2016

How Can We Have A Better Public Transportation System? –An Exploratory Agent Based Model

Boyu Liu
Pomona College

How Can We Have A Better Public Transportation System?
–An Exploratory Agent Based Model

Boyu Liu

In partial fulfillment of a Bachelor of Arts Degree in Environmental Analysis,

2015/16 academic year, Pomona College, Claremont, California.

Readers:

Bowman Cutter, Pomona College

Char Miller, Pomona College

Acknowledgement

I would like to take this opportunity to express my sincere gratitude to Professor Cutter, who guided me through the entire process of writing and provided feedback along the way, and Professor Miller, who kept encouraging and inspiring me.

I would also like to thank my friends who accompanied me and supported me.

I would like to thank SURP for funding my summer research which ultimately led to my thesis.

And finally, thank you to my parents, who made me who I am.

Preface

The inspiration of this piece of work came from an experience I had three years ago. At that time I was a college freshman who had lived in Beijing for 18 years and had just came to the US for college. I was naive enough to choose to take the public transportation to the Getty Center because that was what I had been doing in Beijing, and I didn't have a car. I looked up the route on google map. Three hours, two transfers. Fine, I thought. I had been taking public transportations for at least 10 years and how bad could that be anyways?

It took me 7 hours to get there. When I arrived at the Getty Center, exhausted, it was already 5 PM. I toured two exhibition rooms for an hour and a half, and then left, fearing that I would miss the last Metro link and would have to sleep in the Union Station. The way back went smoothly and I arrived in Claremont three hours later. After 10 hours on the road and only 1.5 hours spent in the Getty Center, I could hardly have had a better introduction to Los Angeles' public transportation system.

The reason that my outbound trip took 7 hours is that I missed the correct stop or the correct direction for 4 times, and each time I had to wait for an hour for the next bus to arrive. Ok I admit I might have been daydreaming a little bit, but the buses did not announce the names to each stop either. However, the real problem is that if I missed a bus in Beijing, it would have taken 10 minutes for the next bus to arrive, or 3 minutes for the next subway. So the trip would have taken less than 4 hours. Standing in the sun for an hour to wait for the next bus is also a memory that cannot be easily forgotten. I also noticed that all the people who were waiting for buses and taking buses with me were poor people, and mostly non-white. It got me into thinking how come LA, the supposedly modern and filthy rich city, does not have an effective public transportation system. What is the reason that people in cities like Beijing and Tokyo rely much more on public transportation than people in LA? Could the situation have been different? One of the hypotheses I was ruminating on was that since so few people use public transportation, it would not make sense for the public transportation authority to increase frequency or improve service. And since the frequency is low, people will have really bad experience taking buses as I did. Therefore, the whole situation is locked-in this vicious cycle.

Then I forgot about all these because I could think of no way to prove my postulation. At the end of last year, I stumbled upon a modelling method called "agent-based modelling". It rekindled my hope and interest in this topic because this method allows for building fairly complex and interactive models to explore the co-evolution of model components. Therefore, I practiced building agent-based models in the spring semester of 2015, and then built an agent-based model on public transportation frequency during the summer, which ultimately evolved into this thesis.

Table of Contents

Introduction

Los Angeles is notorious for its car dependency, clogged freeways, and lack of good public transportation. In particular, improving public transit is often cited as an effective way to solve the other two problems. The problem is not limited to LA either. Despite its meager 2% total transit share, LA's ridership rate is higher than 38 of the 50 largest urban areas in the US[1].

Public transportation is also linked with social equity. According to US Census data, the poor and the minorities are the primary users of the LA public transit system[2]. Activists protest against the billions spent for massive highway reconstruction, and repeated and ongoing efforts to interfere with transit[3]. Not only are the poor the primary users of public transit, they also can benefit the most from improved public transit service. Harvard business professor Rosabeth Moss Kanter argues that public transit may aid social mobility, and points out that half of the top 20 cities for "intergenerational social and economic mobility" are also half of the top 20 cities for public transit[4]. The importance of improving public transportation is widely recognized, but how one should proceed to do that has sparked lots of controversy.

