Quantifying taxi drivers' behaviors with behavioral game theory

Mengyu JI
Singapore Management University, mengyuji.2017@phdis.smu.edu.sg

Yuhong XU
Singapore Management University, yuhongxu@smu.edu.sg

Shih-Fen CHENG
Singapore Management University, sfcheng@smu.edu.sg

Quantifying Taxi Drivers' Behaviors with Behavioral Game Theory

Mengyu Ji[1], Yuhong Xu[1], and Shih-Fen Cheng[1]

***Abstract*— With their flexibility and convenience, taxis play a vital role in urban transportation systems. Understanding how human drivers make decisions in a context of uncertainty and competition is crucial for taxi fleets that depend on drivers to provide their services. As part of this paper, we propose modeling taxi drivers' behaviors based on behavioral game theory. Based on real-world data, we demonstrate that the behavioral game theory model we select is superior to state-of-the-art baselines. These results provide a solid foundation for improving taxi fleet efficiency in the future.**

I. INTRODUCTION

In urban cities like Singapore, New York, and Hong Kong, taxis play a vital role in the public transportation system. They complement extensive transit networks by offering flexible, door-to-door mobility services. Moreover, the taxi industry generates a significant number of employment opportunities. Therefore, enhancing the efficiency of taxi operations is essential, as it not only improves overall transportation effectiveness but also contributes to the income growth of taxi drivers.

In the conventional taxi operation model, when a taxi is vacant, the taxi driver faces a series of decisions regarding the optimal directions to find the next passenger. Interestingly, significant variations in taxi operation efficiency are observed within the same city, wherein certain drivers adopt more effective 'rational' strategies, resulting in higher hourly salaries compared to others. To gain a deeper understanding of the diverse decision-making dynamics among taxi drivers, we have summarized their characteristics as follows:

- Agents (taxi drivers) compete for the same shared resources (passengers).
- Interactions among agents (taxi drivers) occur due to the existence of these common resources (passengers).
- Agents (taxi drivers) don't always act perfectly rational, either due to their inability to sense and collect necessary information for making decisions or limited real-time decision-making capabilities.

If we view taxi fleet operations as a group of rational agents (taxis) competing for common shared resources (passengers), we can utilize game-theoretic models to describe and predict how agents would behave. However, due to the limitation of human rationality, traditional game-theoretic models that assume fully rational agents cannot be used directly.

To address this, we look into behavioral game theory, in particular, the family of the Cognitive Hierarchy (CH) models [1], in which agents are assumed to have different "levels" of reasoning. In a CH model, an agent whose reasoning level is k can optimize against other agents, with the assumption that they are acting with reasoning levels that are from 0 to $k-1$ (the CH model assumes that a Poisson distribution is used to describe the distribution of levels). A level-0 agent does not consider the existence of other agents and makes decisions myopically. The current state-of-the-art CH model is called the Quantal Cognitive Hierarchy (QCH) model [2], in which the agents' optimal decisions are "quantal" (i.e., probabilistic and following a softmax function). The latest variant of the QCH model further relaxes the requirement that the opponents' reasoning levels follow a Poisson distribution, and instead uses a data-driven approach to "learn" the actual distribution of agents' reasoning levels. This approach is called the "Iterative Population Learning" (QCH-IPL) [3], and we will adopt this framework to model the interactions among taxi drivers.

An important part of the implementation of the QCH-IPL framework is to come up with an efficient approach to compute agents' optimal responses against a mixed population of opponents with varying reasoning levels. To achieve this, we follow the formulation and adopt the algorithm proposed by Varakantham et al. (2012) [4]. The model is based on the Markov Decision Process (MDP), where other agents' decisions are aggregated into a vector of state variables. The key to the scalability of this approach is to ensure that the size of the state is invariant to the number of agents.

