Formulation of Nash Equilibrium And Ideal Toll Rates In A Toll Usage Scheme For Public Transport Congestion Deterrence

Mason duBoef

Spring 2022

1 Introduction

Congestion is a struggle that plagues public transport systems around the world. Whether it be in underground subways or above ground bus routes, unexpected influxes in demand can lead to build up of commuters at stops, forcing negative-externalises to arise.

This paper will exclusively refer to a subway environment, and investigate toll usage using real world New York City subway data. However, given that subways and buses share a similar underlying system, collecting commuters at stops along a predetermines route. The learnings and applications outlined in this paper could just as easily be applied to bus routes. Before formulating ideal toll rates I will discuss why the issue of congestion must be mitigated at the toll booth, before it is allowed to accrue inside the station where commuters are attempting to board. In doing so, I will also demonstrate the difficult no-win situation train drivers are placed in when toll-setters allow stations to become congested. it will be shown that no matter what strategy drivers employ disutility cannot be meaningfully mitigated by subway drivers through strategic use of timed departure and door closure.

When a greater than expected number of commuters attempt to board a train drivers are faced with a challenging decision. If drivers keep subway doors open long enough to accommodate the flow of all passengers attempting to enter the vehicle they will likely depart the station later than anticipated, thus, running behind schedule and providing a negative utility for all the commuters who will board at later stations. Here t_s is defined as the scheduled amount of time allowed for boarding at a particular station and t_r is the real time allocated to boarding before the train or bus departs. In this scenario, a driver allows all commuters to board, despite their unexpectedly large numbers. Thus, $t_s < t_r$. Here, negative utility is distributed among all commuters, those who have already boarded and those who will be boarding at later stations. Those onboard when departing the station in question will arrive at their target destination $t_r - t_s$ minutes later than scheduled. Likewise, those waiting at later stations waiting to board will suffer a delay of $t_r - t_s$ minutes, as the train arrives late, after the scheduled arrival time. As such, negative utility for individual effected commuters can be represented as a delay of $t_r - t_s$ minutes. Here, I define c_d as the number of commuters on the train after a late departure and define the number of commuters who will board at later stations as c_l. The total negative utility across all commuters in this situation is represented as the following.

$N_u = (c_d + c_l) * (t_r - t_s)$.

A complete model of real world public transport usage would also account for the potential for missed connections. If any of these commuters who suffer a delay of duration $t_r - t_s$ are forced to miss a connecting train as a result of the delay then their negative utility might be greater than just $t_r - t_s$ minutes. They are forced to wait for the next train or take another less preferable route.

Drivers' alternative choice to this scenario would be to close the doors preemptively, before all commuters at the station are able to board. This would allow the driver to stay on schedule and assure that not delays will be suffered by commuters set to board at later stops. I define the number of commuters who are left behind, unable to board before doors close as c_a. Also define the time till the next train arrives at the station as t_n. The total negative utility across all commuters in this situation is represented as the following.

$N_u = c_a * t_n$.

A compete representation of the problem would also account for the potential that some of the c_a commuters will pursue an alternative route after being left behind instead of waiting for the next train. Trains also have a very limited capacity¿ there are many situations where train cars might be at capacity and unable to fit new incoming commuters even if you have enough scheduled time for them to board

Most public transport systems opt to keep doors open and gather historical data on congestion build up and anticipate periods of high traffic by preemptively increasing t_s. While typically effective, traffic projection will never be perfect. Unforeseen congestion will appear at stations, perhaps at random, perhaps because of a nearby event. Regardless, subway organizers are forced into a situation where an influx of commuters are attempting to enter the station and they have to take action to mitigate this unexpected build up. No matter what approach to door closure the driver takes negative utility will be distributed in significant quantities as a result of unexpected congestion forming at a station. As such, deterrence measures can and should be taken to increase the cost of entering the station, decreasing the number of commuters attempting to board the train. With appropriate deterrence, a train should be able to board all commuters at the station without suffering suffering any delays.