People disagree on how to improve Los Angeles' public transportation. Wendell Cox, who served for three terms on the late Mayor Tom Bradley's Transportation Commission, says the convenience of the automobile trumps in Los Angeles and public transportation simply won't make sense[5]. Others provide a plethora of proposals. Some suggest using mobile apps to improve

[1] Morris, Eric A. "Los Angeles Transportation Facts and Fiction: Transit." Freakonomics. N.p., 03 Mar. 2009. Web. 11 Sept. 2015.
[2] Walker, Chris. "Why Don't White People in L.A. Take the Bus?" L.A. Weekly. L.A. Weekly, 23 June 2014. Web. 11 Sept. 2015.
[3] Ramey, Corinne. "How America's Transportation System Discriminates Against the Most Vulnerable." America's Unfair Rules of the Road. Slate Magazine, 27 Feb. 2015. Web. 11 Sept. 2015.
[4] O'Hara, Mary. "When Poor People Can't Get on Due to Lack of Public Transport." The Guardian, 26 May 2015. Web. 11 Sept. 2015.
[5] Walker, Chris. "Why Don't White People in L.A. Take the Bus?" L.A. Weekly, 23 June 2014. Web. 11 Sept. 2015.

user experience[6], some recommend exclusive bus ways[7], and some suggest frequent bus rapid transit[8] and stress the importance of population density and walkability[9]. According to a survey report of 12,000 people across the country from the Transit Center, total travel time and travel time reliability are consistently ranked the two most important factors affecting their travel mode choice by all age groups, while cost is only ranked the third by the youth group and the 5th by the other two groups[10]. Jarrett Walker, a professional transit planning consultant, argues that frequency matters because it reduces waiting time, connection time, and improves reliability[11]. Besides building rails or designating exclusive lanes for public transit to avoid congestion, these are indeed the only places to reduce total travel time. Due to the huge costs associated with experimenting with each different idea, it would be helpful to be able to simulate the effect of different policies so that we can make better decisions in reality.

To this end, I developed a theoretical agent based model to study the effect of public transit frequency, reliability, and the population constitution on ridership. The model consists of a heterogeneous population choosing between driving and taking public transit and a public transit authority deciding the optimal service frequency. For each frequency, there will be a long run average ridership, and thus also an average profit level. I assume the transit agency has a mandate of maximizing ridership subject to a budget constraint. The result is a stable equilibrium ridership where neither party has the incentive to significantly change their decisions.

---

[6] Brown, Justine. "Can Technology Help Improve Mass Transit Use?" Government Technology, 14 May 2015. Web. 11 Sept. 2015.
[7] Moore II, James E., and Thomas A. Rubin. "Better Transportation Alternatives for Los Angeles." Reason Foundation, 1 Sept. 1997. Web. 11 Sept. 2015.
[8] Walker, Jarrett. "Los Angeles: The next Great Transit Metropolis?" Human Transit, 30 Mar. 2010. Web. 11 Sept. 2015.
[9] Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.
[10] Walker, Jarrett. "What Motivates Mode Choices for Urban Residents?" Human Transit, 2 Oct. 2014. Web. 11 Sept. 2015.
[11] Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.

Specifically, different setups of the model will be simulated to test several hypotheses. These hypotheses were tentative and were proposed in the hope of directing explorations and experiments with the various assumptions of this model.

Hypothesis 1: there could be multiple equilibria, which is defined in this context as the situation where the public transit agency may think it is impossible to increase ridership furthermore without incurring a long term net loss, but actually profit returns positive after a bigger increase in transit frequency. (See Fig. 1 for an example of this. The two red points represent the two local ridership maxima)

Intuition: this is from my personal experience. I grew up in Beijing where most people use public transportation daily. Cities in Japan and some other countries also have the reputation of having great public transportation that can get people to anywhere fast. This is in such a huge contrast to my experience of using public transportation in LA and I wonder if LA could improve its public transit to the same level of efficiency and reliability without having to subsidize it immensely forever. This might be possible if better and more frequent service makes public transit drastically more attractive and thus resulting in a huge ridership that can share the total cost.

