A Congestion Pricing Design with Shared Autonomous Mobility using Deep Reinforcement Learning

Jaehyung Lee* and Michele D. Simoni

Division of Transport and Systems Analysis, KTH Royal Institute of Technology, Sweden

Short summary

Shared autonomous mobility (SAM) may exacerbate congestion by increasing traveled distances, but dedicated congestion management strategies remain relatively underexplored. This study proposes a deep reinforcement learning framework to dynamically adjust congestion charges under SAM. To represent mixed traffic conditions (privately owned and shared autonomous vehicles), we integrate a multi-region Macroscopic Fundamental Diagram with a spatial dial-a-ride model. Building on this environment, a double deep Q-learning algorithm is applied to optimize congestion charges for a fixed cordon. The results show that the proposed framework can efficiently manage congestion. Furthermore, application results indicate that the introduction of SAM leads to higher congestion, as it increases vehicle total traveled distance.

Keywords: Congestion pricing, Deep reinforcement learning, Dial-a-ride model, Macroscopic fundamental diagram, Pricing strategies, Shared autonomous mobility

1 Introduction

With the advancement of autonomous driving technologies, autonomous vehicles are expected to be deployed on the roads in the near future. Due to the high initial cost of automation and potential benefits from reduced operating costs (Bösch et al., 2018), shared autonomous mobility (SAM) is likely to be more prevalent than privately owned vehicles. SAM has already been piloted in several cities worldwide and is expected to be commercialized soon, prompting extensive research on its potential impacts.

The impacts of SAM on traffic congestion are multifaceted. On the transport supply side, SAM is expected to enhance traffic flow by reducing delays and human errors (Bertoncello & Wee, 2015; Park et al., 2021; Yu et al., 2023). On the transport demand side, SAM could improve accessibility particularly in areas with limited access to public transit and for people who are unable to drive, such as adolescents, the elderly, people with disabilities (Fagnant & Kockelman, 2015). In addition, SAM may encourage longer and more frequent trips by reducing the burden of driving (Truong et al., 2017; Paithankar et al., 2025; Milakis et al., 2018). Meanwhile, SAM can replace private car usage (Duarte & Ratti, 2018; Fagnant & Kockelman, 2018), but they also induce additional empty trips, such as traveling to depots and picking up passengers. The overall finding suggests that SAM may worsen road congestion, by increasing total vehicle kilometers traveled by 10% or more (Fagnant & Kockelman, 2015; Litman, 2025).

SAM can alter transportation systems by increasing the number of trips and traveled distances, potentially influencing the effectiveness of the current congestion pricing, one of the most established transport demand management strategies. Nevertheless, congestion pricing under SAM remains underexplored. Previous research has primarily adopted scenario-based approaches, using agent-based simulations to capture complex interactions between SAM and congestion management policies (Lee et al., 2026; Kaddoura et al., 2020; Simoni et al., 2019). These studies emphasize that the need to regulate SAM since it can make current congestion pricing less effective. However, due to the high computational cost of simulation, they are limited in practically implementing dynamic congestion pricing scenarios with SAM.

Although reinforcement learning has been widely applied to optimal control problems, including traffic signal control, vehicle dispatching, and routing, little attention has been paid to its application for dynamic congestion pricing due to complex network dynamics (Vermeulen et al., 2026). Lu et al. (2024) developed a multi-agent transformer-based dynamic tolling system coupled with a microscopic simulation environment. Parishad et al. (2025) proposed a dynamic congestion pricing

