice modeling approach offers improvements in behavioral realism over static discrete choice models that have been applied in similar choice contexts. For instance, Ashkrof et al. (2022) formulate a static discrete choice model of ride acceptance in the context of ridehailing using a stated preference experiment. Our proposed model builds on such work by using revealed preference data from actual delivery driver requests as well as aligning with the dynamic frameworks developed in other recent optimization models that aim to determine optimal driver behaviors in the ride-sourcing context (Shou et al., 2020). This method also allows us to estimate the courier discount factor, which determines the extent to which couriers consider the value of future orders relative to those in the current time-period. The

discount factor has important practical implications for how companies design compensation strategies and provide order offers to maximize order acceptance. Third, from an empirical standpoint, we consider how the trip length, geographic location, time of day (on- or off-peak), request type (pre-book or real-time order), and number of undelivered orders still held by the courier impact the acceptance decision for a new order. This allows us to make detailed recommendations about what types of orders are preferred by couriers and how companies can improve their overall acceptance rate by promoting orders of certain types.

Fourth, from a methodological standpoint, we allow continuous unobserved heterogeneity in driver characteristics, accommodating the fact that the repeated decisions of drivers are likely to be correlated over time. This type of heterogeneity has rarely been considered in DDCM models (see Song et al. (2022) for a detailed discussion), although the failure to control for unobserved heterogeneity has been shown to result in significant biases in many other econometric applications (Heckman and Singer, 1984; Bhat, 1996). Concurrently, we use random normal error components for both unobserved components (period-specific unobserved utility shocks and time-independent individual errors terms). While the vast majority of existing DDCM literature has assumed an extreme value distributed error term, the normal error component avoids the restrictive independence assumption inherent in extreme value distributions. Thus, the use of normal error components provides more flexibility for extending this modeling approach to a multivariate case or joint decision making scenarios with multiple decision outcomes, appropriate for additional routing/destination decisions in the delivery context. Additionally, although this choice introduces some additional computational complexity, the use of matrix-based approximations to the multivariate normal cumulative distribution function (Bhat, 2018) and Halton draws for simulation-based integration (Bhat, 2003) makes such an approach tractable and scalable even for large datasets. Finally, we use two methods to translate the results in ways that provide important insights for policy making. We use average treatment effects to demonstrate how companies can incentivize consumers to place orders of certain types that are more attractive to drivers. Using these results, companies can determine how to influence the mix of orders that are available at different times and locations to best maintain high acceptance rates, system efficiency, and driver satisfaction. Additionally, since companies have control over some aspects of the decision-making process itself (by deciding when and how to provide information to drivers and developing pricing and incentive schemes), we explore how changes to the coefficients in the MDP would influence the overall acceptance rate.

### 3. Methods

#### 3.1. Analytic Framework

The modeling framework adopted for this analysis is a panel version of a longitudinal dynamic binary probit model that accommodates heterogeneity across individuals due to unobserved individual attributes. The model is applied to analyze individual decisions regarding the acceptance of delivery order requests with the option to accept or reject each request. In the model, each decision maker is assumed to accept or reject each request based on the utility they receive from that decision in the current period and their expectations about how that decision will impact their future utilities. To model this expectation of future value, we characterize the sequence of repeated decisions with a discrete-time infinite-horizon MDP. A visual depiction of this MDP is shown in Figure 1 (the notations in the figure refer to the mathematical formulations discussed in the following section). In the figure, the decision maker is presented with a choice to accept or reject an order. At this stage, they know the details of the offered order and their current state in the system, and have a set of expectations about how accepting or rejecting the order would influence their future state and the potential future orders they may be offered. That is, they have some probability of transitioning from their current state to every possible future state (where a state is made of up an offered order and a set of characteristics of the driver within the system), where each of these transition probabilities depends on whether they accept or reject the current offer. Based on these expectations, drivers make decisions based on their total expected long-term reward (in the example in the figure, they choose the gray option of  $k=1$  and then have a real transition to a new state where they are offered their next order). Because the driver is uncertain about what potential future orders they may be offered, this predicted long-term value of their utility when making this initial decision is based on the probabilities of transitioning to all potential

future states rather than the actual state they end up at next. This MDP may then be used to characterize the decision process for each choice made by each individual. So, for the complete set of repeated choices made by the driver, we formulate a new MDP for each choice occasion such that we get their long-term expected utility at each choice occasion for the discrete choice model. This complete framework is shown in Figure 2, where each real transition leads to a new choice occasion for the individual with a complete MDP characterizing the decision process at each choice occasion.

As an example of this process, consider a driver who is offered their first order at the beginning of their shift. They may, of course, have some preferences for accepting or rejecting the order based simply on its immediate characteristics. For instance, they may prefer a longer-distance order that keeps them working continuously for a longer period of time or they may prefer an order in a part of the city where they are more experienced and comfortable with navigating. However, they are also likely to consider the longer-term rewards or costs of accepting or rejecting the order. If the order takes them far away from high-demand areas, they may prefer to reject it because they are afraid that they will not get another order once they complete the current one in that area. Conversely, if the order takes them through a high-demand area, they may be willing to accept it even if it is otherwise relatively unattractive, because they will be more likely to be offered another order along the way. Further, the quality of orders may vary in the future; the driver may prefer to take a trip which is less attractive in the current period if they believe it will lead to a higher probability of getting a high-value trip later. Thus, the accept/reject decision must weigh the immediate utility of each offer against its influence on the future bundle of orders they may receive. Note, drivers never directly compare multiple orders against each other since they are only ever offered one at a time. Thus, the driver can only weigh the current offer against the probability distribution of future orders given the outcome of the current binary decision.

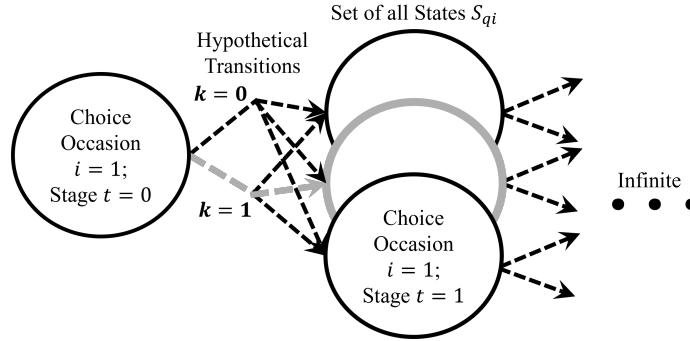


**Figure 1: Diagram of the Markov Decision Process constructed for a single choice occasion.**

As noted above, we use an infinite-time MDP for this process. In reality, drivers are not making decisions according to an infinite-time process, as they will eventually decide to stop taking orders and leave the system. However, the current data includes no information about the activities of drivers after they have left the system or details as to why they stopped working. Therefore, there is no way to know, using the current data, whether drivers make different decisions in the last period to end in a specific location or at a specific time, ending state information that is necessary for the formulation of a finite horizon MDP. Instead, the infinite-time assumption is a good approximation as long as two conditions are met. First, we can only consider decisions while drivers are actively engaged in a sequence of orders, rather than making a single isolated trip or at the end of their sequence. Therefore, we remove the final decision from every driver's trip sequence and only model their decision from the start of their working period until that point (see Section 3.3 for additional discussion). Second, the infinite horizon MDP is a better approximation when decision makers are more myopic (primarily consider their current period rewards). Thus, as long as the discount factor is sufficiently low (discounting future rewards heavily), the infinite horizon MDP is a reasonable approximation (see Lewis and Paul (2019)). Based on our estimation results (see Section 4), we do find a significant discount factor (confirming the dynamic nature of the decision process, as discussed in

more detail later), but the value of 0.59 is low enough to indicate substantially diminishing value for future orders. By the third future period the retained value is approximately  $0.59^3 \approx 0.21$ , and by the fifth period it falls to roughly 0.07, suggesting that the infinite-time assumption is reasonable with the removal of the final decision in each sequence. Even with relatively myopic drivers and the last decision removed, it might still be reasonable to assume that fatigue or end-of-shift effects substantially impact driver behavior. To test this we included several factors external to the MDP to test the impact of work duration. However, none of these resulted in any meaningful changes to behavior, further validating the infinite time assumption (see Section 4).

### 3.2. Mathematical Formulation

For the formulation of the DDCM, we develop an MDP to model the expected future value of each decision, as described above. To do so, we use the index  $q$  ( $q = 1, 2, \dots, Q$ ) to denote the individual and  $i$  ( $i = 1, 2, \dots, I_q$ ) to denote each observed choice occasion of that individual. We use  $k$  ( $k = 0, 1$ ) to denote choice alternatives (reject or accept). Figure 2 shows this framework visually, with each panel representing an individual  $q$  and each circle along the left side representing a choice occasion  $i$ . Then, for the formulation of the MDP at each choice occasion, let  $t$  ( $t = 0, 1, 2, \dots, \infty$ ) represent the stage (the evolution of the MDP between states which starts at zero for each MDP). In Figure 2, this is shown moving to the right from each choice occasion, with state evolution based on the choice of  $k$ . The indices  $i$  and  $t$  must be identified separately since a new Markov chain is formulated for each choice occasion  $i$  and evolves over  $t$ . In each Markov chain, the stage  $t = 0$  is the only observed stage and represents the choice of interest in the DDCM. To be clear, the MDP formulated for each individual  $q$  at choice occasion  $i$  encompasses all possible states that could be reached in the next stage, including the next observed state. This evolution is shown in Figure 2, where the gray line represents the observed transition and the black lines represent the other possibilities that the driver considers before future information is known (the mechanism of this evolution to possible future states will be discussed in more detail below). Finally, designate  $m_{qi}$  to be the actual observed choice of individual  $q$  for choice occasion  $i$  (at the observed period in the corresponding Markov chain,  $t = 0$ ).

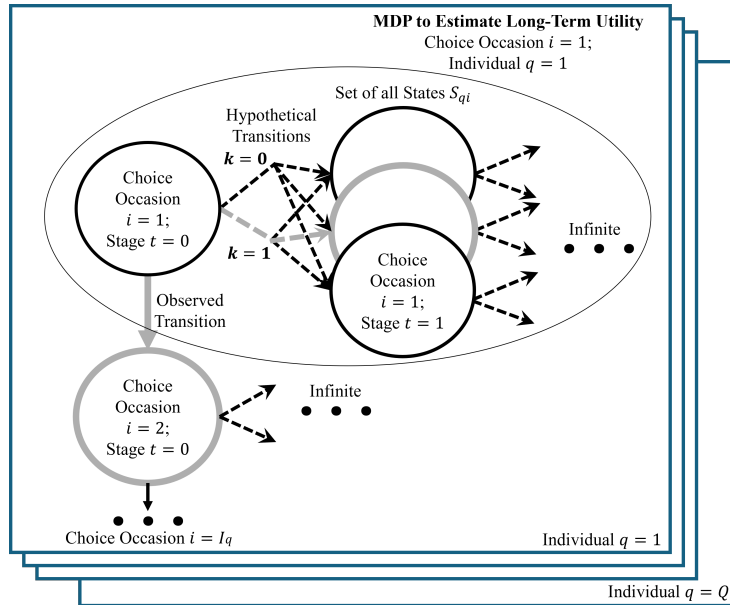


Figure 2: Diagram of the choice process and constructed MDPs for each individual and choice occasion.

