A MULTIOBJECTIVE REINFORCEMENT LEARNING FRAMEWORK FOR EQUITABLE TOLL DESIGN ON EXPRESS LANES

FINAL REPORT

JUNE 2024

AUTHORS

Ridwan Tiamiyu

Venktesh Pandey

Christian Bowens

North Carolina Agricultural and Technical State University

US DEPARTMENT OF TRANSPORTATION GRANT 69A3551747125

DISCLAIMER

1. Report No.	2. Government Accession No.	3. Recipient’s Catalog No.
4. Title and Subtitle A Multiobjective Reinforcement Learning Framework for Equitable Toll Design on Express Lanes		5. Report Date June 24, 2024
		6. Source Organization Code
7. Author(s) Tiamiyu, Ridwan; Pandey, Venktesh; Bowens, Christian		8. Source Organization Report No. CATM-2024-R8-NCAT
9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)
Center for Advanced Transportation Mobility Transportation Institute 1601 E. Market Street Greensboro, NC 27411		11. Contract or Grant No. 69A3551747125
12. Sponsoring Agency Name and Address University Transportation Centers Program (RDT-30) Office of the Secretary of Transportation–Research U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590-0001		13. Type of Report and Period Covered Final Report: Jun 2023– Dec 2023
		14. Sponsoring Agency Code USDOT/OST-R/CATM
15. Supplementary Notes:		

16. Abstract

Express lanes are commonly used to mitigate traffic congestion by providing reliable travel time in exchange for tolls. However, there is a lack of guidance on designing equitable discounts for low-income travelers, creating barriers to equitable transportation systems. The past study on "Equitable Dynamic Pricing for Express Lanes" addressed some of these gaps by providing guidance for differential tolls and analyzing unintended traffic patterns. However, the framework was limited for single-objective optimization of equity. Furthermore, limitations on open-source algorithms for equity optimization hinder accessibility for researchers and practitioners. This project develops a reinforcement-learning-based optimization of express lane discounts considering reinforcement learning frameworks and creates an open-source tool for making previous research findings more accessible. In this research, the team has (a) designed an open-source platform that integrates advances in reinforcement learning literature for designing discounts for express lanes, and (b) tested the usefulness of the designed tools across multiple datasets and development platforms. The research findings indicate that traffic operations for express lanes are multiobjective in nature and require intentional tradeoffs between conflicting objectives such as maximizing equitable benefits and/or improving revenue.

17. Key Words Express lanes, Equitable discounts, Congestion Pricing, Transportation Equity		18. Distribution Statement Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 44	22. Price …

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Express lanes offer reliable travel times for a toll, but equity and fairness issues have prompted arguments for discounts rather than mandatory tolls. Recent studies, like the I-405 analysis, show that low-income travelers use express lanes when needed. However, guidance on designing equitable discounts considering multiple criteria such revenue, total system travel time, and equity is lacking. Lately, deep reinforcement learning (Deep-RL) algorithms have been used to learn the traffic control strategies in the areas of express lane pricing, signal control, and the control of CAVs in mixed autonomy conditions; however, the potential of such algorithms for equity is less explored. The PI's past study on "Equitable Dynamic Pricing for Express Lanes" addressed the gap by providing guidance for differential tolls and analyzing unintended traffic patterns. However, the framework was limited for single-objective optimization of equity.

This report conducts a multiobjective analysis in designing tolls and discounts for equitable benefits. In particular, the report analyzes and implements the latest Deep-RL methods including soft-actor-critic (SAC) and proximal policy optimization (PPO) towards objectives such as minimizing total system travel time (TSTT), maximizing equity metrics (or minimizing inequality measures), maximizing revenue, and increasing the corridor throughput. Our analysis reveals several key findings. For minimizing TSTT, most Deep-RL algorithms performed similarly with the SAC generally outperforming PPO across various objectives in finding toll policies with higher revenues. Revenue was generally higher when no discounts were offered, and the equity gap increased as travelers with lower value of time (VOT) were worse off. When discounts were allowed, the SAC algorithm provided uniform TSTT across VOTs, resulting in an equitable performance. Enforcing discounts reduced revenue, with 25% discounts maintaining a better balance of travelers between toll and non-toll lanes compared to 50% discounts, which encourage too many travelers to use the toll lanes, thereby violating the minimum speed limit constraint. The price of fairness, calculated as the revenue sacrificed to achieve the most equitable tolling, was found to be 91.4%.

The findings from the study suggest that while maximizing revenue and efficiency in managed lanes is a priority for investors, ensuring equitable access is crucial, and implementing personalized tolling can maintain high revenue while reducing travel times for low-income travelers. While the price of fairness was higher in our experimental studies, we recommend agencies conduct sensitivity analyses for “price of fairness” under different discount levels for effective weighing of equity relative to other priorities of the agency. As a limitation of the mesoscopic simulation-based analysis, the study findings can be improved through inclusion of departure time choice decisions and modeling the complexities of lane change behavior using microsimulation.

BACKGROUND

Price managed lanes provide reliable travel times in exchange for a toll and are increasingly being considered across the United States as a means for addressing congestion in a supply-constrained urban infrastructure. With the widespread adoption of managed lanes, equity concerns emerge as it relates to the needs of all travelers. Equity in transportation refers to a fair distribution of costs and benefits associated with a public project (like transportation infrastructure) across all groups and communities, especially such that some groups are not (AbuLibdeh, 2017).

Potential equity concerns emerge from express lanes where economically disadvantaged travelers are left worse off due to express lanes. A few analyses have emerged in this space over the last few years such as Hall (2018), Tan and Gao (2018), Twaddell and Zgoda (2020), Debreczeni (2021), and Xie et al. (2024). However, there is a lack of studies on the design of equitable discounts for low-income travelers, creating barriers to equitable transportation systems. The past study on "Equitable Dynamic Pricing for Express Lanes" (report #CATM-2022-R9-NCAT) addressed some of these gaps by providing guidance for differential tolls and analyzing unintended traffic patterns. However, the framework was limited for single-objective optimization of equity. Furthermore, limitations on open-source algorithms for equity optimization hinder accessibility for researchers and practitioners.

The goal of this implementation-focused research is to develop a multiobjective reinforcement-learning-based optimization of express lane discounts and create an open-source tool for making previous research findings more accessible. Such multiobjective considerations are critical for managed lanes and transportation systems as a whole where agencies are required to consider various goals and objectives for meeting the needs of users and stakeholders. For example, transportation agencies desire traffic control policies that result in best congestion mitigation, without compromising the throughput, safety, and reliability, while generating sufficient revenue and/or benefit to cost ratio for the expected investments. Furthermore, the increasing emphasis on equitable access to transportation systems (Gordon, 2021), and designing resilient infrastructure systems have necessitated careful consideration of the impacts of traffic controls across various metrics and key performance indicators (KPI), which may sometimes conflict with each other (for example, enhancing efficiency on a corridor might warrant compromising the equitable access for all travel groups).

Lately, deep reinforcement learning (Deep-RL) algorithms have been used to learn the traffic control strategies in the areas of signal control and the control of CAVs in mixed autonomy conditions (Peng et al., 2021). While Deep-RL algorithms can be extended for other traffic control applications (Wei et al., 2021), the interpretability of the model and optimization of multiple objectives (such as increasing revenue while enforcing equitable tolls) together remain challenging and have been left unaddressed.

In this research project, three stated research gaps are addressed: (a) multiobjectivity in toll design addresses the tradeoffs between objectives like maximizing revenue and improving fairness, (b) the use of reinforcement learning methods in traffic management further addresses the knowledge gap in implementation of Artificial Intelligence methods for congestion pricing, and (c) creating a framework for an open-source tool addresses the gap of designing open-source algorithms for equity optimization to promote accessibility for researchers and practitioners.

DESCRIPTION OF PROBLEM

This section outlines the multiobjective reinforcement learning problem being studied in this report. Figure 1 shows the different components involved as part of the toll optimization problem. Each component represents a key aspect of the transportation system and contributes to the formulation of a comprehensive model:

1. <u>Traveler Choice Models</u>: These models simulate the decisions travelers make, such as which lane to use, which route to take, and when to depart. Lane choice can be influenced by factors such as the perceived value of time (VOT), which is often modeled using a value distribution or a binary logit model that predicts the

probability of a traveler choosing a particular lane based on the associated cost and time savings.

2. Traffic Flow Model: This is the heart of the traffic simulation, capturing the movement and interaction of vehicles within the lanes, the effects of changes in lane speeds, and the phenomena of queue spillback, where congestion in one area spills over and affects upstream traffic. This model can be microscopic, focusing on individual vehicles and their interactions, or mesoscopic, dealing with aggregate flows of traffic.
3. Demand Model: This component predicts the demand for lane use, which can be modeled as either deterministic, with a fixed number of vehicles expected to use the lanes, or stochastic, with variability in usage based on different factors. The demand can be measured in real-time or estimated based on historical data to understand typical traffic patterns.
4. Toll Pricing Model (Reinforcement Learning Framework): A Deep-RL framework is used to optimize toll pricing. The objective is to minimize the total system travel time (TSTT), maximize revenue, or achieve other specified goals. Constraints such as the minimum and maximum toll rates and minimum speed limits on express lanes ensure that the model's solutions are practical and adhere to policy requirements. The RL framework would learn from the traffic and demand data, adjusting tolls in response to the dynamics of traveler choices and traffic flow to meet its objectives.

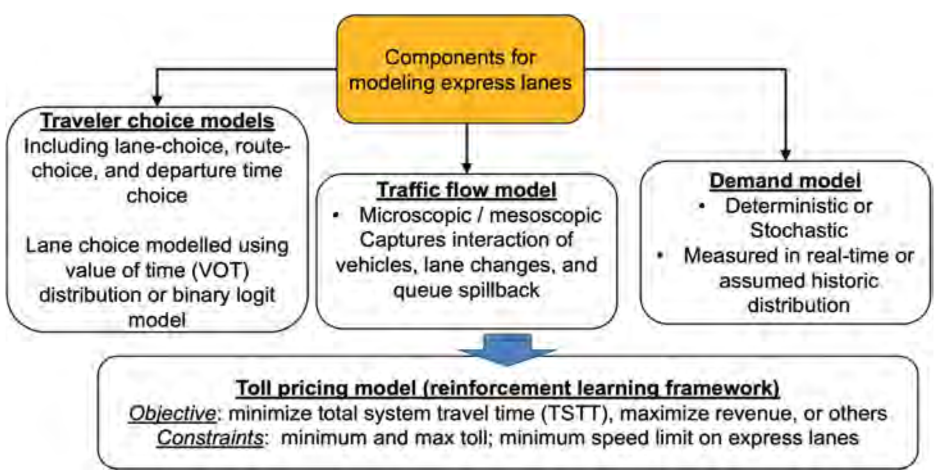


Figure 1 Component models for express lanes

Next, we discuss the multiobjective nature of the problem. Even from an algorithmic point of view many approaches have been proposed for handling multiple objectives in pricing. The study of optimizing toll charges for express lanes encompasses a diverse array of methodologies, including dynamic programming (Wang et al., 2012), model predictive control (Tan and Gao, 2018), optimization grounded in dynamic traffic assignment (Zhang et al., 2018), feedback integral control, analytical calculations based on a single-bottleneck model (Hall, 2018), and deep-reinforcement learning (Pandey et al., 2020). A comprehensive review of these models, as presented by Lombardi et al. (2021), emphasizes the intricate and variable nature of tolling operations. Despite these advancements, there remains a notable gap in the research concerning the equitable distribution of benefits from these tolling designs.

Hall's 2018 study delved into the practical implications of achieving a Pareto improvement by assessing the distributional and overall impacts of congestion pricing. This investigation revealed the feasibility of attaining Pareto efficiency through the implementation of tolls on select lanes, even prior to reallocating the generated revenue for public welfare. This scenario predominantly favors individuals with a higher value of time and greater flexibility. While these economic evaluations offer valuable insights, the applicability of these findings to networks with multiple access points and varying tolling objectives remains ambiguous.

Building on the gaps in the literature, **the key research questions addressed in this report are**:

- How can we integrate multiobjective optimization techniques (Hayes et al., 2022) within reinforcement learning for traffic operations (building off latest RL algorithms)?
- How can we interpret the findings from such techniques and validate the performance of the open-source code on different test networks?
- How can we incorporate equity in the tolling designs?

LITERATURE REVIEW

Overview of Reinforcement Learning

Reinforcement Learning (RL) is a category of machine learning algorithm where an agent learns to make decisions through trial and error. This learning process involves interacting with an environment and adapting based on feedback, similar to the way humans learn from their experiences (Sutton and Barto, 2018). RL is distinguished from other types of machine learning by its focus on learning from interaction and its use of feedback over time, rather than direct instruction. It is often used in scenarios where the right decisions or actions are not known in advance, or the environment is too complex to model explicitly.

