Gig Worker Learning and Algorithmic Recommendations: From Empirical Insights to Algorithmic Design

SUBMISSION 2576

The gig economy has introduced unprecedented flexibility and autonomy for workers but also significant challenges in task selection and performance optimization. As gig platforms increasingly employ algorithmic recommendation systems to enhance worker efficiency, the interaction between human learning and algorithmic guidance remains underexplored. This paper investigates how gig workers on an on-demand retail delivery platform learn and adapt over time, balancing their own decision-making strategies with algorithmic recommendations. Using a dataset of 1.2 million orders completed by 5,000 workers in New York City over a year, we analyze worker performance improvements and decision behaviors through a two-way fixed-effects regression and a multinomial logit model. Our findings reveal diminishing returns to task-specific experience but highlight the transferability of cross-context experience. We further observe a decline in reliance on algorithmic recommendations as workers gain experience, underscoring the need for adaptive, learning-aware recommendation algorithms. Building on these insights, we propose a novel transformer-enhanced Upper Confidence Bound (UCB) framework for on-demand task recommendations, designed to align with workers' evolving capabilities and learning trajectories. This work contributes to the design of human-centered recommendation systems and offers actionable insights for enhancing worker engagement and platform performance in the gig economy.

Contents

Manuscript submitted for review to the 26th ACM Conference on Economics & Computation (EC'25).

1 Introduction

The rapid rise of gig economy platforms has fundamentally reshaped labor markets, offering flexible, task-based employment opportunities across industries such as ride-hailing, delivery, and freelancing. These platforms empower workers with autonomy over task selection, enabling them to optimize their work schedules and earnings. However, this independence comes with challenges. Unlike traditional work environments that provide structured training, mentorship, and peer support, gig workers often rely on trial-and-error learning to navigate task selection, optimize performance, and adapt to the platform's operational dynamics [Yao et al., 2021]. This reliance on self-directed learning can lead to inefficiencies and inconsistent performance, particularly for new workers.

To address these challenges, gig platforms have increasingly turned to algorithmic recommendation systems. These systems suggest tasks to workers based on demand patterns, historical performance, and other contextual factors, with the goal of enhancing operational efficiency and worker productivity. For example, in the retail delivery sector, platforms often recommend task bundling to minimize travel time and maximize earnings. While such recommendations can streamline workflows, they also add complexity to workers' decision-making processes, requiring them to reconcile algorithmic guidance with their own evolving strategies.

This paper investigates the interplay between gig worker learning and algorithmic recommendations, addressing two key questions: (1) How do gig workers improve their performance over time, and what roles do task-specific and cross-context experiences play in this process? (2) How do workers' responses to algorithmic recommendations evolve as they gain experience, and what implications does this have for the design of recommendation systems? To answer these questions, we analyze a dataset of 1.2 million orders completed by 5,000 gig workers on a retail delivery platform in New York City over a 364-day period. Our study combines econometric analysis of worker learning with behavioral modeling of task selection to provide a comprehensive understanding of worker-platform interactions.

The findings from this analysis highlight key dynamics in gig worker behavior. First, we observe significant learning effects, with worker performance improving rapidly through task-specific experience but showing diminishing returns after approximately 100 orders. Importantly, cross-context experience—accumulated from working across different stores—also enhances performance, indicating the transferability of general skills such as navigation and time management. Second, our analysis of task selection behavior reveals that workers initially rely heavily on algorithmic recommendations but gradually develop independent strategies, reducing their dependence on the platform's guidance. This shift underscores the limitations of one-size-fits-all recommendation systems and highlights the need for adaptive algorithms that account for workers' evolving capabilities and preferences.

Building on these insights, we propose a novel transformer-enhanced Multi-armed Bandits (MAB) framework for job recommendations. This framework leverages a tabular transformer architecture to predict service quality with rich contextual data and integrates these predictions into a Upper Confidence Bound (UCB) algorithm which dynamically recommends jobs to gig workers using their learning curve information. By balancing exploration and exploitation, the proposed algorithm adapts task recommendations to workers' individual learning curves, optimizing long-term platform performance while supporting worker growth.

This paper contributes to the literature on gig economy operations, worker learning, and recommendation system design in several ways. First, it provides empirical evidence on the dynamics of gig worker learning, extending prior research on learning curves in traditional and algorithmic work settings [Argote, 2012, Bavafa and Jónasson, 2021]. Second, it highlights the behavioral nuances of task selection in gig platforms, offering insights into how workers balance algorithmic guidance with personal decision-making. Finally, it introduces a conceptual framework for learning-aware recommendation algorithms, bridging the gap between empirical behavioral insights and computational design. Together, these contributions advance our understanding of human-algorithm collaboration in the gig economy and offer actionable implications for platform design.

The remainder of the paper is structured as follows. Section 1.1 reviews the relevant literature and situates our contributions within both empirical and computational research streams. Section 2 describes the dataset and study context. In Section 3, we present our econometric analysis of worker learning and performance improvement. Section 4 examines task selection behavior via a multinomial logit model. Section 5 discusses the implications of our findings and outlines our proposed framework for learning curve-aware job recommendations. Finally, Section 6 summarizes our contributions and limitations as well as highlights future directions.

1.1 Related Work

Our work relates to three major streams of literature: worker learning in general, gig worker decision-making and learning, and online recommendation algorithm.

1.1.1 Worker learning and performance improvement. Worker learning is a topic extensively studied in traditional corporate environments. Comprehensive reviews of this research can be found in the works of [Argote, 2012, Dar-El, 2013]. Given the depth and breadth of studies in this area, we provide only a brief review here. Worker learning has been studied across various workplace settings. For instance, [Fong Boh et al., 2007] explored learning in software development, [Shafer et al., 2001] investigated learning in assembly lines, [Grosse and Glock, 2015] examined manual item-picking processes, and [Bavafa and Jónasson, 2021] focused on emergency service workers. Researchers have found that workers learn through a variety of mechanisms, with experience-based learning being the most common, see [Gorry, 2016, Hermann et al., 2009]. The learning curve is one of the most widely used methods to measure this process [Anzanello and Fogliatto, 2011]. In addition, [Clark et al., 2013] studied how workers learn from customers, while [Akşin et al., 2021] explored learning from other team members. Notably, [Camerer and Hua Ho, 1999] proposed a reinforcement learning model, experience-weighted attraction, to study learning under strategic decision-making and how workers learn from the rewards they receive from past interactions.

While worker learning is well-studied in traditional workplace settings, learning in the gig economy may present unique challenges and opportunities. As [Gerber, 1998] points out, learning mechanisms can vary significantly across different workplace contexts and must be examined accordingly. Our work contributes to this body of literature by exploring worker learning within the gig economy, offering new insights into a rapidly evolving labor landscape.

1.1.2 Gig worker decision-making and learning. As the gig economy continues to evolve, research has also focused on various dimensions of gig worker, including workers' performance, learning, and decision-making. For example, [Allon et al., 2023] identified that gig workers are motivated not only by pay rates but also by internal motivators such as income and time targets. [Knight et al., 2022] showed that AI-guided systems can enhance service quality among gig workers, particularly for novices, although these systems can also extend task completion times due to reliance on AI consultations. [Guha and Corsten, 2023] emphasized the importance of day-to-day experiences in improving both service quality and productivity. [Dai et al., 2022] proposed a model of exploration-exploitation behavior among gig workers, arguing that during the early stages of experience, workers explore new regions, leading to decreased productivity and lower quality outcomes. As workers gain more experience, they tend to batch more orders and achieve better performance. [Benson et al., 2018] present experimental evidence showing that public reputation signals significantly influence worker recruitment and task acceptance in online labor markets. [Chen and Sheldon, 2015] analyze surge pricing on the Uber platform using panel data and structural estimation, revealing how dynamic price signals drive worker behavior. [Lehdonvirta, 2018] empirically investigates how gig workers improve over time on online piecework platforms.