To formalize this process, consider the individual-specific single-period utility function to take the form:

$$U_{qikt} = \beta'_k \mathbf{x}_{qit} + \varepsilon_{qit} + \eta_q, \quad (1)$$

where  $\mathbf{x}_{qit}$  is an  $(A \times 1)$  vector of exogenous attributes (including a constant for one of the two alternatives) defined as part of the state space  $\mathbf{s}_{qit}$  (where  $\mathbf{x}_{qit} \subset \mathbf{s}_{qit}$ ) and  $\beta_k$  is a corresponding  $(A \times 1)$  vector of coefficients capturing the effects of the exogenous variables on the utility of selecting alternative  $k$ . For the delivery order acceptance problem, the state variables define current information about the driver (including their location zone and number of existing orders in their queue) as well as information about the order itself (such as its length, destination, and whether it is a prebooked order). Define the set of all possible states that individual  $q$  could ever reach after choice occasion  $i$  as  $S_{qi}$  (although we do not have any individual-specific characteristics in our dataset, we define the set of possible states to be individual-specific to allow for this possibility). The period-specific error term  $\varepsilon_{qit}$  is assumed to be independent and identically normally distributed across individuals and choice occasions (with a fixed scale), while the individual-specific error term  $\eta_q$  is assumed to be independent and identically normally distributed (with scale  $\sigma^2$ ) across individuals but constant across choice occasions (both observed choice occasions and evolving choice occasions at each stage of the MDP) to generate correlations across the choice occasions of each individual. As stated previously, the ability to handle correlations across the choice occasions of each individual in a panel-formulation is an important contribution, as some individual biases are likely to persist over time, even as the driver moves through different states. However, this flexible structure has important implications for the formulation of the MDP because (a) an individual in a specific state at two different stages will not necessarily make the same decision in each case (because they will have different period-specific error terms  $\varepsilon_{qit}$ ), and (b) the MDPs formulated at each choice occasion  $i$  are not independent, since they will share the term  $\eta_q$  which is specific to the individual but not the choice occasion. While, in general, an infinite-horizon MDP will achieve steady state conditions, this is not the case in the current formulation. Instead, we will continue to keep track of the stage within the MDP (using the index  $t$ ) and demonstrate how the *expected value* of the value function will instead reach steady state, enabling the use of value iteration to solve the MDP.

Using the utility function provided in equation (1), the long-term utility for an individual  $q$  at choice occasion  $i$  can be formulated as an MDP and written as a fixed-point problem using the recursive Bellman's equation (Bellman, 1957):

$$V(\mathbf{s}_{qit}) = \max_k \{U_{qikt} + \theta \mathbb{E}[V(\mathbf{s}_{qit+1}|k)]\}, \quad (2)$$

where  $V(\mathbf{s}_{qit})$  is the total utility (short- and long-term) of choosing the optimal action while in state  $\mathbf{s}_{qit}$ ,  $\theta$  is the discount factor (to be estimated) that determines the extent to which the individual prioritizes the value of future rewards, and  $\mathbb{E}$  represents the expected value taken with respect to the state transition from  $\mathbf{s}_{qit}$  to  $\mathbf{s}_{qit+1}$ . This recursive form implicitly includes the sum over the infinite set of future stages, as the value in stage  $t+1$  is itself a function of the value in stage  $t+2$ , and so on.

To calibrate this MDP, a few additional components are needed. First, a set of transition probabilities govern the evolution of the state between stages in the MDP (defining the expected value in equation (2)). Let  $\pi(\mathbf{s}'|\mathbf{s}_{qit}, k)$  be the probability of an individual transitioning to each possible state (for all  $\mathbf{s}'$  in  $S_{qi}$ ) in the next stage, given their state in the current stage and their selected choice alternative  $k$  (see in Figure 2 that there is a different set of transition probabilities based on the choice  $k$ , which could account for drivers relocating to better locations when unwanted orders are rejected or other behaviors that are dependent on order acceptance). Therefore, the expected value in equation (2) becomes the sum over the value of each possible state in the next stage, multiplied by the corresponding transition probability ( $\mathbb{E}[V(\mathbf{s}_{qit+1}|k)] = \sum_{\mathbf{s}' \in S_{qi}} \pi(\mathbf{s}'|\mathbf{s}_{qit}, k) V(\mathbf{s}'|k)$ ). Second, for the maximum likelihood formulation, the utility components for each alternative are needed (rather than solely the maximum for the selected alternative as in the traditional Bellman equation). That is, the values of  $V_k(\mathbf{s}_{qit=0})$  are needed for both  $k=0$  and  $k=1$ . To accommodate this, we move the maximization inside the expected value function to solve for the total utility of each alternative (maintaining the form of the MDP for all future periods while keeping all utility values in the

starting period for each choice occasion). We rewrite equation 2 with these updates as:

$$V_k(\mathbf{s}_{qit}) = U_{qikt} + \theta \sum_{\mathbf{s}' \in S_{qi}} \pi(\mathbf{s}' | \mathbf{s}_{qit}, k) \max_k \{V_k(\mathbf{s}')\}, \quad (3)$$

where the value function is now defined separately for each choice alternative and the expected value has been replaced with a sum using the definition of the expected value for a discrete random variable.

Next, recall that the individual-specific single-period utility  $U_{qikt}$  contains two error terms  $\eta_q$  and  $\varepsilon_{qit}$ . Since the individual-specific error terms  $\eta_q$  are assumed to be known to the individual and constant over time (and the MDP is solved separately for each individual), they can be subsumed into the state space to get the conditional utility. Essentially, the fixed-point problem can be easily solved, conditional on a given value of  $\eta_q$ , which then simply appears as another constant in the fixed-point problem.

$$V_k(\mathbf{s}_{qit} | \eta_q) = \beta'_k \mathbf{x}_{qit} + \varepsilon_{qit} + \eta_q + \theta \sum_{\mathbf{s}' \in S_{qi}} \pi(\mathbf{s}' | \mathbf{s}_{qit}, k) \max_k \{V_k(\mathbf{s}' | \eta_q)\}, \quad (4)$$

Finally, since the fixed-point problem includes the expected value of the maximum of two random variables (based on the period-specific error terms  $\varepsilon_{qit}$ , the future values of which are not known by the decision-maker in the current period), the formula can be redefined to give the expected conditional utility of each alternative by removing the period-specific error components in the present period and using the expected value of the maximum of two random normally distributed variables (see Clark (1961)). Since the expected value of  $\varepsilon_{qit}$  is constant over time, the problem then becomes stationary with respect to the stage. This gives the fixed-point problem:

$$\begin{aligned} \bar{V}_k(\mathbf{s} | \eta_q) = & \beta'_k \mathbf{x} + \eta_q + \theta \sum_{\mathbf{s}' \in S} \pi(\mathbf{s}' | \mathbf{s}, k) \times \left[ \bar{V}_0(\mathbf{s}' | \eta_q) \Phi\left(\bar{V}_0(\mathbf{s}' | \eta_q) - \bar{V}_1(\mathbf{s}' | \eta_q)\right) \right. \\ & \left. + \bar{V}_1(\mathbf{s}' | \eta_q) \Phi\left(\bar{V}_1(\mathbf{s}' | \eta_q) - \bar{V}_0(\mathbf{s}' | \eta_q)\right) + \phi\left(\bar{V}_0(\mathbf{s}' | \eta_q) - \bar{V}_1(\mathbf{s}' | \eta_q)\right) \right], \end{aligned} \quad (5)$$

where  $\bar{V}_k(\mathbf{s} | \eta_q)$  refers to the expected conditional utility of being in state  $\mathbf{s}$  and selecting alternative  $k$  given a value of the individual specific error term  $\eta_q$ ,  $\Phi(\cdot)$  is the cumulative standard normal distribution function, and  $\phi(\cdot)$  is the standard normal density function. Note that, in this equation, there is no difference in the expected conditional utility for two individuals in the same state beyond their individual specific error terms, and no difference in being in the same state at different points in time (different choice occasions or different stages). Therefore the indices  $q$ ,  $i$ , and  $t$  are dropped from the state and from  $\mathbf{x}$ , which is simply a subset of the state. Thus, the fixed point problem can be solved for every possible state simultaneously using value iteration given a specified value  $\eta_q$ , and the realized long-term utility value for individual  $q$  at choice occasion  $i$  is simply  $\bar{V}_k(\mathbf{s}_{qit=0} | \eta_q)$ . In equation 5, the first term is the observed component of the single-period utility function and the second term is the fixed individual-specific error term. The third term, is the discounted (by  $\theta$ ) sum of the probability of transitioning to each future state multiplied by the long-term value of being in that state. Since this long-term value is the maximum of two random normally distributed values it takes the complex form shown, where the first two components are essentially the probability weighted values of each alternative being the optimal choice (maximum long-term value), while the third component accounts for the uncertainty in the maximum value (since the maximum of two random variables is always greater than the maximum of their expected values) as defined in Clark's formula (Clark, 1961).

Then, the likelihood function can be written separately for each individual using the standard form of the binary probit model for their sequence of decisions. We integrate over the individual-specific error term  $\eta_q$  using a normal density function to account for unobserved heterogeneity that persists across all choice occasions for an individual:

$$L_q(\beta, \theta, \sigma) = \int_{\eta_q = -\infty}^{\infty} \left[ \prod_{i=1}^{I_q} \Phi\left(\bar{V}_{k=m_{qi}}(\mathbf{s}_{qit=0} | \eta_q) - \bar{V}_{k=1-m_{qi}}(\mathbf{s}_{qit=0} | \eta_q)\right) \right] \phi(\eta_q) d\eta_q. \quad (6)$$

This formulation involves the fixed-point problem (equation 5) nested inside the maximum likelihood estimator (equation 6) and simulation over the integral for the individual-specific error term, requiring the fixed-point problem to be solved (using value-iteration) a large number of times. Multiplying the individual-specific likelihood contributions creates the total likelihood. This function is then maximized over the structural parameters ( $\beta$ ,  $\theta$ , and  $\sigma$ ).

The computational complexity of this problem can be reduced using two mechanisms. First, we use the Halton sequence to draw realizations of  $\eta_q$  from the normal distribution, which has been demonstrated to outperform traditional pseudo-Monte Carlo methods (Bhat, 2001, 2003). Additionally, since the solution to the fixed-point problem provides the expected conditional utility at every state, it need only be solved once for each Halton draw (for each individual). Then, for each draw, the likelihood of each choice occasion for that individual can be determined from the same fixed-point solution without re-solving the fixed-point problem for each choice occasion. Since some individuals in our dataset have as many as 705 choice occasions (over an eight day period), this approach greatly reduces the computation time.

### 3.3. Data Description

This study uses data provided by Meituan, an on-demand delivery service based in China (Meituan, 2024). The dataset consists of 654,343 records of orders offered to a driver pool of 4,955 couriers. The data covers an eight-day period from October 17, 2022, to October 24, 2022. Substantial preprocessing was necessary to get this data into a form usable for the DDCM. First, in the dataset, the location of the courier at the time of each order offer was only available for drivers who accepted the ride. To approximate locations of drivers who rejected the ride, we perform a linear interpolation between the locations of the request immediately before and after the rejection decision. Although the result is only an approximation of the location at the time of the offer, locations and distances are grouped into large zones and discrete bins, making this approach sufficient for the selected application. Validating this approach for the driver’s acceptance decision (where the locations are available at the time of the offer), we find that decisions are classified into the correct location bin 90.83% of the time and the correct trip length bin 90.72% of the time. This is a conservative estimate for the classification of courier location for rejected orders, because the distance spread between the known driver locations for unobserved rejected orders is 47.07% lower than that of the observed accepted orders (because, these drivers generally receive another offer more quickly). These classification errors could be avoided in future studies if location information is collected for all decisions rather than just the acceptance decisions.

Meituan employs two types of drivers: “full-time” (zhuansong) and “part-time” (zhongbao). The full-time riders are paid a fixed monthly salary plus a small fee for orders they complete. Part-time drivers have no salary but are paid more per ride. The order selection mechanisms also differ as the full-time drivers follow the traditional process of being offered specific rides which they choose to accept or reject, while the part-time drivers must compete for rides which they select from the system themselves, so there is no “rejection” decision (Zhang, 2019) (note that this terminology distinction between zhuansong and zhongbao refers more generally to labor regime distinctions between structured contracted drivers and crowdsourced drivers rather than legal employment definitions commonly used in the United States). Therefore, as the decision processes are substantially different for the two driver types and only full-time drivers make accept/reject decisions, the part-time drivers were removed. While there was no identification of driver type in the dataset, a geographic identifier for the zone of the sender was present which defined overlapping areas for each driver type. These zones (provided by Meituan) correspond to separate business districts and are used to identify driver movements throughout the region. However, they are anonymized, and no further geographic or land-use characteristics are available for them. Also note that the coordinates provided for all locations have been shifted to preserve anonymity. Therefore the zone shapes shown in Figure 4 and Figure 5 are likely to be somewhat distorted compared with real-world district locations. To separate the drivers, we generated a bipartite graph matching drivers to all zones they served. Using a breadth-first search on this graph, we identified the subset of areas served by part-time drivers and the subset served by full-time drivers, allowing us to remove the part-time drivers from consideration. Although we do not model the behaviors of the part-time drivers further in this paper, it is interesting to note that, in addition to the different order assignment

mechanisms, the part-time drivers are much less likely to move between zones (the part-time drivers were split into 6 distinct geographic regions which never overlapped). In addition, data from two outlying areas that had no trips connecting to the city center were removed from the dataset.

