

Are taxi rides just a game? Analyzing taxi tipping behavior between drivers and passengers

Julia Fausto, Lourdes Lauterborn, Charles Nelson,
Mithun Rameshkumar

Department of Mathematics, Math 30.04 (Evolutionary Game Theory),
Dartmouth College, Hanover, 03755, NH, USA.

Abstract

Taxi tipping in New York City involves a strategic interaction with partially misaligned incentives. Drivers can maximize earnings either by providing efficient service in exchange for a tip (Tip-Max) or by extending trips to inflate the meter fare (Meter-Max), while passengers can choose to tip or not. We model this as a two-population asymmetric game and analyze its dynamics with replicator dynamics, calibrated to approximately 6.5 million credit-card-paid Manhattan rush-hour rides drawn from the NYC Taxi & Limousine Commission trip records spanning February 2023 through February 2026. The model produces two key results. First, the cooperative outcome, which consists of Tip-Max drivers and tipping passengers, is never evolutionarily stable. Meter-Max strictly dominates Tip-Max for drivers because the meter premium enters the fitness difference as a constant additive advantage, independent of passenger behavior. Second, the system exhibits a phase transition that establishes two distinct regimes. When social signaling exceeds the tip cost, the system equilibrates at an outcome where all drivers Meter-Max yet all passengers continue to tip, sustained by reputational incentives alone. Conversely, in the weak-signaling regime, the system collapses to a state of no tipping and no efficient driving. Empirically, the Meter-Max premium exceeds the mean tip in every quarter we examine and NYC's credit-card tip incidence holds near 95%. We interpret NYC's tipping equilibrium as norm-sustained rather than reciprocity-driven, with practical implications for the external scaffolding that sustains the tipping norm.

Keywords: evolutionary game theory, replicator dynamics, tipping behavior, taxi service, NYC TLC data, indirect reciprocity

1 Introduction

Every taxi ride can be seen as a small game. Drivers and riders have individual incentives, which establish various strategic interactions between the two parties. Considering yellow cabs with ticking meters, drivers want to maximize their earnings. They can do that in two ways: drive efficiently and earn a tip, or take longer and run up the meter. On the passenger side, their payoff structure runs in the opposite direction as they want to keep their costs down. A faster trip with a more direct route reduces both the meter fare and the time cost of being in the cab for passengers. Passengers decide between tipping and not tipping, given social norms and pressures of tipping. Passengers who tip give up some money but enjoy reputational benefits from being seen as generous, particularly when riding with other passengers who observe the transaction. It is important to remember that tipping is completely voluntary. This tension of partially aligned incentives is core to what this paper explores, and analyzing this matters because the outcome is not just about a few dollars. It also has direct welfare and economic implications: inefficient routing can have several effects, from congestion to an increase in travel time and to costs in an already overburdened urban environment, while also affecting driver income and broader social norms.

Our project explores the cab scene in New York City and analyzes how the strategic interactions between drivers and passengers play out through an evolutionary game theory lens and real ride data. Questions we grapple with are: Does the cooperative norm of drivers driving efficiently and passengers tipping survive? Does this equilibrium collapse under individual incentives or depend on third-party nudges? To answer these questions, we utilize replicator dynamic models to evaluate how strategic incentives respond to each other in the rush hour scene of Manhattan. We therefore restrict our analysis to 6.5 million credit-card Yellow Taxi trips in Manhattan during weekday rush hours, using data from the NYC Taxi & Limousine Commission’s public trip records [1].

Prior work on tipping has framed our approach. Norris et al. [2] applied static game theory to restaurant tipping, finding that service quality and group size are recognized as factors influencing the customer’s decision-making on tipping amount. This finding informs our model assumptions by suggesting that observable service characteristics affect tipping through social signaling rather than pure reciprocity, which motivates our use of group size as a relevant observable characteristic in our context. Kang [3] argued that tipping in service industries is sustained by social norms. Even when underlying economic incentives shift, people continue to tip out of habit and pressure rather than pure reciprocity. Our core contribution is the dynamic piece, as previous literature does not examine how tipping dynamics survive at the population-level norm. Thus, we formalize taxi tipping as an evolutionary two-population game with explicit driver and passenger strategies and explore how the equilibrium evolves in a taxi cab setting.