Our research extends beyond these existing studies by focusing specifically on the learning processes and strategic decision-making of gig workers as they interact with platforms. In particular, we investigate how heterogeneous strategies among gig workers lead to diverse learning outcomes. Leveraging the advantages of our dataset, we are able to employ a multinomial logit (MNL) model to systematically study the choices gig workers make, particularly in response to platform recommendations.

1.1.3 Multi-armed bandits and its applications on recommendation system. The multi-armed bandit (MAB) framework has long served as a fundamental model for sequential decision-making under uncertainty. Originally introduced by Thompson [Thompson, 1933], the MAB problem captures the trade-off between exploration—gathering information about uncertain options—and exploitation—leveraging known rewards to maximize immediate gain. This foundational concept has spurred decades of research across statistics and computer science, setting the stage for numerous algorithmic developments. Building on these early ideas, Auer, Cesa-Bianchi, and Fischer [Auer et al., 2002] advanced the literature by developing the Upper Confidence Bound (UCB) algorithm, which provides finite-time performance guarantees for balancing exploration and exploitation. This approach has become a cornerstone for adaptive algorithms, and its theoretical insights have been widely adopted in various applications. Comprehensive introductions to reinforcement learning and sequential decision-making, such as the work by Sutton and Barto [Sutton and Barto, 2018], as well as more modern treatments like Lattimore and Szepesvári [Lattimore and Szepesvári, 2020], further underscore the broad applicability of MAB methods.

The MAB framework has found significant applications in recommendation systems and gig platforms, where the challenge is to make sequential recommendations in dynamic environments. For instance, Li et al. [Li et al., 2010] employed a contextual bandit approach to personalized news recommendation, effectively integrating user context into the decision process. Similarly, Chapelle and Li [Chapelle and Li, 2011] provided empirical evidence supporting the use of Thompson Sampling in online recommendation settings, highlighting its capacity to rapidly adapt to new information. [Johari et al., 2016] formulate a dynamic matching problem using MAB where the platform must learn heterogeneous worker types while balancing exploration and exploitation. [Ho et al., 2014] address the repeated principal–agent problem in crowdsourcing markets by leveraging a multi-armed bandit framework to adaptively refine contract offers.

Collectively, this body of literature provides a robust foundation for our work, which first proposes a transformer-enhanced UCB prototype that incorporates worker learning behavioral insights into dynamic job recommendation. Although our computational proposal is currently conceptual and awaits further simulation and experimental validation, it offers a promising direction for future research aimed at integrating empirical behavioral insights with online algorithmic design.

2 Data: US-Based Retail Delivery Platform

We collaborate with an on-demand retail delivery company (hereafter referred to as "the company" or "the platform") to analyze a comprehensive dataset consisting of online retail orders completed in New York City over a 364-day period, spanning from November 2022 to October 2023. This dataset captures a wide range of information, including completed orders by workers, order characteristics, and productivity metrics such as time spent shopping, checkout time, and driving time. Additionally, the dataset provides detailed evaluations of each completed order, such as whether the delivery was on time.

One of the key advantages of this dataset is its granularity, which allows us to observe: (1) the orders recommended to each gig worker by the platform's algorithm, and (2) detailed information about orders that were bundled together by the platform for simultaneous delivery. In the following sections, we provide an overview of the platform's operations, describe the interface through which workers interact with the system, and present descriptive statistics related to the workers and the recommended orders they received. We also outline the supplementary datasets incorporated into our analysis.

2.1 Platform Overview

The company operates as an online retail delivery platform, offering on-demand retail and essential goods delivery services across multiple cities in the United States. Customers place orders through the platform's mobile application or website, with the option to schedule deliveries at flexible times. The platform facilitates prompt delivery by matching customers with gig workers who are responsible for driving to the store, hand-picking the ordered items, and providing real-time communication via chat services for updates on the shopping process. Gig workers then deliver the items directly to the customers' addresses.

Gig workers are compensated on a per-order basis, with payment varying depending on factors such as the size and complexity of the order. In addition to their base earnings, gig workers can receive tips directly from customers, providing an additional source of income. Furthermore, the platform offers bonuses to gig workers for meeting specific performance criteria, such as fulfilling deliveries during high-demand periods.

2.2 Worker Process

To participate on the platform, workers must first undergo a screening process, which includes verifying that they meet certain eligibility criteria such as being over 18 years of age, possessing a valid driver's license, and owning a vehicle. Upon successfully completing this process, workers are officially designated as gig workers. These gig workers can select their preferred working regions and define their working hours daily, typically within the operational timeframe of 7:00 AM to 12:00 AM. The platform utilizes this information, along with historical information of customer experiences with the gig worker such as on-time delivery rate and customer ratings-to generate order recommendations.

A key distinction between this platform and ride-hailing services (e.g., Uber and Lyft) is the ability of gig workers to exercise discretion in selecting orders. Unlike ride-hailing drivers, who are automatically assigned rides and lack the ability to browse available tasks, gig workers on the focal platform are presented with a list of recommended orders generated by an algorithm. Gig workers can browse through these recommendations and make informed decisions based on details such as payment amount, delivery time windows, store and customer locations, and the items included in the order. Additionally, they have the flexibility to bundle multiple orders and fulfill them concurrently, thus optimizing their work efficiency.

Beyond receiving algorithmically recommended orders based on their selected hours and regions, gig workers also have access to a separate section of the platform where they can view non-algorithmically recommended orders. Gig workers are free to select from this pool of available orders, providing them with additional opportunities to maximize their work during their active hours.

2.3 Descriptive Statistics

The dataset consists of approximately 5,000 gig workers who collectively fulfilled around 1.2 million orders across 800 stores. The number of orders completed by each gig worker within a year varies significantly, ranging from individuals who completed only one order before leaving the platform, to highly active gig workers who processed over 6,000 orders in a single year. On average, gig workers completed 230 orders annually, with the 25th, 50th, and 75th percentiles at 5, 28, and 136 orders, respectively.

Similarly, the volume of orders processed by each store shows considerable variation, from as few as one order per year to over 100,000. On average, stores processed 1,600 orders annually, with the 25th, 50th, and 75th percentiles at 5, 13, and 39 orders, respectively.

A significant portion of orders on the platform−approximately 60%−are bundled. The platform employs algorithms to identify similarities between orders based on features such as store, item selection, and delivery destination. Each bundled order consists of two similar orders.

2.4 Supplementary Data: TLC Trip Records and Weather Records

To account for the potential influence of traffic and weather conditions on workers' behaviors, we incorporate two additional datasets into our analysis.

The first dataset is the New York City Taxi and Limousine Commission (TLC) dataset, which provides detailed trip-level records for taxi and ride-hailing services in New York City (NYC). This dataset includes information such as pickup and drop-off locations, timestamps, trip distances, fares, and payment methods, encompassing millions of rides over multiple years. From the TLC dataset, we derive two key traffic-related proxies: the traffic volume for each hour and the average hourly speed of taxis, both serving as indicators of overall traffic conditions in NYC.

The second dataset is sourced from the OpenWeather platform, which offers global meteorological data across a broad range of parameters, including temperature, humidity, wind speed, and precipitation, as well as specialized metrics like air pollution and UV index. We initially extracted over 50 weather variables from this platform. After performing variance inflation factor (VIF) testing to address multicollinearity, we selected three weather parameters—apparent temperature, rainfall, and wind speed—for inclusion in our subsequent regression analyses.