While geographic region IDs were included for sender location, only the (shifted) geographic coordinates were provided for courier location and recipient location. To determine the actual boundaries of these geographic regions, the sender locations were clustered into 100 groups using k-means clustering, which were then assigned the most common geographic ID. These clusters were merged to form the larger geographic regions so recipient and courier locations could be classified. Figure 3 shows the sender locations and zones (on the left) and a Voronoi diagram depicting the clustering method used to generate the zones (on the right). The final zones used in the analysis are shown in Figure 4. To validate, the sender locations were reclassified based on the inferred geographic boundaries, finding a correct classification rate of 97.45%. See Figure 5 which shows the classified locations of the courier and recipient for each order.

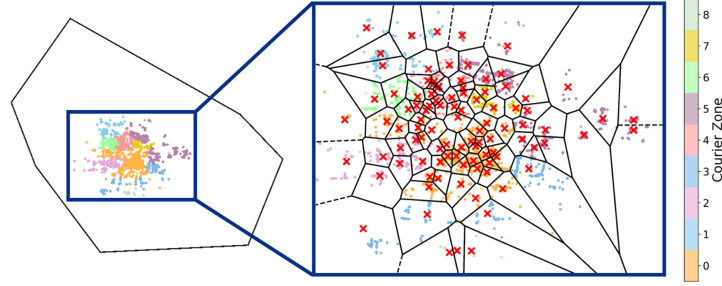


**Figure 3: Sender Locations within convex hull of all sender, courier, and recipient locations (left) and Voronoi diagram of k-means clusters (right)**

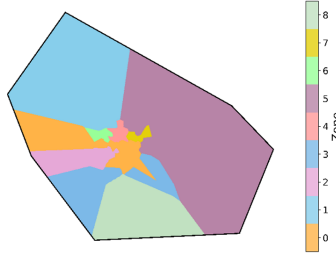


**Figure 4: Zones determined based on the k-means clusters**

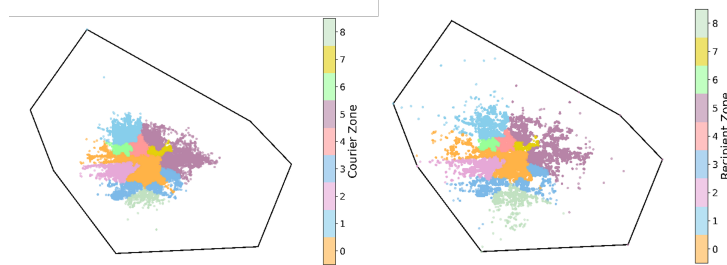


**Figure 5: Courier (left) and recipient (right) zone classification based on the zones in Figure 4**

Finally, as mentioned in Section 3.1, we model the driver decision process as an infinite time MDP.

However, drivers near the end of their working period may be motivated by other concerns, such as ending their trip near a specific location. To avoid these decisions, we first grouped all decisions by a single driver into blocks, such that each block consisted of a string of rides with no more than 60 minutes of down-time between orders. We then removed the last decision of each group, as this decision is most likely to be influenced by external factors not considered as part of the MDP. After data processing, the dataset consisted of a total of 390,151 total decisions made by 3,786 unique full-time drivers, with a rejection rate of 19.85%.

### 3.4. Exogenous Variables and State Space

The state space includes information about the current state of the driver (courier zone, time of day, and number of existing orders in the driver’s queue) and the parameters of the order request that is offered to the driver (trip length, request type, and recipient zone). For the purposes of the dynamic discrete choice method, the state space was simplified to include three times of day, morning peak (10:30 am – 1:00 pm), evening peak (4:30 pm – 8:30 pm) and off-peak (all other times). The time zone used is China Standard Time (GMT+8), and these periods were chosen to match the observed fluctuations in the data (peaks match food delivery windows for lunch and dinner meals). Trip lengths were classified into short (<25,000 units), medium (25,000-50,000 units), and long (>50,000 units). The distance units are arbitrary due to the shifted coordinates provided in the data to preserve privacy, but this classification provides an overall sense of driver preference based on the range of trip lengths available in this empirical context. The request type refers to whether the order was a pre-order (booked in advance) or a real-time request. Courier and recipient zones were classified into 9 areas (see Figure 4), and the number of undelivered orders were classified into four bins (of 0, 1, 2, and 3 or more undelivered orders). Descriptive statistics for the exogenous variables are shown in Table 1. Note that, as discussed above, the characteristics of these zones are unknown. However, the inclusion of zone-specific variables in the model serves to absorb location-specific spatial heterogeneity. As a result, the remaining behavioral parameters are estimated net of these geographic effects, which may offer some insulation against the spatial specificity of this particular dataset. Finally, while the monetary value of a ride is an important part of the driver’s decision process, Meituan did not provide any information about driver compensation. However, Meituan full-time drivers are paid a monthly salary, so the fees for completing a ride are relatively small. In addition, fees for delivery drivers are usually calculated as a combination of a fixed amount for each ride and a variable component based on other variables already included in the state space (such as the trip length), or not shown to the courier at all (Ashkrof et al., 2020). Still, the lack of any details about these payment structures is one significant limitation that may impact our results. Given that these payment structures are unobserved, it should be noted that our estimates relating to time of day, trip length, and geography may be impacted in part by any existing payment/reward mechanisms that may be based on these trip characteristics. Thus, the results of these variables, discussed in the following section, should be interpreted in the context of their implications for monetary compensation in addition to direct preferences couriers have for these attributes (for instance, a preference for a longer trip aggregates the direct preference for that trip type with the overall higher compensation rewarded for longer trips). Were payment information available, these effects could be distinguished, but without it valuable insights can still be drawn.

For the MDP, it is also necessary to compute the probability of transitioning from each state to the next. We do this using the transitions observed in the data within each block of offers made to the same courier. To create the most representative transition matrix, we split the exogenous variables into groups to calculate transition probabilities independently, multiplying them together to get the final transition matrices. Transitions between times of day, request type, and number of accepted undelivered orders were all calculated independently conditional on the accept/reject decision. However, the transitions based on the courier and recipient zone as well as the trip length are all related. We calculate the transition probability from each courier-recipient zone pair to each other pair and then calculate the trip length (category) as a function of the new courier-recipient zone pair (again, conditional on the accept/reject decision). Note that this allows for couriers to transition to new zones (for repositioning or to fulfill existing orders) even if they reject the currently offered order. Also, the trip length is a function of the sender zone as well, although

Table 1: Descriptive Statistics for Exogenous Variables

| Exogenous Variables                   | Frequency | Percent in Sample |
|---------------------------------------|-----------|-------------------|
| Time of day                           |           |                   |
| Off-peak                              | 113,259   | 29.03             |
| Morning peak                          | 138,116   | 35.40             |
| Evening peak                          | 138,776   | 35.57             |
| Trip Length                           |           |                   |
| Short                                 | 137,796   | 35.32             |
| Medium                                | 205,115   | 52.56             |
| Long                                  | 47,240    | 12.12             |
| Request type                          |           |                   |
| Real-time                             | 375,653   | 96.28             |
| Pre-order                             | 14,498    | 3.72              |
| Number of accepted undelivered orders |           |                   |
| 0                                     | 363,745   | 93.22             |
| 1                                     | 12,128    | 3.12              |
| 2                                     | 8,831     | 2.26              |
| 3+                                    | 5,447     | 1.40              |
| Recipient Zone (zone 0)               |           |                   |
| Zone 0                                | 115,802   | 29.68             |
| Zone 1                                | 39,153    | 10.03             |
| Zone 2                                | 32,415    | 8.31              |
| Zone 3                                | 26,996    | 6.92              |
| Zone 4                                | 44,110    | 11.31             |
| Zone 5                                | 81,286    | 20.83             |
| Zone 6                                | 26,523    | 6.80              |
| Zone 7                                | 19,320    | 4.95              |
| Zone 8                                | 4,546     | 1.17              |
| Courier Zone                          |           |                   |
| Zone 0                                | 141,253   | 36.20             |
| Zone 1                                | 26,449    | 6.78              |
| Zone 2                                | 30,164    | 7.73              |
| Zone 3                                | 12,079    | 3.10              |
| Zone 4                                | 73,887    | 18.94             |
| Zone 5                                | 69,091    | 17.71             |
| Zone 6                                | 16,523    | 4.24              |
| Zone 7                                | 18,976    | 4.86              |
| Zone 8                                | 1,729     | 0.44              |

only courier and recipient zones are explicitly included in the state space.

The sequential estimation approach used here to estimate transition probabilities directly from observed sample frequencies prior to estimation of the structural preference parameters is standard practice in related dynamic modeling approaches (Rust, 1987; Aguirregabiria and Mira, 2007; Arcidiacono et al., 2016). Beyond the computational benefit of estimating these parameters directly, it avoids any identification issues between the transition probabilities and structural preference parameters (Aguirregabiria and Mira, 2010). By estimating transition probabilities nonparametrically from observed frequencies, we exploit variation in the data that is purely informative about the transition process, ensuring that the structural parameters are identified from choice behavior conditional on transitions rather than from a conflation of the two. This approach is also consistent with the assumptions of the conditional choice probability (CCP) approach of Hotz and Miller (1993), which similarly estimates the transition probabilities from observed frequencies in the data (Arcidiacono and Miller, 2011; Chou and Derdenger, 2025). However, since the traditionally estimated standard errors do not account for any sampling uncertainty in the transition probabilities, we estimate standard errors using the approach of Akerberg et al. (2012). Also, given that these transition probabilities are estimated from a large pooled sample, they have high precision, and so the effect of this adjustment is relatively minor.

Further, this approach to estimating the transition probabilities implicitly assumes that drivers have rational expectations regarding state transitions. That is, their subjective beliefs about future states align with the objective probabilities observed in the data. This is a necessary identification restriction, as transition perceptions can only be inferred separately from objective transitions if (a) the exact mechanisms behind objective state transitions are known or (b) additional data is collected about participant’s perceptions about transition probabilities (An et al., 2021). Since the agents in this context are experienced couriers making repeated accept/reject decisions, it is likely that they have developed reasonable expectations about how states evolve over the course of their working day. In fact, qualitative evidence from Ashkrof et al. (2020) suggests that couriers do develop these expectations. Nevertheless, future research could examine the quality and accuracy of these expectations directly, for example by eliciting subjective beliefs from couriers and comparing them to the real transition probabilities observed in the data, which would provide valuable evidence on the validity of the rational expectations assumption in this setting. Finally, this framework assumes the transition process governing future offer generation is stationary and not strategically influenced by individual driver behavior. In practice, offer generation is likely algorithmically controlled by the platform and may adapt over time or respond to individual accept/reject decisions, such that the transition probabilities faced by a driver could depend on their own behavioral history (Muldoon and Raekstad, 2023; Kadolkar et al., 2025). To the extent that any such relationship was stable during our observation period (which it is likely to be over the short time-horizon considered), it would be implicitly captured in the empirically estimated transition probabilities. That is, the fact that algorithms may systematically group certain order types together or perform destination matching is captured in the observed transition probabilities (and highlights exactly the importance of considering the dynamic nature of the model). But, we do not capture any potential confounding effects caused by strategic decision making by drivers to influence the transition probabilities themselves (though there is evidence suggesting that these strategies may not play a large role in driver acceptance decisions; see Idug et al. (2023)).