In this investigation, tolling will be utilized as the primary means of deterrence. For commuters to enter a given station and board a train a set monetary toll will need to be paid. This approach has a number of drawback that do need to be addressed. Firstly, a traditional tolling scheme does favor high income individuals. More specifically, commuters who ascribe more monetary value to their time will be more likely to board than commuters who assign less monetary value to the savings in time provided by the subway. A number of approaches have been formulated to solve this potential inequity. In their 2017 paper Lahlou and Wynter outline a tradable credits scheme for congestion deterrence where credits are evenly distributed amongst a population. Base credit allowances aren't enough to afford preferred transport alternatives so commuters are left with the choice to buy or sell their credits depending on their on their wealth and preferences. This way, those who cannot afford the best alternative can at least monetarily gain from their decision at the expense of those who consider it worthwhile to pay for additional credits. Lahlou and Wynter go on the investigate the parameters governing market conditions and formulate Nash equilibria based on this scheme. While this credit scheme proposal seems promising this paper focuses on a traditional tolling scheme for sake of simplicity.

2 Body

2.1 Modeling

I have selected the New York City's 1 Line to model for my investigation after. It is one of the busiest subway lines in New York and includes the single busiest subway station, the 42nd St.-Times Square station. It contains 38 stations arranged North to South, going from the Upper Bronx to Lower Manhattan. My simulation of the line pulls data from the NYC Metro Transport Authority. 2018 data shows how many commuters entered every given station along the 1 Line on an average day. Furthermore, referencing 2018 train schedules shows how many trains were servicing this daily commuter population. By dividing the average daily commuter population by the number of daily train arrivals we can roughly calculate an average train stop at every station along the 1 Line.

These figures are somewhat non-ideal but do service the purpose of my simulation well enough. First, they count the number of commuters who entered the station and paid its respective toll, likely between \$1 and \$3, judging by typical tolling rates. Ideally this sample would count commuters entering with no toll in place. This would provide a baseline level of interest in boarding before tolls are included. Unfortunately, no such data is available at scale. Presumably, the provided data shows fewer commuters than the figure I want, the number of commuters considering entering the station, depending on the toll amount. This figure would be equal to the number of people who would enter the station at the minimum possible toll of \$0.

The other major drawback of this data is that is it counts the number of commuters who enter the station, not the number of commuter who enter the station to take the 1 Line. 20 of the 38 stations in the 1 Line, service at least one other train line. Some larger stations like 42nd Street-Times Square, service numerous stations. The popularity of the 1 Line relative to other less common lines helps mitigate this issue. At almost all stations the 1 Line is the most popular by a wide margin. However, this inaccuracy in data specification means the number of people interested in taking the 1 Line will be inflated because these 2018 figures count those who enter the station and take a different train. Ideally, commuters would only be counted if they boarded a 1 train but no such data is available.

Hopefully, the inflationary effect of stations servicing multiple lines is mitigated by the deflationary effect of tolls being in place when data is collected. Analyzing the data, it seems that the projected number of commuters attempting to board at a average subway stop is realistic and feasible. Figures for large stations, particularly 42nd Street-Times Square, do seem larger than expected but not unfeasible. Considering this is an exploration of how to mitigate severe congestion with tolls, slightly inflated commuter data might even be desirable.