Our analysis yields several key findings. First, the cooperative outcome, where all drivers do not run up the meter and all passengers tip, is never an evolutionarily stable strategy. This is due to the Meter-Max strategy fitness constantly dominating the Tip-Max drivers, meaning the driver population always drifts toward Meter-Max in the long run. Second, we find that our model exhibits a phase transition when the

value of social signaling is equal to the monetary amount of the tip. This threshold finding governs whether tipping survives, splitting our findings into two regimes. In the strong-signaling regime, the system equilibrates at all Meter-Max drivers alongside all tippers, whereas the system collapses in the weak-signaling regime. The Meter-Max premium exceeds the mean tip in every quarter, and the tip incidence holds steady at approximately 95% of credit-card-paid rides. Together, these observations place NYC in the strong-signaling regime, with external scaffolding inflating the observed tipping rate.

The remainder of this paper is organized as follows: the Results section introduces the replicator dynamics model, derives its equilibrium dynamics, and presents the empirical results alongside phase diagrams. The Methods section then details the assumptions and parameter choices. The paper concludes with a discussion of the implications and limitations of our findings.

2 Results

2.1 Model setup and replicator dynamics

We model the driver–passenger interaction as a two-population asymmetric game with replicator dynamics. The payoff matrices and parameter definitions ($\tau, m, k, s_0, \alpha, \beta, n$) are given in Methods (§3.1); modeling assumptions appear in §3.2.

We let x equal the fraction of Tip-Max drivers, making $1 - x$ the fraction of Meter-Max drivers. For the passenger population, we let y equal the fraction of Tippers and $1 - y$ equal the fraction of Non-Tippers. This gives the standard two-strategy replicator equations [4]

$$\dot{x} = x(1 - x) \cdot \Delta f_{\text{driver}}, \quad (1)$$

$$\dot{y} = y(1 - y) \cdot \Delta f_{\text{passenger}}. \quad (2)$$

Plugging in our payoffs gives the driver fitness difference

$$\Delta f_{\text{driver}} = \tau y - (m + \tau y) = -m, \quad (3)$$

and the passenger fitness difference

$$\Delta f_{\text{passenger}} = x \cdot s_0 + (\alpha + \beta) \ln(n) - \tau. \quad (4)$$

Setting $\Delta f_{\text{passenger}} = 0$ and solving for x gives the indifference threshold, which is the point where an individual is indifferent between the two strategic choices:

$$x^* = \frac{\tau - (\alpha + \beta) \ln(n)}{s_0}. \quad (5)$$

From the driver payoff matrix defined in §3.1, the expected fitness of a Tip-Max driver is τy and that of a Meter-Max driver is $m + \tau y$. Subtracting yields $-m$. Since

this is negative and x is the fraction of Tip-Max drivers, Meter-Max dominates Tip-Max as a driver strategy. Both strategies collect the value of a tip, so the τ terms cancel through subtraction, and what's left is the m premium that the Meter-Max driver earns. This means that the driver population will always tend toward Meter-Max, no matter what the passengers do.

For the passenger side, after cancellations, equation (4) represents how the signaling term is always going to pay no matter how many drivers are efficient. This is where the indifference threshold gains importance. For points above x^* tipping pays, but non-tipping pays below that point. If social signaling is strong enough and x^* is negative, then tipping pays for every value of x . The following are the phase diagrams we created using these replicator dynamics.

