Towards Fairness in Transportation Gig Markets: Identifying, Imitating, and Mitigating Algorithm Discrimination via Deep Reinforcement Learning

Zijian Zhao
Department of Civil and Environmental Engineering
The Hong Kong University of Science and Technology
Hong Kong, China
zzhaock@connect.ust.hk

Abstract

Recently, deep Reinforcement Learning (RL) methods have been widely used in labor management within transportation gig markets, such as ride-hailing, food delivery, and express delivery. Compared to traditional rule-based and optimization-based methods, RL can capture more information about long-term uncertainty and environmental dynamics, leading to better and non-myopic strategies. However, deep learning methods have long been criticized for their low interpretability, raising concerns about algorithmic discrimination in gig markets. Currently, most works focus on this issue from the perspective of statistical analysis and surveys. However, the underlying reasons related to the algorithms remain unclear, as most companies do not disclose their algorithms. This lack of transparency can hinder governments from designing efficient management policies to address these problems. To fill this research gap, this thesis proposal aims to develop appropriate RL methods to mimic the labor management behavior of transportation gig platforms and to propose effective policies that protect the rights of gig workers.

Research Goals

This thesis aims to address algorithmic discrimination in on-demand transportation gig platforms and propose effective regulatory policies to protect workers' rights. Our research objectives are threefold:

- **Identify and characterize algorithmic discrimination** in real-world platform data, focusing on order assignment and payment disparities among workers.
- **Develop a novel Reinforcement Learning (RL)-based simulation framework to model platform behavior** and reproduce discriminatory outcomes observed in real data.
- **Evaluate regulatory interventions**—such as Data Permission Rights (DPR)—using the simulation environment, and analyze their impact on stakeholders (platforms, workers, customers).

In the following research, we focus on the on-demand food delivery market, utilizing a public food delivery dataset from Meituan (Meituan 2024) and a private food delivery dataset from Hong Kong, China.

Previous Work

To begin with, we focus on the Meituan dataset for our analysis. This dataset provides information on when an order is assigned to a courier and whether the courier accepts this assignment. We observe significant differences in the order amounts assigned to and accepted by each courier during different time periods. However, the dataset does not include any private information about the couriers, limiting our analysis to behavioral factors, specifically the order rejection rate. Surprisingly, we find that couriers with a higher rejection rate tend to receive more assigned orders but accept a lower number of them, resulting in a lower overall workload. This order rejection rate may be related to the reservation value of each courier, which represents the minimum payment they are willing to accept for a unit of work. Under similar payment conditions, a higher reservation value can lead to a higher order rejection rate. In conclusion, we hypothesize that the reservation value is a crucial factor in algorithmic discrimination, with platforms likely preferring to assign more work to couriers with lower reservation values. This finding is intuitive, as couriers with lower reservation values generally have a lower rejection rate, which is advantageous for platforms seeking better management with reduced uncertainty. However, as most datasets do not disclose worker payments, the issue of payment discrimination remains unclear. Previous survey research indicates that some algorithms do adhere to the principle of "equal pay for equal work", based on feedback from workers (Griesbach et al. 2019; Zhu et al. 2024).

Then, we aim to explore the underlying reasons causing algorithmic discrimination and focus on developing a RL method to (i) replicate the order assignment discrimination observed in the Meituan dataset and (ii) investigate whether such a discriminatory algorithm can also lead to payment discrimination. Since the Meituan dataset has anonymized the location information, we construct the simulation environment using our private food delivery dataset in Hong Kong, China. This research presents two main challenges: (i) designing the state space to enable the agent to differentiate between couriers and (ii) developing an algorithm that simultaneously facilitates order assignment and payment setting. For the first challenge, we incorporate the historical behavior information of each courier into the state space. Specifically, we store the unit payments for orders that each courier accepted and rejected, which can reflect their reservation value to some extent while avoiding an excessively large state space. For the second challenge, we design a hybrid action space RL method, employing Double Deep Q-Networks (DDQN) (Van Hasselt, Guez, and Silver 2016) for order assignment and Proximal Policy Optimization (PPO) (Schulman et al. 2017) for payment setting. Compared to previous hybrid methods (Xu et al. 2023), our approach innovatively utilizes the Q-network of DDQN as the critic for PPO, fostering better cooperation between the two modules and enhancing training performance. Additionally, to address the order dispatch challenge, we propose two alternative methods (Zhao and Li 2025b,c), significantly enhancing efficiency. Through simulations using the real-world dataset, we find that our method successfully replicates the discrimination scenario observed in the Meituan dataset. Additionally, our findings reveal that couriers with higher reservation values can command higher unit payments, but this may be due to their tendency to reject lower-paying orders.