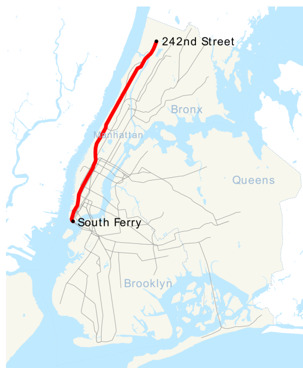

Figure 1: Map of the 1 Line in New York City

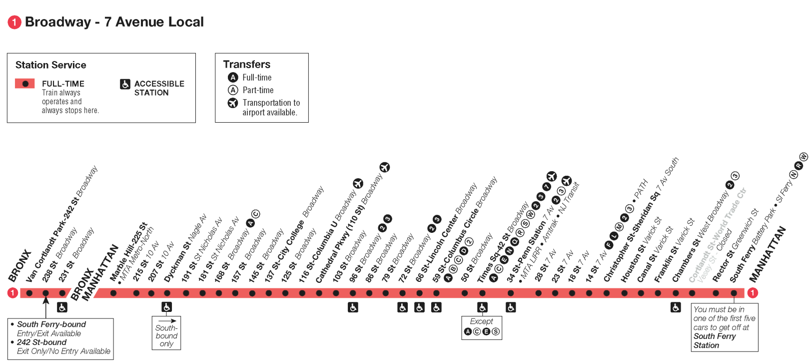

Figure 2: Linear map of all 38 stations on the 1 Line

This investigation models subway congestion as a binary transport game with tolls set at each station, separately. The system as a whole is modeled as a set of stations, S, each with a unique set of commuters, C. A toll setting agent is placed at each station and tasked with setting tolls. For my NE formulation tolls are only available in 50 cent increments from \$0 to \$4. However, in my formulation of ideal toll rates based on simulation any toll rate is available. A flat invariable time of 2 minutes is set as the time between two neighboring stations. This keeps calculation simple and isn't too dissimilar to real life.

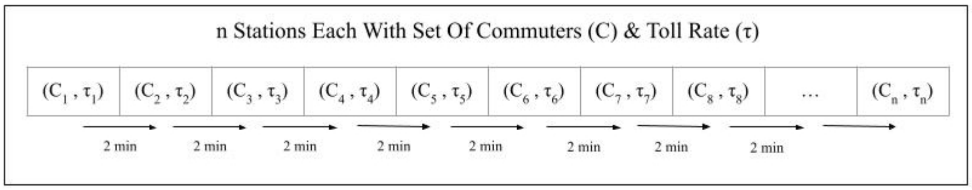

Figure 3: Abstract diagram of array of subway stations

In this project I have set a arbitrary ideal occupancy rate of 80% for the 1 Line. This should service as many commuters as possible while leaving space in case of room is urgently needed for whatever reason. Hopefully, by leaving 20% the subway is able to curtail the risk of reaching full occupancy and having to disallow new commuters hoping to board. The methods outlined in this paper will, however, operate with any desired rate of ideal occupancy.

Furthermore, I have defined a unique set of characteristics that make up the preference profile of each individual commuter. A commuter's value of time is defined as the maximum amount on money a commuter is willing to pay to save 1 minute of transport time. Its inclusion in commuter preference profiles accounts for differentiation in wealth and how one weighs monetary versus temporal negative externalites.

$$V : \text{Value of time}$$

In simulation, a commuter's value of time is randomly selected from range range of values from \$0.20/min to \$0.80/min. This range was selected after referencing NYC wage data from the Bureau of Labor Statistics. According to the BLS, the median hourly wage in NYC was \$23.43 in 2021. This comes out to \$0.39/min. Assuming New Yorkers value their time working similarly to their time in public transport this figure should function well as the median value of time for simulated commuters. The minimum possible V equates closely to the 2021 New York state minimum wage of \$12.50. Meanwhile, the upper end of the simulated spectrum, more than doubles the median and equates to a \$48 hourly wage. While many New Yorkers, particularly those being regularly serviced by the 1 Line, make significantly more than \$48/hour, I chose that cut=off to make simulation run cleanly and to compensate for the fact that, in real life, the distribution of New Yorkers along a wage graph is not linear. Rather, far more people are concentrated in the lower income brackets versus the higher ones, where differences in wage among wealthy individuals can be very large..

Meanwhile, other variables outline the monetary and temporal cost of two different alternatives for that commuter. Alternative A is the subway route being considered while alternative B is the commuters next best option. For example, alternative B might be a taxi cab with a higher cost but shorter transport time than the subway or it could be a walking route, free of charge by much longer in duration.

$$\tau_A : \text{Toll at station S}$$

$$\tau_B : \text{Monetary price of taking alternative B}$$

$$T_A : \text{Time for subway route, A}$$

$$T_B : \text{Time for alternative route, B}$$

In this paper τ_A will be largely unknown, considering it is our independent variable we will be experimenting with and solving for. Meanwhile, all other variables above are assigned randomly set values in a predetermined range. Few, if any, modes of transport chose to reward commuters monetarily. Thus, τ_B can never be negative. Similarly, we are only considering tolls as a tool for deterrence in times of severe congestion. We can ignore any schemes where commuters might monetarily rewarded for taking a given subway route.