The formulation of the MDP using this set of discrete states and the calculation of transition probabilities in this way provides a good approximation of the driver decision making process but entails a simplification of the ways that these variables interact and evolve over time. To better approximate this process, a continuous-time MDP could be used to account for the rate of offers being made and connected vehicle data could be used to track driver location between orders to account more comprehensively for the evolution of each driver’s queue of orders and resulting trip lengths (to maximize overlap and keep track of served orders). However, such a procedure would require (a) additional data on driver location in continuous time, (b) additional routing computations, and (c) an exponentially larger state space entailing longer computation times for the continuous-time MDP. Due to the nested nature of the MDP within the DDCM framework and the need to integrate over the individual specific error term, the much larger state space could be computationally prohibitive. In addition, since many companies limit the information seen by drivers when making these

decisions (for detailed discussion of this issue, see Ashkrof et al. (2020, 2022); Rosenblat and Stark (2016)), it is unclear whether the more complicated routing mechanisms embedded within a continuous time MDP would significantly impact the accept/reject decision process.

## 4. Results

The model specification is presented in Table 2. All variables are in the form of dummy variables due to the constraints of solving the MDP with a discrete state space. The specification was developed using an iterative process of testing the statistical fit of various combinations of exogenous variables. A t-statistic threshold of 1.96 (corresponding to a 0.05 level of significance or 95% confidence level) was used to retain variables in the model. The coefficients shown in the table refer to the impacts of each variable on the underlying utility of accepting an order relative to rejecting. The constants in the first row do not have any meaningful interpretation but are simply estimated to match the observed choice proportions of the sample.

In addition to the variables discussed below, the impact of work duration was considered but found not to result in any meaningful changes in behavior. Specifically, the effects of (a) the number of orders already completed by the driver in their shift, (b) the amount of time elapsed since the beginning of their shift, and (c) the proportion of their shift that they had completed were tested. These work duration effects were considered external to the MDP. That is, the assumption when testing these coefficients was that they influenced the utility of each choice occasion directly, but not the evolution of future orders. The fact that behaviors seem to be relatively stable over the course of a shift also provides additional support for the infinite-time approximation, as no substantial behavior changes were observed as drivers neared the end of their workday.

The results are organized into several sections. Section 4.1 presents the main estimation results, showing the direct effects of each of the exogenous variables on the binary acceptance outcome. In addition to demonstrating the general behavioral tendencies of delivery drivers, these results have important implications for platforms, such as determining how best to match new orders with couriers based on their existing set of orders. Section 4.2 discusses the model fit in relation to a series of additional models that either do not consider the dynamic nature of the decision process or do not consider the correlations between repeated decisions made by a single decision maker. Then, to further explore the direct impacts of each exogenous variable on the acceptance decision, two types of sensitivity analysis are performed. Section 4.3 discusses average treatment effects (ATEs), which represent the effect of changing the value of each exogenous variable on the likelihood of accepting an order, and Section 4.4 presents coefficient sensitivity analysis, which demonstrates the impacts of changing the values of each of the coefficients. The ATEs help reveal how operators can modify their offerings or payment structures to help get preferred order types and improve overall system efficiency. In contrast, the coefficient sensitivity indicates how platforms could modify the decision process itself to influence how their couriers are making decisions in this context.

### 4.1. Main Estimation Results

The main estimation results are presented in Table 2. The effects shown should be interpreted as general tendencies, as there is significant variation in behavior around these general (average) tendencies. As seen in the table, couriers tend to be more selective (they are more likely to reject order requests) during the morning and evening peak periods, compared with off-peak times. The higher rejection rate during peak periods is unsurprising given that higher consumer demand during these periods allows couriers to choose between several competing offers while being confident of rapidly receiving additional offers if they reject the current order. This behavior may have the indirect effect of spreading demand away from peak periods as consumers whose orders are rejected often during the peak period may shift to off-peak periods. This could be beneficial for restaurants and stores which struggle to meet high demands for both in-person and delivery service during peak periods. On the other hand, it may also erode loyalty to the platform both from companies and individuals if orders are never served or are often late.

We find that preferences for orders with different trip lengths depend on the time of day. During the off-peak periods, couriers place a greater utility on shorter trips (see the negative overall effects of trip length

**Table 2: Main Estimation Results**

| <b>Exogenous Variables (base)</b>                         | <b>Parameter</b> | <b>t-statistic</b> |
|-----------------------------------------------------------|------------------|--------------------|
| Constant                                                  | 1.27             | 114.87             |
| Time of day (off-peak)                                    |                  |                    |
| Morning peak                                              | -0.53            | -71.24             |
| Evening peak                                              | -0.43            | -50.72             |
| Trip Length (short)                                       |                  |                    |
| Medium                                                    | -0.12            | -12.75             |
| Long                                                      | -0.21            | -15.98             |
| Interactions                                              |                  |                    |
| Morning peak * Medium trip length                         | 0.22             | 16.13              |
| Morning peak * Long trip length                           | 0.34             | 17.07              |
| Evening peak * Medium trip length                         | 0.18             | 13.85              |
| Evening peak * Long trip length                           | 0.29             | 12.17              |
| Request type (real-time)                                  |                  |                    |
| Pre-order                                                 | 0.31             | 26.84              |
| Number of accepted undelivered orders (0)                 |                  |                    |
| 1                                                         | 0.14             | 10.36              |
| 2                                                         | 0.09             | 5.51               |
| 3+                                                        | -0.10            | -6.36              |
| Recipient Zone (zone 0)                                   |                  |                    |
| Zone 1                                                    | 0.33             | 48.14              |
| Zone 2                                                    | 0.12             | 12.50              |
| Zone 3                                                    | 0.25             | 24.50              |
| Zone 4                                                    | 0.12             | 20.08              |
| Zone 5                                                    | 0.29             | 42.82              |
| Zone 6                                                    | -0.10            | -11.53             |
| Zone 7                                                    | 0.27             | 35.56              |
| Zone 8                                                    | 0.54             | 20.53              |
| Discount factor                                           | 0.59             | 12.27              |
| Standard deviation of unobserved individual heterogeneity | 0.69             | 115.42             |

in the table). This is likely because longer trips may take couriers away from the center of the service area, resulting in more deadheading distance as couriers return. In these off-peak periods, couriers may be less confident that these larger spatial movements through the system will bring them back to the service areas they prefer. During peak periods, however, couriers place a greater utility on longer trips compared with shorter trips. Since monetary payments are often tied to trip length or travel time, accepting a longer trip provides a larger single payment and ensures that the courier is working continuously for a longer period. Thus, couriers may generally prefer to take a single longer trip compared with several shorter trips that may entail more empty waiting time when they are not making money. Concurrently, during peak periods, demand may be sufficient for couriers to find return orders which are not as readily available during off-peak hours. It is worth noting that all the trips in the current dataset seem to be intra-city trips that keep couriers within the contained service area, though zones around the periphery do still have less activity. Therefore, while some longer trips may take couriers further from downtown locations with more regular requests, there is less danger that couriers may need to deadhead long distances to return to the city than if they were considering even longer inter-city trips.

The request type also influences the order acceptance rate, with couriers generally preferring pre-booked orders compared with real-time requests. This preference for pre-orders likely stems from the increased certainty and ability to plan ahead that these orders offer couriers. Since pre-orders generally provide more flexibility in terms of order pickup and drop-off time, they allow couriers to better organize their schedules and continue to adjust as new orders are received along the way. Therefore, platforms should promote pre-orders more heavily to consumers since drivers tend to prefer these orders (and may be willing to accept lower fees in exchange for the increased security these orders offer). Similarly, customers who live in areas with particularly high rejection rates or who have high priority orders at peak hours should consider utilizing pre-orders as this may increase the likelihood of rapid service even if companies do not specifically guarantee better service for such orders.

Similarly, the effect of a courier’s existing set of orders reveals some interesting results. Couriers have the highest propensity of accepting a new order when they already have one accepted order. When a courier has yet to accept any requests, they are likely to be more selective because this first order will impact the future set of orders that they are offered based on where they go in the network. Specifically, the order matching algorithms used by companies will match future offers to this existing route, meaning that the first order in any batch of orders has an outsized impact on the rest of that batch of orders and may be more carefully considered. However, the utility of accepting a new order also declines when the courier already has more than one existing order. This suggests couriers are comfortable managing a small queue of orders but become more selective when their workload becomes too high. These results suggest that delivery platforms should try to maintain drivers with 1-2 open orders, as this maximizes their acceptance rate and keeps them working rather than waiting idle for new requests.

We find that the recipient zone does play a significant role in order acceptance, with drivers showing a clear preference for trips to some zones rather than others. These spatial preferences are a reflection of the expected future opportunities to get orders from these zones as well as factors such as traffic conditions, navigability, courier comfort, and safety, which might affect courier decision making in entering those areas. We discuss these effects in more detail in Section 4.3.

The statistically significant discount factor indicates that couriers consider future opportunities when making order acceptance decisions, supporting the dynamic nature of their decision-making process. The moderate value of the discount factor (0.59) suggests that couriers do place considerable weight on the future conditions that are influenced by their current decision and that they are potentially looking ahead across one or more future orders as they make current decisions, although the perceived value of these future rewards does decline substantially. Finally, the standard deviation of unobserved individual heterogeneity is also statistically significant, suggesting that there is considerable variation in courier preferences and behavior. Specifically, there are unobserved individual characteristics that lead some couriers to be more likely to accept any order they are offered while others are generally more likely to reject. This result indicates a need for delivery platforms to collect more driver characteristics to better understand why some drivers reject orders more often and how best to match specific requests to specific drivers.

## 4.2. Model Fit

The performance of the proposed model can be compared with that of models that ignore either the dynamic nature of the decision process or correlations between the repeated decisions made by the same individuals. Although the significant discount factor and standard deviation of unobserved heterogeneity already point to the importance of including these aspects of the model, the proposed model is compared with three additional models in Table 3. The first column shows the proposed longitudinal dynamic binary probit model, the second column shows a dynamic binary probit model that considers the forward-looking nature of the decision process but does not consider any unobserved individual heterogeneity (the standard deviation of unobserved heterogeneity shown at the bottom of Table 2 is assumed to be zero), the third column shows a longitudinal binary probit model that does accommodate this individual heterogeneity but does not consider the forward-looking behaviors of agents (the discount factor is assumed to be zero), and the fourth column shows a binary probit model that does not consider either individual heterogeneity or the dynamic component.

The relative performance of the four models is assessed by comparing a series of goodness-of-fit metrics. First, the adjusted likelihood ratio index for the proposed model is larger than that of any of the other models, suggesting that the proposed model offers a better overall fit. Second, the proposed model can be compared with the independent model using a Bayesian Information Criterion (BIC) statistic. Since the model with the lower BIC value is preferred, the proposed model outperforms all three other models. Third, the predictive performance of the models can be compared based on the average probability of correctly predicting a rejection decision. This metric was computed because, due to the imbalanced nature of the dataset, significant improvements in predicting rejections may be masked in the overall likelihood-based metrics. Again, the proposed model outperforms each of the other three models by this metric (note that the probability of a correct prediction within this group may seem relatively low with all four models, but does exhibit reasonable improvements over a constants-only model, which would result in an average probability of correctly predicting a rejection equal to the sample share of 0.198). Finally, each model can be compared with the proposed model using a likelihood ratio test. Each of the three likelihood ratio test statistics are much higher than the corresponding chi-squared value at any reasonable level of significance, indicating a superior fit for the proposed model. These results collectively demonstrate the superiority of the proposed model and suggest that individuals do indeed consider the evolution of their future state as they make decisions in the context of delivery acceptance and that there is unobserved individual heterogeneity that impacts how individuals are making these decisions.

Finally, we will note that the statistics presented in Table 3 suggest that, though the proposed dynamic discrete choice approach outperforms all three other models, the second best approach would be the longitudinal binary probit model which considers unobserved individual heterogeneity but not the dynamic nature of the decision process. In fact, the significant improvement in performance of the model when incorporating this longitudinal component demonstrates the importance of considering individual heterogeneity in the choice process, implying that delivery companies should collect data on driver characteristics to appropriately dispatch orders and highlighting the importance of incorporating such behavioral heterogeneity into the DDCM approach. The relatively good performance of this model compared to the others suggests that the incorporation of the longitudinal structure of the model results in larger gains than the consideration of the dynamic nature of the decision alone. However, while the impact is smaller, the incorporation of the dynamic component yields meaningful improvements. All tests agree that the inclusion of the dynamic component is statistically significant, even accounting for the increased complexity of the model, validating the use of the DDCM model. Further, the ability to correctly predict rejection decisions is much improved when the dynamic nature of the decision process is included, suggesting that this component is critical from a prediction standpoint. The smaller performance gains from the dynamic component in terms of likelihood-based improvements suggest that future studies should consider further fine-tuning of the dynamic components of the model, including using continuous-time and finite-horizon frameworks if data can be obtained to support these approaches (as previously discussed in Section 3).