In summary, we aim to make the following contributions in this paper:

- We utilize the Selfish Routing with Transition uncertainty (SRT) framework, as proposed by Varakantham et al. (2012) [4], to model the decision-making process of taxi drivers.
- We present the application of QCH-IPL for taxi drivers using both synthetically-generated data and real-world data from Singapore taxi drivers.
- We formally define the determination of the population reasoning level distribution as a fixed point-

[1] School of Computing and Information Systems, Singapore Management University, mengyuji.2017@phdis.smu.edu.sg, {yuhongxu, sfcheng}@smu.edu.sg

seeking problem within the population dynamics.

- By comparing QCH-IPL with other frameworks, such as QCH and Iterative Bayesian Inference (IBI), we demonstrate that our iterative process can efficiently converge to a more accurate and stable population reasoning level distribution.

II. Literature Review

The decision-making processes of taxi drivers have been a subject of interest in transportation research. One line of research focuses on understanding the spatial and temporal dynamics of taxi operations. For example, Guan et al. (2016) [5] analyzed taxi drivers' decision-making processes in response to changing demand patterns and traffic conditions. Similar studies have investigated the impact of factors like congestion, weather conditions, and economic variables on taxi driver decision-making (Liu et al., 2017 [6]; Wang et al., 2018 [7]).

Another area of research explores the role of competition among taxi drivers. As taxis compete for the same pool of passengers, drivers make decisions on routes, fares, and passenger selection to maximize their earnings. Zheng et al. (2018) [8] investigated the spatial competition among taxi drivers and proposed a model to analyze the effects of competition on their decision-making behaviors.

Furthermore, the bounded rationality of taxi drivers has been a subject of investigation. Models that capture 'limited iterative strategic thinking' include the Cognitive Hierarchy (CH) model [1], which assumes that players have varying levels of strategic thinking ability and that each player believes they understand the game better than other players. The CH model incorporates decision rules that reflect an iterative process of strategic thinking [9], [10]. Models that capture 'cost-proportional errors' include the Quantal Response Equilibrium (QRE) model, where better responses are more likely to be chosen than worse responses, but the best responses are not chosen with certainty [11]. Empirical evidence has shown that CH and QRE models often fit better than random or Nash models in many game situations.

Traditionally, player reasoning levels have been assumed to be fixed. However, recent research by Ho et al. (2021) [12] introduces adaptive learning for level 0 players and sophisticated learning for higher-level players. In our work, we take a different approach by assuming that agents' reasoning level follows a distribution.

While Bayesian methods have been widely used to estimate probability distributions over reasoning levels [12], [4], [13], [14], the traditional Bayesian approach tends to perform poorly in Markov games with complex strategy spaces. To efficiently estimate the population reasoning level distribution, we propose to use the Iterative Population Learning (IPL) method, which allows us to iteratively update and refine the distribution of reasoning levels based on observed behaviors and decision histories. This contributes to the existing literature on using fixed point-seeking mechanisms to estimate population reasoning level distribution in the context of the Markov decision process.

III. Taxi Driver's Decision-Making Problem

Our research aims to address the issue of rationality levels among urban taxi drivers by applying behavioral models to their decision-making processes. By categorizing taxi drivers into different groups based on their rationality levels, we investigate their strategies for passenger-seeking during their free time.

Let's define the n-player game for urban taxi drivers. In this game, we have a set of players (drivers) denoted as N, where $N = \{1, ..., n\}$. The time horizon is represented by $\mathcal{H}$, with $\mathcal{H} = \{0, ..., T-1\}$. The city is divided into various zones, forming the set of states denoted as $\mathcal{S}$, where $\mathcal{S} = \{1, ..., S\}$. Additionally, we define the set of reasoning levels as $\mathcal{L}$, with $\mathcal{L} = \{1, ..., L\}$. For each player (driver) $i \in N$, in each state $s \in \mathcal{S}$ and at each time period $t \in \mathcal{H}$, while using reasoning level $l \in \mathcal{L}$, they have a strategy set denoted as $A_{ilst} = \{a_{ilst1}, ..., a_{ilstJ_s}\}$. Here, J_s represents the number of pure strategies in the strategy set A_{ilst}.We define the action space as $\mathcal{A} = \prod_{i\in N}\prod_{l\in\mathcal{L}}\prod_{s\in\mathcal{S}}\prod_{t\in\mathcal{H}} A_{ilst}$.