3 Learning to Improve: *How Do Gig workers Learn to Improve Performance?*

In this section, we analyze how gig workers learn to improve their performance over time as they accumulate experiences, focusing on two key performance metrics: (1) the on-time percentage (OTP), which represents the proportion of orders delivered no later than the time specified by the platform and serves as the platform's indicator of service quality; and (2) the number of items picked per hour, which reflects the gig worker's productivity. These metrics provide a comprehensive view of both the service reliability and operational efficiency of gig workers. We begin by presenting model-free evidence to identify general trends in performance improvement. This is followed by the introduction of our empirical approach, which employs a two-way fixed effects regression analysis to control for time-invariant characteristics of gig workers and stores. Finally, we present the results and insights derived from the analysis. These insights will serve as a basis for the subsequent section, where we investigate how gig workers learn to adapt strategically to the platform's recommendation algorithms.

3.1 Model-free Evidence of Performance Improvement

We first consider 1,131 gig workers who joined the platform within the duration of the dataset (e.g., after November 1, 2022) and investigate how their performance may improve over time. Figure 1 illustrates the relationship between the number of orders a gig worker has completed (binned into intervals) and their corresponding average on-time delivery rate. The x-axis represents the binned number of orders worked observed in the data, ranging from 0 to 500, divided into intervals of approximately 10 orders. The y-axis denotes the average on-time rate, varying from 0.75 to 0.925. Error bars indicate the confidence intervals for the average on-time rate within each bin.

The trend reveals an initial dip in the on-time performance measure within the lower order bins (0–10 orders), which can be attributed to the platform introducing bundled orders to gig workers after they have completed few deliveries. Following this dip, there is a consistent improvement in the average on-time rate as the number of orders increases. Once gig workers exceed 100 orders, the on-time rate stabilizes with fluctuations around the 0.85–0.9 range. This pattern suggests that, while accumulated experience has a positive impact on a gig worker's performance, such benefit has diminishing returns as experience continues to grow.

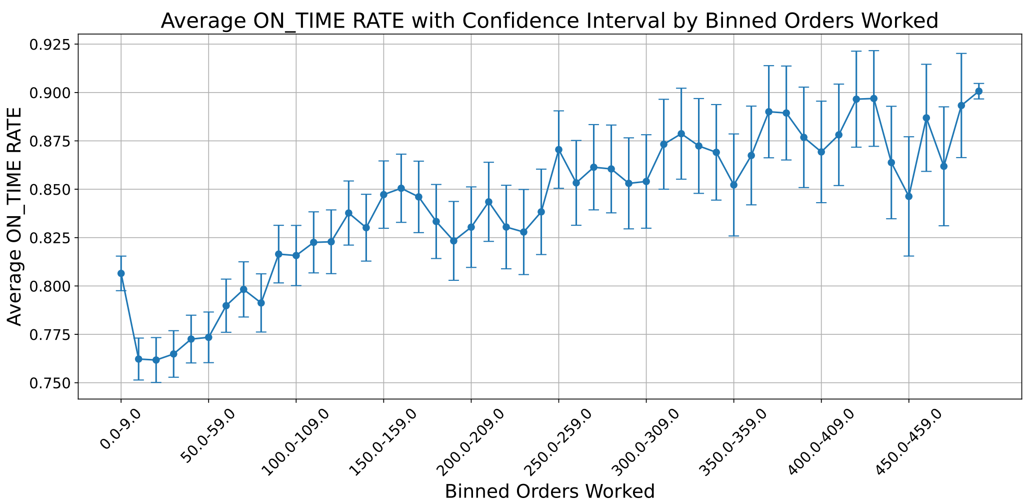

Fig. 1. Average rate of on-time delivery by orders worked

Notes: This figure illustrates the relationship between the number of binned orders worked and the average on-time delivery (OTP) rate, with 95% confidence intervals represented by the error bars. Overall, the figure suggests that while initial experiences yield rapid performance improvements, sustained gains require more effort, and the learning curve eventually flattens as workers reach proficiency.

3.2 Two-way Fixed-Effects Regression Analysis of Worker Performance

To establish a causal relationship between gig worker experience and performance, we perform a two-way fixed effects regression analysis for each of the two performance metrics: on-time delivery rate ($OnTime$) and the number of items picked per hour ($ItemsPerHour$) [Wooldridge, 2010].

$$\begin{aligned} PerformanceMetric_{ist} = \beta_0 &+ \beta_1 OTS_{ist} + \beta_2 OTS_{ist}^2 \\ &+ \beta_3 OOS_{ist} + \beta_4 OOS_{ist}^2 \\ &+ \mathbf{X}'_{ist}\beta + \mu_{is} + \gamma_t + \epsilon_{ist} \end{aligned} \tag{1}$$

where

- $PerformanceMetric_{ist}$ is either the delivery performance of or the number of items picked by gig worker i when shopping at store s at time t.
- OTS_{ist} and OOS_{ist} captures the number of orders that gig worker i has completed at store s and other stores by time t, respectively.
- OTS^2_{ist} and OOS^2_{ist} are the squared terms.
- $\mathbf{X}_{ist}$ is a vector of control variables, including external factors such as weather conditions (e.g., temperature, rain, wind speed), order-specific variables (e.g., total payment, bonuses, requested item quantities, delivery distance), and urban traffic metrics (e.g., taxi volume, average traffic speed).
- μ_{is} represents gig worker-store fixed effects, controlling for unobserved heterogeneity at the gig worker-store level that is constant over time.
- γ_t denotes time fixed effects, capturing any temporal patterns such as day-of-week or seasonal variations that might influence performance.
- ϵ_{ist} is the idiosyncratic error term, assumed to be independently and identically distributed across i, s, and t.

3.2.1 Description of Key Variables.

Dependent variables. $OnTime$ is a binary variable capturing the rate of on-time delivery, which equals to 1 if the delivery was completed on time, or 0 if the delivery was delayed. This measure of service quality is chosen to be the key performance indicator by the platform. Another performance metric considered is $ItemsPerHour$, which is the number of items the gig worker successfully picked per hour. $ItemsPerHour$ therefore serves as a proxy for the gig worker's productivity.

Independent variables. To examine the relationship between a gig worker's accumulated experience with specific stores and their delivery performance, we introduce two key independent variables: OTS ($OrdersThisStore$) and OOS ($OrdersOtherStore$). OTS represents the number of deliveries completed by a gig worker for a particular store. This variable functions as a proxy for the gig worker's familiarity with that store's unique operational environment, such as store layout, inventory management, and staff interactions. We hypothesize that as a gig worker visits a given store more often, their delivery efficiency will improve due to their familiarity with the store. For example, they might spend less time searching for items in the store and make fewer errors. Additionally, familiarity with the store may allow for more effective route optimization, both within the store during item retrieval and externally during the delivery process. To explore potential nonlinear effects of experience, we also include a squared term, OTS^2, which enables us to test whether performance improvements exhibit diminishing returns after a certain threshold of experience, or whether continued experience yields progressively better outcomes.

In contrast, OOS captures the number of deliveries a gig worker has completed for stores other than the focal one. This variable allows us to explore whether broader experience across different store environments translates to enhanced performance in a specific store. Similarly, we also incorporate a squared term, OOS^2 to account for potential nonlinear effects.

Control variables. $StoreId$: A categorical variable controlling for store-specific fixed effects, such as differences in store location, management efficiency, or operational processes, that may influence delivery performance. $WorkerId$: A categorical variable controlling for individual gig worker-specific effects, accounting for heterogeneity in personal efficiency, delivery habits, or experience levels.

Time fixed effects: The delivery timestamp is decomposed into months and weekdays to control for temporal dynamics that may affect delivery performance, such as seasonal demand fluctuations or weekday traffic patterns. *Order characteristics*: We include variables that describe the features of each order, such as financial incentives, items, distances from store to customer to account, delivery time window for the complexity of each order, which are hypothesized to impact the timeliness of deliveries. *Traffic*: Variables like hourly taxi volume and hourly average speed of taxi at NYC are incorporated to control for urban traffic conditions that could delay deliveries. *Weather*: Environmental factors, such as apparent temperature, rain, and wind speed, are included to account for weather conditions that may significantly affect delivery times.