Table 3: Disaggregate Model Fit

| Metric                                                         | Longitudinal<br>Dynamic Binary<br>Probit Model | Dynamic<br>Binary<br>Probit Model | Longitudinal<br>Binary Probit<br>Model | Binary<br>Probit<br>Model |
|----------------------------------------------------------------|------------------------------------------------|-----------------------------------|----------------------------------------|---------------------------|
| Log-Likelihood at Convergence                                  | -166,501                                       | -187,500                          | -166,930                               | -187,567                  |
| Log-Likelihood at Constants                                    | -194,399                                       | -194,399                          | -194,399                               | -194,399                  |
| Number of Parameters                                           | 23                                             | 22                                | 22                                     | 21                        |
| Adjusted Likelihood Ratio Index                                | 0.143                                          | 0.035                             | 0.141                                  | 0.035                     |
| Bayesian Information Criterion                                 | 166,565                                        | 187,561                           | 166,992                                | 187,626                   |
| Average Probability of Correctly<br>Predicting Order Rejection | 0.313                                          | 0.272                             | 0.281                                  | 0.225                     |
| Likelihood Ratio Test                                          | —                                              | 41,998.2                          | 858.3                                  | 42,133.2                  |

#### 4.3. Average Treatment Effects

While the main estimation results presented in Section 4.1 provide important insights into the ways that individuals make decisions in the context of delivery order acceptance, it may be more helpful for delivery platforms to understand the magnitude of effects of each of the exogenous variables. Thus, platforms can determine the impacts of making policy changes to alter the set of orders being offered to improve order acceptance rates. For this, we use average treatment effects (ATEs), which represent the impact of a change in the state of an exogenous variable on the acceptance share of the main outcome. To compute the ATEs, we begin by setting the values of a single exogenous variable to a base state for every observation in the dataset (while maintaining the values of all other exogenous variables) and computing the predicted share of couriers accepting orders. Then, the same procedure is adopted for the treatment level of the same exogenous variable and the difference in the acceptance share is presented as the treatment effect. In the dynamic discrete choice framework, we assume that the future value is secondary (the transition probabilities within the MDP do not change) as the treatment is applied.

Table 4 shows the ATEs for time of day, trip length, request type, and number of undelivered orders. For example, the ATE of -0.0767 in the first row indicates that if 100 orders were offered to couriers during the morning peak, one may expect about 8 fewer orders to be accepted than if the same set of orders were offered during an off-peak period. The ATEs provide insights into the change in the acceptance rate if the company makes service changes. In particular, the negative ATEs during peak periods suggest that order acceptance rates drop significantly during these times. Therefore, delivery platforms should consider what interventions they may be able to make to shift the demand towards off-peak hours when acceptance rates are higher. Strategies such as implementing surge pricing, providing consumer benefits for off-peak orders, or advertising improved service quality during off-peak times may help to improve the overall order acceptance rate. Alternatively, offering higher compensation for orders during peak periods may attract more drivers during these times and reduce these demand effects. The substantial positive ATE for pre-orders indicates the companies can increase order acceptance substantially by promoting advance bookings rather than real-time requests. This may be especially important during peak periods, when additional driver flexibility could make less attractive orders more feasible. Further, expanding pre-order options could also help ensure coverage at specific times and in specific geographic locations that are less attractive to drivers overall. If platforms can bundle pre-orders to prevent deadheading, they may be able to generate package orders that offer long guaranteed active working segments while ensuring that drivers will return from more remote locations. Given our results, these types of bundles may be very attractive to drivers and result in much higher acceptance rates. More generally, given the results of the interactions between time of day and trip length presented in Section 4.1 and the small positive ATE for trip length, it seems likely that bundling orders to return drivers to their starting zone (or some nearby zone with a high order rate) would be attractive.

The ATEs for undelivered orders suggest that companies can optimize the workloads of drivers and maximize order acceptance by offering orders to drivers based on their existing workload. Acceptance rates

Table 4: Average Treatment Effects

| Variable                     | Base      | Treatment    | ATE     |
|------------------------------|-----------|--------------|---------|
| Time of Day                  | Off-peak  | Morning peak | -0.0767 |
|                              | Off-peak  | Evening peak | -0.0597 |
| Trip Length                  | Short     | Long         | 0.0068  |
| Request Type                 | Real-time | Pre-order    | 0.0597  |
| Number of Undelivered Orders | 0         | 1            | 0.0292  |
|                              | 1         | 2            | -0.0107 |
|                              | 2         | 3+           | -0.0407 |

are higher among drivers who already have an open order but lower among those with several open orders. This suggests that carefully designing matching procedures that offer new orders to couriers that already have open orders to nearby locations would be an effective strategy. These systems should be optimized to consider courier queue lengths and could be designed with adaptive pricing techniques to incentivize acceptance for drivers with no existing orders who are otherwise idle. This result also highlights the challenge for platform operators looking to maximize efficiency while keeping the set of on-hand orders for any individual courier low. Since there is significant overlap between some routes, simplifying the assignment process may help make couriers more comfortable with larger queues of open orders, helping to increase profits both for the drivers themselves and the platforms. Grouping orders with nearby senders or recipients (or even identical senders) and offering these orders to couriers as bundles may help to provide large attractive driving segments (as mentioned above) that maximize overlapping segments, while providing simplified options for drivers.

The ATEs for the recipient zone and courier zone at the time of the request are shown in Figure 6. The left plot shows the ATEs based on the recipient zone, with zone six as the base such that all the effects are positive. While we would expect that order acceptance rates would be higher in zones with higher rates of order requests, we find little correspondence between the ATEs based on recipient zone and the request rates by zone (shown in Figure 7). Instead, it seems that preferences for geographic zone of the recipient are based more on distance from the city center, as acceptance rates seem to be higher for trips to zones outside of the city center. This may reflect heuristics used by drivers to try to avoid downtown congestion at peak periods. The right plot shows the ATEs based on the zone of the courier at the time of the request, with zone two as the base. While there are no coefficients for the courier zone in the model (there are no corresponding coefficients in Table 2), ATEs can still be computed since courier location appears as part of the MDP in the transition probabilities. While the ATEs for the courier zone tend to be smaller than those of the recipient zone, the areas with the highest acceptance rates align with the areas with the lowest request rates (shown in Figure 7). Specifically, when couriers are in areas with low request rates, they are more likely to accept any order that is offered to them, while the acceptance rate goes down in areas with more requests being made. These location-based ATE effects suggest that geographic effects do play a significant role in influencing courier behaviors. In the same way, companies should be using these considerations to determine how best to use their assignment processes to move couriers through the system in intentional ways. The recipient zone ATEs reveal that there are significant preferences for travel to specific locations that are not solely influenced by achieving higher order rates. So, if companies can better determine why some areas are more attractive, and which drivers may be more influenced by these geographic effects, they may be better able to align their order matching systems to these preferences. Unfortunately, zonal characteristics are not included with the order data, so we cannot identify if there are common underlying characteristics making some zones more popular than others.

#### 4.4. Coefficient Sensitivity

While the ATE results provide important insights into how companies could use specific policies to achieve higher order acceptance rates by influencing the characteristics of the offered orders, they also have additional control over the decision process itself. For instance, companies can decide what information to show to

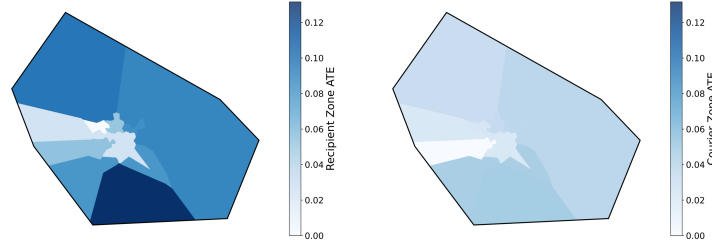


**Figure 6: Location-Based Average Treatment Effects (Left: ATEs for Recipient Zone, Right: ATEs for Courier Zone)**

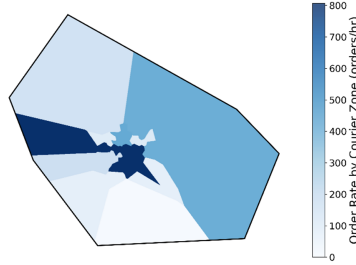


**Figure 7: Location-Based Average Treatment Effects (Rate of Delivery Order Requests by Courier Zone)**

drivers and how to provide it. Companies should consider what pieces of this information are beneficial (as allowing drivers choice over which orders to accept is important for their own satisfaction and will help maintain a strong driver base), what information may be unnecessarily leading to low acceptance rates, and what information could be highlighted to help improve order acceptance. At the same time, it is important to consider that there are many ethical issues related to information provision as the balance of power tends to favor the platform rather than the driver (Rosenblat and Stark, 2016).

To understand how changing the valuation for different factors in the decision process would impact acceptance rates, we plot the overall acceptance rate as a function of the estimated coefficients for a selected set of coefficients (varying parameters independently while holding all others constant), see Figure 8. These coefficient sensitivity effects are influenced by the starting values of the parameters, and the relative number of orders they effect. For instance, the top left plot shows that increasing the constant term will improve acceptance rates with diminishing returns while reducing the constant term will significantly worsen acceptance rates. This is an intuitive result for the constant, suggesting that overall strategies designed to incentivize acceptance of all orders will increase the acceptance rate overall. However, the effects of the other coefficients can be used to develop some insights into how different decisions by the company regarding specific types of orders may impact request acceptance rates.

The top right plot shows the discount factor, suggesting that if drivers were more myopic (had a lower discount factor and considered future value less) they would reject fewer orders. This suggests that many orders that provide sufficient benefits in the current period to be attractive may be rejected because they put drivers in a worse position for future requests. Therefore, delivery companies should be actively considering these impacts as well. As mentioned above, if companies can bundle orders to guarantee that drivers will return from outlying areas without deadheading, then forward-looking couriers will be more willing to accept these orders. Similarly, knowing that experienced drivers will try to maximize their total utility by considering these downstream effects, they can provide benefits to offset less advantageous orders or highlight specific benefits that may be of interest. For instance, highlighting that an order ends in an area with a high request density or that there is a high chance of receiving a second order at the same pick-up location would

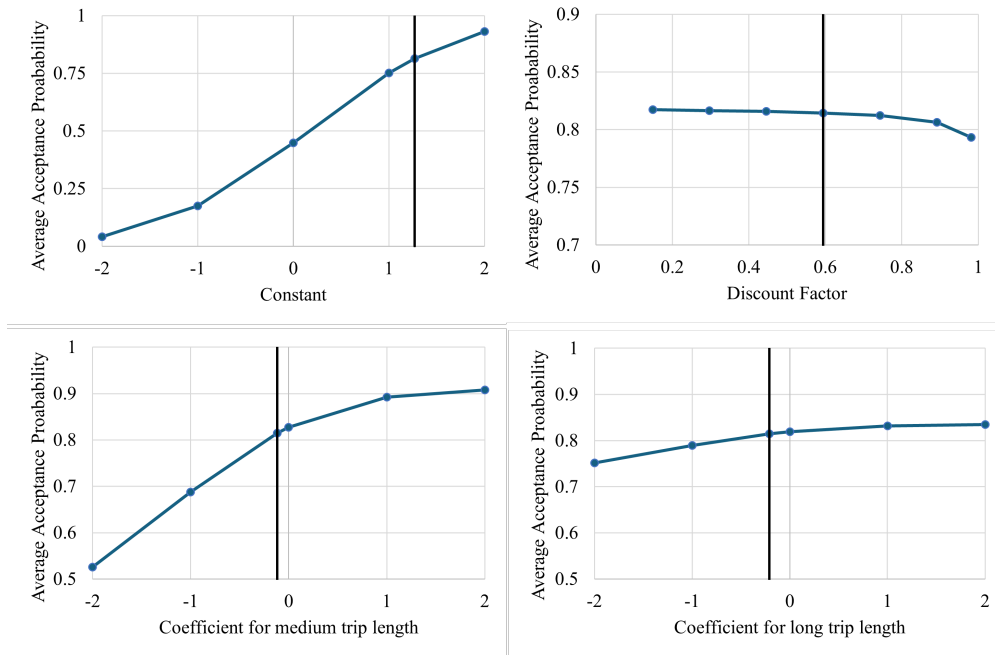


**Figure 8: Coefficient Sensitivity**

give drivers more confidence in the benefits of accepting some orders, essentially reducing the short-term negative effects of a single order.