2.2 Phase diagrams: strong and weak-signaling regimes

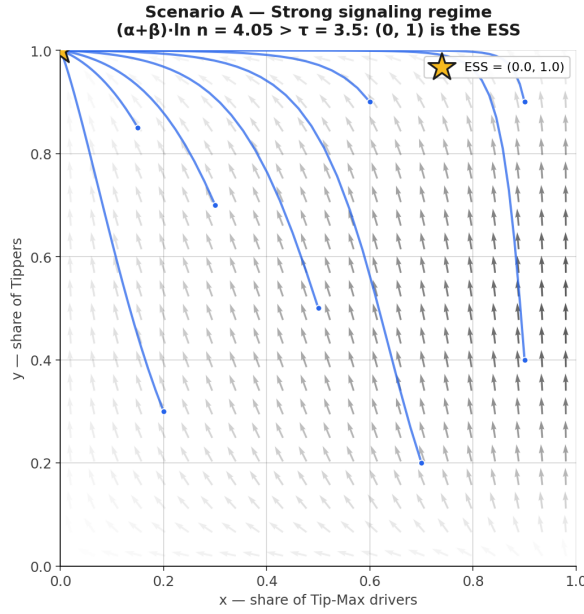


Fig. 1 Strong-signaling regime.

Figure 1 is an example of a strong-signaling regime. The arrows in the background represent the velocity field of the replicator dynamics, and the blue curves are sample trajectories starting from different initial conditions. Tipping pays for every starting point. Because the indifference threshold is negative here and outside the range, there is no attainable point where a rider might change their strategy. Even if every driver elects to Meter-Max, the social reward $(\alpha + \beta) \ln(n)$ is so strong that it overcomes the cost of the tip. The driver payoff is still the same, $-m$, which drags x to 0, but

passenger dynamics push y upward because of the strong social signaling. This leads to every trajectory converging at $(0, 1)$.

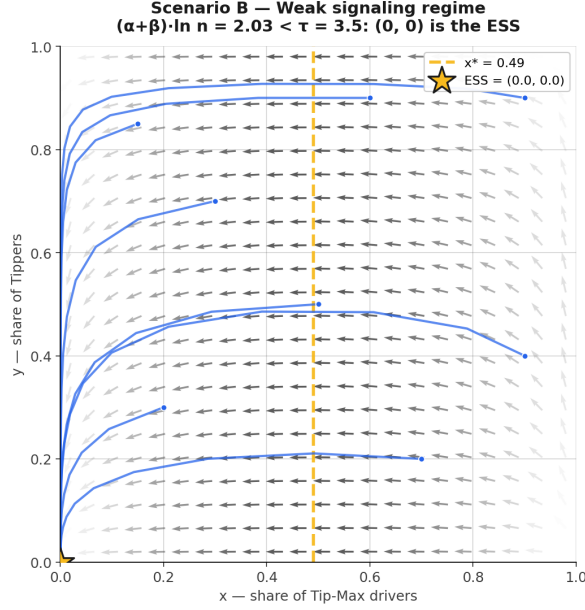


Fig. 2 Weak-signaling regime.

Figure 2 is an example of a weak-signaling regime. This is the scenario where $(\alpha + \beta) \ln(n)$ is less than τ . Now the point x^* shifts toward the interior. In this case, x^* is 0.49, represented by the yellow dashed line. The dynamics here are more nuanced: above x^* , there are enough efficient drivers that the expected social signaling reward from tipping still outweighs the cost of the tip, but below that point, the passenger's payoff flips sign, reflecting how non-tipping dominates. Trajectories that start above the threshold initially rise in y even while x falls, but when they hit the threshold, the passengers' response flips and y crashes. Now all trajectories converge to $(0, 0)$: a full norm collapse, with no one tipping and no one driving efficiently.

Given that the model for the weak regime predicts full collapse to $(0, 0)$, this is what NYC tipping culture would look like if social signaling weakened. However, it is important to flag that the tip incidence in our data hold near 95%. This gap between prediction and observation is informative: real-world tipping in this regime would only persist because of external nudges that our basic model doesn't capture. Examples of these nudges include default 20/25/30% buttons on in-cab payment screens, driver ratings, and social pressure from other passengers [5]. It is these mechanisms that keep the system off the $(0, 0)$ attractor predicted by our model.

Table 1 Nash and ESS classification of the four corner equilibria. The cooperative outcome $(1, 1)$ is never evolutionarily stable. $(0, 0)$ and $(0, 1)$ are each ESS in mutually exclusive parameter regimes separated by the phase transition $(\alpha + \beta) \ln(n) = \tau$.