3.3 Results: *Diminishing Positive Return on Experience*

Table 1 illustrates the impact of gig worker experience on two key performance metrics: $OnTime$ (the rate of on-time delivery) and $ItemsPerHour$ (the number of items picked per hour).

Table 1. The impact of experience on performance among new gig workers

	$OnTime$	$ItemsPerHour$
OTS	6.0552e-05***	9.4281e-03***
	(0.003)	((0.008))
OTS^2	-8.9995e-09***	-5.4681e-06**
	(0.000)	(0.003)
OOS	5.9109e-05***	6.3414e-03*
	(0.000)	(0.022))
OOS^2	-8.8744e-09***	-7.6253e-07
	(0.015)	(0.235)
Fixed Effects controls	✓	✓
Weather controls	✓	✓
Traffic controls	✓	✓
R^2	0.029	0.013
Observations	105543	105543

Our results offer strong evidence that both store-specific and general delivery experience significantly influence on-time delivery performance and overall productivity. First, we observe that store experiences, as measured by OTS_{ist}, plays a critical role in driving improvements in these metrics. The inclusion of the squared term, OTS_{ist}^2, highlights a pattern of diminishing returns: although early interactions with a store lead to notable gains in performance, these improvements plateau after approximately 100 to 150 orders. This suggests that gig workers quickly internalize the store's processes, with additional experience providing only marginal benefits.

Second, experience accumulated from working at other stores also enhances performance, indicating the presence of transferable skills. Gig workers appear to draw on general knowledge—such as navigating various store layouts, managing diverse customer requests, and optimizing in-store operations—which contributes to timely deliveries even in less familiar environments.

Third, while experience from other stores improves both service quality and productivity, we observe that its impact on service quality is more significant, suggesting that cross-store experience helps gig workers adapt to different customer expectations and service standards.

In summary, our findings show that delivery performance improves rapidly with initial store-specific experience but reaches a plateau after a certain point.

4 Responding to Recommendations: *Orders to Select*

While gig workers can freely choose to work on any order available on the platform, the platform typically recommends a number of orders to them based on the demand level and past performance. We denote the orders that are recommended by the platform as *algorithmically recommended orders* and the remaining orders as *non-algorithmically recommended orders* .

In this section, we break down the first 200 orders into 5 periods and apply a multinomial logit (MNL) model [McFadden, 1972] to estimate the choice a gig worker makes in each of these periods. This analysis allows us to investigate how gig workers' decision-making processes evolve over time as they gain more experience and become more familiar with the platform's recommendations. Understanding these dynamics provides insight into the strategic choices gig workers make as they navigate the platform.

4.1 Multinomial Logit Model of Workers' Selected Orders

4.1.1 Description of Key Variables. We describe the variables used in the multinomial logit model to analyze gig worker behavior across different choice sets. The dependent variable is the choice outcome, and the explanatory variables represent the characteristics of the alternatives and the gig workers.

Dependent variable. The dependent variable *CHOSEN* in the model is a binary indicator representing whether a specific alternative was chosen that takes the value 1 if the alternative was chosen by the gig worker, and 0 otherwise.

Independent variables. We document the important independent variables included in the MNL model here. We defer the complete list of variables to Appendix A. These variables represent characteristics of the alternatives and other relevant attributes influencing the decision. Each of these variables contributes to the deterministic component of the utility function, V_{ij}, for each alternative j faced by gig worker i.

- *LIST*: Indicates whether the order is in the recommendation list by the platform's algorithm (1) or it is not (0).
- *REINFORCEMENT*: The total earnings from the gig worker's previous 100 orders at a given store. This variable reflects how exploratory gig workers are, measuring whether they prefer familiar stores or explore new ones. It mimics a reinforcement learning model, as described by [Camerer and Hua Ho, 1999], where it is assumed that people learn from the rewards they receive from past interactions.

Independent variables. We again classify gig workers into 5 groups based on the total number of orders they have completed. We then introduce group-specific effects in the MNL model by converting the group category, into dummy variables using one-hot encoding. The group `300+` was designated as the reference group, and dummy variables were created for the remaining groups. Interaction terms were then generated between these group dummies and key independent variables (e.g., *LIST*), allowing us to capture how the effect of these predictors varied across gig worker groups. Only interaction terms with non-redundant information were included, ensuring efficient model specification. This approach allows us to interpret coefficients in relation to the reference group, revealing group-specific differences in decision-making behavior.

4.1.2 Model specification.

Utility function. The utility function U_{ij} for alternative j and gig worker i is composed of a deterministic component V_{ij} and a stochastic component ϵ_{ij}:

$$U_{ij} = V_{ij} + \epsilon_{ij} \tag{2}$$

The deterministic component V_{ij} is modeled as a linear function of the explanatory variables:

$$V_{ij} = \beta_0 + \beta_1 \cdot \text{Bundled}_{ij} + \beta_2 \cdot \text{List}_{ij} + \ldots \tag{3}$$

Here, β_k represents the coefficient associated with each explanatory variable, and ϵ_{ij} is the error term, assumed to follow a Gumbel distribution.

Choice probabilities. The probability that gig worker i chooses alternative j is given by the following multinomial logit probability function:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{l=1}^{J} \exp(V_{il})} \tag{4}$$

where J is the number of available alternatives in the choice set.

4.1.3 Estimation method. The parameters β are estimated using Maximum Likelihood Estimation (MLE) by maximizing the log-likelihood function:

$$\ln L(\beta) = \sum_{i=1}^{N} \left(V_{iy_i} - \ln \left(\sum_{l=1}^{J} \exp(V_{il}) \right) \right) \tag{5}$$

where y_i denotes the alternative chosen by gig worker i. The MLE process yields estimates of the coefficients β, which quantify the effect of each independent variable on the choice probability.

4.2 Results: *Workers Follow Recommendations Less with More Experience*

We estimate the Multinomial Logit (MNL) choice model across five distinct groups of gig workers categorized by their total completed orders: those with fewer than 10 orders, between 10 and 50 orders, between 50 and 100 orders, between 100 and 300 orders, and those with over 300 orders. This analysis is conducted over five time periods, representing different stages in their shopping history: the first 10 orders, between 10 and 20 orders, between 20 and 50 orders, between 50 and 100 orders, and between 100 and 200 orders.

Figures 2 and 3 present the estimated coefficients and their corresponding confidence intervals for the two primary independent variables *REINFORCEMENT* and *LIST*. The x-axis represents the order period for which the choice model is estimated (e.g., 0-10 indicates the estimated coefficients for a worker's first 0-10 orders). The y-axis displays the coefficient values. The different colored columns correspond to the five distinct groups being analyzed.

To facilitate interpretation, we generated dummy variables in 4.1.1, using the group with yearly orders $\geq$ 300 as the reference category. Consequently, in the two figures presented, only the rightmost bar (in purple) depicts the actual coefficient for this reference group. All other bars represent the differential coefficients relative to the reference group, illustrating the deviation in effect sizes for each corresponding category.