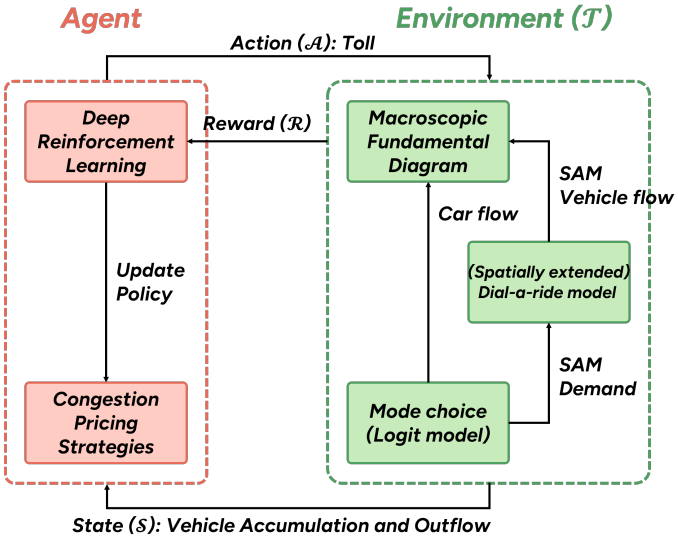


Figure 1: A deep reinforcement learning framework for adjusting congestion pricing

in multi-modal networks based on a trip-based Macroscopic Fundamental Diagram (MFD). They demonstrate the effectiveness of their models. However, they do not account for SAM despite its potential to induce congestion.

This study develops a congestion pricing scheme under SAM using deep reinforcement learning. A multi-region MFD is employed to represent traffic condition. To estimate the traffic flow of SAM vehicles given hourly demand, we adopt a spatial dial-a-ride model proposed by Park et al. (2023). Building on this experimental environment, a deep reinforcement learning agent optimizes congestion charges for a fixed cordon. The proposed framework enables the efficient development of dynamic congestion pricing schemes accounting for SAM while reducing computational burden.

The remainder of this paper is organized as follows. Section 2 presents a deep reinforcement learning framework for cordon-based congestion pricing. Section 3 describes the experimental setup. Section 4 presents results, and Section 5 demonstrates an application of the suggested model. Finally, Section 6 concludes the paper.

2 Modeling framework

Fig. 1 describes a deep reinforcement learning framework for congestion pricing. Multi-modal transport dynamics, including cars, transit, and SAM, are represented through the following processes. First, a logit model estimates mode shares. Then, the dial-a-ride model converts SAM passenger demand into vehicle flows between regions. Finally, the MFD captures traffic dynamics based on car demand and SAM vehicle movements. A deep reinforcement learning agent dynamically adjusts congestion charges at each time step by interacting with the environment.

The dynamic adjustment of congestion charges can be formulated as a Markov decision process, represented by the tuple ($\mathcal{S}$, $\mathcal{A}$, $\mathcal{R}$, $\mathcal{P}$, γ), defined as follows:

- State ($\mathcal{S}$): The state represents the network traffic conditions, including traffic flow and the cumulative number of vehicles in each zone.
- Action ($\mathcal{A}$): The action adjusts congestion charges t each timestep by increasing or decreasing them by a fixed amount, or maintaining them.
- Reward ($\mathcal{R}$): The reward is defined as the weighted sum of the total number of accumulated vehicles at each timestep, with greater emphasis on vehicle accumulation within the cordon. Furthermore, a penalty is imposed on the toll amount at each timestep to prevent excessively high charges. toll amount at each timestep is penalized to prevent excessive toll.
- Transition ($\mathcal{T}$): Transition represents the transport system dynamics based on a macroscopic fundamental diagram and a dial-a-ride model, presented in the following subsections.
- Discount factor (γ): The discount factor is set to 0.75, emphasizing current traffic conditions while still accounting for future conditions.

To solve the Markov decision process and learn optimal congestion pricing strategies, we employed a double deep-Q learning algorithm to mitigate overestimation bias by separating action selection and evaluation (Van Hasselt et al., 2016).

$$y_t^{DDQN} = R_{t+1} + \gamma Q(S_{t+1}, \text{argmax}_a Q(S_{t+1}, a; \theta_t); \theta_t^-) \tag{1}$$

where y^{DDQN} denotes the target values. Q represents the action-value function estimating the expected sum of discounted future rewards. θ and θ^- are the parameters of the online and target network, which determines and evaluates action respectively. To improve stability in training, a replay buffer is employed to store past experiences for sampling during training.