Equilibrium	Description	Nash?	ESS?
$(0, 0)$	All MM, no tipping	Iff $(\alpha + \beta) \ln(n) < \tau$	Same
$(1, 1)$	All TM, all tip	No	No
$(0, 1)$	All MM, all tip (norm-driven)	Iff $(\alpha + \beta) \ln(n) > \tau$	Same
$(1, 0)$	All TM, no tipping	No	No

2.3 Equilibrium classification: Nash and ESS

Table 1 summarizes the Nash and ESS status of each corner equilibrium of the strategy space. The cooperative outcome of efficient drivers and tippers is never an ESS because Meter-Max always invades, given that $\Delta f_{\text{driver}} = -m$ is a constant and is hence unaffected by what the rider does. This reasoning applies to why $(1, 0)$ is also not a Nash equilibrium. As long as Meter-Max drivers gain anything positive from extending trips, the cooperative norm will always fail the ESS test. The cooperative outcome is therefore not merely fragile, but unattainable through reciprocity alone, with MM invading for as long as $m > 0$.

As illustrated in Figure 2, all-MM with no tipping is Nash *and* ESS only when signaling is weak, because x collapses and drags y down too once it falls below x^* . However, as shown in Figure 1, when signaling is strong, tipping persists because the social signaling benefit $(\alpha + \beta) \ln(n)$ is greater than the tip cost τ . Tipping then survives independent of what strategy the driver picks, making $(0, 1)$ a norm-driven ESS. This sets up our policy lever: strengthen signaling to stabilize tipping; weaken it and the norm collapses.

2.4 Effect of the 2025 Congestion Fee

On January 5, 2025, New York City instituted a flat \$1.50 congestion fee on trips entering the central business district. This policy is used as a natural experiment to analyze how drivers and riders respond to an actual change in their payoff parameters. Under the new congestion fee, drivers no longer need to extend trips to capture that extra revenue because it is now charged automatically.

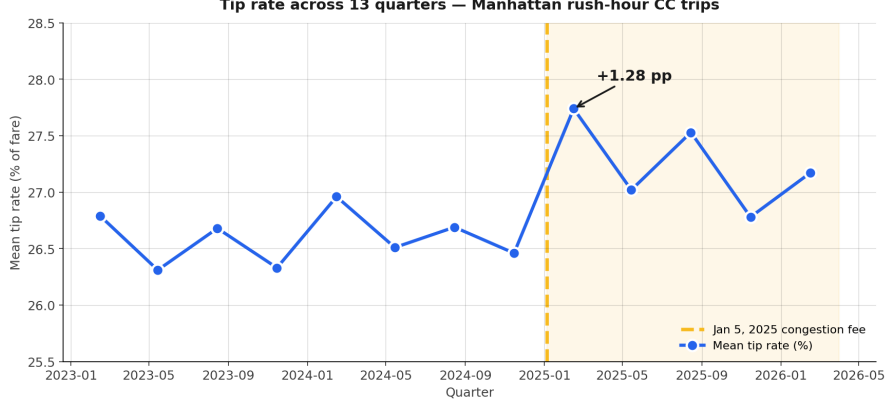


Fig. 3 Mean credit-card tip rate across 13 quarterly snapshots, February 2023 through February 2026. The January 5, 2025 NYC congestion fee is marked as the yellow dashed line.

Figure 3 plots the mean credit-card tip rate across the 13 quarters we looked at, with the congestion fee implementation date marked. From the graph, average tipping rates were stable at around 26.5% through 2023 and 2024. After the congestion pricing fee was implemented, average tip rates increased by about 1.28 percentage points in the quarter following the fee and stayed consistently elevated afterward.

Considering our model, our estimate of m fell substantially over the same interval, specifically from \$5.04 in November 2024 to \$3.46 in February 2025. The flat congestion fee means there are smaller opportunities for discretionary meter-maximizing behavior, which lowers the m payoff in our model. Our model interprets this compression in m as pushing the equilibrium dynamics toward more cooperative outcomes.

