Inequality in Congestion Games with Learning Agents

Extended Abstract

Dimitris Michailidis
University of Amsterdam
The Netherlands
d.michailidis@uva.nl

Sennay Ghebreab
University of Amsterdam
The Netherlands
s.ghebreab@uva.nl

Fernando P. Santos
University of Amsterdam
The Netherlands
f.p.santos@uva.nl

ABSTRACT

Who benefits from expanding transport networks? While designed to improve mobility, such interventions can also create inequality. We show that disparities arise not only from the structure of the network itself but also from differences in how commuters adapt to it. We model commuters as reinforcement learning agents who adapt their travel choices at different learning rates, reflecting unequal access to resources and information. We introduce the Price of Learning (PoL), a measure of inefficiency during learning. We analyze both a variation of the Braess Paradox original network and an abstraction of the Amsterdam metro network. Our simulations show that expansions can simultaneously increase efficiency and amplify inequality, especially when faster learners benefit from new routes before others adapt. We highlight that transport policies must account not only for equilibrium outcomes but also for the heterogeneous ways that commuters adapt.

KEYWORDS

Congestion Games, Reinforcement Learning, Multi-Agent Systems, Fairness, Transport Planning

ACM Reference Format:
Dimitris Michailidis, Sennay Ghebreab, and Fernando P. Santos. 2026. Inequality in Congestion Games with Learning Agents: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/PRIV8934

1 INTRODUCTION

Expanding transport networks is often seen as an unambiguous good: more roads, more transport lines, more options means faster travel. Yet, in practice, such interventions can backfire. Sometimes, adding infrastructure slows traffic [5]. Other times, it benefits advantaged areas, widening inequalities [10]. Consider a city rolling out a new public transportation line intended to reduce crowdedness. What happens when some commuters adapt faster than others? A transport upgrade that deepens divides.

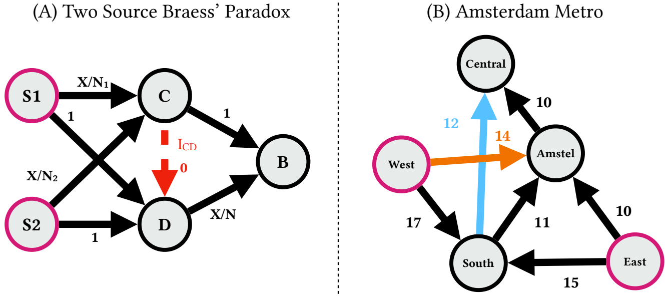

Figure 1: (A) Two-Source Braess' Paradox with fast-lane extension (red). (B) Amsterdam Metro with North-South (blue) and West-Amstel (orange) extension.

Classical game theoretic models, like the Braess's Paradox [5], elegantly show that adding new routes can increase travel times. Empirically, the phenomenon has been observed in New York, Boston, and Seoul [1, 16]. These models, however, suffer from a critical blind spot: they assume homogeneous agents, fully rational and with perfect information, ignoring how real commuters experiment and adapt. In reality, commuters ignore detailed traffic conditions and balance exploration (trying new routes) with exploitation (relying on known ones) [3, 14], through processes resembling Reinforcement Learning (RL). Such dynamics do not always converge to Nash Equilibria[15], suggesting that learning models provide a more realistic assessment of system performance over time.

While research into RL-based adaptation in route-choice games is growing [3, 6, 14], current work disregards the role of agents' learning heterogeneity in the emergence of unfair outcomes. In reality, however, empirical evidence reveals that wealthier individuals exhibit higher risk tolerance and more exploratory behavior [4, 8], which suggests models where agents differ in their adaptation capacity. In this paper, we investigate how exploration-driven learning shapes distributional outcomes. We tackle the question: **How do differences in learning affect fairness and efficiency in congestion games?**

We introduce a framework combining multi-agent RL with fairness analysis: We propose: (i) the Price of Learning (PoL): a dynamic analogue to the Price of Anarchy that quantifies inefficiency during learning, (ii) a two source Braess' Network: a proof of concept for how heterogeneous learning amplifies inequalities, and (iii) the Amsterdam Metro: a real-world abstraction to examine commuter adaptation under network expansions [1].

We show that interventions can worsen inequality: faster-learning groups capture greater benefits, leaving slower learners disadvantaged, even when overall efficiency improves. We highlight the need to move beyond static analysis to consider how heterogeneity shapes long-term outcomes.

---

[1] GitHub code: https://github.com/dimichai/fairness-braess

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/PRIV8934

2 FRAMEWORK: LEARNING AND FAIRNESS

Price of Learning. In congestion games, the *Social Optimum (SO)* is the strategy profile that minimizes total cost: $a^\star \in \arg\min_{a\in\mathcal{A}} C(a)$.