Hypothesis 2: a small increase in train speed will not significantly increase ridership, especially at low train frequencies.

Intuition: more people spent the majority of the time waiting, especially at low frequencies. This is also proposed to explore the sensitivity of the model to the assumption about train speed.

Hypothesis 3: bigger road capacity will make public transportation less attractive but would not affect the likelihood of having multiple equilibria.

Intuition: this will reduce congestion overall, but the reduction will not be significantly bigger at some frequencies over other frequencies. Therefore, there might be an overall decrease in ridership but the reduction will be smooth and will not affect the shape of the profit curve significantly. This is also proposed to explore the sensitivity of the model to the assumption about road capacity.

Hypothesis 4: a change in how heterogeneous the passengers' preferences are will change the likelihood of having multiple equilibria.

Intuition: it makes sense that the population's preference of driving versus taking trains will determine ridership. This is also proposed to explore the sensitivity of the model to the assumption about population preference distribution.

Hypothesis 5: less reliable train arrival time will severely reduce ridership.

Intuition: based on the survey by the Transit Center mentioned before, travel time reliability is one of the most important factors in determining ridership.

Hypothesis 6: there will be large random variations of long term average profit between different simulations (thus hypothetical worlds).

Intuition: passengers in this model are not assumed to be "rational" but base their decision on experiences shaped in part by random factors. Therefore, ridership will have random variations that may last.

Hypothesis 7: lower car ownership will make multiple equilibria less likely to occur

Intuition: if all potential passengers are absolutely dependent on public transit and will use it with no regards to time, then there will be no multiple equilibria at all since the entire population would be already taking public transit at extremely low frequencies and the profit at increased frequencies could only be dropping.