To capture the drivers' decision-making, we introduce the set Δ_{ilst}, which represents the set of probability measures on A_{ilst}. Elements of Δ_{ilst} can be expressed as $p_{ilst} : A_{ilst} \to \mathbb{R}$, where $\sum_{a_{ilstj}\in A_{ilst}} p_{ilst}(a_{ilstj}) = 1$, and $p_{ilst}(a_{ilstj}) \geq 0$ for all $a_{ilstj} \in A_{ilst}$. We denote $p_{ilstj} = p_{ilst}(a_{ilstj})$, making Δ_{ilst} isomorphic to the J_s-dimensional simplex. Mathematically, $\Delta_{ilst} = p_{ilst} = (p_{ilst1}, ..., p_{ilstJ_s}) : \sum_j p_{ilstj} = 1, p_{ilstj} \geq 0$.

IV. Behavioral Models for Drivers

Driver's behavioral model draws from both transportation research and game theory, which are detailed next.

A. Markov Games

The Markov game framework is a valuable tool for modeling the decision-making process of multi-agent games, particularly in the presence of transitional uncertainty. A Markov game is defined by the tuple $< \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{H} >$, where $\mathcal{S}$ represents the set of states that agents (taxi drivers) can be in, $\mathcal{A}$ represents the set of actions or decisions that can be taken by agents in each state. $\mathcal{H}$ represents the time horizon. $\mathcal{T} : \mathcal{S} \times \mathcal{A} \times \mathcal{H} \to \Delta(\mathcal{S})$ is the transition probability function, which determines the probability of transitioning from one state to another when taking a specific action at a given time. It satisfies the constraint $\sum_{s'} \mathcal{T}^t(s, a, s') = 1, \forall s, a, t$. $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{H} \to \mathbb{R}$ represents the reward function, which assigns a real value to each state-action pair at each time step.

Various solution techniques exist for the Markov game [15], [16]. In this paper, we employ Backward Induction. Our objective is to obtain a policy $\pi : \mathcal{S} \times \mathcal{H} \rightarrow \Delta(\mathcal{A})$, representing the quantal best response (QBR) derived from the expected utilities.

B. Nash Equilibrium and the CH Model

Nash equilibrium, as described by Fudenberg and Tirole [17], helps predict how agents would collectively behave. However, Nash equilibrium assumes agents are fully rational, and when human decision-makers, who are at best partially rational, are involved in a game, it provides poor outcome predictions [18]. To address this limitation, the 'cognitive hierarchy' (CH) framework proposed by Camerer et al. (2004) [1], aims to explicitly model the limited rationality of human agents and allows for the explicit specification of different levels of rationality among agents in a game.

C. The QCH Model

In this study, we extend the CH model by incorporating QBR, enabling multiple levels of reasoning and probabilistic strategies. Denote payoff function for each $i \in N$ in each state $s \in \mathcal{S}$ and each period $t \in \mathcal{H}$ when using reasoning level $l \in \mathcal{L}$ as r_{ilst}, $r_{ilst}(a_{ilstj}, \pi^{-i})$ represents the reward obtained by agent i on taking jth strategy given opponets' joint strategy π^{-i}. The QBR for agent i specifies the probability for playing each strategy and is given by:

$$QBR_i(\pi^{-i}) = \pi_i,$$
$$\text{where } \pi_i(a_{ilstj}) = \frac{e^{\lambda r_{ilst}(a_{ilstj}, \pi^{-i})}}{\sum_{a' \in A_{ilst}} e^{\lambda r_{ilst}(a', \pi^{-i})}} \quad (1)$$