A *policy* is the function that defines an agent's action at each timestep t. Let $\pi_t = (\pi_t^1, \dots, \pi_t^n)$ denote the *joint policy*. The expected social cost is $C(\pi_t) = \mathbb{E}_{A_t\sim\pi_t}[C(A_t)]$. We introduce the *Price of Learning (PoL)*:

$$\mathrm{PoL}(t) = \frac{C(\pi_t)}{C(a^\star)} \geq 1. \tag{1}$$

Measuring PoL captures the gap between static equilibrium results and the outcome of learning-based dynamics.

Fairness. We assess fairness through *source disparity*, the difference in average cost between player groups originating from different sources:

$$\mathrm{SD}(s_1, s_2) = \mathrm{AvgCost}(s_1) - \mathrm{AvgCost}(s_2). \tag{2}$$

A positive SD favors s_2, negative favors s_1, and zero indicates perfect fairness.

Reinforcement Learning. Each player i seeks to minimize experienced cost (travel time plus congestion) by updating estimates of expected cost for each strategy $a_i \in \mathcal{A}_i$, expressed through a Q-value table $Q_i(a_i)$, and updated in every timestep:

$$Q_i(a_i) \leftarrow Q_i(a_i) + \alpha\left[r_i(t) + \gamma \max_{a'_i\in\mathcal{A}_i} Q_i(a'_i) - Q_i(a_i)\right], \tag{3}$$

where $\alpha \in (0, 1]$ is the learning rate, $\gamma \in [0, 1]$ the discount factor, and the observed reward is $r_i(t) = -\sum_{e\in a_i(t)} f_e(x_e(a_t); K_e)$, with $a_i(t)$ the strategy chosen at time t and a_t the realized profile [6, 7]). Agents follow an ϵ-greedy policy to balance exploration and exploitation.

3 ENVIRONMENTS

Two-Source Braess Network: An extension of the classical paradox with two sources (S_1, S_2) and one destination. We test a fast lane intervention (I_{CD}), as shown in Figure 1 (A).

Amsterdam Metro Network: An abstraction of the real network through three phases: pre-2018, current North-South line, and future West-Amstel expansion, as shown in Figure 1 (B). The social cost for each edge is calculated as $f_e(x_e; K_e) = t_e^0\left(1 + \frac{x_e}{K_e}\right)$, where t_e^0 is the free-flow travel time and $K_e = N$ is the edge capacity. Free-flow times are calculated from real-world schedules [9, 13].

4 RESULTS

Figure 2 (A) shows that equal learning rates lead to fair outcomes (SD $\approx$ 0). However, unequal learning rates (Figure 2 - B, Figure 3 -B) allow faster learners to achieve persistent lower travel times, even if the system's overall PoL remains low.

Network expansions interact with adaptation to exacerbate disparities. Figure 3 shows that in the Amsterdam Metro (Phase C), unequal learning rates can block convergence to the Nash Equilibrium entirely, creating lasting inefficiencies and inequities that static analysis fails to predict.

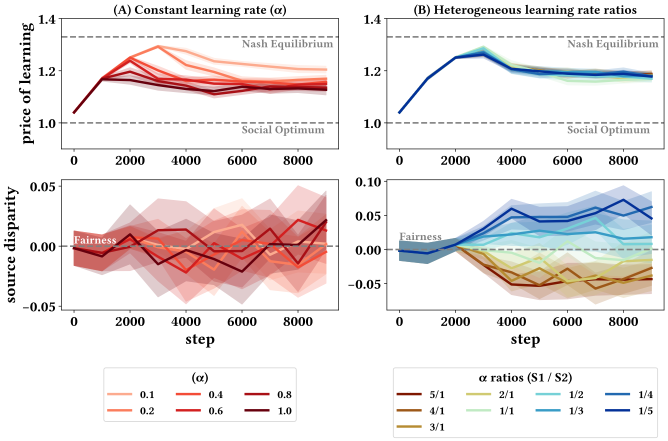

Figure 2: With equal learning rates, source disparity fluctuates around a fair zero. With unequal rates, agents with higher learning rates gain a persistent advantage.

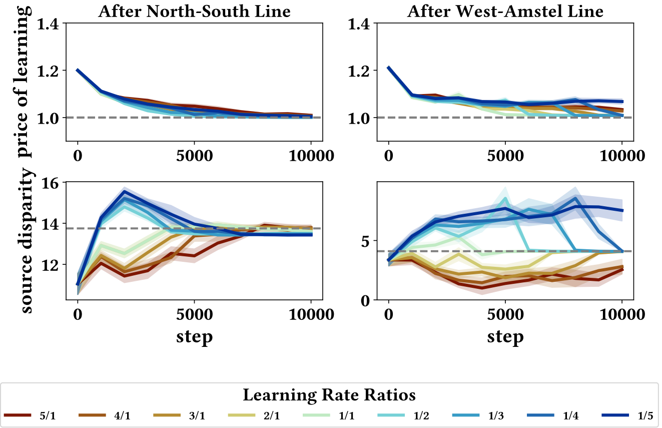

Figure 3: Unequal learning rates produce persistent inefficiencies and systematic disparities, particularly after network expansions, demonstrating that faster-adapting commuters gain a consistent advantage over slower learners.