An additional key point worth noting in Figure 3 is that m exceeds τ in every quarter, meaning that the meter bonus consistently exceeds the tip reward and thus confirms that the cooperative outcome $(1, 1)$ cannot be sustained. However, the observed tip incidence holds near 95% throughout, despite the model’s prediction of the driver population drifting toward all Meter-Max. We can therefore conclude that NYC sits in the strong-signaling regime at the $(0, 1)$ equilibrium. Ultimately, the congestion fee’s compression of m changes the path toward this equilibrium through slower driver defection toward MM, giving the reciprocity channel more time to operate and reinforce the tipping norm. This finding is consistent with our model predictions and the observed 1.28 percentage-point jump.

3 Methods

3.1 Payoff matrix and parameters

The two payoff matrices (Tables 2 and 3) encode the full strategic structure of the model. Each driver chooses between two pure strategies. A Tip-Max (TM) driver

provides efficient service by taking direct routes, minimizing dwell time, and avoiding fare-padding behaviors. They are paid the metered fare plus a tip τ when paired with a tipping passenger, or the metered fare with no tip otherwise. A Meter-Max (MM) driver extends the trip in ways that inflate the meter, earning an additional m above the baseline efficient fare. Driver payoffs are measured relative to the efficient-fare baseline so that they reflect only the strategic premium above what either strategy guarantees through the meter itself.

Table 2 Driver payoff matrix. Rows are driver strategies; columns are passenger strategies. Payoffs are measured relative to a baseline efficient-service ride.

	Passenger Tips	Passenger No-Tip
Tip-Max (TM)	τ	0
Meter-Max (MM)	$m + \tau$	m

Each passenger chooses between Tip (T) or No Tip (NT). The customer payoffs include the tip cost, the time cost k incurred when paired with an inefficient driver, the baseline reciprocity term s_0 when paired with an efficient driver and the passenger tips, and the social signaling terms $\alpha \ln(n)$ and $\beta \ln(n)$, where n is the number of passengers in the cab.

Table 3 Customer payoff matrix. Rows are customer strategies; columns are driver strategies.

	Driver TM	Driver MM
Tip	$s_0 + \alpha \ln(n) - \tau$	$\alpha \ln(n) - k - \tau$
No Tip	$-\beta \ln(n)$	$-\beta \ln(n) - k$

We use seven parameters. The tip amount τ is the dollar value of a tip in a tipping event. The Meter-Max premium m is the additional fare a Meter-Max driver collects above the efficient-route baseline. The time cost k captures the opportunity cost a passenger incurs when paired with an inefficient driver. The baseline reciprocity term s_0 is the personal benefit a passenger receives from tipping a driver who provided efficient service. The social signaling parameters α and β are, respectively, the per-observer reward for being seen tipping and the per-observer cost of being seen withholding a tip. The number of passengers n enters through $\ln(n)$, capturing the standard observation that observability scales sub-linearly with audience size: doubling the audience

does not double the reputational stakes. Because $\ln(1) = 0$, solo rides remove the signaling terms entirely. This isolates reciprocity as the only psychological motive for tipping in private.

3.2 Modeling assumptions

Our model rests on four substantive assumptions about how the parameters enter the dynamics.

First, we treat passenger strategies as *unconditional*: a Tip strategy pays τ regardless of driver behavior, and a No-Tip strategy pays nothing regardless of driver behavior. This is a behaviorally realistic choice. In a real cab, a passenger’s tipping decision is heavily anchored by default tip prompts on the payment screen, which present 20%, 25%, and 30% buttons before the passenger has time to fully evaluate the trip [5]. We therefore treat “tip” as a stable type rather than a per-ride decision conditional on observed service.

Second, social signaling operates *unconditionally on driver type*: the reward $\alpha \ln(n)$ for tipping and the penalty $\beta \ln(n)$ for stiffing fire whether or not the driver provided efficient service. The behavioral intuition is that observers in the cab cannot reliably distinguish a meter-maxed trip from one delayed by genuine traffic; what they see is the act of tipping itself. Only the reciprocity term s_0 , which represents the passenger’s own internal sense of fair return for service, depends on whether the driver was efficient.