In RL, several key components interact to facilitate the learning process. The agent, central to RL, makes decisions and learns from the outcomes within an environment encompassing everything the agent interacts with. The state represents the current situation in the environment, informing the agent's decision-making process. Actions are the choices available to the agent that can change the state, and each action leads to a reward, a feedback mechanism indicating the success or failure of the action, which is usually numeric. The policy, a strategy defining the agent's behavior, dictates the actions based on the current state. The value function helps the agent assess the potential of different states and actions based on expected cumulative rewards. These components work synchronously to enable the agent to adapt and optimize its behavior in dynamic settings.

An RL agent engages with its environment sequentially. Mathematically, at every time step t, the agent observes a state s_t from a set of possible states S and chooses an action a_t from a set of possible actions A. This choice is guided by a policy $\pi(a_t|s_t)$, which is essentially the agent's strategy, defining how it selects actions a_t in response to the state s_t (Li, 2017). Following this, the agent receives a numerical reward r_t and moves to the next state s_{t+1}. This transition is determined by the environment's dynamics, specifically the reward function $R(s, a)$ and the probability of transitioning to a new state $P(s_{t+1}|s_t, a_t)$. We also have an action value, $Q(s, a)$, which is a numerical estimate of the expected cumulative reward that an agent can expect to receive starting from a given state, s and taking an action, a, following a policy, π.

$$Q^{\pi}(s_t, a_t) = \sum_{s_{t+1} \in S} P(s_t, a_t) \left[R(s_t, a_t) + \gamma \sum_{a_{t+1} \in A} \pi(s_{t+1}) Q^{\pi}(s_{t+1}, a_{t+1}) \right] \quad (1)$$

Eq (1) shows the formula of Bellman's Equation for Q-learning (Sutton and Barto, 2018) where the action value of an agent starting from state, s, taking an action, a, and following a policy, π, at time step, t is updated using future rewards (here $0 < \gamma \leq 1$ is the discounting factor that lowers the reliance on future reward values in the current objective).

There has been a wide scale adoption of RL in applications such as aviation policy (Nikki Lijing Kuang, 2019), logistics (Gemmink, 2019), medicine (Hajar, et al., 2023), energy management (Karl Mason, 2019), traffic signal control (Chu, et al., 2019), and toll pricing (Pandey et al., 2020). It has empowered machines to master playing video games solely from visual screen input to outperform champion human players in Go using the AlphaGo program, and to accomplish a range of tasks encompassing movement and basic problem-solving (Francois Belletti, 2017).

In a recent infrastructure deployment and training, RL was also used to optimize energy management for grid-interactive efficient buildings, utilizing the City Learn framework to effectively balance energy consumption, cost, and emissions (Vázquez-Canteli et al., 2019). The study investigated RL methods in grid-interactive efficient buildings using the framework aiming to optimize energy use, cost, emissions, and demand management. First, a Rule-Based Controller (RBC) was used to generate sample policies and then Q-learning algorithms are used for insights and optimal policy. Initial findings with an RBC revealed limitations in Tabular Q-learning (TQL) due to its dependence on a single observation and computational demands. The Soft Actor Critic (SAC) algorithm, employing neural networks for Q-value approximation, was introduced to address these issues and showed superior performance in training efficiency and energy optimization. Challenges in SAC's reward function led to the creation of a custom reward function, enhancing SAC's effectiveness. This improved SAC demonstrated better battery and solar energy management, with potential for further gains from extended training.

Reinforcement Learning in Transportation Systems

Various reinforcement learning methods have been used and discussed in the transportation literature over the past few years. Table 1 summarizes the state, action, reward, and policy algorithms used in the literature. We also identify whether multi-objective tradeoffs were considered, noting that there are limited studies in this area.

Table 1 Summary of RL components in the recent literature on infrastructure control and optimization

Paper	State	Action	Reward	Policy learning	Algorithm	Multiobjective tradeoffs (if any)
Recent Advances in Reinforcement Learning for Traffic Signal Control: A Survey of Models and Evaluation (Wei et al., 2021)	Description of the environment (e.g., signal phase, waiting time of cars, queue length of cars, and positions of cars, lane-level queue length and phase)	Changing to a certain phase in the single intersection scenario	Queue length (single intersection), Pressure(multi-intersection)	Model free, Actor-critic, RNN	DDPG, DQN	-
Improving Deep Reinforcement Learning-Based Perimeter Metering Control Methods With Domain Control Knowledge (Zhou and Gayah, 2023).	Regional congestion levels, traffic demands	Controlling the proportion of traffic passing through each perimeter	Cumulative trip completion in a time step	Model-free off-policy Actor–critic	DDPG	-
EMVLight: A multi-agent reinforcement learning framework for an emergency vehicle decentralized routing and traffic signal control system (Su et al., 2023)	The number of vehicles on each outgoing lanes and incoming lanes, the distance of the EMV to the intersection, the estimated time of arrival and the next link the EMV will be routed to.	Switching phase.	Weighted sum of pressure	-	A2C (MA2C)	-
Multiadvisor Reinforcement Learning for Multiagent Multiobjective Smart Home Energy Control (Tittaferrante and Yassine, 2021)	Temperature (Indoor, outdoor, and set), battery charge level and energy consumption	Directing the energy to different appliances.	Appropriate temperature, range anxiety, consumption reduction	Off policy action-value	DQN	Multiadvisor agents in each appliances and their associated reward functions

A Multiobjective Reinforcement Learning Framework for Microgrid Energy Management (Liu et. al., 2023)	The temperature, wind speed, solar radiation, stream flow of the local river that provides hydropower and the real-time location marginal price of previous day at hour t.	The decision vector that determines the power output of each device for the next time step.	Optimal expected cost, expected emission, and expected heat waste	Model free	Borg Multi-Objective Evolutionary Algorithm	A weighted sum method to convert the original multiple objective formulations into a single objective representation that tacitly infers that the specified weights capture all stakeholders' preferences.
A Framework for Automated Multiobjective Factory Layout Planning using Reinforcement Learning (Klar et al., 2022)	Permissible and impermissible area on the grid, the geometry and location of functional unit placed, load bearing capacity for each position.	Placement of the functional unit by choice.	Functional units in permissible area and higher for better solution and optimized locations.	-	Deep Q-learning	-
A Multiobjective Reinforcement Learning Approach to Trip Building (Dos Santos and Bazzan, 2022)	Each intersection in the traffic network	Link to take at an intersection.	Time it takes to travel and its emission.	Model free	Q-learning (Adapted Pareto Q-Learning)	Accumulation of the reward after factoring for penalty and bonus
User Preference-Based Demand Response for Smart Home Energy Management Using Multiobjective Reinforcement Learning (Chen et al., 2021)	The status of the home appliances and battery system, current electricity price.	Whether to switch the appliances on or off.	Electricity cost, user comfort.	-	Q-learning.	Weighted sum method of the Q-tables.
Reinforcement Learning for Solving the Vehicle Routing Problem (Nazari et al., 2018)	The locations and demands of the customer nodes.	Selecting which customer node to visit next or whether to return to the depot.	The negative of the total distance traveled.	Model based, MDP	A3C	-

Multiobjective RL

Multiobjective Reinforcement Learning (MORL) is defined as a specialized branch of reinforcement learning that focuses on optimizing multiple, often conflicting objectives simultaneously (Hayes et al., 2022). Unlike traditional reinforcement learning, which typically aims to maximize a single cumulative reward, MORL deals with a vector of rewards, each representing a different objective. This approach is particularly relevant in complex environments where decisions must balance trade-offs between competing goals, such as cost, efficiency, safety, or environmental impact. MORL algorithms are designed to navigate these trade-offs, seeking policies that provide an optimal compromise among the multiple objectives, often leveraging techniques like Pareto efficiency to identify solutions that cannot be improved on one objective without degrading another.

Advances in Personalized Tolling

Recent advancements in transportation management and consumer behavior research have highlighted innovative approaches to toll pricing and demand management, as well as the implications of big data in consumer pricing strategies. A study by Zhang (2019) presents a proactive toll pricing framework using a Dynamic Traffic Assignment (DTA) system, which dynamically optimizes toll rates based on real-time traffic predictions. This system, enhanced through online calibration for accurate traffic predictions, has shown superior performance in managing traffic compared to static and reactive models. Building upon this, Zhang extends the framework to incorporate a personalized toll pricing system, utilizing electronic toll collection and vehicle identification data to develop individualized route choice models. This personalized approach, offering targeted discounts based on traveler preferences, effectively improves revenue optimization over non-personalized systems.

Parallel to these developments, Azevedo et al. (2018) explores incentive-based demand management strategies in transportation. These strategies, including dynamic pricing and quantity control, are becoming increasingly accepted as alternatives to traditional methods like congestion charging. The study emphasizes the importance of personalizing incentives, such as cash rewards or High Occupancy Vehicle (HOV) passes, to impact commuting decisions and optimize travel demand management systems. Azevedo et al. (2018) also highlights the necessity of integrating predictive models and real-time behavioral adjustments into transportation management systems for efficiency, potentially leading to significant energy savings and reduced congestion.

Big Data's Role in Consumer Pricing and Ethical Considerations

In the domain of consumer behavior, the impact of big data on pricing strategies is explored by Shiller (2013). They investigate the shift from traditional demographic data to web-browsing data in predicting consumer behavior, using Netflix as a case study. Shiller's model, which integrates economic theory with machine learning and uses nearly 5,000 web-browsing variables, demonstrates a potential profit increase of 12.2% with big data-driven personalized pricing, marking a significant improvement over demographic-based predictions. The study also discusses the broader implications of first-degree price discrimination through big data, highlighting potential equity concerns and consumer responses to such pricing strategies.

The ethical and legal concerns in pricing, particularly online personalized pricing have also been highlighted (van der Rest et al., 2020a; van der Rest et al., 2020b), and suggest that the legal system may not provide immediate solutions. It emphasizes the need to examine ethical and legal implications in pricing, including indirect price discrimination through psychological pricing and neuromarketing.

Shiller's research concludes with simulated scenarios that show how using large amounts of data (big data) can improve pricing strategies that are tailored to individual customers. These simulations help demonstrate the potential benefits of using detailed data to set prices more effectively for each person. This leads to a broader discussion on the evolving nature of price discrimination in the digital era and the need for ongoing research in this dynamic field. Overall, these studies collectively highlight the potential of real-time data analysis and personalization in both transportation management and consumer pricing strategies, marking a significant shift towards more nuanced and data-driven approaches in these fields.

Personalized Pricing and Fairness: Balancing Profitability and Social Implications

Kallus and Zhou (2021) delve into the complexities of personalized pricing in the modern data-rich environment. Their study extends beyond traditional demographic data, incorporating detailed behavioral data for more accurate consumer behavior predictions and willingness to pay. A critical aspect of their research is exploring fairness in personalized pricing, particularly the challenges in ensuring equitable pricing strategies that avoid disproportionately impacting certain demographic groups.

Kallus and Zhou (2021) develop a model that seeks to balance profitability with fairness and welfare considerations. A key finding is the need for careful management of personalized pricing to prevent unfair price burdens on specific groups, especially in sectors with profound social implications like healthcare and finance. The authors emphasize the necessity of equitable access to essential goods and services, discussing how personalized pricing can enhance access by extracting more revenue from higher-valuation groups to subsidize lower-valuation ones. However, this raises concerns about the fairness of pricing across different demographic groups.

The paper also examines the long-term dynamics of personalized pricing strategies. The authors suggest that while immediate financial benefits are evident, the long-term effects on customer loyalty, market dynamics, and social welfare require careful consideration. The study highlights the importance of algorithmic designs that take into account these broader implications, ensuring that pricing strategies are not only profitable but also equitable over time.

Advancements in App-Based Recommender Systems: A Hierarchical Bayes Approach

The article by Danaf et al. (2019) presents an innovative framework for estimating and updating user preferences in app-based recommender systems. Utilizing a Hierarchical Bayes procedure, the study accounts for both inter-consumer and intra-consumer heterogeneity, capturing random taste variations among individuals and across different choice situations. This framework involves three levels of preference parameters: population-level, individual-level, and menu-specific, estimated periodically offline and updated in real-time as users make choices.

The study addresses two significant gaps in discrete choice models in recommender systems. Firstly, it proposes an online estimation methodology for individual preferences, tackling computational constraints associated with offline methods. Secondly, it acknowledges a more advanced level of heterogeneity, both inter- and intra-consumer, thus enhancing prediction quality and recommendation accuracy. This approach is particularly useful for systems where alternative attributes vary over time, like travel recommendations.