While many of the current benefits that are offered to help boost acceptance rates are targeted at improving acceptance overall, the two plots at the bottom of the figure demonstrate that targeting features of specific trips may be particularly effective. For instance, many companies have implemented programs that provide special privileges to drivers who have high overall order acceptance rates (essentially targeting an improvement of the constant parameter by increasing the utility of accepting any order). However, some strategies targeting specific order types that have low acceptance rates may have a big impact. For instance, providing additional incentives to drivers who accept medium length orders is likely to have a larger impact than providing the same incentives for longer orders (see the greater slope of the sensitivity plot for medium trip length compared with long trip length). These results could be included in existing scoring systems that give drivers ratings based on their acceptance rate and prioritize order offers to high scoring drivers. Giving more credit to drivers for medium- and long-distance trips during off-peak hours, for example, would help give drivers an incentive to accept more of these less desirable orders. Finally, the sensitivity results present a different option for the platform to increase the driver acceptance rate besides changing information provision or offering additional incentives. The model results presented here represent the specific behavior of Meituan’s current driver fleet. However, the sensitivity results suggest the potential increases in efficiency the company could get if it actively recruited drivers with slightly different characteristics. For instance, specifically recruiting drivers interested in working in specific underserved zones or who are more willing to accept longer trips may improve the overall acceptance rate for these orders. At the same time, considering how these preferences differ among drivers would be valuable for matching algorithms. Although the lack of individual characteristics in the current dataset prohibits modeling these behavioral differences among drivers, this analysis motivates the need for the platform to collect this information to potentially recruit a more effective fleet of drivers who might be more willing to accept rides that the current fleet finds undesirable.

## 5. Conclusion

This study provides a comprehensive examination of delivery driver order acceptance behaviors through a dynamic discrete choice modeling approach. By utilizing a Markov decision process framework, we have illuminated the complex, forward-looking decision-making process that drivers employ when accepting or rejecting order requests. This approach can reveal actionable insights to platforms, helping them to improve efficiency, increase profits, maintain strong customer and business loyalty, and remain competitive with competing platforms. As these platforms compete for both customers and orders, driver satisfaction and timeliness of order arrival are key to successful operations.

Our model demonstrates that delivery drivers make nuanced decisions based on a sophisticated evaluation of immediate and future opportunities. The substantial variation in acceptance rates across different times, trip lengths, and geographic zones underscores the need for more adaptive and flexible delivery platform strategies. Based on our results, we suggest that platforms can increase efficiency by:

- Promoting pre-book orders which provide increased flexibility and planning, and drivers consistently prefer. Platforms can highlight pre-book options more prominently in the app interface and provide estimated delivery time guarantees to increase customer confidence in these options to help promote the use of these options.
- Developing more targeted incentives for trips that would otherwise be less desirable (such as those during peak hours, with longer distances, or ending in low-demand zones) by offering bonuses or priority matching for drivers that accept these orders. For instance, existing demand-based pricing strategies could be supplemented to provide extra compensation for accepting long-distance orders during off-peak hours or for serving low-demand zones and highlighting these incentives as part of the order offer to drivers.
- Actively managing driver workloads by assigning new requests to couriers with 1–2 active orders, maximizing acceptance rates and minimizing idle time. Matching should also consider route compatibility for offers made to active couriers, as well as prioritizing offering orders to active couriers that may be less desirable as the first order in a sequence.
- Bundling orders strategically to reduce deadheading and improve route efficiency, especially for longer trips or trips to peripheral zones. Bundling can also help create guaranteed return paths, increasing driver willingness to accept outbound orders.

Beyond these recommendations, the significant unobserved individual heterogeneity revealed in our model also suggests that platforms should track detailed driver characteristic data. This could enable more personalized matching and incentive strategies, allowing platforms to tailor order assignments and incentives based on driver characteristics, individual geographic preferences, and time-of-day effects. Using this detailed individual-level data, platforms could improve on these existing overall suggestions, potentially improving overall system performance and driver satisfaction.

Future research should extend this work by incorporating more detailed driver characteristics, as well as zonal characteristics to better describe which drivers are accepting which offers. In addition, more advanced decisions processes could be developed using finite-horizon and continuous-time MDPs which could better track and update the order queue and account for order overlap as new orders are assigned and orders are completed (particularly if additional data is available). Concurrently, as this type of analysis would increase the size of the state space, alternative solution strategies to quickly solve both the MDP and maximum likelihood should be explored. Specifically, recent advances in machine learning and artificial intelligence could be leveraged to efficiently solve the MDP and enable more rapid convergence of the maximum likelihood procedure. The use of such procedures to improve estimation efficiency would also allow for the incorporation of additional taste heterogeneity, including heterogeneity in the discount factor, which may capture meaningful differences in how drivers trade off immediate versus future earnings. Finally, while we focus here on the binary accept/reject decision, the probit-based flexible error specification used here allows for the specification of multinomial choices with correlated choice outcomes, allowing future studies employing

these methods to examine more complex driver decisions such as routing and relocation behaviors alongside order acceptance.

## Acknowledgments

This research was supported by data provided by Meituan for the First INFORMS TSL Data-Driven Research Challenge (2024-2025 TSL-Meituan).

## References

- D. Akerberg, X. Chen, and J. Hahn. A Practical Asymptotic Variance Estimator for Two-Step Semiparametric Estimators. *The Review of Economics and Statistics*, 94(2):481–498, May 2012. ISSN 0034-6535. doi:10.1162/REST\_a.00251. URL [https://doi.org/10.1162/REST\\_a\\_00251](https://doi.org/10.1162/REST_a_00251).
- V. Aguirregabiria and P. Mira. Sequential Estimation of Dynamic Discrete Games. *Econometrica*, 75(1):1–53, 2007. ISSN 1468-0262. doi:10.1111/j.1468-0262.2007.00731.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0262.2007.00731.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2007.00731.x>.
- V. Aguirregabiria and P. Mira. Dynamic Discrete Choice Structural Models: A Survey. *Journal of Econometrics*, 156(1):38–67, May 2010. ISSN 0304-4076. doi:10.1016/j.jeconom.2009.09.007. URL <https://www.sciencedirect.com/science/article/pii/S0304407609001985>.
- G. Allon, M. C. Cohen, and W. P. Sinchaisri. The Impact of Behavioral and Economic Drivers on Gig Economy Workers. *Manufacturing & Service Operations Management*, 25(4):1376–1393, July 2023. ISSN 1523-4614. doi:10.1287/msom.2023.1191. URL <https://pubsonline.informs.org/doi/abs/10.1287/msom.2023.1191>.
- Y. An, Y. Hu, and R. Xiao. Dynamic Decisions Under Subjective Expectations: A Structural Analysis. *Journal of Econometrics*, 222(1, Part C):645–675, May 2021. ISSN 0304-4076. doi:10.1016/j.jeconom.2020.04.046. URL <https://www.sciencedirect.com/science/article/pii/S0304407620302414>.
- P. Arcidiacono and R. A. Miller. Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. *Econometrica*, 79(6):1823–1867, 2011. ISSN 1468-0262. doi:10.3982/ECTA7743. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7743>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7743>.
- P. Arcidiacono, P. Bayer, J. R. Blevins, and P. B. Ellickson. Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition. *The Review of Economic Studies*, 83(3):889–931, July 2016. ISSN 0034-6527. doi:10.1093/restud/rdw012. URL <https://doi.org/10.1093/restud/rdw012>.
- P. Ashkrof, G. H. d. A. Correia, O. Cats, and B. van Arem. Understanding Ride-Sourcing Drivers’ Behaviour and Preferences: Insights from Focus Groups Analysis. *Research in Transportation Business & Management*, 37, Dec. 2020. ISSN 2210-5395. doi:10.1016/j.rtbm.2020.100516. URL <https://www.sciencedirect.com/science/article/pii/S221053952030078X>.
- P. Ashkrof, G. Homem de Almeida Correia, O. Cats, and B. van Arem. Ride Acceptance Behaviour of Ride-Sourcing Drivers. *Transportation Research Part C: Emerging Technologies*, 142:103783, Sept. 2022. ISSN 0968-090X. doi:10.1016/j.trc.2022.103783. URL <https://www.sciencedirect.com/science/article/pii/S0968090X22002121>.

- J. Bai, K. C. So, C. S. Tang, X. M. Chen, and H. Wang. Coordinating Supply and Demand on an On-Demand Service Platform with Impatient Customers. *Manufacturing & Service Operations Management*, 21(3):556–570, July 2019. ISSN 1523-4614. doi:10.1287/msom.2018.0707. URL <https://pubsonline.informs.org/doi/10.1287/msom.2018.0707>.
- M. W. Barbosa, J. V. de Oliveira Miranda, and L. K. de Oliveira. Identification of Antecedents of Risky Driving Behavior of Food Delivery Riders: An Analysis During the COVID -19 Pandemic in Brazil. *Case Studies on Transport Policy*, 20:101403, June 2025. ISSN 2213-624X. doi:10.1016/j.cstp.2025.101403. URL <https://www.sciencedirect.com/science/article/pii/S2213624X25000409>.
- R. Bellman. *Dynamic Programming*. Princeton University Press, New Jersey, 1957.
- S. Berry, J. Levinsohn, and A. Pakes. Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841, July 1995. ISSN 00129682. doi:10.2307/2171802. URL <https://www.jstor.org/stable/2171802?origin=crossref>.
- C. R. Bhat. A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity. *Transportation Research Part B: Methodological*, 30(3):189–207, June 1996. ISSN 0191-2615. doi:10.1016/0191-2615(95)00029-1. URL <https://www.sciencedirect.com/science/article/pii/0191261595000291>.
- C. R. Bhat. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B: Methodological*, 35(7):677–693, Aug. 2001. ISSN 0191-2615. doi:10.1016/S0191-2615(00)00014-X. URL <https://www.sciencedirect.com/science/article/pii/S019126150000014X>.
- C. R. Bhat. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. *Transportation Research Part B: Methodological*, 37(9):837–855, Nov. 2003. ISSN 0191-2615. doi:10.1016/S0191-2615(02)00090-5. URL <https://www.sciencedirect.com/science/article/pii/S0191261502000905>.
- C. R. Bhat. New Matrix-Based Methods for the Analytic Evaluation of the Multivariate Cumulative Normal Distribution Function. *Transportation Research Part B: Methodological*, 109:238–256, Mar. 2018. ISSN 0191-2615. doi:10.1016/j.trb.2018.01.011. URL <https://www.sciencedirect.com/science/article/pii/S0191261517306847>.
- J. Bunting. Continuous Permanent Unobserved Heterogeneity in Dynamic Discrete Choice Models, Feb. 2024. URL <http://arxiv.org/abs/2202.03960>.
- G. P. Cachon, K. M. Daniels, and R. Lobel. The Role of Surge Pricing on a Service Platform with Self-Scheduling Capacity. *Manufacturing & Service Operations Management*, 19(3):368–384, July 2017. ISSN 1523-4614. doi:10.1287/msom.2017.0618. URL <https://pubsonline.informs.org/doi/10.1287/msom.2017.0618>.
- H. A. Caparas. Investigating the Performance of Food Delivery Riders Through Work System and Stress-Coping Approach. *Ergonomics*, 0(0):1–14, Feb. 2026. ISSN 0014-0139. doi:10.1080/00140139.2026.2621897. URL <https://doi.org/10.1080/00140139.2026.2621897>. eprint: <https://doi.org/10.1080/00140139.2026.2621897>.
- A. H. Chen, J. Z. Lee, and Y.-L. Ho. Influential Factors for Online Food Delivery Platform Drivers’ Order Acceptance. *Information Discovery and Delivery*, 54(1):113–128, Jan. 2025. ISSN 2398-6247. doi:10.1108/IDD-03-2024-0038. URL <https://www.sciencedirect.com/org/science/article/abs/pii/S2398624725000056>.