$$\tau_A \geq 0$$

$$\tau_B \geq 0$$

For the purposes of simulating real world congestion, the cost of transport alternatives are meant to be roughly similar to those of the New York City subway system. Their range is outlined below.

$$\tau_B : \$0 \text{ through } \$5$$

Travel duration must be expressed relative to the number of stations remaining in the subway line, before the end. We can define r to be the number of stations remaining in the subway line. After reach the end, subways reverse direction and go along the same route in the opposite direction. If train will turn around and do the opposite direction after r stops it would make no sense for a commuter to get off after $r+1$ stops. They could have gotten off the train $r-1$ stops from when they boarded, arrived at their same destination and, saved 4 minutes of travel time. As such, it is never efficient to stay on the train for more than r stops after boarding, where r is defined to be the number of stations remaining in the line before the train will turn around. If commuters wanted to go North they would simply take the North-bound train, they would never board a South-bound train and wait for it to turn around and go North. This will never be faster than waiting for South-bound train. As such, we can conclude it never makes sense for commuter to be on the train when it turns around at the end of the line. To express this, a hard limit has been placed on the range of possible travel durations. The duration of alternative A can never be longer than the time it takes to reach the furthest away station. It can also never be less than 2 minutes since that is the time it takes to travel one single stop.

$$T_A : 2 \text{ through } 2r\text{, in 2 minute increments}$$

The range for possible values for T_B is a little more arbitrary. After tweaking the values the following parameters produced consistently realistic simulated results, when comparing to T_A. There is likely other, better means of simulating T_B to produce realistic sums but this has worked well for this simulation.

$$T_B : 1 \text{ through } 3r\text{, in 1 minute increments}$$

Below you can see formulations for the utility of each alternative. Utility, here is expressed in terms of minutes and will never be less than zero since cost and transport time deliver negative utility to commuters. Value of time is used to translate the disutility of transport cost into units of minutes in accordance with that individuals preference. In doing so, we are able to express the utility of alternatives A and B in terms of minutes as seen below.

$$u_A = -T_A - \frac{\tau_A}{V}$$

$$u_B = -T_B - \frac{\tau_B}{V}$$

Furthermore, I have chosen to define a best case scenario version of alternative A where the toll is set to the minimum possible quantity, \$0. As shown, this simplifies to the opposite of the duration of alternative A.

$$u'_A = -T_A - \frac{0}{V} = -T_A$$

We define this best case utility, u'_A for the sake of valid preference generation. If the utility of alternative B is better than the best case utility of alternative A then the commuter will never even consider entering the subway station. They would have no reason to check and see the toll price before making their decision. They would simply take alternative B. So, if we define $\hat{C}$ to be set of all commuter traveling in the area then the set of commuters that concerns this investigation is the subset of $\hat{C}$ where the best case utility of alternative A is at least as good as alternative B. In our simulation of real world subway congestion we randomize the time and cost for each alternative. In situations where $u'_A < u_B$ variables are re-randomized until valid.

$$\hat{C} : \text{Set of all commuters near station S}$$

$$C = \{x \in \hat{C} | u'_A \geq u_B\}$$

The below variables declarations represent commuter choice and, as such, are binary, either 1, taking the specified route, or 0, not taking the specified route. δ_A represents a commuter's choice in picking alternative A while δ_B represents the opposite, their choice picking alternative B. In cases of tied utility commuters will default to choosing alternative A.

$$\delta_A = \begin{cases} 1, & \text{if } u_A \geq u_B \\ 0, & \text{otherwise.} \end{cases}$$

$$\delta_B = \begin{cases} 1, & \text{if } u_A < u_B \\ 0, & \text{otherwise.} \end{cases}$$

By defining these variables we are able to express overall commuter utility as one expression, regardless of the alternative they choose.

$$U = \delta_A * u_A + \delta_B * u_B$$

2.2 Formulation Of Ideal Toll Rates

The maximum toll a subway organiser can charge a given commuter and expect them to board the subway is reached when the utility of alternative A is equal to the utility of alternative B. If the rate is raised any further, the disutility of A be greater than he disutility of B and the commuter will choose alternative B.

$$u_A = u_B$$

Using the definition of u_A, this expression can be expanded. The expression can then be rearranged to express a maximum toll rate τ_{Amax} in terms of known randomized variables.

$$-T_A - \frac{\tau_{Amax}}{V} = u_B$$

$$-\frac{\tau_{Amax}}{V} = u_B + T_A$$

$$\tau_{Amax} = -V(u_B + T_A)$$

Comprised of a combination of randomized variables, each commuter will likely have a unique τ_{Amax}. This means commuters at the same station will have different thresholds for the price at which they will refuse to play the toll and instead take their own alternative B.