Mode choice

The transition of the Markov decision process consists of three sequential steps, capturing the evolution of transport dynamics given the current network conditions, an origin-destination matrix, and congestion pricing control at each timestep. First, a logit-based choice model estimates an origin-destination matrix for each mode. Second, a dial-a-ride model estimates SAM vehicle flows using SAM demand. Finally, the multi-region MFD estimates network conditions given the car demand and SAM vehicle flows. The utility functions of car, transit, and SAM are expressed as follows, similar to Parishad et al. (2025):

$$\begin{cases} U_{car} = \tau(\beta_{car}^{ivt} \cdot ivt_{car} + \beta_{car}^{cost} \cdot cost_{car}) \\ U_{transit} = \tau(\beta_{transit}^{ivt} \cdot ivt_{transit} + \beta_{transit}^{wait} \cdot wait_{transit} + \beta_{transit}^{cost} \cdot cost_{transit}) \\ U_{SAM} = \tau(\beta_{SAM}^{ivt} \cdot ivt_{SAM} + \beta_{SAM}^{wait} \cdot wait_{SAM} + \beta_{SAM}^{cost} \cdot cost_{SAM}) \end{cases} \tag{2}$$

where ivt, $wait$, and $cost$ represent in-vehicle time, waiting time, and travel cost, respectively. β denotes the utility parameters for each mode and attribute. τ indicates a scale parameter associated with the variance of unobserved utility.

Dial-a-ride model

Dial-a-ride model is a generalized model of on-demand mobility services describing operation status of on-demand mobility services using a queuing network, firstly proposed by Daganzo & Ouyang (2019). Park et al. (2023) extended this model considering spatial heterogeneity in travel demand and balancing demand and supply including relocation problem in the dial-a-ride model.
In the spatially extended dial-a-ride model proposed by Park et al. (2023) (Fig. 2), each vehicle belongs to one of four workload states within each zone and transitions between these states over time: idle, assigned, serving, and relocation. These states can be expressed by a vector describing the number of serving passengers, the number of assigned passengers, relocation state, and the zone index.

- Idle (0,0,0,i): Vehicle is wandering without serving any passengers and in zone i
- Assigned (0,1,0,i): Vehicle is assigned to a passenger and moving to pick-up location in zone i
- Serving (1,0,0,i): Vehicle is serving a passenger originating from zone i
- Relocation (0,0,1,i): Vehicle is relocating and moving from zone i

A spatially dial-a-ride model is generally used to determine the optimal fleet size to achieve waiting time T_w^i, given the area and demand density of each zone and their distribution, also providing the vehicle flow between regions. Table 1 presents the input and output variables of the spatially extended dial-a-ride model.
Based on the flow conservation of each zone, the steady-state solution can be derived from Little's formula as follows:

$$\begin{cases} n_{00}^i \approx \kappa R_i / (v T_w^i)^2 \\ n_{01}^i = p_{01}^i \times T_w^i \\ n_{10}^{ij} = d_{10}^{ij}/v = \lambda_i R_i \mu_{ij} l_{ij}/v \\ n_r^{ij} = r_{ij} l_{ij}/v \end{cases} \tag{3}$$

For further details, see Daganzo & Ouyang (2019) and Park et al. (2023).