One of the challenges discussed is calibrating choice models at the individual level, especially with limited data per user. The proposed online estimation procedure mitigates this by leveraging robust priors from the offline Hierarchical Bayes estimator, making it efficient and practical for real-time applications. The methodology's effectiveness is demonstrated through Monte-Carlo simulations and real data, showing that online parameter updates significantly improve real-time estimates. This framework marks a notable advancement in personalized recommendation strategies, boosting the capability of recommender systems to provide accurate, user-specific suggestions in dynamic, real-time environments.

Bridging Theory and Practice: The Challenges and Prospects of Reinforcement Learning in Dynamic Traffic Control

Han et al. (2023) provides an insightful resource in understanding the applications and challenges of Reinforcement Learning (RL) in dynamic traffic control systems. The comprehensive survey begins by examining the evolution and current state of RL-based traffic control strategies. One of the core focuses of the article is the practical challenges faced in implementing RL-based strategies in real-world scenarios. These challenges are broadly categorized into two areas: the learning costs associated with online RL methods and the transferability issues concerning offline RL methods. The paper provides a critical analysis of how online training methods, despite their potential, are impeded by high exploration and learning costs. Similarly, offline RL methods, although beneficial in theory, are heavily dependent on the accuracy of training simulators, raising concerns about their effectiveness in real-world applications.

To further elucidate these challenges, the authors present detailed simulation experiments. These experiments are designed to assess both the learning costs of online RL methods and the transferability of offline RL methods. The findings reveal that online RL methods suffer significantly due to the costs incurred during the exploration phase. Moreover, the performance of offline RL methods is largely dependent on the reliability of the simulation environment to real-world traffic conditions. This insight is critical, as it highlights the gap between theoretical RL models and their practical applications in traffic systems.

Addressing these challenges, the article proposes several research directions, which include integrating physical traffic flow models into RL (physics-informed RL), learning from demonstration, meta-reinforcement learning (meta-RL), combining RL with traditional traffic control methods, and exploring adversarial reinforcement learning (ARL). Each of these approaches aims to mitigate the identified challenges, either by reducing learning costs, enhancing transferability, or improving the overall robustness of RL strategies in traffic control. It also reveals some results on RL strategy with different degrees of mismatch which highlighted the critical importance of precise training simulator accuracy for the efficacy of the RL-based ramp control strategy, as the strategy's performance degrades significantly when there is a substantial mismatch between training and testing environments.

APPROACH AND METHODOLOGY

This section outlines the modeling approach to quantifying and addressing equity gaps between groups using a mesoscopic cell-transmission-based traffic flow model with a trapezoidal fundamental diagram and using the framework to obtain values for different objectives. Furthermore, we assume that travelers do not equilibrate their route or time of departure.

Equity Evaluation Framework

Considering the five-step equity analysis framework for public engagement and building off the past study in Pandey et al. (2022), we first determine the relevant factors for equity analysis (similar to frameworks in Bills and Walker (2017), Litman (2021), and Twaddell and Zgoda (2020)):

- **Population for Analysis**: We consider the population of travelers who wish to travel from an origin to a destination using their personal vehicle. These travelers can be further grouped based on their need for travel. The primary criterion used for grouping is the value of travel time (VOT) defined as the dollar amount value that an individual is willing to sacrifice in order to save a unit of travel time. We measure VOT in $/hr. It has been shown that VOT is correlated with an individual's income.

- **Needs and concerns**: Transportation systems connect individuals to their destination. We consider that the primary need of a traveler is to arrive at their destination as quickly as possible by minimizing the travel time they incur while driving on the roadway system. Given this need, an equity concern emerges when certain groups of travelers are forced to spend extra time on travel relative to other groups. As we demonstrate later, the travel time spent by low VOT travelers is commonly higher and if VOT correlates with travelers' income, then low-income travelers suffer a higher burden of congestion relative to high-income travelers. It is worth noting that transportation is one of the basic needs of travelers and is not a luxury item that only individuals with a high willingness to pay need to access.
- **Measuring impacts of proposed options**: Express lanes charge tolls, and we measure how these lanes create equity differences by considering a modeling perspective described later in this section. Broadly, we consider the interaction between supply and demand side models for travelers using the corridor.
- **Determine disparities**: In our analysis, disparities are measured by quantifying the average delay per person experienced by a traveler in each VOT group and using the maximum absolute difference of delay differential as the equity metric. We chose this metric since the purpose of managed lanes is to provide reliable travel time. This equity metric is defined later in terms of the modeling parameters.
- **Develop strategies to mitigate inequities**: Once we quantify equity, we determine strategies that can address this gap. In particular, we consider various discounting methods for travelers in different VOT groups.

Component Models

Broadly, component models for express lanes can be categorized as shown in Figure 1. Next, we discuss how these component models were implemented in this study.

Consider a managed lane network in Figure 2 given by a directed graph $G = (N, A, Z)$ where N is the set of all nodes, A is the set of all links, and Z is the set of all zones where trips begin or end. Let T be the set of toll gantries where tolls are collected. We assume that these gantries are located on links. Let $A_T \subset A$ be the subset of links that charge a dynamic toll. Without loss of generality, the tolled links are selected such that the tail node of the links is a diverge node where travelers make a choice between express lane and general-purpose lanes. Assessing the impacts of various congestion pricing metrics is done by integrating models for interactions between transportation demand and supply, which we define next.

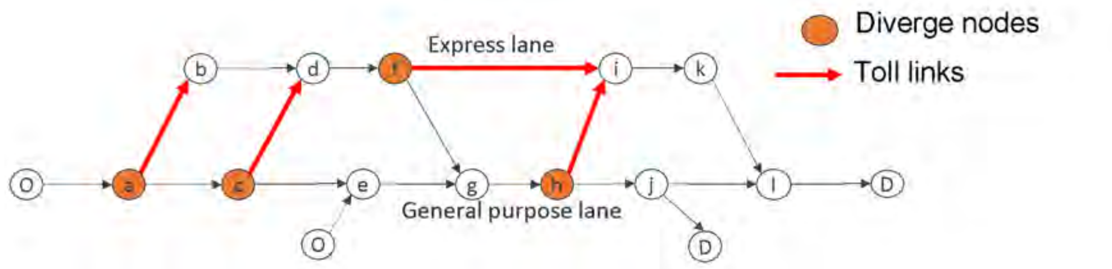


Figure 2 An express lane network comprising of links that are part of express lanes and general-purpose lanes. Toll gantries are assumed to be located on highlighted toll links. Travelers enter the corridor through origin nodes and exit through destination nodes

Demand characteristics: We consider deterministic time-dependent demand using the corridor between an origin-destination pair. Let $d_{rs}(t)$ be the demand entering the corridor at time t at origin node $r \in Z$ traveling towards destination $s \in Z$. For simplicity of the model, we assume that travelers do not adapt their departure time in response to tolls and thus $d_{rs}(t)$ is assumed known apriori (estimated using the historical usage of the tolled facility). We group the travelers by their value of time (VOT) modeled using a discrete VOT distribution. Let α_k represent the VOT for travelers in group $k \in K$ where K is the set of all groups. Without loss of generality, we order travelers such that $\alpha_1 > \alpha_2 > \cdots > \alpha_{|K|}$. As discussed earlier, such distributions can be estimated using historical travel patterns based on the income distribution for travelers using the corridor.

Traffic flow model (Supply-side characteristics): Models for traffic flow determine the variation of traffic density for different times and locations expressed as a partial differential equation. The Lighthill-Whitham-Richards (LWR) assumes a deterministic relationship between density and flow expressed as the fundamental diagram. In our analysis, we model the evolution of traffic on the corridor using a macroscopic multiclass cell transmission model (Daganzo, 1995) which numerically solves the LWR equation by dividing the space into cells and time into discrete time-intervals (let T be the set of all time intervals). The multiclass CTM model relates flow and density on each link using a trapezoidal fundamental diagram. For brevity, we refer the reader to prior literature on multiclass CTM model for further details (Tan and Gao, 2018; Pandey and Boyles, 2018).

Lane choice model: At each diverge point, a traveler from a VOT group $k \in K$ compares the utility across different lane alternatives. Utility of travelers making a decision at toll gantry $g \in T$ is determined as a linear combination of travel time and toll on managed lane and GPL options. We assume that the information about the current travel time is provided by measuring instantaneous travel time with no time lag, and that all travelers' utility are calculated only using the instantaneous travel time and toll information. For a given diverge point, we consider two routes connecting the current diverge with the first exit from the managed lane if a traveler were to enter the managed lane then. For example, in Figure 2, the routes considered at diverge node a are $[a, b, d, f, g]$ and $[a, c, e, g]$. The utility on a route for a traveler in group $k \in K$ is then given as the linear combination of travel time and toll: $U = \alpha_k t + \tau$ where t and τ are travel times and tolls for the route.

Toll Optimization using Reinforcement Learning

Once the interaction between supply and demand is established, the toll optimization problem can be formulated as the choice of toll $\tau_{ML}^{l}(t)$ for toll link $l \in A_T$ at different time intervals $t \in T_{toll} \subset T$ (since tolls are updated less frequently than traffic updates). We consider two objectives for toll optimization: maximizing revenue and minimizing total system travel time (TSTT).

Building on the open-source reinforcement learning framework in Pandey et al. (2020), we optimize the toll using a reinforcement learning framework. The components of the Markov decision process associated with the reinforcement learning problem are outlined below:

- **Timestep**: Tolls are optimized over a finite time horizon for each time interval $t \in T_{toll}$.
- **State**: The traffic state at any given time is characterized by number of vehicles of each group $k \in K$ across all cells in the network.
- **Action**: The action in any state is the toll charged on each toll links $l \in A_T$, where the toll is considered bounded $\tau_{ML}^{l}(t) \in (\tau_{min}, \tau_{max})$

- **Transition function**: The transition from a given state and the chosen action is governed by the multiclass cell transmission model
- **Reward**: The reward in each state after taking an action is governed by the tolling objective. For revenue maximization, the reward is the immediate revenue obtained in that time-step. For TSTT minimization, the reward is the equal to the total number of vehicles present in the network multiplied by minus 1 (to accommodate the minimization objective).

RL Algorithms Considered in this Study

For the experiments, we employed three algorithms: Soft-Actor Critic (SAC), PPO Proximal Policy Optimization (PPO) and Advantage Actor Critic (A2C) for the macroscopic simulations. SAC is an off-policy algorithm that optimizes a stochastic policy. It uses a technique called the clipped double-Q trick, which helps improve its performance. Additionally, because SAC's policy naturally involves some randomness, it also gains advantages similar to target policy smoothing. SAC is designed to perform well in continuous action spaces and has been successful in tasks ranging from robotic control to complex locomotion. PPO is an on-policy gradient method for RL. There are two main variants of PPO: one that clips the policy ratio and another that adds a penalty based on the KL divergence between the old and new policies. PPO's ease of implementation and efficient use of data have made it a popular choice for many RL tasks. It strikes a balance between data efficiency, ease of tuning, and final performance. A2C maintains a policy (actor) that suggests the next action to take given the current state of the environment, and a value function (critic) that evaluates the chosen action. The "advantage" part of the name refers to how the algorithm estimates the relative value of each action in a given state, which helps in reducing the variance of the updates. A2C works by simultaneously updating the actor and critic using gradients that aim to maximize expected future rewards.

We use the OpenAI-gym RL environment for macroscopic simulation provided by Pandey et al. (2020) and use the open-source implementation of the Soft-Actor Critic algorithm (Hou et al., 2020) to find tolls that maximize the reward over the time horizon. The SAC algorithm has been shown to converge to optimal tolls for the express lane pricing problem (Pandey et al., 2020) and in our experiments, we observe the same pattern of convergence.

Overview Algorithm

The reinforcement learning algorithm proceeded in following steps (See Pandey et al. (2020) for additional mathematical details on the parameters and functions):

<u>Step 1</u>: Initialize policy parameters and value function parameters.

<u>Step 2</u>: Simulation N Training Epoch (N=500 for our experiments)

for index ID ∈ {1,2….N} **do**
  Collect set of trajectories by running policy for ten 2-hour simulations on network
  Compute rewards to go for the defined reward value
  Compute advantage estimates using rewards-to-go
  Update Policy parameters using A2C, PPO, or SAC updates
  Update value function approximation parameters
end for

<u>Step 3:</u> Report and analyze the performance statistics over the training period

FINDINGS AND RESULTS

For our analysis, four abstract simulation networks were considered from the literature (Pandey et al., 2020), as depicted in Figure 3. The data for Loop-1 express lanes were derived from the regional model.