τ_{Amax} must be calculated for every commuter in the aforementioned set of all commuters considering entering the station, C. The commuters in this set can then be reordered by their τ_{Amax} values. Those with the lowest τ_{Amax} and closest to turning away alternative A are placed at the highest indices while those with the highest τ_{Amax} values are place at the lowest indices.

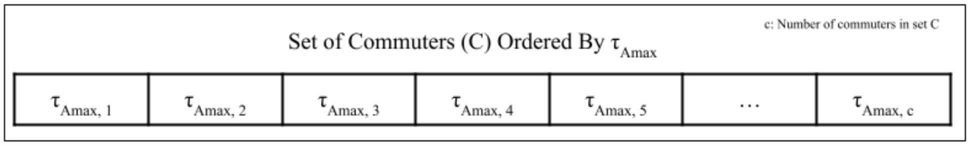

Figure 4: Commuter array C ordered by τ_{Amax} values

s, the number of commuters that would need to be board a train for it to reach its ideal occupancy rate is expressed below.

$$t_c : \text{Train capacity}$$

$$i_o : \text{Ideal occupancy ratio}$$

$$0 < i_o \leq 1 \text{ in all cases, } i_o = 0.8 \text{ for purposes of simulation}$$

$$o : \text{Current train occupancy, number of commuters on train before boarding}$$

$$s = t_c * i_o - o$$

Subways along to 1 Line in NYC have 10 train cars, each capable of carrying 42 passenger, giving the train a total capacity of 420. If we imagine the train is carrying 332 commuters before people begin to board we can fill in terms for the expression of s.

$$s = 420 * 0.8 - 332$$

$$s = 4$$

With the knowledge that we need to fill 4 seats we can revisit the array of commuters ordered by their τ_{Amax} values. To "select" 4 commuters to board the toll-setting agent can set the toll at that station, τ_s, equal to the τ_{Amax} of the fourth commuter in the array.

$$\tau_s = \tau_{Amax,4}$$

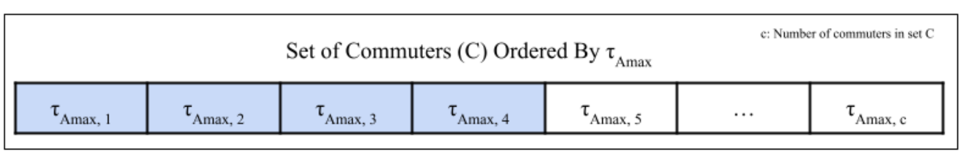

Figure 5: Selection of 4th term of commuter array C ordered by τ_{Amax} values

If $s > c$ where c in the number of commuters in set C, then τ_s, the station's toll, will be set to $\tau_{Amax,c}$ instead of $\tau_{Amax,s}$.

For my simulation of the 1 Line, this same process was used to calculate the ideal toll rate for each station. The simulation began with a empty train starting at the South Ferry stop and traveling North until reaching 242nd Street-Van Cortlandt Park, the final station in the line. At each station it calculated the number seats to be filled, a_s, as the following.

$$a_s = \begin{cases} c, & \text{if } c \leq o_{goal} - o \\ o_{goal} - o, & \text{otherwise.} \end{cases}$$

$$c : \text{Number of commuters considering entering station } s$$

$$o : \text{Current number of passengers occupying of train}$$

$$o_{goal} : \text{Number of passengers occupying of train when at ideal occupancy rate}$$