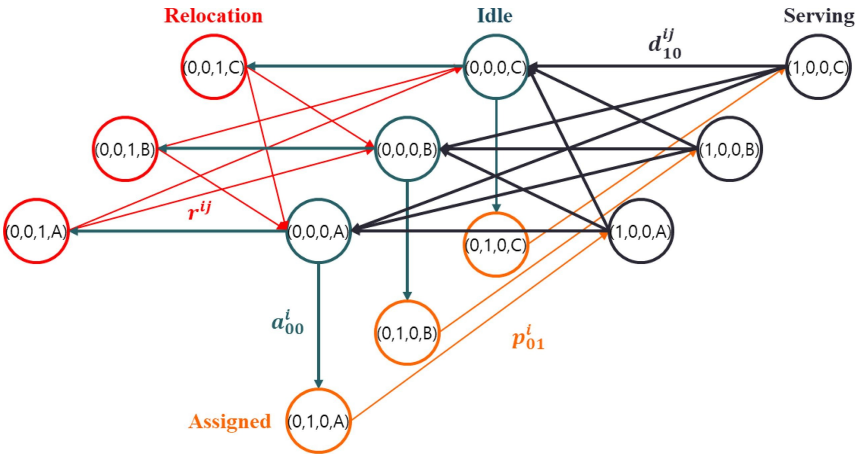


Figure 2: Workload transition network for on-demand mobility services (Source: Park et al. (2023))

Table 1: Input and output variables of workload transition network for on-demand mobility services

Category	Variable	Unit	Definition
Input	R_i	km^2	Area of zone i
	λ_i	trips/(h·km^2)	Demand density of zone i
	μ_{SAM}^{ij}	-	Ratio of trips with origin i and destination j to total trips from i
	T_w^i	h	Average waiting time of zone i
	v	km/h	Average vehicle speed i
	κ	-	Network shape constant i
Output	a_{00}^i	veh/h	Transition rate from "idle" to "assigned" state in zone i
	p_{00}^i	veh/h	Transition rate from "assigned" to "serving" state in zone i
	d_{00}^i	veh/h	Flow of serving vehicles moving from origin i to destination j
	r_{00}^i	veh/h	Flow of relocation vehicles moving from origin i to destination j
	n_{00}^i	veh	Number of idle vehicles in zone i
	n_{01}^i	veh	Number of assigned vehicles in zone i
	n_{10}^{ij}	veh	Number of serving vehicles moving from origin i to destination j
	n_r^{ij}	veh	Number of relocation vehicles moving from origin i to destination j

Multi-region macroscopic fundamental diagram

The MFD–motivated by fundamental diagram–represents the network traffic condition at the regional level by aggregating links with homogeneous traffic context (Godfrey, 1969). The MFD can be extended to multiple heterogeneous regions by accounting for inter-regional transfer flows (Geroliminis et al., 2012). The temporal dynamics of vehicle accumulation in each region can be expressed as follows:

$$\begin{cases} \frac{dn^{ii}(t)}{dt} = q^{ii}(t) - M_i^{ii}(t) - \sum_{h\in V^i} M_h^{ii}(t) + \sum_{h\in V^i} M_i^{hi}(t) \quad (\text{for } i=j,\ i,j\in R) \\ \frac{dn^{ij}(t)}{dt} = q^{ij}(t) - \sum_{h\in V^i} M_h^{ij}(t) + \sum_{h\in V^i} M_j^{hj}(t) \quad (\text{for } i\neq j,\ i,j\in R) \end{cases} \tag{4}$$

where $q^{ij}(t)$ and $N^{ij}(t)$ represent traffic flow and vehicle accumulation generating in region i with destination j at time t. $V^i(t)$ denotes the set of neighbors of region i at time t. $M_i^{ii}(t)$ indicates the internal trip completion rates in region i at time t. $M_h^{ij}(t)$ represents the transfer flow originating from region i and destining for region j at time t.
Internal trip completion rates and transfer flows are derived as follows:

$$M_i^{ii}(t) = \theta_i^{ii} \cdot \frac{N^{ii}(t)}{N^i(t)} \cdot \frac{F^i(N^i(t))}{L^{ii}(t)} \tag{5}$$

$$M_h^{ij}(t) = \min\left[\theta_h^{ij} \cdot \frac{N^{ij}(t)}{N^i(t)} \frac{F^i(N^i(t))}{L^{ih}(t)}, C^{hi}(N^h(t)) \cdot \frac{N^i j(t)\cdot\theta_h^{ij}}{\sum_{k\in R} N^{ik}(t)\cdot\theta_h^{ik}(t)}\right] \tag{6}$$