- L. Chen, A. H. Valadkhani, and M. Ramezani. Decentralised Cooperative Cruising of Autonomous Ride-Sourcing Fleets. *Transportation Research Part C: Emerging Technologies*, 131:103336, Oct. 2021. ISSN 0968-090X. doi:10.1016/j.trc.2021.103336. URL <https://www.sciencedirect.com/science/article/pii/S0968090X21003399>.
- R. Chen and M. W. Levin. Dynamic User Equilibrium of Mobility-on-Demand System with Linear Programming Rebalancing Strategy. *Transportation Research Record*, 2673(1):447–459, Jan. 2019. ISSN 0361-1981. doi:10.1177/0361198118821629. URL <https://doi.org/10.1177/0361198118821629>.
- X. Chen, M. W. Ulmer, and B. W. Thomas. Deep Q-Learning for Same-Day Delivery with Vehicles and Drones. *European Journal of Operational Research*, 298(3):939–952, May 2022. ISSN 0377-2217. doi:10.1016/j.ejor.2021.06.021. URL <https://www.sciencedirect.com/science/article/pii/S0377221721005361>.
- Y. Chen, Y. Qian, Y. Yao, Z. Wu, R. Li, Y. Zhou, H. Hu, and Y. Xu. Can Sophisticated Dispatching Strategy Acquired by Reinforcement Learning? - A Case Study in Dynamic Courier Dispatching System, Mar. 2019. URL <http://arxiv.org/abs/1903.02716>.
- C. Chou and T. Derdenger. CCP Estimation of Dynamic Discrete Choice Demand Models with Segment Level Data and Continuous Unobserved Heterogeneity: Rethinking EV Subsidies vs. Infrastructure. *Marketing Science (INFORMS)*, 44(5):1163–1187, Sept. 2025. ISSN 0732-2399. doi:10.1287/mksc.2024.0860. URL <https://research.ebsco.com/linkprocessor/plink?id=44c193ec-539b-3356-a064-909db1833ede>.
- C. Cirillo and R. Xu. Dynamic Discrete Choice Models for Transportation. *Transport Reviews*, 31(4):473–494, July 2011. ISSN 0144-1647. doi:10.1080/01441647.2010.533393. URL <https://doi.org/10.1080/01441647.2010.533393>.
- C. Cirillo, R. Xu, and F. Bastin. A Dynamic Formulation for Car Ownership Modeling. *Transportation Science*, 50(1):322–335, Feb. 2016. ISSN 0041-1655. doi:10.1287/trsc.2015.0597. URL <https://pubsonline.informs.org/doi/10.1287/trsc.2015.0597>.
- C. E. Clark. The Greatest of a Finite Set of Random Variables. *Operations Research*, Apr. 1961. doi:10.1287/opre.9.2.145. URL <https://pubsonline.informs.org/doi/abs/10.1287/opre.9.2.145>.
- E. Fatemi Pour, S. A. Madnanizdeh, and H. Joshaghani. Accept or Reject a Ride? This Is the Problem. *Journal of Economic Studies*, 50(7):1346–1374, Jan. 2023. ISSN 0144-3585. doi:10.1108/JES-12-2021-0617. URL <https://doi.org/10.1108/JES-12-2021-0617>.
- X. Feng and M. Wang. Strategic Driver’s Acceptance-or-Rejection Behavior and Cognitive Hierarchy in on-Demand Platforms. *Transportation Research Part E: Logistics and Transportation Review*, 176, Aug. 2023. ISSN 1366-5545. doi:10.1016/j.tre.2023.103175. URL <https://www.sciencedirect.com/science/article/pii/S1366554523001631>.
- A. Fielbaum and A. Tirachini. The Sharing Economy and the Job Market: The Case of Ride-Hailing Drivers in Chile. *Transportation*, 48(5):2235–2261, Oct. 2021. ISSN 1572-9435. doi:10.1007/s11116-020-10127-7. URL <https://doi.org/10.1007/s11116-020-10127-7>.
- M. Fosgerau, E. Frejinger, and A. Karlstrom. A Link Based Network Route Choice Model with Unrestricted Choice Set. *Transportation Research Part B: Methodological*, 56:70–80, Oct. 2013. ISSN 0191-2615. doi:10.1016/j.trb.2013.07.012. URL <https://www.sciencedirect.com/science/article/pii/S0191261513001276>.
- J. Gao, X. Li, C. Wang, and X. Huang. A Pricing Mechanism for Ride-Hailing Systems in the Presence of Driver Acceptance Uncertainty. *IEEE Access*, 10:83017–83028, 2022. ISSN 2169-3536. doi:10.1109/ACCESS.2022.3196684. URL <https://ieeexplore.ieee.org/document/9851427>.

- Y. Gao, D. Jiang, and Y. Xu. Optimize Taxi Driving Strategies Based on Reinforcement Learning. *International Journal of Geographical Information Science*, 32(8):1677–1696, Aug. 2018. ISSN 1365-8816. doi:10.1080/13658816.2018.1458984. URL <https://www.tandfonline.com/doi/full/10.1080/13658816.2018.1458984>.
- M. Gebresselassie. Labor Issues from the Perspective of Drivers on the Uber and Lyft Apps and the Impact on Riders Who Use Wheelchairs. *Travel Behaviour and Society*, 38:100891, Jan. 2025. ISSN 2214-367X. doi:10.1016/j.tbs.2024.100891. URL <https://www.sciencedirect.com/science/article/pii/S2214367X24001546>.
- G. Gowrisankaran and M. Rysman. Dynamics of Consumer Demand for New Durable Goods. *Journal of Political Economy*, 120(6):1173–1219, Dec. 2012. ISSN 0022-3808. doi:10.1086/669540. URL <https://www.journals.uchicago.edu/doi/full/10.1086/669540>.
- H. Guda and U. Subramanian. Your Uber Is Arriving: Managing On-Demand Workers Through Surge Pricing, Forecast Communication, and Worker Incentives. *Management Science*, 65(5):1995–2014, May 2019. ISSN 0025-1909. doi:10.1287/mnsc.2018.3050. URL <https://pubsonline.informs.org/doi/10.1287/mnsc.2018.3050>.
- Y. He, C. Sun, and F. Chang. The Road Safety and Risky Behavior Analysis of Delivery Vehicle Drivers in China. *Accident Analysis & Prevention*, 184:107013, May 2023. ISSN 0001-4575. doi:10.1016/j.aap.2023.107013. URL <https://www.sciencedirect.com/science/article/pii/S000145752300060X>.
- J. Heckman and B. Singer. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, 52(2):271–320, 1984. ISSN 0012-9682. doi:10.2307/1911491. URL <https://www.jstor.org/stable/1911491>.
- V. J. Hotz and R. A. Miller. Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies*, 60(3):497–529, July 1993. ISSN 0034-6527. doi:10.2307/2298122. URL <https://doi.org/10.2307/2298122>.
- S. Hou, C. Wang, and J. Gao. Reinforced Stable Matching for Crowd-Sourced Delivery Systems Under Stochastic Driver Acceptance Behavior. *Transportation Research Part C: Emerging Technologies*, 170:104916, Jan. 2025. ISSN 0968-090X. doi:10.1016/j.trc.2024.104916. URL <https://www.sciencedirect.com/science/article/pii/S0968090X24004376>.
- Y. Hu and M. Shum. Nonparametric Identification of Dynamic Models with Unobserved State Variables. *Journal of Econometrics*, 171(1):32–44, Nov. 2012. ISSN 0304-4076. doi:10.1016/j.jeconom.2012.05.023. URL <https://www.sciencedirect.com/science/article/pii/S0304407612001479>.
- Y. Idug, S. Niranjan, I. Manuj, D. Gligor, and J. Ogden. Do Ride-Hailing Drivers’ Psychological Behaviors Influence Operational Performance? *International Journal of Operations & Production Management*, 43(12):2055–2079, Jan. 2023. ISSN 0144-3577. doi:10.1108/IJOPM-06-2022-0362. URL <https://doi.org/10.1108/IJOPM-06-2022-0362>.
- Y. Jiao, X. Tang, Z. T. Qin, S. Li, F. Zhang, H. Zhu, and J. Ye. Real-World Ride-Hailing Vehicle Repositioning Using Deep Reinforcement Learning. *Transportation Research Part C: Emerging Technologies*, 130:103289, Sept. 2021. ISSN 0968-090X. doi:10.1016/j.trc.2021.103289. URL <https://www.sciencedirect.com/science/article/pii/S0968090X21003004>.
- I. Kadolkar, S. Kepes, and M. Subramony. Algorithmic Management in the Gig Economy: A Systematic Review and Research Integration. *Journal of Organizational Behavior*, 46(7):1057–1080, Sept. 2025. ISSN 0894-3796. doi:10.1002/job.2831. URL <https://onlinelibrary.wiley.com/doi/full/10.1002/job.2831>.

- H. Kasahara and K. Shimotsu. Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices. *Econometrica*, 77(1):135–175, 2009. ISSN 1468-0262. doi:10.3982/ECTA6763. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA6763>.
- J. Ke, C. Wang, X. Li, Q. Tian, and H.-J. Huang. Equilibrium Analysis for on-Demand Food Delivery Markets. *Transportation Research Part E: Logistics and Transportation Review*, 184:103467, Apr. 2024. ISSN 1366-5545. doi:10.1016/j.tre.2024.103467. URL <https://www.sciencedirect.com/science/article/pii/S1366554524000577>.
- N. D. Kullman, M. Cousineau, J. C. Goodson, and J. E. Mendoza. Dynamic Ride-Hailing with Electric Vehicles. *Transportation Science*, 56(3):775–794, May 2022. ISSN 0041-1655. doi:10.1287/trsc.2021.1042. URL <https://pubsonline.informs.org/doi/abs/10.1287/trsc.2021.1042>.
- X. Lai, Z. Yang, J. Xie, and Y. Liu. Reinforcement Learning in Transportation Research: Frontiers and Future Directions. *Multimodal Transportation*, 3(4):100164, Dec. 2024. ISSN 2772-5863. doi:10.1016/j.multra.2024.100164. URL <https://www.sciencedirect.com/science/article/pii/S2772586324000455>.
- B. Larsen, F. Oswald, G. Reich, and D. Wunderli. A Test of the Extreme Value Type I Assumption in the Bus Engine Replacement Model, Dec. 2011. URL <https://papers.ssrn.com/abstract=1989027>.
- D. A. Lazar, E. Biyik, D. Sadigh, and R. Pedarsani. Learning How to Dynamically Route Autonomous Vehicles on Shared Roads. *Transportation Research Part C: Emerging Technologies*, 130:103258, Sept. 2021. ISSN 0968-090X. doi:10.1016/j.trc.2021.103258. URL <https://www.sciencedirect.com/science/article/pii/S0968090X21002709>.
- J. Leong, N. Nassir, S. S. Mohri, and M. Sarvi. A Dynamic Discrete Choice Modelling Approach for Forward-Looking Travel Mode Choices. *Transportation Research Part A: Policy and Practice*, 190, Dec. 2024. ISSN 0965-8564. doi:10.1016/j.tra.2024.104272. URL <https://www.sciencedirect.com/science/article/pii/S0965856424003203>.
- M. E. Lewis and A. Paul. Uniform Turnpike Theorems for Finite Markov Decision Processes. *Mathematics of Operations Research*, 44(4):1145–1160, Nov. 2019. ISSN 0364-765X. doi:10.1287/moor.2017.0912. URL <https://pubsonline.informs.org/doi/abs/10.1287/moor.2017.0912>.
- K. Lin, R. Zhao, Z. Xu, and J. Zhou. Efficient Large-Scale Fleet Management via Multi-Agent Deep Reinforcement Learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD ’18, pages 1774–1783, New York, NY, USA, July 2018. Association for Computing Machinery. ISBN 978-1-4503-5552-0. doi:10.1145/3219819.3219993. URL <https://doi.org/10.1145/3219819.3219993>.
- Y. Liu and C. Cirillo. A Generalized Dynamic Discrete Choice Model for Green Vehicle Adoption. *Transportation Research Part A: Policy and Practice*, 114:288–302, Aug. 2018. ISSN 0965-8564. doi:10.1016/j.tra.2018.01.034. URL <https://www.sciencedirect.com/science/article/pii/S0965856417308947>.
- Y. Liu, F. Wu, C. Lyu, S. Li, J. Ye, and X. Qu. Deep Dispatching: A Deep Reinforcement Learning Approach for Vehicle Dispatching on Online Ride-Hailing Platform. *Transportation Research Part E: Logistics and Transportation Review*, 161:102694, May 2022. ISSN 1366-5545. doi:10.1016/j.tre.2022.102694. URL <https://www.sciencedirect.com/science/article/pii/S1366554522000862>.
- W. Mao, L. Ming, Y. Rong, C. S. Tang, and H. Zheng. Faster Deliveries and Smarter Order Assignments for an On-Demand Meal Delivery Platform, Oct. 2019. URL <https://papers.ssrn.com/abstract=3469015>.