When all passengers considering entering can fit comfortably, a_s is equal to the number of commuters considering entering. When not everyone can be fit comfortably the number of available seats is the difference between the current train population and the train population at ideal occupancy. The below figure shows the results of the simulation, utilizing this technique to find the ideal toll rate.

```
Station 0 : 18/18 : $2.07
Station 1 : 7/7 : $5.74
Station 2 : 24/24 : $0.01
Station 3 : 13/13 : $1.6
Station 4 : 4/4 : $3.99
Station 5 : 15/15 : $2.26
Station 6 : 16/16 : $1.74
Station 7 : 19/19 : $3.09
Station 8 : 27/27 : $0.91
Station 9 : 62/62 : $0.06
Station 10 : 26/26 : $2.08
Station 11 : 22/22 : $0.45
Station 12 : 32/32 : $1.2
Station 13 : 18/18 : $0.2
Station 14 : 36/36 : $1.28
Station 15 : 30/30 : $0.73
Station 16 : 30/30 : $0.9
Station 17 : 86/86 : $0.21
Station 18 : 46/46 : $0.82
Station 19 : 36/36 : $0.32
Station 20 : 94/94 : $0.01
Station 21 : 44/55 : $3.88
Station 22 : 30/175 : $15.68
Station 23 : 63/63 : $0.41
Station 24 : 69/490 : $14.7
Station 25 : 36/204 : $12.52
Station 26 : 35/35 : $0.23
Station 27 : 37/37 : $0.08
Station 28 : 21/21 : $0.89
Station 29 : 116/116 : $0.16
Station 30 : 23/23 : $0.42
Station 31 : 37/37 : $0.55
Station 32 : 15/15 : $0.2
Station 33 : 14/14 : $1.65
Station 34 : 51/51 : $0.48
Station 35 : 8/8 : $0.31
Station 36 : 21/21 : $0.4
Station 37 : 81/81 : $0.06
```

Figure 6: Results from simulation, calculating ideal toll rate for stations along the 1 Line

In the left column the index number of the station is given. The further North along the line a station sits, the higher its index number. The middle column shows two quantities. The first, on the left, is a_s, the number of seats to be filled at the stop and, thus, the number of commuters who boarded at station s. The second term in the middle column is the number of commuters who considered entering the station, regardless of whether they chose to pay the toll and enter the station. In the third column is t_s, the ideal toll rate calculated to get the train closest to its ideal rate of occupancy given the preference profiles of commuters at station s.

2.3 Formulation Of Nash Equilibria

By modeling large scale republic transport as a series of games between two actors, a commuter and the toll-setter, each with a discrete number of available options, we are able to calculate Nash equilibria for every pairing of commuter and toll-setter. In this section I demonstrate such while outlining the steps for doing so. In Lahlou and Wynters' 2017 paper they demonstrate the use myopic best responses in the formulation of NE. I've adapted this approach to function in a more traditional toll setting scheme in place of their credit allocation scheme.

In order to discretize the options of the toll setting agent I have limited them to choosing one of nine available toll rates. Meanwhile, the second agent, representing a single commuter in set C, has a binary choice between alternatives A and B. The utility of the commuter, U, is defined by the previously derived general utility function.

$$U = \delta_A * u_A + \delta_B * u_B$$

Meanwhile, the disutility of the toll-setting agent is defined as the difference between the train's train population and the train population at the ideal occupancy rate. Thus, the toll-setter would like to board the commuter if below the ideal train population and turn away the commuter if at or above the ideal occupancy ratio. The toll-setter's utility, U_t, is formally given below.

$$U_t = -|o - o_{goal}|$$

Given a situation where the train population is one commuter short of reaching a the ideal occupancy rate of 80%, the binary transport game looks like the following. Here, the τ_{Amax} of the commuter is $2.