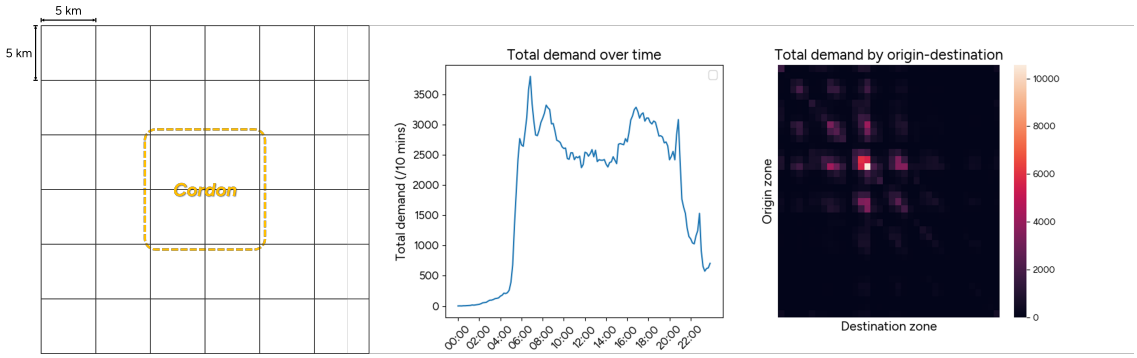


Figure 3: Grid network and demand distribution

where $F^i(\cdot)$ is the production MFD of region i, a function of the vehicle accumulation in region i, $N^i(t) = \sum_{j \in R} N^{ij}(t)$. L indicates the average trip length corresponding to travel between superscripts. θ indicates the ratio of outflow from region i with destination j to transfer neighbor region h. Multi-region MFD estimates network traffic conditions based on the car flow and SAM vehicle flow derived from mode choice and the dial-a-ride model.

3 Experiments

A hypothetical 6 × 6 grid network is considered, with the cordon defined as 2 × 2 grid in the center, where congestion charges are applied when crossing the cordon. An origin-destination matrix with a 10 minute time interval is generated to account for urban traffic patterns, including congestion within the cordon and peak-hour patterns. A grid network and demand distribution is represented in Fig. 3. We assumed that the impacts of public transit on traffic flow is negligible, given its high capacity.

Each traveler considers in-vehicle time, waiting time, and fare. In this experiment, the waiting time for car is set to zero, while the waiting time for transit and SAM is set to 5 minutes. The transit fare is fixed at €2.5. The SAM fare consists of a fixed fare of €4 and a distance-based cost of €0.4/km. Travel speeds for private cars and SAM are determined based on the MFD, whereas the travel speed for transit is fixed at 25 km/h. Mode choice parameters used in this study are presented in Table 2, with the value of travel time for cars set to €12.5/h. The parameters associated with SAM are based on (Huda et al., 2023).

Table 2: Mode choice parameters

Parameter	**Transport mode**	**Value**
β^{ivt}	Car	-0.1
	Transit	-0.08
	SAM	-0.07
β^{wait}	Car	-
	Transit	-0.15
	SAM	-0.15
β^{cost}	Car	-0.04
	Transit	-0.04
	SAM	-0.04
Scale parameter	Generic	0.1

For the dial-a-ride model applied to the grid network, each zone has an area of $25km^2$, and κ is set to 0.63. The target average waiting time is 5 minutes. MFD production function is set to a three-degree polynomial as $5.0 \times 10^{-7} n^3 - 0.15n^2 + 50.0n$ $(veh \cdot km/h)$, where n indicates the accumulated vehicles in region. Hyperparameter setting for the double deep-Q learning is presented in Table 3, similar to Parishad et al. (2025).