Given the QCH model, the probabilistic strategy for an agent i taking each strategy, a_{ilstj} is defined as follows:

$$\pi_i(a_{i0stj}) = \frac{1}{|A_{ilst}|},$$
$$\pi_i(a_{ilstj}) = QBR_i(\lambda, \hat{\Delta}_l, \pi_0^{-i}, \pi_1^{-i} ... \pi_{(l-1)}^{-i}) \quad (2)$$

where λ represents the precision or sensitivity of an agent to the actual utility value. When it approaches 0, agents choose strategies uniformly; as it approaches infinity ($\lambda \rightarrow \infty$), agents tend to choose the exact best responses [19]. $\hat{\Delta}_l$ is the estimator of the normalized proportion of other agents for level l agents. Players with higher levels of reasoning are more likely to make better decisions since they have a better estimate of the actual agent distribution. We enhance the CH model by allowing players to make mistakes, enabling players to make decisions based on QBR's probabilistic strategies.

V. Iterative Population Learning (IPL)

This section provides a comprehensive overview of the IPL method for estimating the population's reasoning level distribution. We will provide detailed insights and prove the existence of fixed-point. Some useful notations used in the algorithm are presented in the table below:

TABLE I
NOTATION USED IN ALGORITHM 1.

	Description
l_{max}	Maximum level of reasoning.
w	Moving window used in IPL.
π_l^e	Level-l policy in iteration e.
$d_h^t(s)$	Taxi driver distribution for level h while at time t and state s.
$\vec{\boldsymbol{d}}_e^0$	Agents distribution from $t = 0$ to $T - 1$ in iteration e.
Δ_l	Proportion of level-l agents in the population, $\Delta(l) = \frac{\sum_i Pr_i(l)}{N}$.
$\Delta_l(h)$	Proportion of level-h agents in the population as computed by a level-l agent, $\Delta_l(h) = \frac{\Delta(h)}{\sum_{k=0}^{l-1} \Delta(k)}$, $\forall h < l$.

Algorithm 1 Iterative Population Learning

$e \leftarrow 0$
$\Delta_l = Uniform(0, l_{max}), \forall l$
$\overline{\Delta_l^e}$: Δ_l at iteration e
$\overline{\Delta_l^e} = mean(\Delta_l^{e-w}, ..., \Delta_l^{e-1}), \forall l$
$\delta \leftarrow$ a very small number
$\boldsymbol{d}^0 \leftarrow UniformDistribution$
while $|\overline{\Delta_l^e} - \Delta_l^e| > \delta, \forall l$, **or** $e = 0$ **do**
  $\pi_0 = GetLevel0Strategy()\{uniform\}$
  for $\forall l \leq l_{max}$ **do**
    $\vec{\boldsymbol{d}}_e^0 = GetDist(T, \overline{\Delta_l^e}, \boldsymbol{d}^0, < \pi_0, .., \pi_{l-1} >)$
    $\pi_l^e = GetQBR(\lambda, T, \vec{\boldsymbol{d}}_e^0)$
  end for
  $\Delta_l^e = CLR(\pi_0^e, ..., \pi_{l_{max}}^e, transition), \forall l$
  $e = e + 1$
end while

A. Supply Distribution Simulation

The supply distribution function takes the following form:

$$\begin{aligned}
&GetDist(T, \Delta_l, \boldsymbol{d}^0, \{\pi_0, ..., \pi_{l-1}\}) \\
&\quad d_h^0(s) = \Delta_l(h) \cdot d^0(s) \\
&\quad d_h^c(s) = \sum_{s', a} \pi_h(s', a) \cdot d_h^{c-1}(s') \cdot \mathcal{T}_h^t(s', a, s, \boldsymbol{d}^{c-1}), \\
&\quad \forall\, c \in [1, T-1] \\
&\quad d^c(s) = \sum_{h<l} d_h^c(s), \forall\, c \in [1, T-1]
\end{aligned} \quad (3)$$