Situation: 1 short of 80% occupancy, τ_{Amax} = $2

	$0.00	$0.50	$1.00	$1.50	$2.00	$2.50	$3.00	$3.50	$4.00
A	$(u_{A,\ \$0.00}, 0)$	$(u_{A,\ \$0.50}, 0)$	$(u_{A,\ \$1.00}, 0)$	$(u_{A,\ \$1.50}, 0)$	$(u_{A,\ \$2.00}, 0)$	$(u_{A,\ \$2.50}, 0)$	$(u_{A,\ \$3.00}, 0)$	$(u_{A,\ \$3.50}, 0)$	$(u_{A,\ \$4.00}, 0)$
B	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$

Figure 7: Binary transport game between commuter and toll-setter

We can start to find potential NE candidates by finding and highlighting each actors best responses to the potential actions of the other. If the commuter choices alternative A the utility of the toll-setter will be 0 regardless of what they set the toll to. Similarly, if the commuter chooses alternative B the toll-setters utility will be -1 no matter what toll they set. Thus all tolling options are considered myopic best responses on the part of the toll-setting agent.

Situation: 1 short of 80% occupancy, τ_{Amax} = $2

	$0.00	$0.50	$1.00	$1.50	$2.00	$2.50	$3.00	$3.50	$4.00
A	$(u_{A,\ \$0.00}, 0)$	$(u_{A,\ \$0.50}, 0)$	$(u_{A,\ \$1.00}, 0)$	$(u_{A,\ \$1.50}, 0)$	$(u_{A,\ \$2.00}, 0)$	$(u_{A,\ \$2.50}, 0)$	$(u_{A,\ \$3.00}, 0)$	$(u_{A,\ \$3.50}, 0)$	$(u_{A,\ \$4.00}, 0)$
B	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$

Figure 8: Binary transport game with toll-setter's best responses highlighted in blue

Moving on to the commuter agent, when the toll is set to less than $2 the commuter's best response would be alternative A, the subway. This is because the toll is under the commuter's τ_{Amax}, meaning that the utility of alternative A will be higher than alternative B given the toll rate. When the toll rate is $2, precisely equal to the commuter's τ_{Amax}, the utility of A and B are equal, making both options a valid best response to a \$2 toll. When the toll is greater than \$2 it is over the commuter's τ_{Amax} value. The utility of B will be higher than A in these situations, meaning the commuter best response would be to choose alternative B, avoiding the subway. Nash equilibria can be defined as outcomes where both agents chose the best possible option. Thus, outcomes highlighted as the best response by both the commuter and toll-setting agent are pure strategy Nash equilibria.

Situation: 1 short of 80% occupancy, τ_{Amax} = \$2

	\$0.00	\$0.50	\$1.00	\$1.50	\$2.00	\$2.50	\$3.00	\$3.50	\$4.00
A	$(u_{A,\$0.00}, 0)$	$(u_{A,\$0.50}, 0)$	$(u_{A,\$1.00}, 0)$	$(u_{A,\$1.50}, 0)$	$(u_{A,\$2.00}, 0)$	$(u_{A,\$2.50}, 0)$	$(u_{A,\$3.00}, 0)$	$(u_{A,\$3.50}, 0)$	$(u_{A,\$4.00}, 0)$
B	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$	$(u_B, -1)$

Figure 9: Binary transport game with NE highlighted in purple

3 Conclusion

In this paper I have created original means of modeling the relationship between toll rates and train congestion. I created a multifaceted preference profile guided by 2021 NYC wage data. This allowed me to randomize commuter preference along a realistic range. Combining this preference randomization with real world station population and train scheduling data allowed me to construct a lifelike simulation of a single train, dynamically dealing with congestion along the 1 Line in New York City. This simulation could be used by NYC toll-setters to project commuter demand and how those commuter will respond to price hikes as a means of congestion deterrence.

My formulation of Nash equilibrium demonstrates how public transport can be modeled as binary transport game. This analysis of myopic best responses to calculate NE with discretized toll options is relatively computationally inexpensive, allowing for PSNE to be calculated at a large scale with hundreds of simulated commuters placed at a single station to inform real-world tolling decisions.

Analyzing the simulation output brings to light a number of learnings about toll setting¿ It also highlights clear steps toll-setters at the Metro Transport Authority can take to improve service along the 1 Line. Currently its typical for toll rates in NYC to sit between \$1 and \$3. They differ by small amount station to station and hardly ever step outside that \$2 range. My simulation shows that if system administrators want to prioritize trains operating at manageable levels of occupancy than there is no reason for tolls at small, unpopular station to go beyond a few dozen cents. Typically those stations are nearly empty, and don't have the potential to instantly, radically congest the train like the larger stations do. As such, I see no reason why small stations, which see less than 40 commuters considering the subway at a typical stop, are given \$1 to \$3 tolls, around the same price as the tolls placed on large stations. When arriving at a small station with a train population far below the optimally I see no reason why the toll wouldn't be lowered to accommodate all commuters even those whose τ_{Amax} is near \$0. With little to no risk of congestion toll-setters should seek to lower cost and thus provide maximum utility to commuters.