Table 3: Hyperparameters for the double deep-Q network

Parameter	**Value**
Number of iterations	1,000
Replay buffer size	1,440
learning rate	1×10^{-4}
Batch size	64
Discount factor	0.9

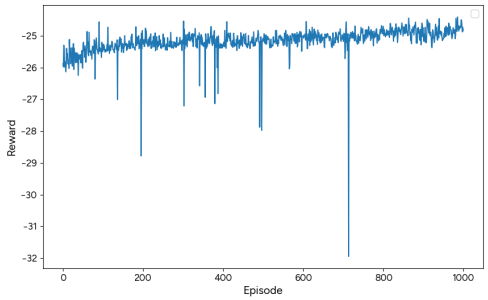


Figure 4: Episode reward

4 Results

Fig. 4 shows the training results of the agent, showing that the learning process becomes relatively stable after 800 iterations. After training, the optimal congestion pricing policy derived by the deep reinforcement learning agent is presented in Fig. 5, yielding a lower reward compared to the scenario without congestion pricing (Fig. 6). Although peak hour demand is slightly higher in the morning, the longer duration of the peak period leads to higher congestion charges compared to the morning peak.

5 Applications: Impacts of SAM on congestion pricing

Fig. 7 presents the travel demand for SAM and the corresponding vehicle flow in the network, estimated by the dial-a-ride model. To maintain an average waiting time of less than 5 minutes, about half of vehicles are deployed compared to demand. In particular, during the morning peak hour, when demand is most concentrated, vehicle relocation becomes more active to efficiently meet demand while limiting the increase in total number of deployed SAM fleets.

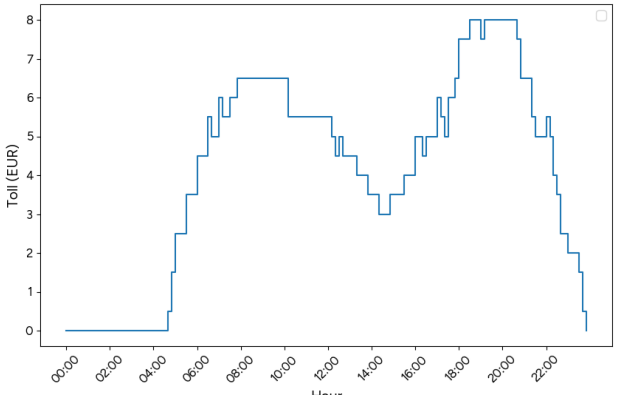


Figure 5: Congestion charges by time-of-day

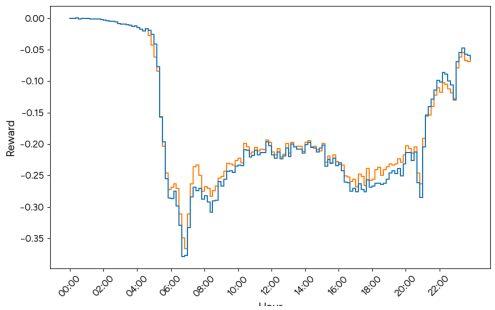


Figure 6: Comparison of reward under scenarios with and without congestion charges by time-of-day

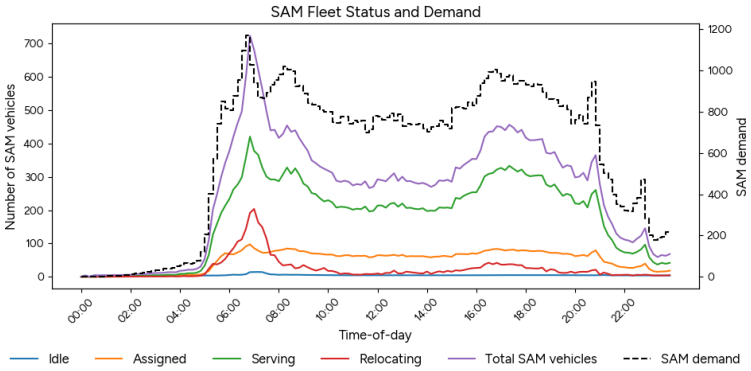


Figure 7: The number of SAM vehicles by fleet status and demand

Fig. 8 describes the average mode share and the total traveled distance of cars and SAM under scenarios with and without SAM. Since the introduction of SAM absorbs transit trips, total traveled distance increases under SAM, consistent with (Lee et al., 2026). Furthermore, although the number of SAM vehicles is lower than SAM demand, SAM vehicles are continuously operating unlike private cars, thereby affecting network conditions more substantially than private cars (Fig. 9).