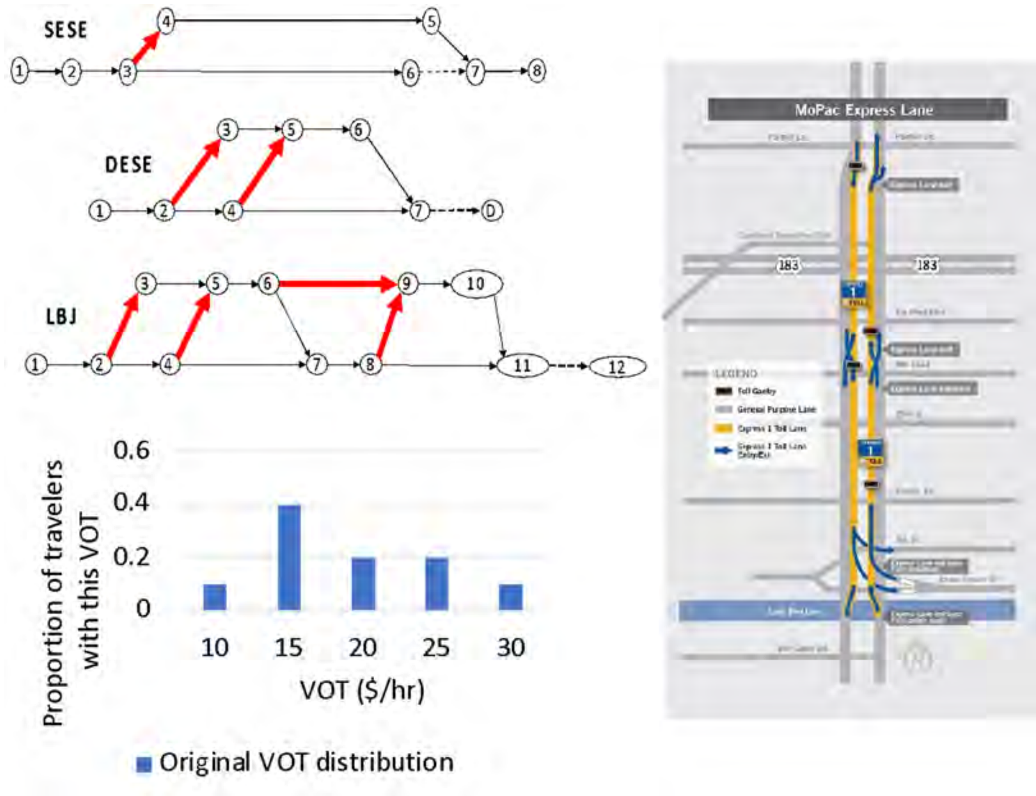


Figure 3 Test Networks Considered for Training of Reinforcement Learning Algorithms

A discrete value of time distribution was assumed over a three-hour peak period, as shown in Figure 3. Broadly, this distribution follows the Burr distribution observed in the literature. For any five-minute time period, the maximum toll was set at $8, and the minimum toll was set at 10 cents. We used the simulation model for the corridor created in an earlier study, which utilizes the multiclass cell transmission model for traffic simulation. Lane choice is governed by binary logit and decision-route models. Toll profiles were updated in five-minute intervals. Tolls were analyzed on the Pareto frontier of multiple objectives, such as total system travel time (TSTT), revenue, and equity.

To model the discount, we optimize, for any time period t, the discount $d_k(t)$ given to group $k \in K$. Our preliminary model scales the discounts linearly by controlling the maximum discount offered to travelers with minimum value of time (VOT). This is shown in Figure 4(a) as restricted discount case. Our secondary model allows the discounts to be any value in the specific range between the maximum and minimum discount (Figure 4(b) or unrestricted discount). As discussed, next, the revenue maximizing cases result in no discount for all travelers, while when the equity gap across travelers is minimized, then discount offered can be positive anywhere in the green rectangular box shown in Figure 4(b).

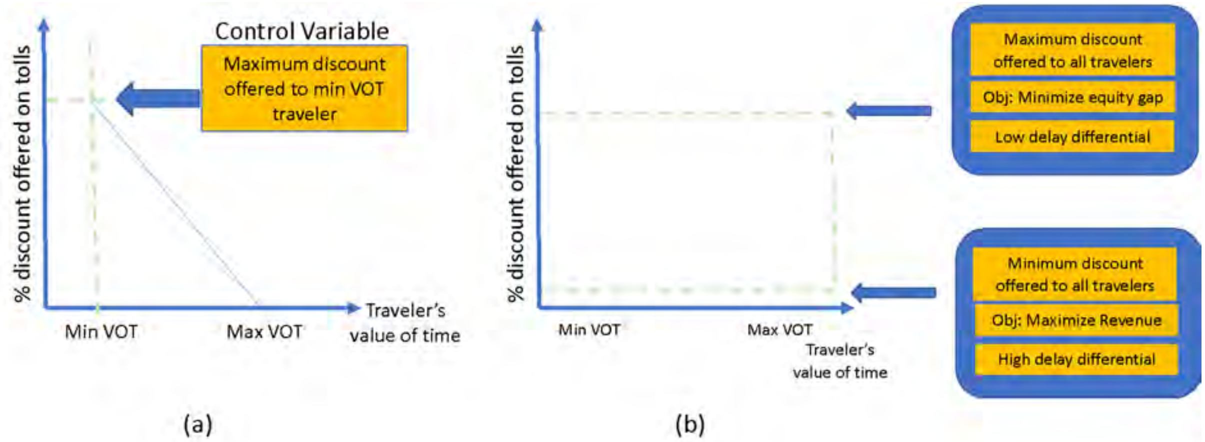


Figure 4: (a) Discount case 1 where travelers with minimum VOT receive the maximum discount which is the control variable in our experiments (Restricted Discount), and (b) Discount case 2 where travelers can receive any discount personalized to their and operators' needs (Unrestricted Discount).

The A2C, PPO, and SAC reinforcement learning algorithms were trained on all test networks considering various objectives. The findings will be discussed next.

First, we see that for minimizing TSTT, the algorithms obtained similar reward as a function of the number of training iterations was fairly consistent, except for a noticeable improvement for PPO algorithm on LBJ network, which suggests that PPO is robust when learning because of the intricacies of LBJ network. But looking at the performance on other objectives (such as maximizing revenue), we see that SAC outperforms PPO in obtaining revenues in fewer training steps (as shown in Figure 5). This might suggest each algorithm has networks that they perform and objectives they perform better on, which is indicative of the experimental nature of reinforcement learning algorithms as indicated in the literature.

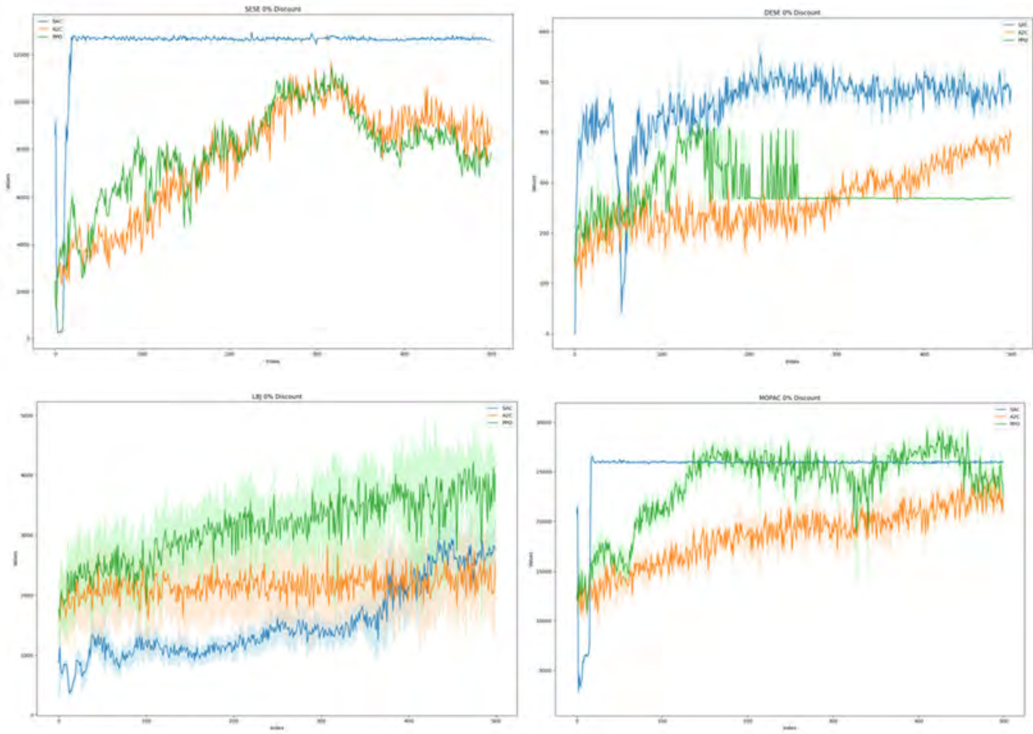


Figure 5: Reward profile (y-axis) as a function of number of training indices (x-axis) for four networks at 0% discount: SESE (top-left), DESE (top-right), LBJ (bottom-left), and Mopac (bottom-right). The three algorithms considered are SAC (blue), PPO (green), and A2C (orange).

Second, we see the max throughput values were similar in each network for all algorithms and the values were high, it might be the case that all the algorithms learned that well and got the highest values for every one of them. Finally, it can be seen that revenue was maximized in the scenario where there is no discount than where there is, with the exception of Mopac network where we obtained similar revenues under 25% discount as without discount. Details for the algorithmic runs are shown in Tables in Appendix 2.

Initial results showed that all algorithms learned similarly overall, although some algorithms performed better than the others on different networks, this varying performance could be due to specific network intricacies, more effective hyperparameter tuning or varying strengths in handling the complexity and size of the network data. However, offering discounts to induce equity did not yield favorable results as the algorithms all learned that to maximize revenue, it is better off minimizing discount for all groups and when the general-purpose lane experiences high density, then travelers opt for the managed lanes.

We then considered forcing the discounts to enable the algorithms to provide discount to low-income group. We also change the objectives to have minimize equity gaps among all VOT group travelers. The key findings are discussed next.

When striving to maximize revenue, Episode TSTT (EpTSTT) was the highest for VOT $10/hr. and progressively decreased for higher VOTs as depicted in Figure 6. This outcome aligns with expectations, as individuals with higher VOTs are more inclined to use toll lanes to expedite their travel times. Additionally, since lower VOT is commonly proportional to lower income group, we can argue that travelers with low VOT are worse off under pricing that maximizes revenue, thus the equity gap increase. Furthermore, SAC initially exhibits similar EpTSTT across all VOTs before diverging noticeably, whereas A2C starts with varying values and generally escalates over epochs. This trend suggests that while the algorithm behave differently initially, they eventually showed the delay differentials with the lowest VOT experiencing the highest delay. However, the delay reduced minimally with higher discounts for all groups.

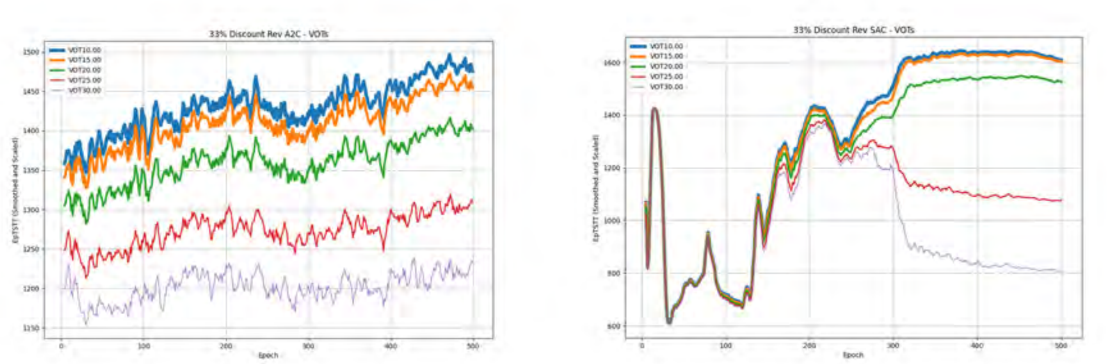


Figure 6: Revenue Optimizing VOT profiles at 33% Discount. A2C plot (left) showing rising travel times for VOTs with higher increase for lower VOTs. SAC (right) learning to give discounts more to higher VOT to incentivize them to use the ML and thus reduce travel time.

SAC initially exhibits similar EpTSTT across all VOTs before diverging noticeably, whereas A2C starts with varying values and generally escalates over epochs. This trend suggests that the algorithm learned to give discounts more to higher VOTs as they are likely to take the tolls. The growth also amplified with higher discounts as shown in Figure 6.

When prioritizing the minimization of equity gaps and using restricted discounts, a decreasing revenue pattern is observed. This is accompanied by lower EpTSTT for 25% discounts compared to 50% across both algorithms as seen in Figure 7. This could be attributed to 25% being a conservative discount value that maintains an optimal balance of travelers in both hot lanes and GPL. Conversely, offering high discounts like 50% proves ineffective as it prompts more individuals to opt for the hot lane, consequently increasing TSTT for all. Unlike the revenue maximization scenario, by the end of the training, the algorithms learned to offer uniform TSTT to all travelers irrespective of VOT.