The simulation process begins with the first formula, using the input $\Delta_l(h)$ and the initial supply distribution $\boldsymbol{d}^0$ (which is uniformly distributed). This lets us obtain the supply distribution of level h computed by level l agents. Next, we simulate the supply distribution over the time horizons based on the corresponding level policies and transition probabilities. Finally, in the third formula, we sum up the expected distribution of all lower levels. The complete process will be run several times until the generated distribution is stable.

B. Constrained Linear Regression (CLR)

In Algorithm 1, we propose the use of CLR to estimate the reasoning level distribution of taxi drivers based on aggregated transitions. CLR is a variant of regular linear regression that offers several advantages, including its intuitive nature, simplicity, and low data requirement. In CLR, the response variable is the aggregated action frequency. The independent variables are the QBRs at different reasoning levels. By fitting the actual observations to the predictions from all the reasoning levels, we can estimate the parameters that represent the distribution of reasoning across these levels.

The CLR incorporates two additional constraints: the first constraint ensures that the sum of the parameters equals 1, reflecting the fact that the reasoning level distribution is a probability distribution. The second constraint enforces the non-negativity of the parameters. The CLR can be represented by the following formulas:

$$\min \sum_{a_0 \in A_0} \left(y(a_0) - \sum_{l \in L} \alpha_l \pi_l(a_0) \right)^2 \tag{4}$$

s.t.

$$\sum_{l} \alpha_l = 1, \alpha_l \geq 0, \forall l \in [0, l_{max}].$$

C. Fixed Point of the QCH-IPL Process

Algorithm 1 decomposes the QCH-IPL process into two main parts. First is the 'level strategy generation' where strategies for each level are generated based on the population reasoning level distribution. Second is the 'population distribution' part, which identifies the population reasoning level distribution based on the level strategies.

We can combine all the steps in Algorithm 1 functionally into a single composite function. Define $\Phi(\cdot)$ as the CLR in line 13 of Algorithm 1. Further define Ψ as the computation of QBR (π_0, …, $\pi_{l_{max}}$) in line 11 of Algorithm 1. We can examine our QCH-IPL approach from a mathematical perspective:

$$(\pi_0, \ldots, \pi_{l_{max}}) = \Psi\Big(\Phi\Big(\big((\pi_0, \ldots, \pi_{l_{max}}), tr\big)\Big)\Big) \tag{5}$$

Equation 5 is a strategy and distribution combined iterative process and $(\pi_0, \ldots, \pi_{l_{max}})$ is the fixed point of this iterative process. The IPL process is seeking a vector $(\pi_0, \ldots, \pi_{l_{max}})$, which would be returned as output when given as input.

Theorem 1 *A fixed point exists in the above policy and distribution combined iterative process [eq.(5)] (Theorem 1 from [3])*

proof: Firstly, we represent the set of strategies in topology form. Next, we establish the compactness and convexity properties of the strategy set. We also establish the continuity property of the strategy and reasoning level distribution combined iteration $\Psi(\Phi(\cdot, tr))$. To complete the proof, we can apply *Schauder-Tychonoff fixed point theorem*. More details can be found in [3]. To make sure the successive approximation executed by the loop algorithm converges to a fixed point from any starting point, and this fixed point is unique, we can establish contraction property.

VI. Numerical Studies

To prove that our method can be applied in games with more complex strategy space, we apply QCH-IPL to study the reasoning level of taxi drivers. Motivated by Varakantham et al. (2012) [4], we incorporate the SRT framework. We assume that taxi drivers compete for passengers across Singapore, and each driver faces uncertainty when deciding to move from their current location. In QBR, we employ the same transition and reward functions as [4] functions (4) and (5) shows. And all symbols we use in the following are also from the SRT framework in [4] section 4.