6 Conclusions

This paper proposes a deep reinforcement learning framework to dynamically adjust congestion pricing under SAM. To capture mixed traffic conditions between private cars and SAM, we integrated a MFD with a dial-a-ride model. The result shows that the proposed framework can effectively adjust congestion pricing. Furthermore, application results reveal that SAM influences congestion pricing by increasing the number of vehicles, consistent with previous studies based on agent-based simulations with higher model complexity and computational cost (Lee et al., 2026; Kaddoura et al., 2020; Simoni et al., 2019).

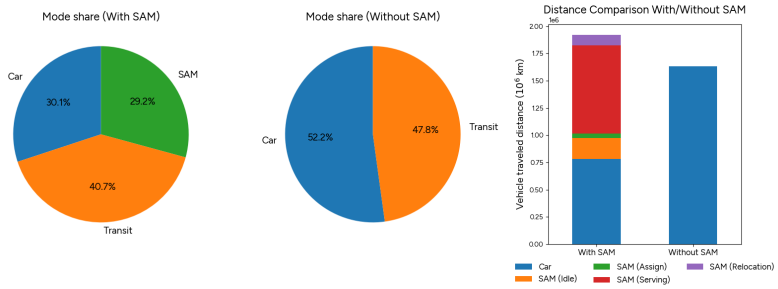


Figure 8: Mode share and traveled distances under scenarios with and without SAM

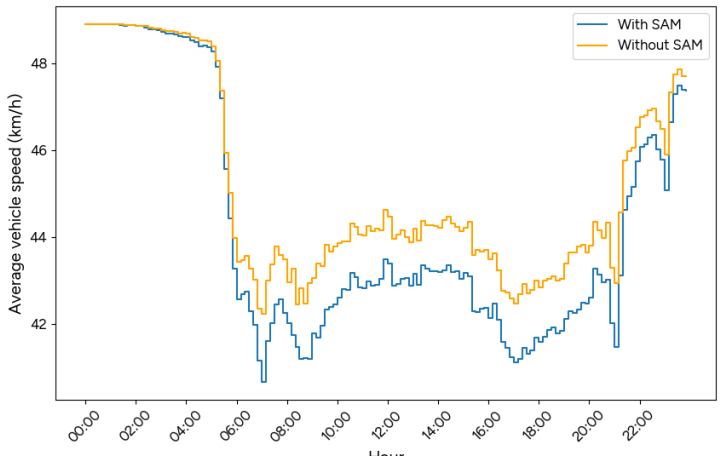


Figure 9: Average vehicle speed under scenarios with and without SAM

Despite these contributions, the proposed framework has some limitations. First, since the model is not a path-based, it is challenging to fully capture the effects of congestion pricing. Furthermore, the proposed framework share the limitations of the dial-a-ride model. In particular, the spatially-extended dial-a-ride model is limited to a passenger capacity of one due to the complexity of accommodating multiple passengers. These limitations can be addressed in future research. First, a trip-based MFD could better represent network condition by explicitly modeling vehicle arrivals. Second, the dial-a-ride model could be extended to accommodate multiple passengers.

References

Bertoncello, M., & Wee, D. (2015). *Ten ways autonomous driving could redefine the automotive world* (Tech. Rep.). McKinsey & Company.