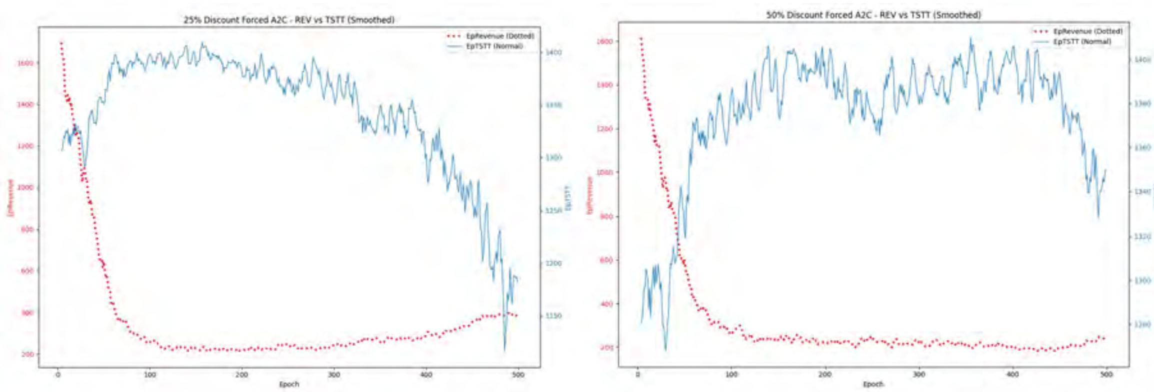


Figure 7: Revenue vs TSTT Minimizing Equity Gap (Forced discount) 25% discount showing decreasing trend in travel time with low revenue, 50% showing similar but lower revenue and higher travel time.

When aiming to minimize equity gaps without enforcing discounts, revenue also experiences a decline as the objective could only be achieved by giving discounts to lower VOT groups. The TSTT and revenue remain unaffected by the maximum discount offered as shown in Figure 8, yielding similar results with variations stemming from exploration (where agents attempt to search for an optimal result). Similar to the scenario with enforced discounts, the algorithms provide consistent TSTT to all travelers at the end of training.

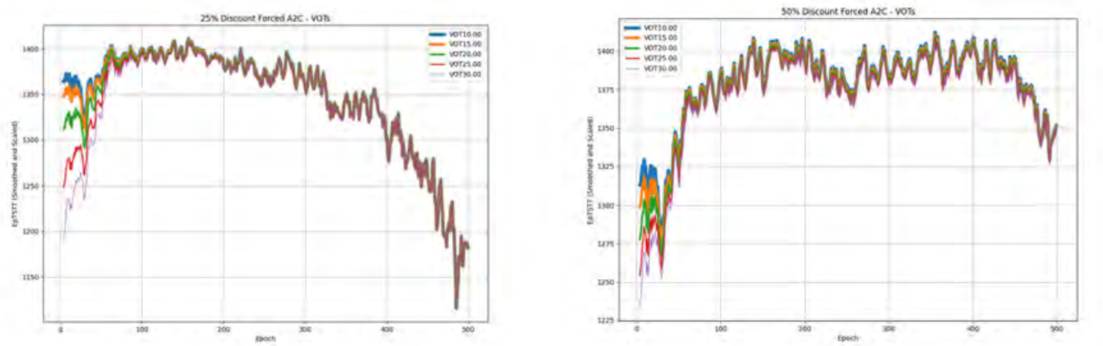


Figure 8: Equity Gap Plots for VOTs. 25% discount showing decreasing trend in travel time and no equity gap with ongoing training, 50% discount showing same but higher travel time for all which suggests the value is not the optimal discount.

Comparing the revenue from the most equitable profile to the max revenue enables us to compute the price of fairness (Bertsimas et al., 2011).

$$Price\ of\ Fairness = \frac{Revenue\ at\ no\ discount - Revenue\ at\ lowest\ equity\ gap}{Revenue\ at\ no\ discount}$$

$$Price\ of\ Fairness = \frac{4502 - 386}{4502} = .914$$

This means that in order to provide the most equitable tolling; we sacrifice 91.4% of our revenue for the LBJ network.

CONCLUSIONS, RECOMMENDATIONS, AND FUTURE WORK

This report analyzed reinforcement learning algorithms in the context of multi-objective optimization for express lane pricing. First of our findings indicates that commonly used objectives for managing express lane systems such as improving revenue, reducing total system travel time (TSTT), and enhancing equitable access can be at conflict with each other. Specifically, we observe that optimizing for revenue might only widen equity gap. In terms of the usefulness of RL algorithms, SAC, A2C, and PPO algorithms each showed distinct abilities to optimize rewards and reduce total system travel time (TSTT) when using personalized tolling strategies.

The report also reviews the critical issue of ensuring equity in managed lanes highlighting the importance of equitable access to transportation infrastructure. Separating travelers based on their VOT, we see that we can reduce the equity gap among the travelers, although it comes with a 91% loss on revenue, which may not be acceptable in practical cases. Offering discounts on toll has also been seen to minimize the equity gap and reduce TSTT for all. Given the unique characteristics of each network, including their dynamics and the diverse groups of travelers they serve, the outcomes of these analyses can vary (such as the extent to which the objectives are at a conflict). Therefore, tailored tolling approaches may be necessary to accommodate the specific needs of each system.

Recommendations:

1. While maximizing revenue and efficiency in managed lanes is a priority for investors, ensuring equitable access to these infrastructures is crucial. Instead of solely aiming to minimize equity gaps, implementing personalized tolling can maintain high revenue while also reducing travel times for low-income travelers.
2. Since equity remains a significant concern for many stakeholders, and achieving it can be expensive, partnerships with government bodies or nonprofit organizations focused on enhancing urban mobility and access should be considered. This collaboration can help ensure shorter travel times for all travelers, regardless of their value of time.
3. The effectiveness of these strategies on mobility should be regularly assessed, especially the improvements in throughput and/or corridor travel time efficiency and reliability. Using the insights from these evaluations can further optimize the system's efficiency.

Future Work

Based on the insights from analyzing reinforcement learning algorithms in transportation system optimization, future research should investigate comprehensive multiobjective frameworks. Such frameworks should not only aim to minimize Total System Travel Time (TSTT), optimize revenue, and minimize equity gaps but also consider broader objectives like environmental impacts, user satisfaction, social equity, and particularly fairness over time.

One key limitation of the proposed work is that it is based on mesoscopic simulations which do not consider departure time choice and/or weaving operations around toll entrance. Moreover, conducting extensive simulations and real-world pilot studies across various network types will be crucial in evaluating the practical effectiveness of these strategies. A significant area of focus should be the long-term fairness and sustainability of access to managed lanes, especially for low-income travelers. It is essential to monitor if these travelers continue to use managed lanes after exhausting any provided credits or subsidies, ensuring that the tolling strategies remain equitable over time and do not inadvertently exclude vulnerable populations. This comprehensive method involves working closely with a wide range of people, such as government officials, those who plan transportation, and local communities. The goal is to understand their requirements and evaluate the broader effects and rules related to introducing advanced tolling systems and inclusive approach promises to advance our understanding and implementation of equitable, efficient, and sustainable transportation solutions.

REFERENCES

AbuLibdeh, A. (2017). *Traffic Congestion Pricing: Methodologies and Equity Implication* (Rep.). IntechOpen. doi:10.5772/66569 https://www.intechopen.com/books/urban-transport-systems/traffic-congestion-pricing-methodologies-and-equity-implications

Azevedo, C. L., Seshadri, R., Gao, S., Atasoy, B., Akkinepally, A. P., Christofa, E., ... & Ben-Akiva, M. (2018, January). Tripod: sustainable travel incentives with prediction, optimization, and personalization. In *Proceedings of the Transportation Research Record 97th Annual Meeting*.

Belletti, F., Haziza, D., Gomes, G., & Bayen, A. M. (2017). Expert level control of ramp metering based on multi-task deep reinforcement learning. *IEEE Transactions on Intelligent Transportation Systems*, *19*(4), 1198-1207.

Bertsimas, D., Farias, V. F., & Trichakis, N. (2011). The price of fairness. *Operations Research*, *59*(1), 17-31.

Bills, T. S., & Walker, J. L. (2017). Looking beyond the mean for equity analysis: Examining distributional impacts of transportation improvements. *Transport Policy*, *54*, 61-69.

Chen, S. J., Chiu, W. Y., & Liu, W. J. (2021). User preference-based demand response for smart home energy management using multiobjective reinforcement learning. *IEEE Access*, *9*, 161627-161637.

Chu, T., Wang, J., Codecà, L., & Li, Z. (2019). Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, *21*(3), 1086-1095.

Daganzo, C. F. (1995). The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79-93.

Danaf, M., Becker, F., Song, X., Atasoy, B., & Ben-Akiva, M. (2019). Online discrete choice models: Applications in personalized recommendations. *Decision Support Systems*, *119*, 35-45.

Debreczeni (2021). Low-Income Toll Program Study for I-405 SR 167 Express Toll Lanes. URL: https://wstc.wa.gov/wp-content/uploads/2021/08/2021-WSTC-Tolling-Equity-Report.pdf. Last accessed: Aug, 2022.

dos Santos, G. D., & Bazzan, A. L. (2022). A Multiobjective Reinforcement Learning Approach to Trip Building. In *ATT@ IJCAI* (pp. 160-174).

Gemmink, M. W. T. (2019). *The adoption of reinforcement learning in the logistics industry: a case study at a large international retailer* (Master's thesis, University of Twente).

Gordon, A. (2021). "The broken algorithm that poisoned American transportation." URL: https://www.vice.com/en/article/v7gxy9/the-broken-algorithm-that-poisoned-american-transportation-v27n3. Last accessed: Aug, 2022.

Hajar, M. S., Kalutarage, H., & Al-Kadri, M. O. (2023, January). RRP: A reliable reinforcement learning based routing protocol for wireless medical sensor networks. In *2023 IEEE 20th Consumer Communications & Networking Conference (CCNC)* (pp. 781-789). IEEE.

Hall, J. D. (2018). Pareto improvements from Lexus Lanes: The effects of pricing a portion of the lanes on congested highways. *Journal of Public Economics, 158*, 113-125.

Han, Y., Wang, M., & Leclercq, L. (2023). Leveraging reinforcement learning for dynamic traffic control: A survey and challenges for field implementation. *Communications in Transportation Research, 3*, 100104.

Hayes, C. F., Rădulescu, R., Bargiacchi, E., Källström, J., Macfarlane, M., Reymond, M., ... & Roijers, D. M. (2022). A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems, 36(1)*, 1-59.

Hou, Z., Zhang, K., Wan, Y., Li, D., Fu, C., & Yu, H. (2020). Off-policy maximum entropy reinforcement learning: Soft actor-critic with advantage weighted mixture policy (sac-awmp). *arXiv preprint arXiv:2002.02829*.

Kallus, N., & Zhou, A. (2021, March). Fairness, welfare, and equity in personalized pricing. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 296-314).

Klar, M., Langlotz, P., & Aurich, J. C. (2022). A framework for automated multiobjective factory layout planning using reinforcement learning. *Procedia CIRP*, *112*, 555-560.

Kuang, N. L., & Leung, C. H. (2019). Leveraging Reinforcement Learning Techniques for Effective Policy Adoption and Validation. In *Computational Science and Its Applications–ICCSA 2019: 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019, Proceedings, Part II 19* (pp. 311-322). Springer International Publishing.

Li, Y. (2017). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*. https://arxiv.org/abs/1810.06339 (Last accessed, Jun 2024)

Litman, T. (2021). *Evaluating Transportation Equity*. Victoria Transport Policy Institute. https://www.vtpi.org/equity.pdf

Liu, M. V., Reed, P. M., Gold, D., Quist, G., & Anderson, C. L. (2023). A Multiobjective Reinforcement Learning Framework for Microgrid Energy Management. *arXiv preprint arXiv:2307.08692*.

Lombardi, C., Picado-Santos, L., & Annaswamy, A. M. (2021). Model-based dynamic toll pricing: An overview. *Applied Sciences*, *11*(11), 4778.

Mason, K., & Grijalva, S. (2019). A review of reinforcement learning for autonomous building energy management. *Computers & Electrical Engineering*, *78*, 300-312.

Nazari, M., Oroojlooy, A., Snyder, L., & Takác, M. (2018). Reinforcement learning for solving the vehicle routing problem. *Advances in Neural Information Processing Systems*, *31*.

Pandey, V., & Boyles, S. D. (2018). Dynamic pricing for managed lanes with multiple entrances and exits. *Transportation Research Part C: Emerging Technologies*, *96*, 304-320.

Pandey, V., Wang, E., & Boyles, S. D. (2020). Deep reinforcement learning algorithm for dynamic pricing of express lanes with multiple access locations. *Transportation Research Part C: Emerging Technologies*, *119*, 102715.