A. Data Extraction

Under our defined settings, the variable Γ represents levels of reasoning, ranging from 0 to 3. To form the transition ($\mathcal{T}$) and reward ($\mathcal{R}$) functions used for computing QBRs, we require four essential types of data: passenger flow data, taxi drivers distribution data, fare data, and costs data. Here is an explanation of how each type of data is extracted and utilized:

- Passenger Flow Data: We utilize the taxi trip dataset from 2009 to 2012 and find the number of trips within each hour between zones i and j.
- Taxi Drivers Distribution Data: We set the total number of taxi drivers and assign them a uniform distribution across the zones. This distribution is then incorporated into the Supply Simulation.
- Fare Data: We apply the fare charged for each trip to get the normalized value over all tuples of origins and destinations.
- Costs Data: We utilize the Google API to extract the distance between any two zones. Then we calculate the gasoline costs of the trip distance.

With the aforementioned data and a Backward Induction framework, we can successfully compute QBRs. However, we also require observations of taxi drivers' transition behavior at free status to run CLR. Here's a breakdown of getting transition data:

1. Time Slicing: We divide the time into 20-minute intervals.
2. Origin Selection: For each 20-minute interval, we identify the zone ID with the longest duration in free status as the driver's origin zone.
3. Transition Probability Distribution: The free status duration over zones in the next 20-minute interval represents the drivers' transition probability distribution.
4. Removal of Transition Between 6 am and 8 am: Drivers tend to wander around their homes during this period.
5. Aggregation: To increase the number of transition observations, we aggregate the 20-minute transitions into 1-hour intervals from 8 am to 5 pm each day over the course of three years.

B. Synthetic Experiments

Before applying the QCH-IPL algorithm to real-world taxi drivers, we compare it with two baseline methods: Iterative Bayesian Inference (IBI) and the QCH method with synthetically generated movement data. IBI is based on iteratively Bayesian inference. It applies a Bayesian update (as formula 6 shows) iteratively for each driver's movement.

$$Pr_i(l|a_{t,i}) = \frac{Pr_i(a_{t,i}|l) \times Pr_i(l)}{\sum_k^{l_{max}} Pr_i(a_{t,i}|k) \times Pr_i(k)} \quad (6)$$

To evaluate the performance of QCH-IPL, we compare its results with those of the IBI and QCH baselines. We focus on population reasoning levels. In the case of IBI, to obtain the proportion of drivers belonging to different levels in the population, we aggregate the probability distributions $Pr_i(l)$ for all drivers i and all levels l. The parameters used in our synthetic experiments are summarized below:

- supply = 18000,
- discount factor = 0.8,
- $\lambda = \{1, 2, 3, 4\}$,
- number of zones = 145,
- number of horizons = 9,
- moving window size = 1.

The initial population level distribution is generated using a truncated Poisson distribution. The parameter τ is varied from 0.1 to 2.0 with an increment of 0.1. The focus of the experiments is on population levels of reasoning, and the CLR (Constrained Least Squares) method is solved using the *Gurobi* optimization solver.

Three ground truth population reasoning level distributions are assumed for evaluation purposes. These distributions are: (1). $[0.3, 0.4, 0.3, 0]$; (2). $[0, 0.3, 0.4, 0.3]$; (3). $[0.25, 0.15, 0.4, 0.2]$. Besides simulated transition, all other input data used in the experiments, are real-world datasets mentioned in the previous part.

As mentioned in step 3, the performance of different methods is evaluated using several metrics. The results, summarized in Table II, indicate the following observations:

- Compared with the IBI baseline, the CLR approach employed in QCH-IPL outperforms the Bayesian update process by a very large percentage. For example, for ground truth (3), QCH-IPL can improve EMD by around 76% and KL-Divergence by around 98%.
- Compared with the QCH baseline, the QCH-IPL method demonstrates improved performance, reducing the EMD by 17% to 68%, KLD by 30% to 93%. This suggests that the IPL aids in better inferring the distribution of reasoning levels.