Figure 2 illustrates the extent to which different groups of gig workers exhibit exploratory behavior. The variable measures the total dollar amount a gig worker has earned at the same store over their last 100 orders. Higher coefficient values suggest that gig workers are placing more weight on stores they have previously visited when making choices. This behavior indicates a tendency toward exploitation rather than exploration. The figure demonstrates that, for gig workers with fewer than 10 orders, only the group that exits the platform early within 10 orders–represented by the blue bar–shows a negative coefficient relative to the reference group. Notably, the absolute value

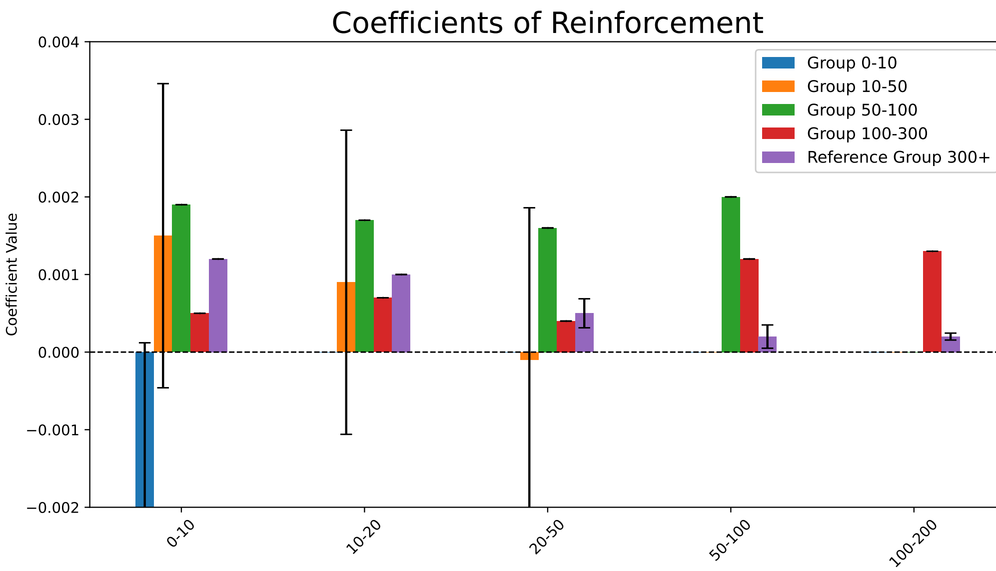

Fig. 2. Estimated coefficients of REINFORCEMENT, defined by total earnings in the chosen store in the last 100 orders, across worker groups and levels of experiences

Notes: This figure illustrates the coefficients of the *Reinforcement* variable across different order ranges (0–10, 10–20, 20–50, 50–100, and 100–200) for five distinct worker groups based on the total number of orders completed. The *Reinforcement* variable represents the total earnings from the worker's previous 100 orders at a given store, serving as a proxy for exploratory behavior. Larger coefficients suggest a preference for familiar stores (exploitation), while smaller or even negative coefficients indicate tendencies toward exploring new stores. Error bars represent 95% confidence intervals. The reference group (300+ orders) is represented in purple, with other groups compared relative to this baseline. Notably, the Group 0–10 (blue bar) shows a negative coefficient in the early stage, indicating excessive exploration, which correlates with early dropout rates. In contrast, high-performing workers (reference group) display consistently positive coefficients, suggesting a more balanced exploration-exploitation strategy. Note that since other groups' coefficients are relative values comparing to the reference group, the reference group has the highest exploration rate among other groups except for the early dropouts.

of this negative coefficient exceeds that of the reference group's (purple bar). This finding suggests that early dropouts are the only group that exhibits negative weighting for this variable, implying that they rarely return to previously visited stores and may engage in excessive exploration during the initial stages of platform use.

Another important observation is that early dropouts, including those who leave within 10 orders (blue bar) and within 50 orders (yellow bar), exhibit high variance in their coefficients compared to other groups. This suggests that early dropout behavior is linked to inconsistent exploratory store choices. In other words, users who do not settle into a consistent store selection pattern — through repeated visits to familiar stores —are more likely to abandon the platform.

In addition, we observe that the best performers—gig workers who have completed more than 300 orders per year and exhibit the highest performance metrics, as demonstrated in previous sections—while placing positive weights on the variable, actually have the lowest coefficient value

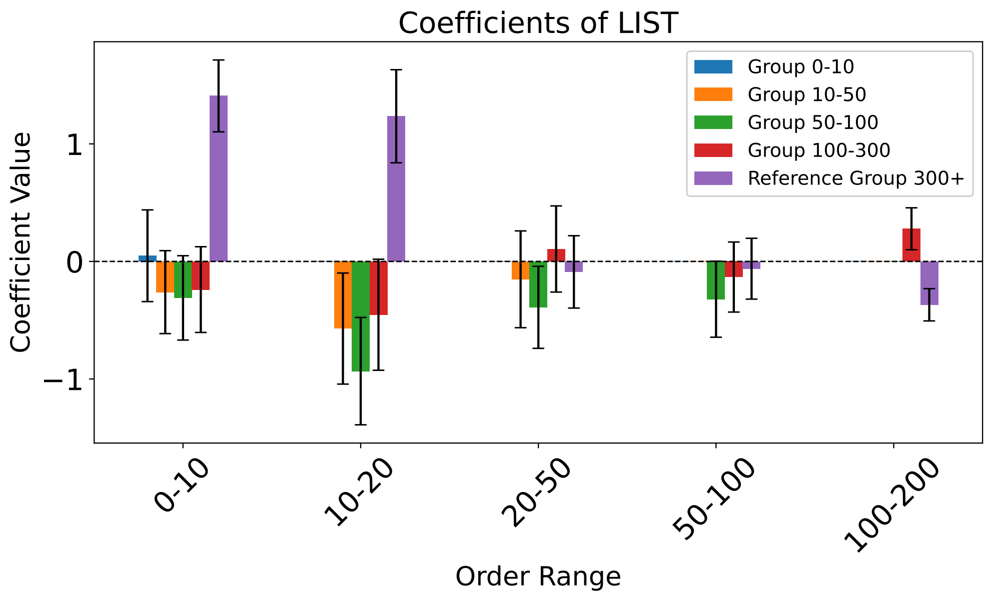

Fig. 3. Estimated coefficients of LIST, defined by whether recommended by the platform's algorithms, across worker groups and levels of experiences

Notes: This figure displays the estimated coefficients of the *LIST* variable across different order ranges (0–10, 10–20, 20–50, 50–100, and 100–200) for five distinct worker groups categorized by the total number of orders completed. The *LIST* variable represents whether an order was recommended by the platform's algorithm (1) or not (0). Positive coefficients suggest a higher likelihood of selecting recommended orders, while negative coefficients indicate a preference for non-recommended orders. The reference group (300+ orders) is represented in purple, with other groups compared relative to this baseline. The results show that the gig workers with best performance (reference group) initially rely heavily on recommendations but gradually shift towards making independent decisions, as reflected by the decline in coefficients over time. In contrast, other workers show less reliance on recommendations in the beginning phase. Error bars represent 95% confidence intervals, highlighting uncertainty in the estimates, particularly among less experienced workers.

among all groups except for early dropouts (noting that the other groups' coefficients in the figure represent differences relative to the reference group). This suggests that a moderate level of exploration may be beneficial. As discussed in Section 3, gaining experience in other stores can also contribute to improved performance, implying that balancing exploration and exploitation is key to long-term success on the platform.

Figure 3 illustrates how different groups respond to the platform's recommendation algorithm over time. A higher coefficient value indicates that the group is more likely to choose recommended orders compared to other orders. We can observe that: (1) the best performers tend to follow the platform's recommendations at the beginning but gradually shift towards ignoring them or even placing negative weights on choosing recommended orders. (2) Other groups, while initially following fewer recommendations than the top-performing group, also tend to decrease their reliance on the platform's recommendations over time. This trend suggests that although gig workers initially tend to follow recommendations, over time, they become more selective or independent in their order selection, likely as they gain more experience and confidence in making their own choices, reducing their reliance on algorithmic suggestions.

5 Managerial Implications

5.1 Implications for Designing More Effective Recommendation Algorithms

Our empirical findings reveal key insights into gig worker behaviors, suggesting multiple ways to enhance recommendation algorithms on gig economy platforms to better support workers and improve overall performance.