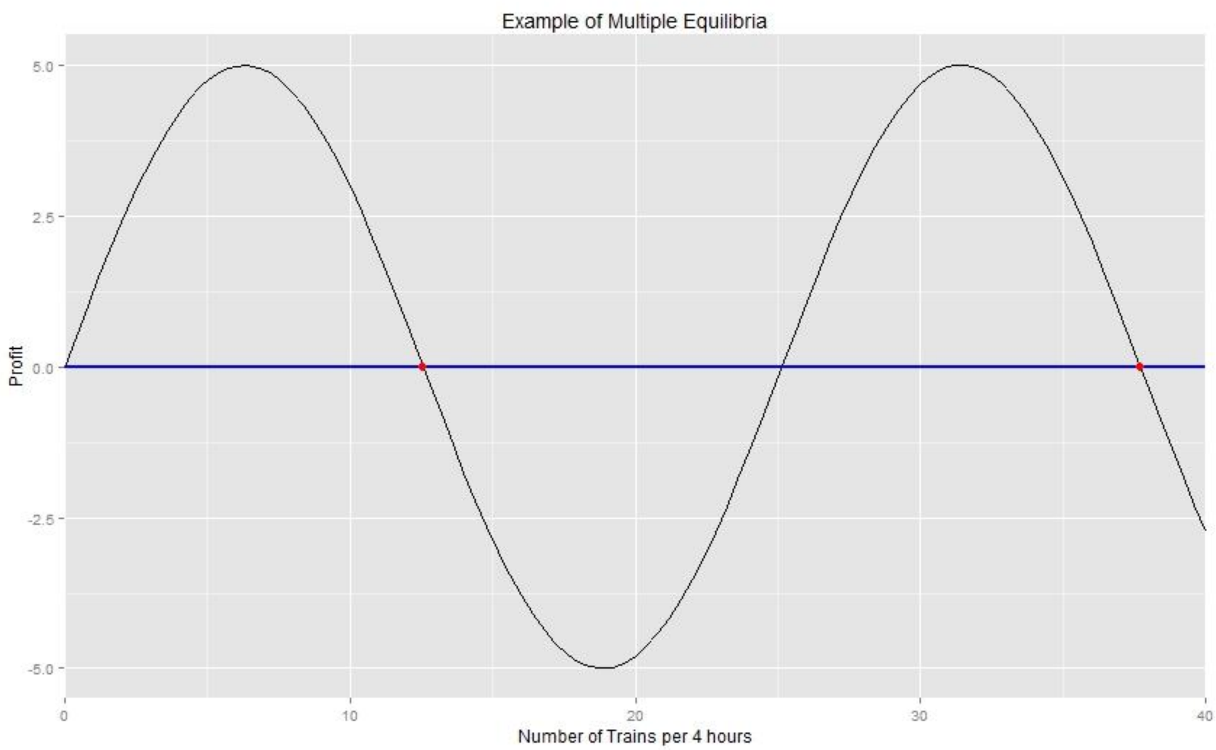


Figure 1. an example profit curve showing multiple equilibria. The two red dots represent two possible outcomes that maximize ridership. The one on the left is the one I have in mind for LA, where the public transportation agency has already maximized ridership (locally) subject to maintaining non-negative profit. An increase in frequency will incur net loss, but increasing it further makes public transit more attractive over driving for so many people that profit returns positive. The global optimum is closer to the red dot on the right, where profit hits zero again but at a much higher ridership level.

The rest of the paper proceeds in the following order: it first describes the modelling approach used for this simulation study and all of its components. Then, it described the method being used and the model calibration. Results follows and are organized by the seven hypotheses I posed earlier. In the end, the conclusion of this study, the limitation of the methods used, and possible future directions of this study are discussed.

Theoretical Model

Components

In order to test for existence of multiple equilibria (see Fig. 1), we would need a curve corresponding to the profit level at each train frequency so that we could identify frequencies where the public transit agency could choose to operate in the long run. Therefore, for each frequency within a reasonable range, the profit level is calculated by averaging long term ridership and subtracting the total cost of operation.

The cost of adding each extra train into service is assumed to be constant to reflect the idea of average levelized cost, which spreads the total present value of the initial investment and future operating costs to the entire life cycle to obtain an index for long term average cost per day of operation. Adding one more train is similar to building one more electricity generating plant in that they both have a large initial investment and long term operation cost that could fluctuate. Levelized cost is the primary way of measuring cost of generating electricity from different sources, thus it can be useful here as well.

Moreover, since this model is theoretical rather than empirical, all parameter values are important only in a relative manner. Therefore, I adopted a more explorative approach by adjusting cost so that the profit level will cross the x axis and become negative within the frequency range of 1 to 39 that I simulated for all cases. In this way, the "cost" is actually an indication of the profitability of the train service so that we can compare the profits at different frequencies. Also, this will have the added benefit of giving the profit curve interesting behaviors, which is the goal of my exploration.

The revenue will be proportional to ridership under the assumption of fixed ticket price. Ridership is a result of a population with heterogeneous preferences choosing between taking trains and driving for commuting. Each individual makes decisions according to a well-defined rule and is affected by both one's own preference and the choices of other people. More people taking trains will make it more crowded and thus less attractive. How much people can tolerate longer waiting time and travel time will depend on the preference values. More details on how decisions are actually made will follow in the sections below.

The time frame and the space

The model has discrete time. The time step is one day. In each day, all individuals make their travel mode decisions first, commute to work in the mode chosen, then update their decisions the next day based on their experience. We are assuming no weekends, holidays, or other special circumstances that can exogenously cause cyclical fluctuation in the number of travelers.

The space is a hypothetical world where there is only one road and two train stops. All people live near one of the stops and work near the other stop. This is a simplified setup, but it is nonetheless illustrative of the effect of train frequency and passenger preference. The setup resembles the case of daily commute to and from work better and is especially useful for that case. A more realistic road network will boost the importance of frequency[12], and will only make the phenomenon studied here more prominent.

Individuals

In computational models, it is not uncommon to model agents with bounded rationality and make decisions using different heuristics or rule of thumb strategies[13]. These models were shown to perform reasonably well – simple models can explain important observed stylized facts in financial time series, such as excess volatility, high trading volume, temporary bubbles and trend following, sudden crashes and mean reversion, clustered volatility and fat tails in the returns distribution[14].