- Meituan. Meituan INFORMS TSL Research Challenge, 2024. URL <https://github.com/meituan/Meituan-INFORMS-TSL-Research-Challenge/tree/main>.
- J. Muldoon and P. Raekstad. Algorithmic domination in the gig economy. *European Journal of Political Theory*, 22(4):587–607, Oct. 2023. ISSN 1474-8851. doi:10.1177/14748851221082078. URL <https://doi.org/10.1177/14748851221082078>.
- M. Nazari, A. Oroojlooy, L. V. Snyder, and M. Takáč. Reinforcement Learning for Solving the Vehicle Routing Problem, May 2018. URL <http://arxiv.org/abs/1802.04240>.
- Y. Oyama and E. Hato. A Discounted Recursive Logit Model for Dynamic Gridlock Network Analysis. *Transportation Research Part C: Emerging Technologies*, 85:509–527, Dec. 2017. ISSN 0968-090X. doi:10.1016/j.trc.2017.10.001. URL <https://www.sciencedirect.com/science/article/pii/S0968090X17302747>.
- P. Papantoniou, D. Pavlou, E. Michelaraki, and V. Amprasi. Investigating the Safe Driving Behaviour of Professional Drivers in Last-Mile Delivery Services in Greece. *Transportation Research Part F: Traffic Psychology and Behaviour*, 114:246–269, Oct. 2025. ISSN 1369-8478. doi:10.1016/j.trf.2025.06.002. URL <https://www.sciencedirect.com/science/article/pii/S1369847825002086>.
- C. Popan. Embodied Precariat and Digital Control in the “Gig Economy”: The Mobile Labor of Food Delivery Workers. *Journal of Urban Technology*, 31(1):109–128, Jan. 2024. ISSN 1063-0732. doi:10.1080/10630732.2021.2001714. URL <https://doi.org/10.1080/10630732.2021.2001714>. eprint: <https://doi.org/10.1080/10630732.2021.2001714>.
- W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, Inc., 2011.
- Q. Qian, J. He, and J. Shi. Analysis of Factors Influencing Aberrant Riding Behavior of Food Delivery Riders: A Perspective on Safety Attitude and Risk Perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, 100:273–288, Jan. 2024. ISSN 1369-8478. doi:10.1016/j.trf.2023.12.007. URL <https://www.sciencedirect.com/science/article/pii/S1369847823002723>.
- J. Robbennolt and M. W. Levin. Maximum Throughput Dispatch for Shared Autonomous Vehicles Including Vehicle Rebalancing. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–15, 2023. ISSN 1558-0016. doi:10.1109/TITS.2023.3268158.
- J. Robbennolt, J. Mohammadi, and S. D. Boyles. Shared Autonomous Electric Vehicle Dispatch with Mobility and Distribution System Resilience Benefits. In *2024 IEEE Texas Power and Energy Conference (TPEC)*, pages 1–6, Feb. 2024.
- A. Rosenblat and L. Stark. Algorithmic Labor and Information Asymmetries: A Case Study of Uber’s Drivers. *International Journal of Communication*, 10(0):27, July 2016. ISSN 1932-8036. URL <https://ijoc.org/index.php/ijoc/article/view/4892>.
- J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987. ISSN 0012-9682. doi:10.2307/1911259. URL <https://www.jstor.org/stable/1911259>.
- F. Safira. Preliminary Study on the Dynamics of Food Delivery Driver. *Journal of Emerging Business Management and Entrepreneurship Studies*, 5(1):66–81, May 2025. ISSN 2797-3115. doi:10.34149/jebmes.v5i1.186. URL <https://jebmes.ppmschool.ac.id/index.php/jebmes/article/view/186>.

- J.-D. Saphores, L. Xu, and B. Park. COVID-19 and Food Shopping: Results from California and Comparisons with China and South Korea. In A. Loukaitou-Sideris, A. M. Bayen, G. Circella, and R. Jayakrishnan, editors, *Pandemic in the Metropolis: Transportation Impacts and Recovery*, pages 95–111. Springer International Publishing, Cham, 2023. ISBN 978-3-031-00148-2. URL [https://doi.org/10.1007/978-3-031-00148-2\\_7](https://doi.org/10.1007/978-3-031-00148-2_7).
- M. W. Savelsbergh and M. W. Ulmer. Challenges and Opportunities in Crowdsourced Delivery Planning and Operations. *4OR - A Quarterly Journal of Operations Research*, 20(1):1–21, Mar. 2022. ISSN 1614-2411. doi:10.1007/s10288-021-00500-2. URL <https://doi.org/10.1007/s10288-021-00500-2>.
- P. Schiraldi. Automobile Replacement: A Dynamic Structural Approach. *The RAND Journal of Economics*, 42(2):266–291, 2011. ISSN 1756-2171. doi:10.1111/j.1756-2171.2011.00133.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1756-2171.2011.00133.x>.
- Z. Shou, X. Di, J. Ye, H. Zhu, H. Zhang, and R. Hampshire. Optimal Passenger-Seeking Policies on E-Hailing Platforms Using Markov Decision Process and Imitation Learning. *Transportation Research Part C: Emerging Technologies*, 111:91–113, Feb. 2020. ISSN 0968-090X. doi:10.1016/j.trc.2019.12.005. URL <https://www.sciencedirect.com/science/article/pii/S0968090X18316577>.
- A. Singh, A. O. Alabbasi, and V. Aggarwal. A Reinforcement Learning Based Algorithm for Multi-hop Ride-sharing: Model-free Approach. 2019. URL <https://www.semanticscholar.org/paper/A-Reinforcement-Learning-Based-Algorithm-for-Singh-Alabbasi/6ad04d31eba23cc0724424a22bf30c3374ae0fd2>.
- Y. Song, D. Li, D. Liu, Q. Cao, J. Chen, G. Ren, and X. Tang. Modeling Activity-Travel Behavior Under a Dynamic Discrete Choice Framework with Unobserved Heterogeneity. *Transportation Research Part E: Logistics and Transportation Review*, 167, Nov. 2022. ISSN 1366-5545. doi:10.1016/j.trc.2022.102914. URL <https://www.sciencedirect.com/science/article/pii/S1366554522002915>.
- G. Sun, Y. Kim, Y. R. Tan, and G. G. Parker. Dinner at Your Doorstep: Service Innovation via the Gig Economy on Food Delivery Platforms. *Information Systems Research*, 35(3):1216–1234, Sept. 2024. ISSN 1047-7047. doi:10.1287/isre.2022.0119. URL <https://pubsonline.informs.org/doi/full/10.1287/isre.2022.0119>.
- Y. Sun, Z. Ding, Z. Hu, and W.-J. Lee. Risk-Aware Operation Modeling for Ride-Hailing Fleet in Order Grabbing Mode: A Distributional Reinforcement Learning Approach. *IEEE Transactions on Smart Grid*, 14(5):3913–3926, Sept. 2023. ISSN 1949-3061. doi:10.1109/TSG.2023.3237574. URL <https://ieeexplore.ieee.org/abstract/document/10018486>.
- W. Tan, Y. Sun, Z. Ding, and W.-J. Lee. Fleet Management and Charging Scheduling for Shared Mobility-on-Demand System: A Systematic Review. *IEEE Open Access Journal of Power and Energy*, 9:425–436, 2022. ISSN 2687-7910. doi:10.1109/OAJPE.2022.3215865. URL <https://ieeexplore.ieee.org/abstract/document/9924240>.
- Y. Tu, M. M. Oshanreh, M. Khaloei, N. A. Khan, and D. MacKenzie. Effect of Trip Attributes on Ride-hailing Driver Trip Request Acceptance. *International Journal of Sustainable Transportation*, 0(0):1–12, Nov. 2025. ISSN 1556-8318. doi:10.1080/15568318.2025.2580366. URL <https://doi.org/10.1080/15568318.2025.2580366>. eprint: <https://doi.org/10.1080/15568318.2025.2580366>.
- B. Turan, R. Pedarsani, and M. Alizadeh. Dynamic Pricing and Fleet Management for Electric Autonomous Mobility on Demand Systems. *Transportation Research Part C: Emerging Technologies*, 121:102829, Dec. 2020. ISSN 0968-090X. doi:10.1016/j.trc.2020.102829. URL <https://www.sciencedirect.com/science/article/pii/S0968090X20307336>.

- J. Urata, Z. Xu, J. Ke, Y. Yin, G. Wu, H. Yang, and J. Ye. Learning Ride-Sourcing Drivers' Customer-Searching Behavior: A Dynamic Discrete Choice Approach. *Transportation Research Part C: Emerging Technologies*, 130, Sept. 2021. ISSN 0968-090X. doi:10.1016/j.trc.2021.103293. URL <https://www.sciencedirect.com/science/article/pii/S0968090X21003028>.
- L. Wang, S. Webster, and E. Rabinovich. Structural Estimation of Attrition in a Last-Mile Delivery Platform: The Role of Driver Heterogeneity, Compensation, and Experience. *Manufacturing & Service Operations Management (M&SOM) (INFORMS)*, 27(2):516–534, Mar. 2025. ISSN 1523-4614. doi:10.1287/msom.2021.0367. URL <https://research.ebsco.com/linkprocessor/plink?id=8c7388cf-94bf-3864-8d37-2b303c830c4b>.
- K. Xu, L. Sun, J. Liu, and H. Wang. An Empirical Investigation of Taxi Driver Response Behavior to Ride-Hailing Requests: A Spatio-Temporal Perspective. *PLOS ONE*, 13(6):e0198605, June 2018. ISSN 1932-6203. doi:10.1371/journal.pone.0198605. URL <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0198605>.
- B. Yildiz and M. Savelsbergh. Provably High-Quality Solutions for the Meal Delivery Routing Problem. *Transportation Science (INFORMS)*, 53(5):1372–1388, Sept. 2019. ISSN 0041-1655. doi:10.1287/trsc.2018.0887. URL <https://research.ebsco.com/linkprocessor/plink?id=43146709-5799-3d19-9e79-2cf118baa94e>.
- X. Yu, S. Gao, X. Hu, and H. Park. A Markov Decision Process Approach to Vacant Taxi Routing with E-Hailing. *Transportation Research Part B: Methodological*, 121:114–134, Mar. 2019. ISSN 0191-2615. doi:10.1016/j.trb.2018.12.013. URL <https://www.sciencedirect.com/science/article/pii/S0191261518303837>.
- L. Zha, Y. Yin, and Y. Du. Surge Pricing and Labor Supply in the Ride-Sourcing Market. *Transportation Research Procedia*, 23:2–21, Jan. 2017. ISSN 2352-1465. doi:10.1016/j.trpro.2017.05.002. URL <https://www.sciencedirect.com/science/article/pii/S2352146517302703>.
- J. Zhang. The Fast and Frustrating Lives of China's Food Delivery Drivers. *Sixth Tone*, July 2019. URL <https://www.sixthtone.com/news/1004343>.
- N. Zuniga-Garcia, M. Tec, J. G. Scott, N. Ruiz-Juri, and R. B. Machemehl. Evaluation of Ride-Sourcing Search Frictions and Driver Productivity: A Spatial Denoising Approach. *Transportation Research Part C: Emerging Technologies*, 110:346–367, Jan. 2020. ISSN 0968-090X. doi:10.1016/j.trc.2019.11.021. URL <https://www.sciencedirect.com/science/article/pii/S0968090X18310684>.