5.1.1 Early dropouts: inconsistent behaviors and excessive exploration. The group of gig workers who exited the platform within their first 10 to 50 orders demonstrated a distinctively inconsistent decision-making pattern. Our results indicate a lack of store loyalty or a tendency to avoid returning to previously visited stores of these early dropouts. Additionally, they exhibited high variance in their store choices. Such inconsistency suggests that the lack of store exploitation and stable shopping routines early on is strongly correlated with dropping out early from the platform.

To mitigate early dropout rates caused by these factors among new workers, platforms should recommend tasks from stores where they have previously visited can help induce familiarity and reduce indecisiveness in the early stage. By supporting a healthy level of exploitation early on, platforms can improve worker welfare, consequently promoting worker retention and better integration into the cooperative work environment.

5.1.2 Optimal exploration among high performers. In contrast, gig workers who achieved the highest performance and remained active on the platform for the longest time were those who engaged in the moderate level of exploration. Among the remaining groups besides early dropouts, the high performers still explored more. This underscores the importance of finding an optimal balance in exploration behaviors, where excessive exploration may hinder performance, while a strategic level of exploration enhances success.

Recognizing that high-performing workers engage in optimal levels of exploration, recommendation algorithms should balance familiar and new task suggestions. Platforms can design systems that encourage strategic exploration without overwhelming workers, tailoring recommendations based on individual performance histories and preferences. This approach supports workers in developing effective strategies and enhances their learning processes over time.

5.1.3 Declining selection of platform-recommended orders. Furthermore, the data indicates that, over time, gig workers tend to select fewer platform-recommended orders. This shift away from algorithmic recommendations indicates that these gig workers might perceive platform suggestions as less beneficial or aligned with their evolved needs, pointing to the need for adaptive recommendation algorithm. As workers become more proficient, platforms should adjust their algorithms to provide greater flexibility, allowing experienced workers to align task selections with their personal strategies. By incorporating worker feedback mechanisms, platforms can further personalize recommendations, ensuring they remain relevant and beneficial as workers' needs evolve over time.

Overall, these implications emphasize the importance of human-centric recommendation systems that adapt to workers' learning trajectories and experience levels. By aligning algorithmic recommendations with the evolving strategies and preferences of workers, platforms can enhance collaborative dynamics, improve performance outcomes, and foster long-term engagement within the gig economy.

5.2 Designing Learning Curve-Aware Recommendation System

In this subsection, we give an example of how we could model and design a worker learning curve-aware job recommendation system. In our dataset context and many gig economy platforms, the supply of available workers far exceeds the volume of orders, prompting a need to shift the focus of recommendation systems from conventional metrics such as click-through rates toward service quality. Unlike traditional recommendation systems that primarily optimize for engagement or conversion, our objective is to dynamically assign orders to workers in a manner that maximizes cumulative service quality. This quality is measured by critical performance indicators such as on-time delivery, order accuracy, and overall customer satisfaction. By centering on service quality, our approach aligns with the operational needs of gig platforms where maintaining high service standards is essential for long-term customer retention and brand reputation.

5.2.1 Model. We consider the problem of dynamically assigning grocery delivery orders to a set of workers in order to maximize the long-term cumulative service quality. Each grocery order is associated with a context (e.g., store id, items ordered, delivery distance). A key feature of the problem is that workers learn and improve over time. We propose to use a tabular transformer model to predict the base service quality from the worker's historical data and the current order context, and then incorporate this prediction into a UCB framework that is adjusted for the known learning (improvement) trend.

Notation and Data.

- Let $T \in \mathbb{N}$ denote the total number of orders/time stamps.
- The set of workers is $\mathcal{K} = \{1, 2, \ldots, K\}$.
- For each order $t = 1, 2, \ldots, T$, let $x_t \in \mathbb{R}^d$ denote the context vector (order information such as store, items, delivery distance, etc.).
- For each worker $k \in \mathcal{K}$, denote by $H_k(t)$ the history of orders (context and observed service quality) completed by worker k up to time t:

$$H_k(t) = \{(x_{k,1}, r_{k,1}), (x_{k,2}, r_{k,2}), \ldots, (x_{k,N_k(t)}, r_{k,N_k(t)})\},$$

  where $N_k(t)$ is the number of orders assigned to worker k up to time t.
- Worker availability at time t is given by $A_k(t) \in \{0, 1\}$ (with $A_k(t) = 1$ if worker k is available, and 0 otherwise).

Worker Performance and Learning. Each worker k is assumed to have an intrinsic (base) ability $\mu_k^0 > 0$. However, as a worker completes more orders, their service quality improves according to a known *trend function* $D_k : \mathbb{N} \to \mathbb{R}_+$, which is assumed to be monotonically increasing. Thus, if worker k has completed n orders, the true expected service quality when processing their $(n+1)$-th order is modeled by

$$\mu_k(n+1) = \mu_k^0 \, D_k(n+1). \tag{6}$$

For example, one might have

$$D_k(n) = 1 + \alpha_k \log(1+n) \quad \text{or} \quad D_k(n) = 1 + \frac{L_k - 1}{1 + e^{-\gamma_k (n - n_{0,k})}},$$

with known parameters $\alpha_k, L_k, \gamma_k, n_{0,k}$ that capture the learning speed of worker k.

Observation Model. When worker k processes their i-th order, the observed service quality is given by

$$r_{k,i} = \mu_k^0 \, D_k(i) + \epsilon_{k,i}, \tag{7}$$

where $\{\epsilon_{k,i}\}$ is an i.i.d. sequence of zero-mean noise (e.g., $\epsilon_{k,i} \sim \mathcal{N}(0, \sigma^2)$).

5.2.2 A Transformer-enhanced UCB Job Recommendation Algorithm Prototype. The multi-armed bandit framework is a natural choice for our work because it offers a rigorous approach to sequential decision-making under uncertainty, particularly in settings where rewards are nonstationary and the optimal action evolves over time. In our application, gig workers' service quality improves with experience, so the reward associated with assigning an order to a worker is not static but changes as more orders are completed. By adopting a MAB framework, we explicitly capture the exploration–exploitation trade-off. The algorithm exploits workers who are currently delivering high service quality while simultaneously "investing" in those with high learning potential, even if their current performance is modest.

Recent advancements in MAB algorithms, such as Neural UCB [Zhou et al., 2020] and multi-armed bandits with known trends [Bouneffouf and Féraud, 2016], provide robust mechanisms to incorporate rich contextual information (e.g., order complexity) and to account for known learning trends into the decision process. These approaches yield theoretical guarantees on regret minimization, ensuring that our dynamic assignment strategy converges toward an optimal policy over time.

Moreover, our study leverages the transformer architecture (introduced in [Vaswani et al., 2017]), and in particular, the TabTransformer [Huang et al., 2020], to address the challenges inherent in modeling heterogeneous, tabular data typical of gig work settings. Transformers are celebrated for their self-attention mechanism, which adeptly captures complex, non-linear dependencies within sequential data. This quality is essential for modeling the evolution of gig workers' performance, as their historical records—comprising order contexts (e.g., store ID, items ordered, delivery distance) and observed service quality—exhibit intricate temporal patterns that standard architectures might miss. The TabTransformer, in particular, extends the conventional transformer framework by effectively embedding categorical variables through multi-head self-attention, thereby producing rich contextual representations from diverse order features. These representations are critical for accurately predicting base service quality and for deriving uncertainty estimates used in our UCB strategy.

In summary, the integration of a transformer-enhanced MAB framework is crucial for our approach. It delivers the computational efficiency and modeling flexibility needed to capture both the complex interactions among contextual features and the dynamic, nonstationary learning behavior of gig workers. This combined framework serves as the foundation for our transformer-enhanced UCB algorithm, which dynamically assigns orders to maximize cumulative service quality over time.