Third, we adopt a loss-aversion framing in which $\beta > \alpha$: the social cost of being seen withholding a tip exceeds the social reward for being seen tipping. This is consistent with a behavioral literature on loss aversion in social judgment. This assumption does not affect the threshold formula (5) because α and β enter additively. However, it does influence the interpretation that norms work more through punishment than through praise.

Fourth, we adopt the standard replicator-dynamics assumption of a *well-mixed population*: each driver is randomly paired with each passenger, and each population evolves according to relative fitness [4]. We acknowledge that real NYC ride dispatch is far from well-mixed; we flag this in §4.3.

These assumptions situate our model within the framework of indirect reciprocity: passengers tip not for direct return but because the act of tipping is observable and reputation matters. The $\alpha \ln(n)$ and $\beta \ln(n)$ terms are the formal expression of this idea, with the logarithm capturing the diminishing marginal effect of additional observers.

3.3 Data, filtering, and parameter estimation

We use the NYC Taxi & Limousine Commission public Yellow Taxi trip records [1]. Files are distributed monthly in Apache Parquet format. We downloaded the February, May, August, and November files for 2023, 2024, and 2025, plus the February 2026 file. There are 13 files total, which are quarterly-spaced to capture seasonal variation while keeping the data volume manageable.

We apply four filters in sequence. First, we restrict to credit-card payments (`payment_type == 1`) because the TLC data only records tips paid by credit-card, so

cash tips are not captured. Second, we restrict to trips that are internal to Manhattan, defined as rides whose pickup and dropoff location IDs both fall within the 69 Manhattan zones in the TLC taxi-zone lookup. Third, we restrict to weekday rush hour: Monday–Friday, 7:00–10:00 AM or 4:00–7:00 PM. Fourth, we exclude airport flat-rate trips (`RatecodeID` 2 or 3), since flat-rate fares decouple the meter from driver effort. We also drop rows with invalid values: zero or negative fare amounts, zero passenger counts, and trip durations or distances that imply implausible speeds (less than 0.5 mph or greater than 60 mph). After filtering, the sample contains approximately 6.5 million rides distributed across the 13 quarters.

We estimate τ as the mean of the `tip_amount` field across tipped credit-card trips within each quarter. This is straightforward because the TLC data records tip amounts directly.

Estimating the Meter-Max premium m is more involved because the TLC data does not flag which trips were meter-maxed. Our approach defines an efficient-fare benchmark from the fastest-pace quartile of trips within each distance bin and measures how much each trip’s actual fare exceeds this benchmark. Specifically: (1) within each quarter, we bin trips by `trip_distance` into 20 equal-width bins; (2) within each bin, we identify the fastest 25% of trips by pace (mph) and compute the median fare in that subset as the bin’s expected efficient fare; (3) for every trip, we compute excess fare as actual minus expected; (4) the Meter-Max premium m is the mean excess fare on trips with excess greater than \$0.50. We also implemented an alternative estimator that uses the slowest-25% versus fastest-25% fare gap. The qualitative result $m > \tau$ holds in 13 of 13 quarters under both specifications.

Phase diagrams are produced by numerically integrating the replicator system (1)–(2) from a grid of initial conditions using SciPy’s `odeint` solver (a wrapper around the LSODA algorithm). For each phase portrait we integrate from eight starting points spread across the unit square for up to 200 time units. Vector fields are computed analytically and rendered as normalized arrow grids. The indifference threshold x^* is drawn as a dashed line whenever it falls inside the unit interval. All data processing was performed in Python 3 using `pyarrow` for Parquet I/O, `pandas` and `NumPy` for filtering and aggregation, SciPy for ODE integration, and `matplotlib` for figure generation.

4 Discussion

4.1 Summary of contributions

This paper formalizes taxi tipping as a two-population evolutionary game and derives the conditions under which tipping collapses or survives as a population-level norm. Our central contribution is demonstrating that the cooperative outcome (efficient drivers and tipping passengers) can never be an evolutionarily stable strategy. Meter-Max strictly dominates Tip-Max for drivers at every point in the strategy space. Ultimately, the analysis of which dynamics persist separates our evolutionary framework from traditional game theory. We are not asking whether cooperation is a Nash equilibrium in a one-shot interaction, but instead whether, over time, it can resist

invasion and persist across a population of adapting agents. We can see that it cannot, at least not through reciprocity alone.