In this model, individuals are assumed to have bounded rationality; specifically, they would not know what decisions would be the best for others or what would be chosen by others. Thus, they would have to rely on their past experiences. They would expect the time of driving to work and the time of taking trains to work is equal to a combination of their most recent five experiences. The expectation of travel time in mode i of person j is given by a partial geometric lag[15]:

$$T\_exp_{i,j} = t_{0,i,j} + \gamma * t_{1,i,j} + \gamma^2 * t_{2,i,j} + \gamma^3 * t_{3,i,j} + \gamma^4 * t_{4,i,j}$$

Next time

$$T\_exp_{i,j} = t_{0,i,j} + \gamma * t_{1,i,j} + \gamma^2 * t_{2,i,j} + \gamma^3 * t_{3,i,j} + \gamma^4 * t_{4,i,j}$$

Initially, all five memories are the same and set equal to the time of driving or taking trains without congestion or waiting. Then, after each round, $t_{k+1,i,j}$ will be substituted by $t_{k,i,j}$ (j=0, 1,

[12] One of the primary reasons frequency matters to ridership is that it makes connections easier and faster. Therefore, frequency matters more when there are connections to make. See Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.
[13] Hommes, Cars H. "Heterogeneous agent models in economics and finance." Handbook of computational economics 2 (2006): 1109-1186.
[14] *Ibid.*
[15] t_0 means the most recent experience, and it will become t_1 the next day; i takes on two values, train (tr) or car (c), j denotes each person, and γ is the time discount factor, currently set to 0.9

2, 3), and t_{0i} will be updated to the new experience if i equals the current travel mode. Memory of the other travel mode remains unchanged.

Then, individuals will minimize their expected travel time subject to certain constraints. Individuals are assumed to minimize their travel time because, according to the survey by the Transit Center, total travel time is consistently ranked the most important consideration by all age groups when people make decisions on whether to use public transit (table 1).

Table 1. Potential Drivers of Transit Ridership by Age[16]

I WOULD RIDE TRANSIT MORE IF...	UNDER 30 (RANK)	30–60 (RANK)	OVER 60 (RANK)
it took less time	1	1	1
stations/stops were closer to my home/work	4	2	2
it were clearly the less expensive transportation option	3	3	3
the travel times were more reliable	2	4	4
there were different transit modes available	7	5	5
it ran more frequently	8	6	6
the stops/stations were safer	6	7	7
the buses/trains were cleaner/nicer	5	8	8
the hours of operation were extended	10	9	11
there were more parking available at the station	12	10	9
the seats were more comfortable	11	11	10
it offered reliable access to Wi-Fi/cellular	9	12	12

For modelling purposes, a single preference value is used to summarize all other considerations and constraints, such as distance to train stations, cost of driving and taking trains, safety, amenities on trains, difficulty of parking, relative importance of convenience, the ability to work

[16] *Who's on Board: Mobility Attitudes Survey*. Rep. The Transit Center, 2014. Web. 1 Sept. 2015. <http://transitcenter.org/wp-content/uploads/2014/08/WhosOnBoard2014-ForWeb.pdf>. 22.

while on trains, and environmental benefits. This method of modelling was draw from Granovetter's seminal paper *Threshold Models of Collective Behavior*[17]. In this paper, Granovetter used a single "threshold" value for each person to represent all characters that can affect how likely that person will participate in a collective action. The threshold is simply that point where the perceived benefits to an individual of doing the thing in question (joining the riot in his paper, in this case taking public transit) exceed the perceived costs. The idea is especially useful when there are many often unquantifiable factors that can affect a person's decision and allows a straightforward way of modelling the dynamics of group behavior. This method also allows for modelling continuous changes in characteristics and was shown to be able to generate interesting emerging results that could not have come from the crude dichotomy of splitting people into homogeneous groups.