TabTransformer. We assume that a tabular transformer model $T_\theta : \mathbb{R}^d \times \mathcal{H} \to \mathbb{R}$ (with parameters θ) is used to predict the base service quality for a worker given the order context and the worker's historical data. Formally, for a new order with context x_t and for worker k (with history $H_k(t)$), we have

$$\hat{r}_k(x_t, H_k(t)) = T_\theta\big(x_t, H_k(t)\big). \tag{8}$$

Uncertainty Estimation. Similar to the Neural UCB algorithm [Zhou et al., 2020], we derive an uncertainty estimate from the network's output. Assume that via linearization (or the neural tangent kernel approximation) we obtain a feature representation $\phi_k(x_t, H_k) \in \mathbb{R}^{d_\phi}$ (e.g., the last-layer features of the network). Let

$$V_k = \lambda I + \sum_{s=1}^{N_k} \phi_k(x_{k,s}, H_k^{(s)}) \phi_k(x_{k,s}, H_k^{(s)})^\top$$

be the design matrix aggregated over worker k's history (with regularization parameter $\lambda > 0$). Then, the uncertainty for worker k on the new order is given by

$$\sigma_k(x_t, H_k) = \sqrt{\phi_k(x_t, H_k)^\top V_k^{-1} \phi_k(x_t, H_k)}. \tag{9}$$

UCB-Based Worker Selection. To balance exploration and exploitation, we adopt an Upper Confidence Bound (UCB) approach. For each worker k at time t, we first compute an empirical estimate of their performance. One common choice is the sample average:

$$\hat{\mu}_k(t) = \frac{1}{N_k(t)} \sum_{i=1}^{N_k(t)} r_{k,i}. \tag{10}$$

However, here we rely on the transformer prediction (8) for the base prediction and then add an exploration bonus.

Following approach in multi-armed bandits with known trend [Bouneffouf and Féraud, 2016], we now incorporate the adjusted prediction and the uncertainty estimate into a UCB-type index. Define the index for worker k at time t as

$$I_k(t) = [\tilde{r}_k(x_t, H_k) + \beta\, \sigma_k(x_t, H_k)] \cdot D_k\big(N_k + 1\big), \tag{11}$$

where:

- $\tilde{r}_k(x_t, H_k)$ is given by (8),
- $\sigma_k(x_t, H_k)$ is given by (9),
- $\beta > 0$ is a tunable exploration parameter,
- The multiplication by $D_k(N_k+1)$ in the uncertainty term reflects that the effect of exploration should also be scaled by the expected improvement.

Worker availability is incorporated via an indicator function $A_k(t)$; thus, the final selection rule is

$$a(t) = \arg\max_{k \in \mathcal{K}} \Big\{ I_k(t) \cdot A_k(t) \Big\}. \tag{12}$$

Overall Algorithm Framework. The overall procedure is as follows:

(1) **Initialization:** For each worker $k \in \mathcal{K}$, initialize the history $H_k(0)$ (which may be preloaded) and set $N_k(0) = 0$.
(2) **For each time step** $t = 1, 2, \ldots, T$**:**
  (a) A new order with context x_t arrives.
  (b) For each worker k with $A_k(t) = 1$ (i.e., available), compute the transformer prediction
  $$\tilde{r}_k(x_t, H_k(t)) = T_\theta\big(x_t, H_k(t)\big)$$
  (c) Compute the UCB index
  $$I_k(t) = [\tilde{r}_k(x_t, H_k) + \beta\, \sigma_k(x_t, H_k)] \cdot D_k\big(N_k + 1\big),$$
  (d) Select the worker
  $$a(t) = \arg\max_{k \in \mathcal{K}} \Big\{ I_k(t) \cdot A_k(t) \Big\}.$$
  (e) Assign order t to worker $a(t)$ and observe the service quality r_t.
  (f) Update the history for worker $a(t)$:
  $$H_{a(t)}(t+1) = H_{a(t)}(t) \cup \{(x_t, r_t)\},$$
  and update $N_{a(t)}(t+1) = N_{a(t)}(t) + 1$. For all $k \neq a(t)$, set $H_k(t+1) = H_k(t)$ and $N_k(t+1) = N_k(t)$.
  (g) Optionally, update the transformer model T_θ using the new data.

Incorporating Worker Task Selection Behavior. While our current transformer-enhanced UCB framework assigns orders to workers based on predicted service quality and learning potential, it does not explicitly model the possibility that workers might reject or delay a recommended order. In our formulation, we assume that once an order is assigned, the worker accepts it and processes it accordingly. This simplifying assumption allows us to concentrate on the core challenge of balancing exploration and exploitation in a nonstationary environment where worker performance improves with experience. In practice, however, gig workers may sometimes reject an order due to personal preferences, scheduling conflicts, or other strategic considerations like in our MNL choice model analysis that our current model does not capture.

Algorithmic Modifications to Account for Order Rejection. To address this limitation, the algorithm can be extended by incorporating an acceptance probability function, denoted as $p_{\text{accept},k}(x_t, H_k)$, which estimates the likelihood that worker k will accept an order with context x_t given their history H_k. This function can be learned from historical order acceptance data using logistic regression or neural network models. The modified UCB index can then be expressed as

$$I'_k(t) = I_k(t) \cdot p_{\text{accept},k}(x_t, H_k),$$

where $I_k(t)$ is the original UCB index defined in Equation (11). By integrating $p_{\text{accept},k}(x_t, H_k)$ into the selection rule, the algorithm effectively balances the potential gain in service quality with the risk of order rejection.

6 Concluding Remarks

Our paper offers novel insights into the dynamics of gig worker learning and the role of algorithmic recommendations in shaping task selection strategies. Using a rich dataset of 1.2 million orders completed by gig workers over a year, we observed that workers exhibit significant learning effects, with rapid performance improvements through task-specific experience. However, these gains diminish after approximately 100 tasks, emphasizing the plateauing nature of skill acquisition in repetitive environments. Furthermore, cross-context experience—gained by working across diverse stores—proved critical for developing transferable skills, highlighting the value of varied exposure in promoting adaptability.

Our analysis revealed a notable decline in reliance on platform recommendations as workers gain experience, with experienced workers prioritizing self-developed strategies over algorithmic guidance. This highlights the need for adaptive recommendation algorithms that evolve alongside workers' learning trajectories. Specifically, platforms must balance exploration (i.e., encouraging workers to take on new tasks) and exploitation (i.e., leveraging familiarity) to optimize worker engagement and task efficiency. Early-stage workers benefit from familiar tasks that reduce cognitive load, while experienced workers thrive when presented with opportunities for strategic exploration.

Building on these empirical findings, we proposed a conceptual multi-armed bandit framework for a learning curve-aware job recommendation algorithm. By integrating a tabular transformer with an Upper Confidence Bound (UCB) approach, our framework aims to dynamically balance exploration and exploitation in a non-stationary environment where workers' performance evolves over time. Although conceptual, this framework offers a promising direction for future research, bridging the gap between behavioral insights and algorithmic design.

Our paper contributes to the literature on gig economy operations and human-centered algorithm design in several ways. First, it empirically demonstrates the interaction between worker learning and platform recommendations. Second, it highlights the behavioral nuances of gig workers' task selection processes. Finally, it introduces a novel algorithmic framework that incorporates worker learning curves into job recommendation systems.

However, our research has limitations. The dataset is limited to a single city and platform, potentially restricting generalizability. Additionally, we focus on individual worker learning, leaving broader social dynamics and external influences unexamined. Future research should expand these findings across diverse geographic and operational contexts, explore the role of team dynamics, and validate algorithmic proposals through simulations and field experiments.

In conclusion, our work underscores the importance of designing adaptive, human-centered recommendation systems that evolve in tandem with workers' learning trajectories. By aligning platform strategies with workers' development, gig economy platforms can foster sustainable worker engagement, enhance productivity, and improve operational efficiency.