Pandey, V., Alamri, B. G., Khoury, H., & Bakre, M. (2022). Equitable Dynamic Pricing for Express Lanes. Center for Advanced Transportation Mobility, Technical Report#13. URL: https://digital.library.ncat.edu/catm/13/

Peng, B., Keskin, M. F., Kulcsár, B., & Wymeersch, H. (2021). Connected autonomous vehicles for improving mixed traffic efficiency in unsignalized intersections with deep reinforcement learning. *Communications in Transportation Research*, *1*, 100017.

Shiller, B. R. (2013). *First Degree Price Discrimination using Big Data* (p. 32). Brandeis Univ., Department of Economics.

Su, H., Zhong, Y. D., Chow, J. Y., Dey, B., & Jin, L. (2023). EMVLight: A multi-agent reinforcement learning framework for an emergency vehicle decentralized routing and traffic signal control system. *Transportation Research Part C: Emerging Technologies*, *146*, 103955.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). The MIT Press.

Tan, Z., & Gao, H. O. (2018). Hybrid model predictive control based dynamic pricing of managed lanes with multiple accesses. *Transportation Research Part B: Methodological*, *112*, 113-131.

Tittaferrante, A., & Yassine, A. (2021). Multiadvisor reinforcement learning for multiagent multiobjective smart home energy control. *IEEE Transactions on Artificial Intelligence*, *3*(4), 581-594.

Tittaferrante, A., & Yassine, A. (2022). Multiadvisor reinforcement learning for multiagent multiobjective smart home energy control. IEEE Transactions on Artificial Intelligence, 3(4):581-594. doi: 10.1109/TAI.2021.3125918.

Twaddell, H., & Zgoda, B. (2020). *Equity Analysis in Regional Transportation Planning Processes, Volume 1: Guide* (No. Project H-54).

van der Rest, J. P., Wang, L., & Miao, L. (2020a). Ethical concerns and legal challenges in revenue and pricing management. *Journal of Revenue and Pricing Management*, *19*, 83-84.

van der Rest, J. P. I., Sears, A. M., Miao, L., & Wang, L. (2020b). A note on the future of personalized pricing: Cause for concern. *Journal of Revenue and Pricing Management*, *19*, 113-118.

Vázquez-Canteli, J. R., Kämpf, J., Henze, G., & Nagy, Z. (2019). CityLearn v1.0: An OpenAI Gym environment for demand response with deep reinforcement learning. In

Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (pp. 356-357). Association for Computing Machinery. https://doi.org/10.1145/3360322.3360998

Wang, X., Yang, H., Zhu, D., & Li, C. (2012). Tradable travel credits for congestion management with heterogeneous users. *Transportation Research Part E: Logistics and Transportation Review*, *48*(2), 426-437.

Wei, H., Zheng, G., Gayah, V., & Li, Z. (2021). Recent advances in reinforcement learning for traffic signal control: A survey of models and evaluation. *ACM SIGKDD Explorations Newsletter*, *22*(2), 12-18.

Xie, Y., Seshadri, R., Zhang, Y., Akinepally, A., & Ben-Akiva, M. E. (2024). Real-time personalized tolling for managed lanes. *Transportation Research Part C: Emerging Technologies*, *163*, 104629.

Zhang, Y. (2019). Real-time personalized toll optimization based on traffic predictions. PhD thesis, Massachusetts Institute of Technology.

Zhang, Y., Atasoy, B., & Ben-Akiva, M. (2018). *Calibration and optimization for adaptive toll pricing* (No. 18-05863).

Zhou, D., & Gayah, V. V. (2023). Improving Deep Reinforcement Learning-Based Perimeter Metering Control Methods With Domain Control Knowledge. Transportation Research Record, 2677(7), 384-405. https://doi.org/10.1177/0361198123115246.

APPENDIX – 1

Publications, presentations, posters resulting from this project:

Tiamiyu, R., Bowens, C. and Pandey, V., (2024). Model-Driven Equitable Toll Design for Express Lanes. Poster Presented at the NCA&T College of Engineering, Graduate Student Symposium.

Tiamiyu, R., Bowens, C. and Pandey, V., (2024). A Reinforcement Learning Framework for Equitable Toll Design in Express Lanes. Presented at the *2024 CATM Symposium*.

PDFs for the poster and presentation are attached separately to the report.

APPENDIX – 2 SIMULATION RESULTS AS TABLES

The tables show the best obtained reward for revenue maximization objective for different algorithms under varying levers of unrestricted discounts.

NO DISCOUNT					
	Algorithm	LBJ	SESE	MOPAC	DESE
TSTT at maximum revenue	SAC	1417	4744	15425	462
	A2C	1420	4819	14643	2289
	PPO	1422	4750	16208	2284
Maximum Revenue	SAC	3600	13053	26878	597
	A2C	4502	11725	26457	414
	PPO	4875	11428	30901	437
Throughout at Maximum Revenue	SAC	7616	11475	54325	11431
	A2C	7616	11475	54212	11427
	PPO	7616	11475	53225	11166

25% DISCOUNT					
	Algorithm	LBJ	SESE	MOPAC	DESE
TSTT at maximum revenue	SAC	1421	4695	15770	564
	A2C	1389	4748	13518	747
	PPO	1421	4824	15842	1798
Maximum Revenue	SAC	2992	13777	26474	566
	A2C	4493	11395	23112	374
	PPO	4728	12285	31545	399
Throughout at Maximum Revenue	SAC	7616	11475	53569	11433
	A2C	7616	11475	54584	11430
	PPO	7616	11475	53515	11265

50% DISCOUNT					
	Algorithm	LBJ	SESE	MOPAC	DESE
TSTT at maximum revenue	SAC	1341	4628	15805	464
	A2C	1386	4711	12053	734
	PPO	1389	4755	15174	1321
Maximum Revenue	SAC	4785	13987	26816	567
	A2C	4228	10784	19797	396
	PPO	4824	11776	27467	386
Throughout at Maximum Revenue	SAC	7616	11475	53652	11432
	A2C	7616	11475	55159	11430
	PPO	7616	11475	53769	11430

70% DISCOUNT					
	Algorithm	LBJ	SESE	MOPAC	DESE
TSTT at maximum revenue	SAC	1421	4730	15825	1570
	A2C	1298	4687	11118	634
	PPO	1318	4764	14410	751
Maximum Revenue	SAC	4263	13050	26619	678
	A2C	4000	10359	16955	444
	PPO	4462	10775	26586	450
Throughout at Maximum Revenue	SAC	7616	11475	53578	11434
	A2C	7616	11475	55517	11429
	PPO	7616	11475	54332	11430

A Reinforcement Learning Framework for Equitable Toll Design in Express Lanes

Researchers: Ridwan Tiamiyu, Christian Bowens, Dr. Venktesh Pandey (PI)

Department of Civil, Architectural, and Environmental Engineering

N.C. A&T State University

Overview

- Background and Research Questions
- Research Methodology
- Findings
- Conclusion

Need for Transportation Equity

- Equitable transportation services provide **equal levels of access to affordable and reliable transportation options** based on the needs of the population being served and distribute the costs and benefits fairly across the groups[1]

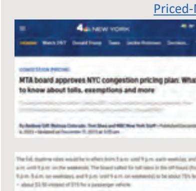
Decisions in transportation systems involving equity consideration	Examples
Provision of facilities and services	Provision of sidewalk ramps for people with disabilities
Usage fees and regulations	**Congestion pricing**; parking fees; **toll discounts**
Accessibility across population groups	Quality and frequency of transit services through poor neighborhoods
Compensation for externalities that disadvantage certain groups more than the others	Displacement of population groups; Excessive noise and air pollution for certain neighborhoods

[1]National Academies of Sciences, Engineering, and Medicine (NASEM) (2020). Equity Analysis in Regional Transportation Planning Processes, Volume 1: Guide. The National Academies Press.

Priced managed lanes as a congestion mitigation tool

- Congestion pricing through express lanes
  - Internalize the congestion externality and provide reliable travel-time alternative; More than 60 operational express lanes in 2023
- Do these facilities leave the economically-disadvantaged travelers worse off?
- Real-world case studies show:
  - Income, residential location, and urgency of travel impact express lane usage[1,3]
  - Express lanes offer a choice to pay or not pay the toll, hence preferred over standard toll roads
  - I-405 lanes: Per trip, low-income households may benefit more (due to more frequent peak-period travel)[2]
  - Nevertheless, high tolls and resistance against toll lane exists

Source: https://www.hntb.com/Projects/I-15-Priced-Managed-Lanes


MTA board approves NYC congestion pricing plan: What to know about tolls, exemptions and more


The Washington Post

Virginia to tweak 66 Express Lanes pricing to address tolls that have topped $47

[1]FHWA (2020). Urban Partnership Agreement Low-Income Equity Concerns of U.S. Road Pricing Initiatives. *Federal Highway Administration*.
[2]Hallenbeck, M., & Iverson, V. (2019). I-405 Express Toll Lanes Usage, Benefits, and Equity. *Washington Department of Transportation*.
[3]Patterson, T., & Levinson, D. M. (2008). Lexus lanes or corolla lanes? Spatial use and equity patterns on the I-394 MnPASS lanes. Report: University of Minneapolis, MN.

Complexities posed by emerging priced managed lanes

- Networks of priced managed lanes are being considered by several cities across the United States
  - Complex access structures
  - Complex tolling mechanisms driven by real-time sensor information

Austin region managed lane

Bay area managed lane

Model -based Evaluation Components

Components for modeling express lanes

Traveler choice models
Including lane-choice, route choice, and departure time choice

Traffic flow model
- Microscopic / mesoscopic
Captures interaction of vehicles, lane changes, and queue spillback

Demand model
- Deterministic or Stochastic
- Measured in real-time or assumed historic distribution

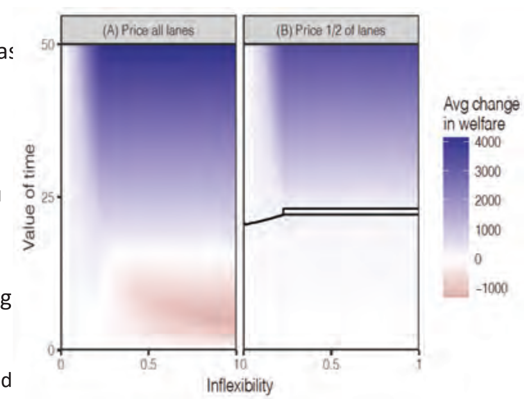
Toll pricing model (reinforcement learning framework)
Objective: minimize total system travel time (TSTT), maximize revenue, or others
Constraints: minimum and max toll; minimum speed limit on express lanes

Research goal

- How can we quantify and address the long-term equity impacts of express lanes?
  - Equity from land use/planning perspective
    - If not managed lanes, what infrastructure alternative will replace it and who will benefit from using it?
    - Tolling partial lanes may be better than tolling all lanes.
  - **Equity from pricing/operations perspective**
    - Certain travel groups might be expected to take the burden
- In this research, we focus on equitable design of tolls using a modeling focus
  - Through the lens of resource allocation, where the resource being distributed is access to low-travel time on the corridor.

Current literature on models for express lane pricing

- Various pricing methods have been proposed based on techniques such as
  - dynamic programming (Yang et al., 2012),
  - feedback integral control (Jin et al., 2020),
  - single-bottleneck model based analytical calculations (Hall, 2018),
  - dynamic traffic assignment-based optimization (Zhang et al., 2019),
  - model predictive control (Tan and Gao, 2020), and
  - deep-reinforcement learning (Pandey et al., 2020)
- Hall (2021) showed that Pareto efficiency is possible by tolling a portion of lanes even before spending the revenue towards public welfare
  - "No traveler is worse off"
  - Travelers with higher value of time and higher flexibility benefit the most



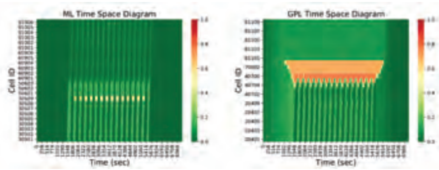
Hall, J. D. (2021). Can tolling help everyone? Estimating the aggregate and distributional consequences of congestion pricing. *Journal of the European Economic Association*, *19*(1), 441-474

Research objectives

- Model in Hall (2021) assumes single entrance and exit, and optimization objective of minimizing costs, that may not hold for current express lanes
- Revenue-maximization tolls may exhibit "jam-and-harvest" behavior[2]
- **Research questions:**
  - Quantify the unintended consequences and differential impacts of dynamic tolls and
  - Assess design of dynamic tolls and policies that prevent such unintended behavior

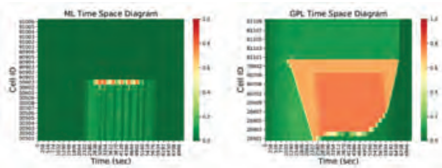
TSTT minimizing toll profile

Revenue=$1203.68, TSTT=1018.7 hr



Revenue maximizing profile

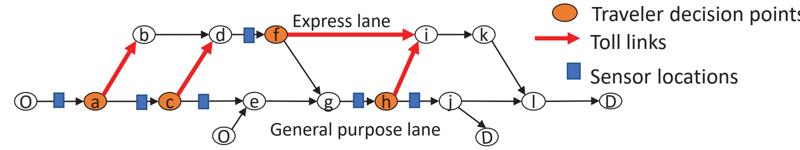
Revenue=$4106.03, TSTT=1421.05 hr



[1]Hall, J. D. (2021). Can tolling help everyone? Estimating the aggregate and distributional consequences of congestion pricing. *Journal of the European Economic Association*, *19*(1), 441-474

[2]Pandey, V., Wang, E., & Boyles, S. D. (2020). Deep reinforcement learning algorithm for dynamic pricing of express lanes with multiple access locations. Transportation Research Part C: Emerging Technologies, 119, 102715.