My simulation also shows the value of very high toll rates at very populous stations that risk bringing the train to capacity. In real life, a toll greater than \$10 would be unheard of. However, by keeping tolls around \$2, toll-setters risk unleashing more people into the station than can realistically board the trains. With inappropriately small tolls trains will be brought to capacity and have to turn away random commuters. Furthermore, with the time it takes to board and deboard congested train nearing capacity, delays could arise for that train's scheduled arrivals. By setting prices high enough to rule out all by a select, manageable number of commuters, not only would toll-setter relieve congestion but, also, select for the commuters who need to get on the train the most. Clearly, wealth/value of time impact commuters τ_{Amax}. However, in stations with hundreds of commuters considering boarding, like #24 and #25 (see simulation results), the commuters with the highest τ_{Amax} almost exclusively need to board the train because of the severe disutility associated with their alternative B. If you want to minimize the disutility of the whole set of commuters, C, you're making a mistake by setting a toll price low enough that commuters with a viable alternatives and low τ_{Amax} values might board before other who really need alternative A. Those with low τ_{Amax} values could bring the train to capacity and force those with high τ_{Amax} values into very unfavorable alternative routes. If worries of economic inequity worry toll-setters they could also adapt my formulation of ideal toll rates to work with a tradable credits scheme like that of Lahlou and Wynter. Whatever the means of payment, the price of entering a small station being serviced by a mostly empty train should be about an order of magnitude cheaper than a populous station with trains near capacity.

4 References

GitHub repository with simulation code:
https://github.com/mduboef/subwayTollSim

Lahlou, Salem, and Laura Wynter. "A Nash Equilibrium Formulation of a Tradable Credits Scheme for Incentivizing Transport Choices: From next-Generation Public Transport Mode Choice to Hot Lanes." Transportation Research Part B: Methodological, Pergamon, 21 Apr. 2017, https://www.sciencedirect.com/science/article/pii/S0191261516306774.

Dixit, Vinayak, and Laurent Denant-Boemont. "Is Equilibrium in Transport Pure Nash, Mixed or Stochastic? Evidence from Laboratory Experiments." Transportation Research. Part C, Emerging Technologies, 14 Jan. 2015, https://halshs.archives-ouvertes.fr/halshs-01103472.

"File:NYCS Map 1.SVG." Wikimedia Commons, https://commons.wikimedia.org/wiki/File:NYCS_map_1.svg.

"1 Train Stops: MTA NYC Metro 1 Train Stops amp; Schedule." 1 Train Stops — NYC Metro 1 Train Schedule — MTA 1 Train, https://new-york.metro-map.net/nyc-subway-map-1-train-stops/.

"Occupational Employment and Wages in New York-Newark-Jersey City - May 2020 : New York–New Jersey Information Office." U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, 15 June 2021, https://www.bls.gov/regions/new-york-new-jersey/news-release/occupationalemploymentandwages_newyork.htm#tablea.f.1.

"New York - May 2021 OEWS State Occupational Employment and Wage Estimates." U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, 31 Mar. 2022, https://www.bls.gov/oes/current/oes_ny.htm.

"Average Weekday Subway Ridership." Mta.info, http://web.mta.info/nyct/facts/ridership/ridership_sub.htm.

New York City Transit Subway Timetable - Home — MTA. http://web.mta.info/nyct/service/pdf/t1cur.pdf.

Area, New York Metropolitan. "MTA Daily Ridership Data: Beginning 2020: State of New York." MTA Daily Ridership Data: Beginning 2020 — State of New York, 29 Apr. 2022,
https://data.ny.gov/Transportation/MTA-Daily-Ridership-Data-Beginning-2020/vxuj-8kew.