References

Zeynep Akşin, Sarang Deo, Jónas Oddur Jónasson, and Kamalini Ramdas. 2021. Learning from many: Partner exposure and team familiarity in fluid teams. *Management Science* 67, 2 (Feb. 2021), 854–874.

Gad Allon, Maxime C Cohen, and Wichinpong Park Sinchaisri. 2023. The impact of behavioral and economic drivers on gig economy workers. *Manufacturing & Service Operations Management* 25, 4 (2023), 1376–1393.

Michel Jose Anzanello and Flavio Sanson Fogliatto. 2011. Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics* 41, 5 (2011), 573–583.

Linda Argote. 2012. *Organizational learning: Creating, retaining and transferring knowledge* (2 ed.). Springer, New York, NY.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. 2002. Finite-time Analysis of the Multi-Armed Bandit Problem. *Machine Learning* 47, 2-3 (2002), 235–256. https://link.springer.com/article/10.1023/A:1013689704352

Hessam Bavafa and Jónas Oddur Jónasson. 2021. The Variance Learning Curve. *Management Science* 67, 5 (May 2021), 3104–3116.

Alan Benson, Aaron Sojourner, and Akhmed Umyarov. 2018. Can Reputation Discipline the Gig Economy? Experimental Evidence from an Online Labor Market. In *Proceedings of the ACM Conference on Economics and Computation*. ACM, 41–60.

Djallel Bouneffouf and Raphaël Féraud. 2016. Multi-armed bandit problem with known trend. *Neurocomputing* 205 (2016), 16–21.

Colin Camerer and Teck Hua Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 4 (1999), 827–874.

O Chapelle and Lihong Li. 2011. An Empirical Evaluation of Thompson Sampling. In *Advances in Neural Information Processing Systems 24*. 2249–2257. https://papers.nips.cc/paper/2011/hash/0a912c511ef3a3b9b0c9ca0e6b0248a4-Abstract.html

Ming Chen and Myrto Sheldon. 2015. Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform. In *Proceedings of the ACM Conference on Economics and Computation*. ACM, 61–80.

Jonathan R Clark, Robert S Huckman, and Bradley R Staats. 2013. Learning from customers: Individual and organizational effects in outsourced radiological services. *Organization Science* 24, 5 (Oct. 2013), 1539–1557.

Hongyan Dai, Jayashankar M Swaminathan, and Yuqian Xu. 2022. Leveraging the Experience: Exploration and Exploitation in Gig Worker Learning Process. (May 2022).

Ezey M Dar-El. 2013. *HUMAN LEARNING: From learning curves to learning organizations* (2000 ed.). Springer, New York, NY.

Wai Fong Boh, Sandra A Slaughter, and J Alberto Espinosa. 2007. Learning from experience in software development: A multilevel analysis. *Management Science* 53, 8 (Aug. 2007), 1315–1331.

Rod Gerber. 1998. How do workers learn in their work? *The learning organization* 5, 4 (1998), 168–175.

Aspen Gorry. 2016. Experience and worker flows. *Quantitative Economics* 7, 1 (March 2016), 225–255.

Eric H Grosse and Christoph H Glock. 2015. The effect of worker learning on manual order picking processes. *Int. J. Prod. Econ.* 170 (Dec. 2015), 882–890.

Reeju Guha and Daniel Corsten. 2023. The Role of Within-Day Learning on Gig Workers' Performance and Task Allocation: Evidence from an On-demand Platform. *Available at SSRN* (2023).

Katharina Hermann, Wim Van Damme, George W Pariyo, Erik Schouten, Yibeltal Assefa, Anna Cirera, and William Massavon. 2009. Community health workers for ART in sub-Saharan Africa: learning from experience–capitalizing on new opportunities. *Human resources for health* 7 (2009), 1–11.

Tony Ho, Adam Slivkins, and Jonathan Vaughan. 2014. Adaptive Contract Design for Crowdsourcing Markets: Bandit Algorithms for Repeated Principal-Agent Problems. In *Proceedings of the ACM Conference on Economics and Computation*. ACM, 21–40.

Po-Sen Huang et al. 2020. TabTransformer: Tabular Data Modeling Using Contextual Embeddings. *arXiv preprint arXiv:2012.06678* (2020).

Rakesh Johari, Sachin Kamble, and Kalyan Kanoria. 2016. Matching while Learning. In *Proceedings of the ACM Conference on Economics and Computation*. ACM, 1–20.

Benjamin Knight, Dmitry Mitrofanov, and Serguei Netessine. 2022. The impact of ai technology on the productivity of gig economy workers. *Available at SSRN 4372368* (2022).

Tor Lattimore and Csaba Szepesvári. 2020. *Bandit Algorithms*. Cambridge University Press. https://tor-lattimore.com/downloads/book/book.pdf

Vili Lehdonvirta. 2018. Flexibility in the Gig Economy: Managing Time on Three Online Piecework Platforms. *New Technology, Work & Employment* 33, 1 (2018), 13–29.

Lihong Li, Wei Chu, John Langford, and Robert E Schapire. 2010. A Contextual-Bandit Approach to Personalized News Article Recommendation. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 661–670. https://dl.acm.org/doi/10.1145/1772690.1772753

Daniel McFadden. 1972. Conditional logit analysis of qualitative choice behavior. (1972).

Scott M Shafer, David A Nembhard, and Mustafa V Uzumeri. 2001. The Effects of Worker Learning, Forgetting, and Heterogeneity on Assembly Line Productivity. *Management Science* 47, 12 (Dec. 2001), 1639–1653.

Richard S Sutton and Andrew G Barto. 2018. *Reinforcement Learning: An Introduction*. MIT Press. http://incompleteideas.net/book/the-book.html

William R. Thompson. 1933. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* 25, 3-4 (1933), 285–294. https://academic.oup.com/biomet/article/25/3-4/285/233264

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All You Need. In *Advances in Neural Information Processing Systems*.

Jeffrey M Wooldridge. 2010. *Econometric analysis of cross section and panel data*. MIT press.

Zheng Yao, Silas Weden, Lea Emerlyn, Haiyi Zhu, and Robert E Kraut. 2021. Together but alone: Atomization and peer support among gig workers. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.

Dongruo Zhou, Lihong Li, and Quanquan Gu. 2020. Neural contextual bandits with ucb-based exploration. In *International Conference on Machine Learning*. PMLR, 11492–11502.

A Full Independent Variables in the MNL Model

- **LIST**: Indicates whether the order is in the recommendation list by the platform's algorithm (1) or it is in the common pool where everyone can take an order (0).
- **BUNDLED**: Indicates whether the order is part of a bundle (1) or not (0).
- **PREV_100_TOTAL_DOLLARS**: The total dollars earned from the shopper's previous 100 orders in this store. This variable indicates how exploratory shoppers are (whether they always choose stores they have usually visited or explore new stores). This is a variable mimicking reinforcement like [Camerer and Hua Ho, 1999], where we assume people learn from the rewards they receive from past interactions.
- **ORDER_TYPE_ID**: The type of order associated with the alternative (e.g., delivery or pickup).
- **MAX_CONTRIBUTIONS_CAT_PCT**: The maximum percentage of contributions from a specific item category in the alternative.
- **MILES_DISTANCE_STORE_CUST**: The distance (in miles) between the store and the customer's location.
- **REQUESTED_ITEMS**: The number of items requested in the order.
- **DOLLARS_BONUS**: The dollar amount of any bonuses offered for completing the order.
- **DOLLARS_PAY**: The total payment offered for completing the order, excluding bonuses.
- **LOCAL_DELIVERY_WINDOW**: The delivery time window of the order.
- **PCT_DAILY_NONFOOD_ITEMS**: The percentage of daily non-food items in the alternative.
- **PCT_EXPANDED_FOOD_ITEMS**: The percentage of expanded food items in the alternative.
- **PCT_GENERAL_MERCH_ITEMS**: The percentage of general merchandise items in the alternative.