4.1 Key Takeaway

This paper contributes to the ongoing research on fair transport network design [2, 10–12], by emphasizing the importance of considering sources of heterogeneity that dictate mobility behaviour.

Our results show that static Nash equilibrium analysis overlooks important dynamics: unequal learning rates introduce persistent disparities between groups, that persist after interventions. A low Price of Anarchy does not guarantee equitable outcomes: faster-adapting commuters can systematically benefit from new routes compared to slower learners. This suggests that infrastructure expansions should be assessed not only on their equilibrium efficiency but also on adaptation dynamics. Planners should support slower-to-adapt groups—for example through targeted communication, real-time travel information, or phased rollouts—so that efficiency gains from new infrastructure are distributed evenly.

REFERENCES

[1] Linda Baker. 2009. Removing roads and traffic lights speeds urban travel. *Scientific American* 1 (2009).

[2] Hamid Behbahani, Sobhan Nazari, Masood Jafari Kang, and Todd Litman. 2019. A conceptual framework to formulate transportation network design problem considering social equity criteria. *Transportation Research Part A: Policy and Practice* 125 (2019), 171–183. https://doi.org/10.1016/j.tra.2018.04.005

[3] Eran Ben-Elia and Yoram Shiftan. 2010. Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice* 44, 4 (2010), 249–264. https://doi.org/10.1016/j.tra.2010.01.007

[4] Melusine Boon-Falleur, Nicolas Baumard, and Jean-Baptiste André. 2024. The Effect of Income and Wealth on Behavioral Strategies, Personality Traits, and Preferences. *Perspectives on psychological science : a journal of the Association for Psychological Science* (01 2024), 17456916231201512. https://doi.org/10.1177/17456916231201512

[5] Dietrich Braess, Anna Nagurney, and Tina Wakolbinger. 2005. On a Paradox of Traffic Planning. *Transportation Science* 39, 4 (2005), 446–450. http://www.jstor.org/stable/25769266

[6] Cesare Carissimo. 2024. Counter-Intuitive Effects of Q-Learning Exploration in a Congestion Dilemma. *IEEE Access* 12 (2024), 15984–15996. https://doi.org/10.1109/ACCESS.2024.3358608

[7] Caroline Claus and Craig Boutilier. 1998. The dynamics of reinforcement learning in cooperative multiagent systems. *AAAI/IAAI* 1998, 746-752 (1998), 2.

[8] Xiaowen Dong, Alfredo J Morales, Eaman Jahani, Esteban Moro, Bruno Lepri, Burcin Bozkaya, Carlos Sarraute, Yaneer Bar-Yam, and Alex Pentland. 2020. Segregated interactions in urban and online space. *EPJ Data Science* 9, 1 (2020), 20.

[9] Zheng Li and David A. Hensher. 2013. Crowding in Public Transport: A Review of Objective and Subjective Measures. *Journal of Public Transportation* 16, 2 (2013), 107–134. https://doi.org/10.5038/2375-0901.16.2.6

[10] Karel Martens. 2016. *Transport justice: Designing fair transportation systems*. Routledge.

[11] Dimitris Michailidis, Sennay Ghebreab, and Fernando P Santos. 2023. Balancing fairness and efficiency in transport network design through reinforcement learning. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 2532–2534.

[12] Dimitris Michailidis, Willem Röpke, Sennay Ghebreab, Diederik M Roijers, and Fernando P Santos. 2023. Fairness in transport network design—A multi-objective reinforcement learning approach. In *Proc. Adapt. Learn. Agents Workshop AAMAS*. 1–7.

[13] Redactie O&S. 2023. *Monitor Sociale Veiligheid Openbaar Vervoer: Tussenmeting 2023*. Technical Report. Gemeente Amsterdam, Onderzoek & Statistiek (O&S). https://openresearch.amsterdam/nl/page/115463/monitor-sociale-veiligheid-openbaar-vervoer-tussenmeting-2023 Behandeld in Commissie Algemene Zaken 5 december 2024.

[14] R. Selten, T. Chmura, T. Pitz, S. Kube, and M. Schreckenberg. 2007. Commuters route choice behaviour. *Games and Economic Behavior* 58, 2 (2007), 394–406. https://doi.org/10.1016/j.geb.2006.03.012

[15] Michael Wunder, Michael L Littman, and Monica Babes. 2010. Classes of multiagent q-learning dynamics with epsilon-greedy exploration. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 1167–1174.

[16] Hyejin Youn, Michael T Gastner, and Hawoong Jeong. 2008. Price of anarchy in transportation networks: efficiency and optimality control. *Physical review letters* 101, 12 (2008), 128701.