Based on the developed discriminatory algorithm, we further aim to explore how regulatory policies can protect the rights of couriers (Zhao and Li 2025a). We first focus on DPR, similar to the GDPR in Europe (Union 2018), and examine their detailed influence on stakeholders. Specifically, DPR allows couriers to decide whether to permit the platform to collect and utilize their behavioral data and enables them to change this decision at any time. We assume that couriers can choose whether to work on the platform and whether to provide their individual information at the beginning of each period based on their own preferences. Furthermore, we model the couriers' decision-making as a Contextual Multi-Armed Bandit (CMAB). To find the equilibrium between the platform and the couriers, we propose a novel two-stage training method. First, we train the platform to learn the optimal policy under various given courier policies. Then, we fix the platform policy to train the courier agent to identify their optimal policy. Interestingly, our experimental results reveal that DPR benefits all stakeholders. By allowing couriers to control their own data rights, more high-reservation couriers choose to join the platform without providing their individual information, thus gaining more work opportunities. The increased volume of active couriers enables the platform to serve more orders simultaneously, improving customer satisfaction by reducing delivery times. Additionally, we observe that more low-reservation couriers opt to work on the platform than high-reservation couriers, which aligns with our findings from the Meituan dataset. In conclusion, DPR creates a win-win scenario, enhancing the profits of all stakeholders.

Future Work

For my Ph.D. future work, I plan to focus on the following three areas:

Long-term Discrimination: In my previous work, we primarily addressed the short-term discrimination problem within individual time periods, such as 1 hour. However, some research has highlighted the importance of long-term discrimination scenarios. For instance, platforms may manipulate labor dynamics by providing more work opportunities to those who commit to working longer on the platform (Chen, Luo, and Yuan 2022). Conducting long-term simulations, however, requires more computational resources and time, presenting a significant research challenge.

Influence of Multiple Platforms: Current research often considers workers as being tied to a single platform. However, datasets like the Chicago Public Passenger Vehicle Chauffeur Survey (Portal 2021) indicate that workers can operate across multiple platforms simultaneously. It is important to investigate how this scenario might influence discrimination outcomes, as different platforms may adopt more worker-friendly strategies to attract labor. The primary challenge here will be identifying the main factors influencing workers' decisions about which platforms to engage with.

Adverse Relationships Between Platforms and Government: While my previous work explored the efficacy of regulatory policies like the DPR, it remains essential to examine how governments enforce supervision and whether there are potential loopholes that platforms could exploit to circumvent regulation.

References

Chen, Y.; Luo, Y.; and Yuan, Z. 2022. *Driving the drivers: algorithmic wage-setting in ride-hailing*. SSRN.

Griesbach, K.; Reich, A.; Elliott-Negri, L.; and Milkman, R. 2019. Algorithmic control in platform food delivery work. *Socius*, 5: 2378023119870041.

Meituan. 2024. Meituan INFORMS TSL Research Challenge.

Portal, C. D. 2021. Public Passenger Vehicle Chauffeur Survey - 2021.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Union, E. 2018. General Data Protection Regulation.

Van Hasselt, H.; Guez, A.; and Silver, D. 2016. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30.

Xu, Y.; Wei, Y.; Jiang, K.; Chen, L.; Wang, D.; and Deng, H. 2023. Action decoupled SAC reinforcement learning with discrete-continuous hybrid action spaces. *Neurocomputing*, 537: 141–151.

Zhao, Z.; and Li, S. 2025a. The impacts of data privacy regulations on food-delivery platforms. *Transportation Research Part C: Emerging Technologies*, 181: 105364.

Zhao, Z.; and Li, S. 2025b. One step is enough: Multi-agent reinforcement learning based on one-step policy optimization for order dispatch on ride-sharing platforms. *arXiv preprint arXiv:2507.15351*.

Zhao, Z.; and Li, S. 2025c. Triple-BERT: Do We Really Need MARL for Order Dispatch on Ride-Sharing Platforms? *arXiv preprint arXiv:2510.03257*.

Zhu, G.; Huang, J.; Lu, J.; Luo, Y.; and Zhu, T. 2024. Gig to the left, algorithms to the right: A case study of the dark sides in the gig economy. *Technological Forecasting and Social Change*, 199: 123018.