Modeling components for dynamic tolling



1. Demand model capturing traveler choices

- Choices: Travel or not (elastic demand); departure time; lane choice at each diverge
- Factors influencing choices:
  - Alternative specific: **toll, travel time** (revealed in real time)
  - Individual specific: Income (**value of time (VOT)**);

2. Supply model capturing interaction of demand with the infrastructure

- Incorporating queue spillback and traffic flow theory

- **Toll optimization:**
  - Dynamic tolls over a fixed period ● Real-time data driven ● Max and min toll constraints
  - Optimization objective: system optimal, revenue maximization

Our assumptions for equitable toll analysis

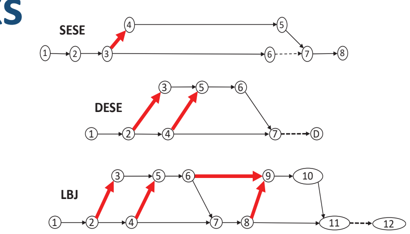
Public Engagement

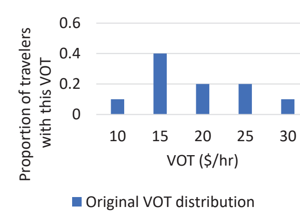
Identify population for analysis → Identify needs & concerns → Measure impacts of proposed projects → Determine disparities → Develop strategies to mitigate inequities

Identify population for analysis	Identify needs & concerns	Measure impacts of proposed projects	Determine disparities	Develop strategies to mitigate inequities
Travel groups with differing values of time	Travelers desire to have low travel time and spend least amount of time driving	Using a mesoscopic traffic simulation model for evaluating the impact of project	Measure the differential of delays across travel groups	Potentially through design of tolls, discounts, subscription-based coupons, and differential prices

Test Networks

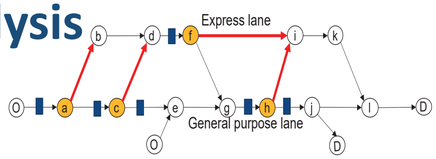
- Four abstract simulation networks considered
- Loop-1 express lanes data derived from the regional model
- Discrete value of time distribution assumed over 3 hours of peak period
  - Broadly, follows the Burr distribution observed in the literature
- For any 5-minute time period Max toll: $8, Min toll: 10 cents

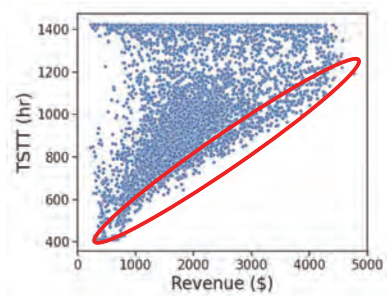




Simulation-based analysis

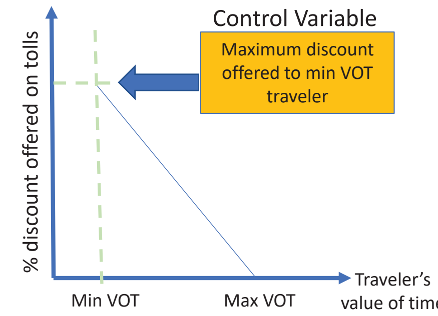
- Create a simulation-model for the corridor
  - Multiclass cell transmission model for traffic simulation
  - Lane choice governed by Binary Logit and decision-route models
- Simulate toll profiles varying in 5-minute intervals
  - Demand > Bottleneck capacity result in inevitable congestion
  - Different tolls choices may underutilize corridor capacity and result in additional congestion
- Analyze tolls on the Pareto frontier of Total system travel time (TSTT) and Revenue plots



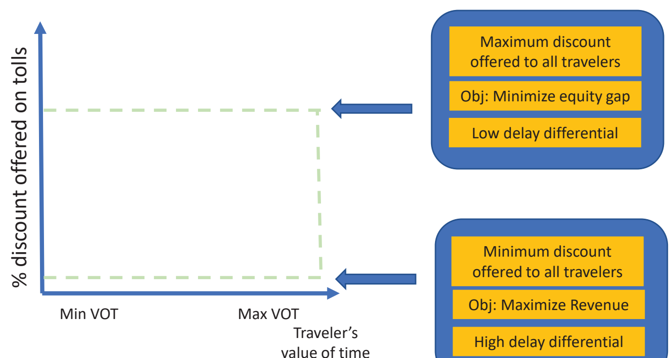


Modeling Discounts

- Broadly, the model for discount that we consider is the following:
  - » For any time-period t, select the discount $d_k(t)$ to give to group k
- Our preliminary model scales the discounts linearly by controlling the maximum discount offered to travelers with minimum VOT
- For demonstrations we narrow our focus to two groups of travelers with VOTs ($\alpha_{\max}$ and $\alpha_{\min}$)

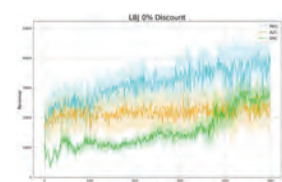


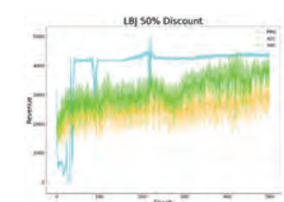
Modeling Discounts



We maximize revenue by minimizing discounts

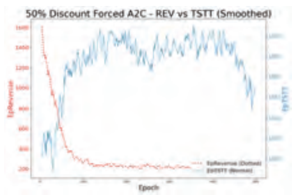
- When maximizing revenue, algorithm learns to offer little/no discount to everyone as that is an effective way to maximize revenue.
- No discount and 50% discount yielded similar revenue. We fix this by enforcing the discounts on the algorithm.





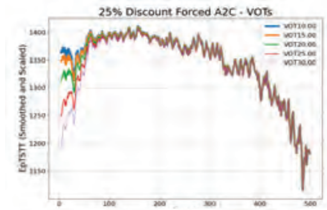
There exists a discount break-even point for efficiency

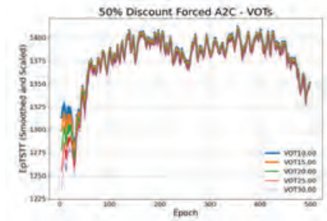
- Objectives minimizing equity gaps were suboptimal for revenue as algorithms aimed to minimize delay differential for all.
- Offering high discounts up to 50% proves ineffective on minimizing travel time as it prompts more individuals to opt for the hot lane, consequently increasing TSTT for all.



There exists a discount break-even point for efficiency

- For the same number of epochs, 25% discount experienced lower TSTT than 50%.
- This could be attributed to 25% being a conservative discount value that maintains an optimal balance of travellers in both hot lanes and GPL.





Approximating Price of Fairness

- Price of fairness defined as the approximate loss of "system efficiency" to allow for fair outcomes[1]
  - Measure it in terms of loss of revenue or increase of total system delay
- $POF_{Revenue} = \dfrac{\text{Loss of revenue for fair outcomes}}{\text{Best Possible Revenue}}$
  - From our experiments, the values is 0.91
- $POF_{TSTT} = \dfrac{\text{Increase in system delay for fair outcomes}}{\text{Lowest Possible Delay}}$
  - In our experiments, the values range from 0.69—0.79

How can we translate these findings to the real-world?

- How should these discounts be implemented?
  - Indirect and more acceptable alternatives can be chosen
  - Examples:
    - I-10 Express Lanes: low-income travelers receive $25 credit on transponders and monthly fees waived;
    - Public buses can use express lanes for free (I-95 Miami, FL)
    - "24-hours-free" discount publicity (NTE/LBJ Dallas, TX)
- Other real-world challenges:
  - Data-driven quantification of traveler's VOT
  - Extending these findings for real-world implementations of express lanes

Conclusion and Future Work

- **Conclusions:**
  - Through simulation-based analysis, we argue that the **choice of dynamic tolls impact the delay differentials** across different groups
  - Focusing solely on revenue optimization might worsen equity gap. Equity gap could be narrowed when optimizing for fairness but lead to significant revenue losses, which may be unsustainable.
  - Equitable access within ML is crucial, and VOT-based discounts could bridge equity gap

Ongoing work: Derive **numerical bounds on price of fairness** as a function of express lane characteristics, and integrate data-driven quantification of traveler's VOT in discount design

Questions/Comments?

Contacts:
Ridwan Tiamiyu: rotiamiyu@aggies.ncat.edu
Venktesh Pandey: vpandey@ncat.edu

Model Driven Equitable Discount Design for Express Lanes

Students(s): Ridwan Tiamiyu (MS), Christian Bowens (BS)
Advisors: Dr. Venktesh Pandey

Cross-Disciplinary Research Area: Reinforcement Learning

Overview

Source: https://www.hntb.com/Projects/I-15-Priced-Managed-Lanes

- **Background**:
  - Express lanes mitigate traffic congestion and provide reliable travel time by using the existing capacity of the roadway
  - As of January 2024, there are 70+ Express lane projects across the US
  - Currently deployed methods rely on look-up tables (based on speed, density, or occupancy measurements) or other heuristics to update toll with time, which may not be optimal per different objectives
  - Determining dynamic prices is a complex optimal control problem with multiple locations of trance/exits, complex driver behavior, and uncertainties in demand and travel time.
- **Literature on design of tolls:**
  - Various pricing methods have been proposed based on techniques such as dynamic programming (Yang et al., 2012), feedback integral control (Jin et al., 2020), single-bottleneck model (Hall, 2018), dynamic traffic assignment (Zhang et al., 2019), model predictive control (Tan and Gao, 2020), and deep-reinforcement learning (Pandey et al., 2020)
  - Göçmen et al. (2015) and Pandey et al. (2020) showed that certain toll profiles may have unintended consequences such as jam-and-harvest nature of revenue maximizing tolls
- **Research question:** How can we quantify and address the long-term equity impacts of express lanes from pricing/operations perspective?