Bösch, P. M., Becker, F., Becker, H., & Axhausen, K. W. (2018). Cost-based analysis of autonomous mobility services. *Transport Policy*, *64*, 76–91.

Daganzo, C. F., & Ouyang, Y. (2019). A general model of demand-responsive transportation services: From taxi to ridesharing to dial-a-ride. *Transportation Research Part B: Methodological*, *126*, 213–224.

Duarte, F., & Ratti, C. (2018). The impact of autonomous vehicles on cities: A review. *Journal of Urban Technology*, *25*(4), 3–18.

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167–181.

Fagnant, D. J., & Kockelman, K. M. (2018). Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in austin, texas. *Transportation*, *45*(1), 143–158.

Geroliminis, N., Haddad, J., & Ramezani, M. (2012). Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. *IEEE Transactions on Intelligent Transportation Systems*, *14*(1), 348–359.

Godfrey, J. (1969). The mechanism of a road network. *Traffic Engineering & Control*, *8*(8).

Huda, F. Y., Currie, G., & Kamruzzaman, M. (2023). Understanding the value of autonomous vehicles – an empirical meta-synthesis. *Transport reviews*, *43*(6), 1058–1082.

Kaddoura, I., Bischoff, J., & Nagel, K. (2020). Towards welfare optimal operation of innovative mobility concepts: External cost pricing in a world of shared autonomous vehicles. *Transportation Research Part A: Policy and Practice*, *136*, 48–63.

Lee, J., Jonsson, D., & Simoni, M. D. (2026). Revisiting congestion pricing under shared autonomous mobility: Efficiency and distributional implications. *Available at SSRN 6187659*.

Litman, T. (2025). *Autonomous vehicle implementation predictions: Implications for transport planning* (Tech. Rep.). Victoria Transport Policy Institute.

Lu, J., Hong, C., & Wang, R. (2024). Magt-toll: A multi-agent reinforcement learning approach to dynamic traffic congestion pricing. *PLOS ONE*, *19*(11), 1-23.

Milakis, D., Kroesen, M., & van Wee, B. (2018). Implications of automated vehicles for accessibility and location choices: Evidence from an expert-based experiment. *Journal of Transport Geography*, *68*, 142–148.

Paithankar, P., Vellimana, M., & Kockelman, K. M. (2025). Impact of autonomous vehicles on long-distance travel mode and destination choices in texas. *Research in Transportation Economics*, *110*, 101521.

Parishad, N., Yildirimoglu, M., & Hickman, M. (2025). Congestion pricing in multi-modal networks: An application of deep reinforcement learning. *Transportation Research Part C: Emerging Technologies*, *177*, 105166.

Park, J., Byun, W., Kim, Y., Ahn, H., & Shin, D. (2021). The impact of automated vehicles on traffic flow and road capacity on urban road networks. *Journal of Advanced Transportation*, *2021*(1), 8404951.

Park, J., Lee, J., Kim, J., & Chung, J.-H. (2023). An optimization model of on-demand mobility services with spatial heterogeneity in travel demand. *Transportation Research Part C: Emerging Technologies*, *153*, 104229.

Simoni, M. D., Kockelman, K. M., Gurumurthy, K. M., & Bischoff, J. (2019). Congestion pricing in a world of self-driving vehicles: An analysis of different strategies in alternative future scenarios. *Transportation Research Part C: Emerging Technologies*, *98*, 167–185.

Truong, L. T., De Gruyter, C., Currie, G., & Delbosc, A. (2017). Estimating the trip generation impacts of autonomous vehicles on car travel in victoria, australia. *Transportation*, *44*(6), 1279–1292.

Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 30).

Vermeulen, O., Siebes, A., & Velegrakis, Y. (2026). Reinforcement learning for road pricing: a review and future directions. *Artificial Intelligence Review*, *59*(2), 65.

Yu, Q., Wu, L., Zhang, H., Lei, L., & Wang, L. (2023). The impact of automated vehicles on road and intersection capacity. *Applied Sciences*, *13*(8), 5073.