Social signaling is what sustains tipping in the real world according to our model. The phase transition at $(\alpha + \beta) \ln(n) = \tau$ divides the system into two qualitatively different regimes. In the strong-signaling regime, $(0, 1)$ is the ESS: all drivers meter-max, yet all passengers tip because the social cost of not tipping exceeds the monetary savings. In the weak-signaling regime, the system collapses entirely to $(0, 0)$. The natural congestion pricing experiment supports this: the compression of m from \$5.04 to \$3.46 following the January 2025 fee aligns with the observed 1.28 percentage-point increase in tip rates, which is consistent with equilibrium dynamics shifting toward more cooperative outcomes.

4.2 Broader implications

Our findings can speak to an even bigger question within the social sciences: when do informal norms substitute for formal contracts, and when can they collapse? We have provided a concrete mechanism through our model that determines whether a norm can be self-sustaining or needs external scaffolding to survive. This is relevant to any gig-economy setting where workers and clients interact without binding contracts. Settings such as food delivery, freelance platforms, and peer-to-peer services all feature asymmetric information, voluntary compensation, and reputational dynamics that shape behavior. The EGT framework developed here can be applied directly to these settings. The automatic 20/25/30% tip options on the payment screens [5], driver rating systems, and clear fare information help strengthen social expectations. Thus, these findings suggest matters like default settings and transparency stabilize the system against defection from both parties and are essential to the system.

4.3 Limitations

Several limitations should be noted within our model. First, the model utilizes binary strategies where drivers either Tip-Max or Meter-Max, and passengers either tip or do not. This is not realistic, as real behavior is continuous. Drivers choose how much to deviate from the optimal route and passengers can choose how much to tip. A continuous-strategy extension would better capture the marginal incentives. Second, replicator dynamics assume a well-mixed population where any driver can be matched with any passenger. However, matching is shaped by neighborhood, time of day, driver reputation, and app algorithm. These are confounding variables that introduce network structures that our model does not take into account. Third, our data only covers credit-card-paying passengers, meaning that we ignore cash tips entirely. Lastly, the key parameters (m , s_0 , α , and β) are estimated through proxies rather than being directly observed. We have not conducted a formal sensitivity analysis to test how much our conclusions would change if these values were different.

4.4 Future directions

Three extensions stand out as the most promising. First, replacing the deterministic replicator with a Moran process would allow for finite-population effects and drift,

which matter when social groups are small. Second, extending the model to rideshare platforms such as Lyft and Uber would incorporate driver rating systems, algorithmic dispatch, and enforcement mechanisms that interact with the informal norm. Finally, the phase transition result—that tipping survival depends on crossing a single threshold in social signaling strength—invites empirical testing. An experiment that varies passenger group size or default tip options could help determine where real New York City rides fall in relation to the x^* threshold, and whether these changes can consistently move the system from one outcome to another.

References

- [1] NYC Taxi and Limousine Commission: TLC trip record data. Public data set (2024). <https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page>
- [2] Norris, C.L., Taylor Jr., S., Taylor, D.C., Snipes, M.: Is tipping just a game? applying game theory to restaurant tipping behavior. *International Journal of Hospitality Management* **114**, 103585 (2023) <https://doi.org/10.1016/j.ijhm.2023.103585>
- [3] Kang, S.: The tipping point: The subtle psychology and economics of taxi fares. Stanford Graduate School of Business Insights (2021). <https://www.gsb.stanford.edu/insights/tipping-point-subtle-psychology-economics-taxi-fares>
- [4] Hofbauer, J., Sigmund, K.: *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge (1998)
- [5] Haggag, K., Paci, G.: Default tips. *American Economic Journal: Applied Economics* **6**(3), 1–19 (2014) <https://doi.org/10.1257/app.6.3.1>