TABLE II

QCH-IPL METRICS COMPARED WITH QCH AND IBI

Ground Truth	(1)	(2)	(3)
QCH-IPL Model			
W-distance (EMD)	0.0433	0.0168	0.0489
KL-Divergence	0.0608	0.0028	0.0443
Mean Squared Error	0.0146	0.0151	0.0144
QCH Model			
W-distance (EMD)	0.0526	0.0524	0.0815
KL-Divergence	0.0873	0.0422	0.1126
Mean Squared Error	0.0149	0.0153	0.0147
IBI Model			
W-distance (EMD)	0.2481	0.1246	0.2002
KL-Divergence	2.3482	0.7182	1.8854

TABLE III

QCH-IPL EMPIRICAL PERFORMANCE COMPARED WITH QCH

τ value	0.1	0.2	average overall τ
QCH-IPL Model			
KLD	26.20	26.20	26.20
EMD	0.0555	0.0555	0.0555
MSE	0.0196	0.0196	0.0196
MAE	0.0681	0.0681	0.0681
RMSE	0.1292	0.1292	0.1292
log-likelihood	-37220	-37220	-37220
QCH Model			
KLD	27.29	26.98	26.86
EMD	0.0575	0.0572	0.0570
MSE	0.0212	0.0208	0.0206
MAE	0.0701	0.0696	0.0693
RMSE	0.1336	0.1326	0.1318
log-likelihood	-37840	-37660	-37590

C. Empirical Experiments

After conducting the synthetic experiments and obtaining promising results, the next step is to apply the QCH-IPL method to the empirical dataset. The following additional procedures are performed for the empirical dataset:

First of all, it is 'filtering day-shift working drivers'. For each month from 2009 to 2012, the list of taxi driver IDs who work 80% of their time during the 8 am to 5 pm period is identified. Secondly, we need to do 'aggregated transition filtering'. Only tuples of (t, s) that have observations exceeding the lower bound are included in the CLR process.

The next important step is to tune the parameters to obtain optimal results. The parameters include supply, discount factor, lower bound for transition data, and the parameter lambda (λ) that influences QBR generation. In this case, the supply parameter is set to 18,000, and the moving window in the iteration is set to 3. QCH-IPL is then run with different values for the discount factor (0.7 and 0.8), lower bound (100 and 150), and λ values ranging from 3 to 10.

Finally, the chosen scenario is $\lambda = 10$, discount factor = 0.8, and lower bound = 150. By running QCH-IPL with these selected parameters, IPL converges within 15 iterations. The final population-level distribution obtained is $[0.2105, 0.2347, 0.4281, 0.1266]$, representing the distribution of drivers across the reasoning levels.

In order to compare the empirical results of QCH-IPL with QCH, we conducted a comparative analysis. For QCH, we set $\lambda = 10$ to match the performance of QCH-IPL. Additionally, we generated a truncated Poisson distribution for τ values ranging from 0.1 to 0.4 with an increment of 0.1. To evaluate the performance of QCH-IPL and QCH, we used metrics such as EMD, KL-Divergence, MSE, MAE, RMSE, and log-likelihood. The comparison was made by calculating the mean improvement overall τ values. Table III presents the value of the mean metrics from $\tau = 0.1$ to 0.4, as well as the values for selected τ values.

Upon comparing the results, it is evident that the QCH-IPL approach outperforms QCH across all evaluation metrics. Specifically, QCH-IPL exhibits notable improvements, such as a 4% to 8% reduction in MSE, and a 2% to 4% decrease in EMD and KL-Divergence.