The preference value will determine the traffic mode decision i_j of person j in the following way:

$$^{18}\; i_j = \begin{cases} train\ if\ \dfrac{T\_exp_{tr,j}}{T\_exp_{c,j}} \leq p_j \\ drive\ if\ \dfrac{T\_exp_{tr,j}}{T\_exp_{c,j}} > p_j \end{cases}$$

In other words, if $\frac{T\_exp_{tr,j}}{T\_exp_{c,j}} \leq p_j$, the relative benefit of taking trains is considered greater than the relative benefit of driving for person j. For example, if someone is really poor and can only afford taking trains, then his/her preference value will be infinity, which means however long the train will take, he/she will always choose the train. A preference value of 1 means the person is totally indifferent and will choose the faster mode. A preference value of 0 represents avoidance of trains, due to preference for comfort, long distance to train stations, or other reasons. A value

[17] Granovetter, Mark. "Threshold models of collective behavior." American journal of sociology (1978): 1420-1443.
[18] i_j is the travel mode choice of person j. p_j is the preference value of person j.

of 2 means the person will not switch to driving as long as expected travel time of taking trains is less than twice of the time of driving.

The Population

There are N residents in the model. It is assumed they live in the same neighborhood around a train station, have the same schedule, and work at places close to another train station. Therefore, if they were to choose taking trains for their daily commute, they will all want to take the same train. They live at different distances to the train station. They also have different attributes such as wealth, car ownership, health status, valuation of time, parking distance and fees at destination, and personal preferences between taking trains and driving. Relative price of gas and train ticket will also affect people's choices. These attributes would be summarized by the "preference" value. It does not take into account anything related to the state of the world, such as train frequency. The exact distribution of preference in the population is unknown, thus many possible distributions are simulated to determine the effect of the distribution. Granovetter used a continuous distribution of "thresholds" in his model, thus, to start with, conventional continuous probability density functions were used to model the distribution of preferences in the population. The beta distribution is chosen first because it cannot be negative and can be doubled so that its peak is at 1.

[19] $$\frac{p_j}{2} \sim \ Beta\ (20{,}20)$$

However, one should note that the beta distribution is biased in favor of driving. The need to have a quotient rule requires the probability density function of preferences to have a certain

[19] Each person's preference value comes from a certain distribution. Many distributions are explored, including gamma, beta, uniform, and homogenous (no distribution).

property to ensure unbiasedness, which is $f(x) = f(\frac{1}{x})$. Also, since 1 stands for indifference towards the two transportation modes, an ideally disinterested preference distribution will have half of its value smaller than 1, half of its value bigger than 1, and its peak be at 1. So we also need $F(x) = 0.5$. One way to get this ideal distribution is to transform the beta distribution.

$$\frac{p_j}{2} \sim Beta\,(20, 20)$$

$$p_j = \begin{cases} p_j \; if \; p_j < 1 \\ \dfrac{1}{2 - p_j} \; if \; p_j \geq 1 \end{cases}$$

However, it turns out the ideal distribution does not noticeably change the result. So I kept the beta distribution for convenience.

For faster model run time, the population is set at 300.

The Road, Cars, and Trains

It is assumed that trains have a separate rail from the road, which is used by cars. More cars on the road will result in congestion. The most authoritative approximation of the relationship between congestion and traffic load is the BPR function, named after the U.S. Bureau of Public Roads[20]. It has been cited as among the best performing models[21]. Its formula is:

$$T_{final} = T_{initial} * \left(1 + \alpha \left(\frac{V}{C}\right)^{\beta}\right)$$ [22]

[20] US. Federal Highway Administration. Office of Environment, Planning, and Realty. Travel Model Validation and Reasonableness Checking Manual. Federal Highway Administration, 24 Sept. 2010. Web. 1 Oct. 2015.
[21] Saberi Kalaee, Meead. Investigating freeway speed-flow relationships for traffic assignment applications. Portland State University, 2010.
[22] Travel Model Validation and Reasonableness Checking Manual. Federal Highway Administration. P9-22

Where:

T_{final} is the final, congested travel time on a link;

$T_{initial}$ is the initial, or starting, travel time on a link (without congestion);

V is the assigned volume on a link;

C is the capacity of the link (at level of service E); and

α and β are model coefficients

A summary of calibrated values of the coefficients α and β is reported below[23]:

Table 2. Range of Reported BPR Function Assignment Parameters

Facility Type	**α**		**β**	
	Minimum	Maximum	Minimum	Maximum
Freeways	0.10	1.20	1.90	10.00
Arterials	0.15	1.00	2.10	4.00

The function used in this model is a simplified version of it because many of the adjustment parameters related to road conditions were not relevant as only the model would only have one road segment. The average travel time needed to get to the destination, t_c, is comprised of the time needed without congestion T_c and the congestion time, which is determined by the number of cars on the road N_c and two coefficients a and b:

$$t_c = T_c + \left(\frac{N_c}{a}\right)^b$$

And there will be variability in how long it takes each person to arrive, so the actual car travel time of each driver comes from a normal distribution of mean t_c and standard deviation σ^2.