Model-based Evaluation Components

Components for modeling express lanes

Traveler choice models
Including lane-choice, route-choice, and departure time choice

Lane choice modelled using value of time (VOT) distribution or binary logit model

Traffic flow model
- Microscopic / mesoscopic
Captures interaction of vehicles, lane changes, and queue spillback

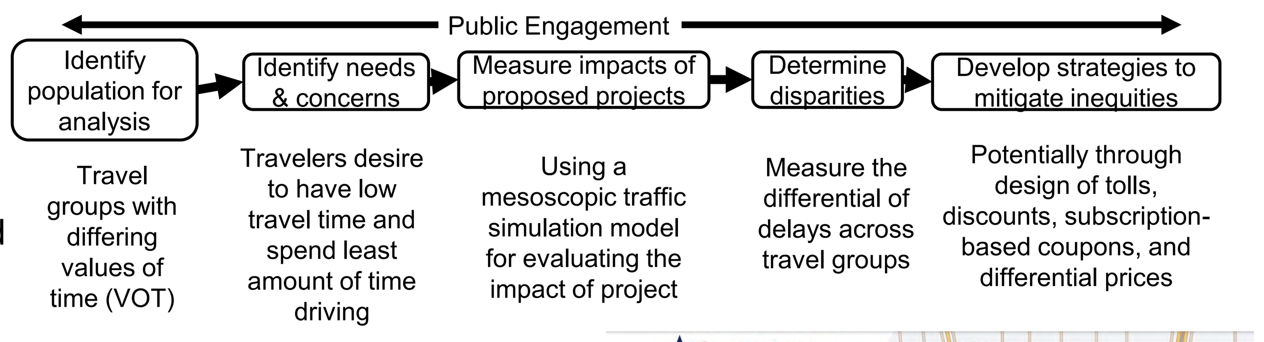
Demand model
- Deterministic or Stochastic
- Measured in real-time or assumed historic distribution

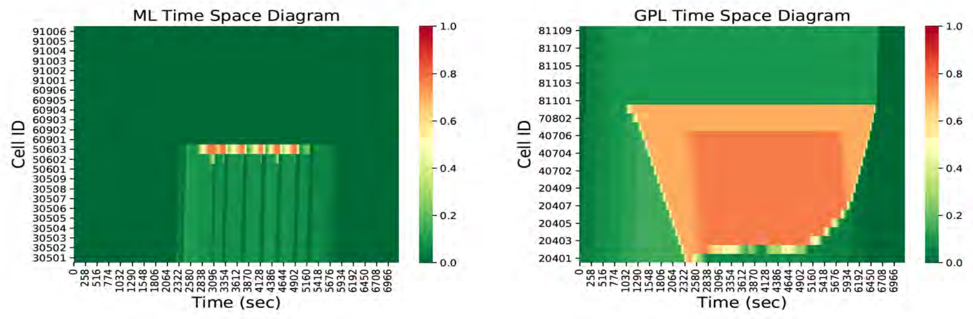
Toll pricing model (reinforcement learning framework)
Objective: minimize total system travel time (TSTT), maximize revenue, or others
Constraints: minimum and max toll; minimum speed limit on express lanes

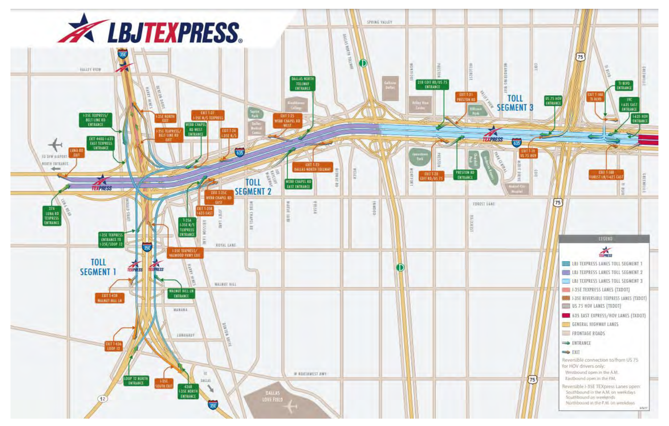
- **Modeling assumptions:** We consider a mesoscopic cell-transmission based traffic flow model with trapezoidal fundamental diagram which ignores impacts of lane changes. Furthermore, we assume that travelers do not equilibrate their route or time of departure. In this study, we conduct a simulation-based analysis of dynamic tolls. Relaxing some of these limiting assumptions is part of the ongoing work.

Equity Considerations

- NCHRP's five step procedure for analyzing equity concerns in transportation projects is shown along with components considered in this project







- "Jam-and-harvest" (JAH) nature of revenue maximizing tolls can contribute to inequities

Observations

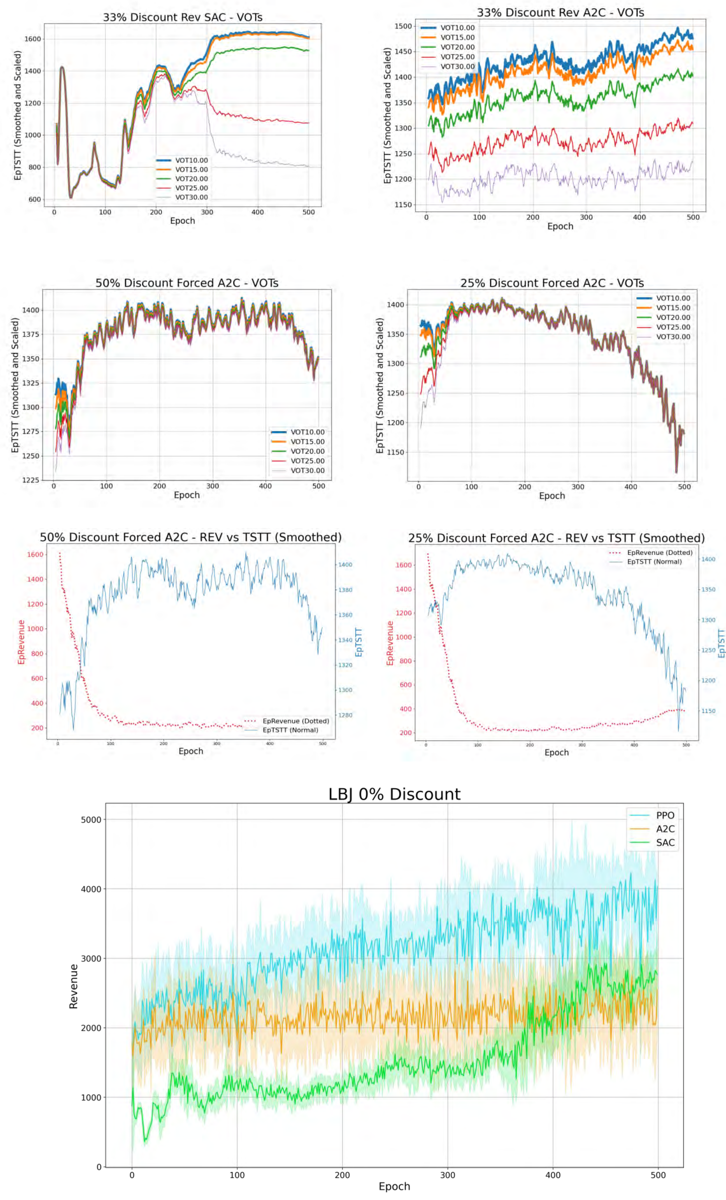
➢ **Maximizing revenue increases total system travel time (TSTT)** as express lanes aren't being used to its capacity and congestions build up.

➢ Reinforcement Learning dynamics and performance
- Soft Actor Critic (SAC) algorithm showed uniform TSTT across VOTs initially, diverging later, whereas A2C demonstrated variability, suggesting a learned preference for discounting higher VOTs as higher VOT travellers are more likely to use ML.
- Variable outcomes across different algorithms suggest influences from unique network characteristics, hyperparameter tuning, or algorithmic strengths in data complexity.

➢ Objectives minimizing equity gaps were suboptimal for revenue as algorithms aimed to minimize delay differential for all. This is because the algorithms prioritize an egalitarian approach to service levels.

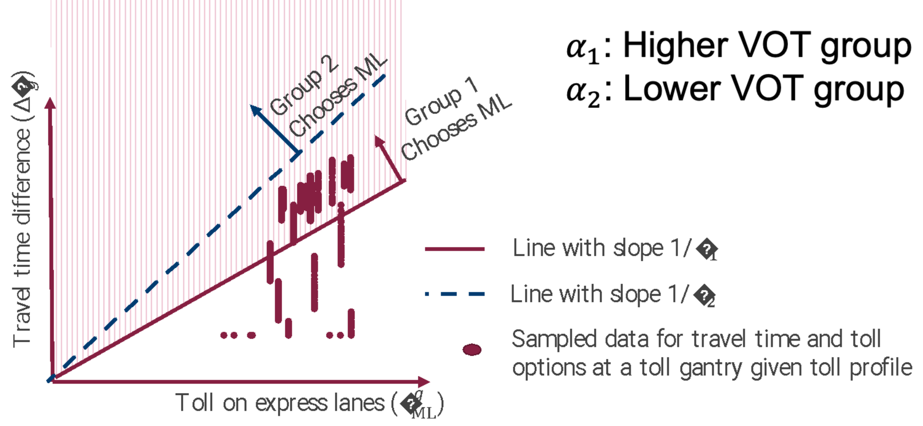
➢ **Offering very high discount encourages more travelers to use the ML thus increasing travel time for all.** A more conservative reduces TSTT and equity gap simultaneously.

➢ When maximizing revenue, the RL algorithm chose to minimize discount for all unless mandatory discounts were enforced. This is intuitive as giving discount to no one can maximize revenue and potentially benefit from JAH. However, when optimizing for equity, enforcing discounts doesn't have much effect.



Equitable discount design

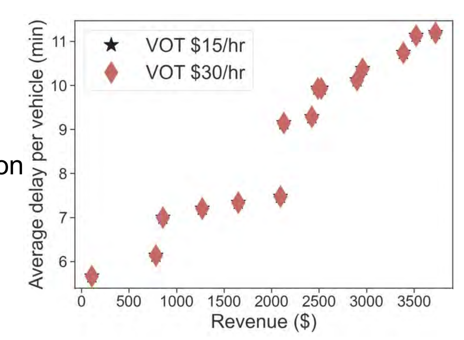
- We plot sampled data on travel time difference and toll on express lane observed for <u>revenue-maximizing tolls</u>
  - the realized values in the space only correspond to tolls that allow travelers in group 1 to choose ML, and travelers in group 2 can never access ML for the set value of tolls (which are too high for their range)
- Hence, we argue that by simply setting the discounts in proportion to travelers' VOT, the delay differentials across the groups are equitable.
  - If we assume that the traveler with the highest VOT receives no discount, then the following proposition establishes the discount for other VOTs.
- We can show that linear-varying discounts address delay differentials across groups
  - Such tolls can be made possible through technologies such as phone apps or personalization of tolls

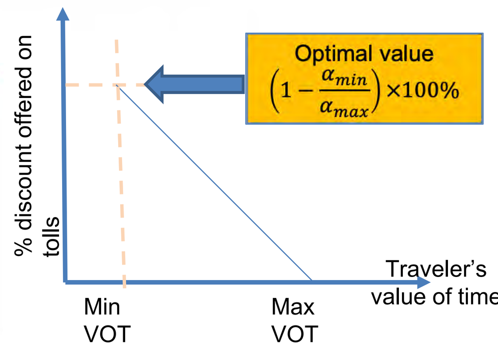


Proposition 1. *If we group travelers in increasing order of value of travel time, such that $\alpha_1 \geq \alpha_2 \geq \ldots \alpha_{|K|}$, then, offering a traveler in each group $k \in K$ a discount of*

$$d_k = \left(1 - \frac{\alpha_k}{\alpha_1}\right) \times 100\%$$

minimizes the delay differential.



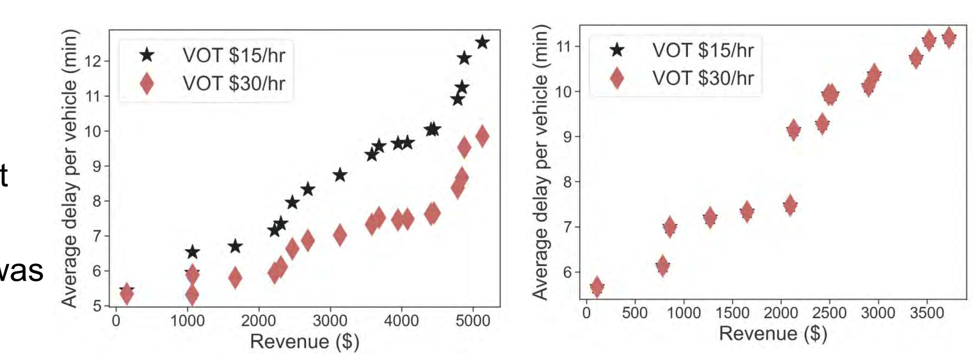


- Price of fairness (POF) defined as the approximate loss of "system efficiency" to allow for fair outcomes
- Measure it in terms of loss of revenue or increase of total system delay

$$POF_{Revenue} = \frac{\text{Loss of revenue for fair outcomes}}{\text{Best Possible Revenue}}$$

$$POF_{TSTT} = \frac{\text{Increase in system delay for fair outcomes}}{\text{Lowest Possible Delay}}$$

- In our experiments, the price of fairness wrt revenue was 0.91, while wrt TSTT was the most equitable value was $0.69 - 0.79$. This means that 91% of the expected revenue was lost to ensure equity. The travel time, also increased by 69% - 79%.



Conclusions and Ongoing Work

1. Conclusions:
   - Through simulation-based analysis, we argue that the **choice of dynamic tolls impact the delay differentials** across different groups
   - Focusing solely on revenue optimization might worsen equity gap. Equity gap could be narrowed when optimizing for fairness but lead to significant revenue losses, which may be unsustainable.
   - Equitable access within ML is crucial, and VOT-based discounts could bridge equity gap.
2. Implementation suggestions: Balanced approach to pricing.
   - For example: Establish a pricing strategy that balances the need for revenue generation with the goal of minimizing equity gaps, possibly by using a tiered system where discounts are based on VOT and income levels; Implementing a cap on minimum toll while using equity gap minimization model.
   - <u>Ongoing work</u>: Derive **numerical bounds on price of fairness** as a function of express lane characteristics, and integrate data-driven quantification of traveler's VOT in discount design

Acknowledgement: Support for this research is provided by the Center for Advanced Transportation Mobility, University Transportation Center. If you'd like any additional feedback or questions, please reach at rtiamiyu@aggies.ncat.edu