VII. Conclusions

This study examines the decision-making process of taxi drivers in urban transportation systems as a whole. In order to gain insights into drivers' rationality during the passenger-seeking phase, we adopt the QCH-IPL approach that combines QBR strategies with IPL. We demonstrate the effectiveness of our approach by comparing it against other frameworks and analyzing both synthetic and real-world data. This current research lays the foundation for further research on how to improve the efficiency of taxi fleets.

Acknowledgement

This research is supported by the Ministry of Education, Singapore, under its Social Science Research Thematic Grant (Grant Number MOE2020-SSRTG-018).

References

[1] C. F. Camerer, T.-H. Ho, and J.-K. Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, vol. 119, no. 3, pp. 861–898, 2004.

[2] J. R. Wright and K. Leyton-Brown, "Predicting human behavior in unrepeated, simultaneous-move games," *Games and Economic Behavior*, vol. 106, pp. 16–37, 2017.

[3] Y. Xu, S.-F. Cheng, and X. Chen, "Improving quantal cognitive hierarchy model through iterative population learning," in *Twenty-Second International Conference on Autonomous Agents and Multiagent Systems*, 2023, pp. 2718–2720.

[4] P. Varakantham, S.-F. Cheng, G. Gordon, and A. Ahmed, "Decision support for agent populations in uncertain and congested environments," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012, pp. 44–53.

[5] H. Guan, Z.-C. Li, and L. Yu, "Taxi drivers' decision-making processes under real-time traffic information," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 431–446, 2016.

[6] R. Liu, D. Yang, J. Xu, and H. Huang, "Impact of weather conditions on taxi demand: A spatiotemporal analysis in shanghai," *Journal of Transport Geography*, vol. 62, pp. 23–31, 2017.

[7] D. Wang, J. Chen, and D. Sun, "The influence of economic variables on taxi travel demand: A panel data analysis in china," *Transport Policy*, vol. 63, pp. 1–11, 2018.

[8] S. Zheng, L. Yao, X. Zhang, and Z. Ma, "Spatial competition in taxi markets: Evidence from shenzhen, china," *Transportation Research Part A: Policy and Practice*, vol. 111, pp. 142–155, 2018.

[9] K. Binmore, "Modeling rational players: Part ii," *Economics & Philosophy*, vol. 4, no. 1, pp. 9–55, 1988.

[10] R. Selten, "Features of experimentally observed bounded rationality," *European Economic Review*, vol. 42, no. 3-5, pp. 413–436, 1998.

[11] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, no. 1, pp. 6–38, 1995.

[12] T.-H. Ho, S.-E. Park, and X. Su, "A Bayesian level-k model in n-person games," *Management Science*, vol. 67, no. 3, pp. 1622–1638, 2021.

[13] A. Ahmed, P. Varakantham, and S.-F. Cheng, "Uncertain congestion games with assorted human agent populations," in *Twenty-Eighth Conference on Uncertainty in Artificial Intelligence*, vol. 26, no. 1, 2012, pp. 1471–1477.

[14] J. R. Wright and K. Leyton-Brown, "Behavioral game theoretic models: a bayesian framework for parameter analysis." in *Eleventh International Conference on Autonomous Agents and Multiagent Systems*, 2012, pp. 921–930.

[15] R. A. Howard, *Dynamic programming and Markov processes.* John Wiley, 1960.

[16] A. G. Barto, S. J. Bradtke, and S. P. Singh, "Learning to act using real-time dynamic programming," *Artificial Intelligence*, vol. 72, no. 1-2, pp. 81–138, 1995.

[17] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.

[18] J. K. Goeree and C. A. Holt, "Ten little treasures of game theory and ten intuitive contradictions," *American Economic Review*, vol. 91, no. 5, pp. 1402–1422, 2001.

[19] J. R. Wright and K. Leyton-Brown, "Beyond equilibrium: Predicting human behavior in normal-form games," in *Twenty-Fourth AAAI Conference on Artificial Intelligence*, 2010.