[23] Travel Model Validation and Reasonableness Checking Manual. Federal Highway Administration. P9-22

$$t_{c,j} \sim Normal\ (t_c, \sigma^2)$$

A graphical representation of their relationship is shown in figure 1.

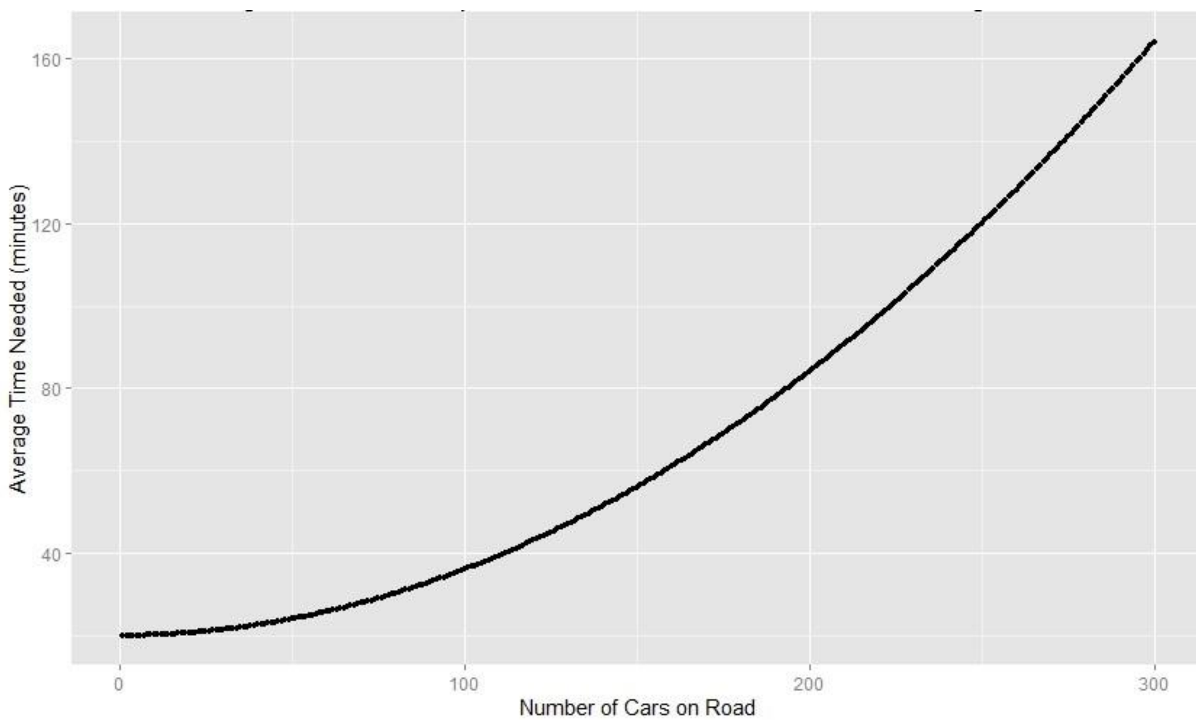


Figure 2. Relationship between Number of Cars on road and Congestion when a=25, b=2

Trains arrive at fixed intervals determined by the number of trains in service. More precisely, the interval, $T_{tr\_interval}$, is assumed to be the total period of time under consideration T_{total} divided by the number of trains in service during this period, N_{trains}. A graphical representation of their relationship is shown below in figure 2.

$$T_{tr\_interval} = \frac{1}{f} = \frac{T_{total}}{N_{trains}}$$

f denotes frequency of service, so its inverse represents the time interval between two trains.

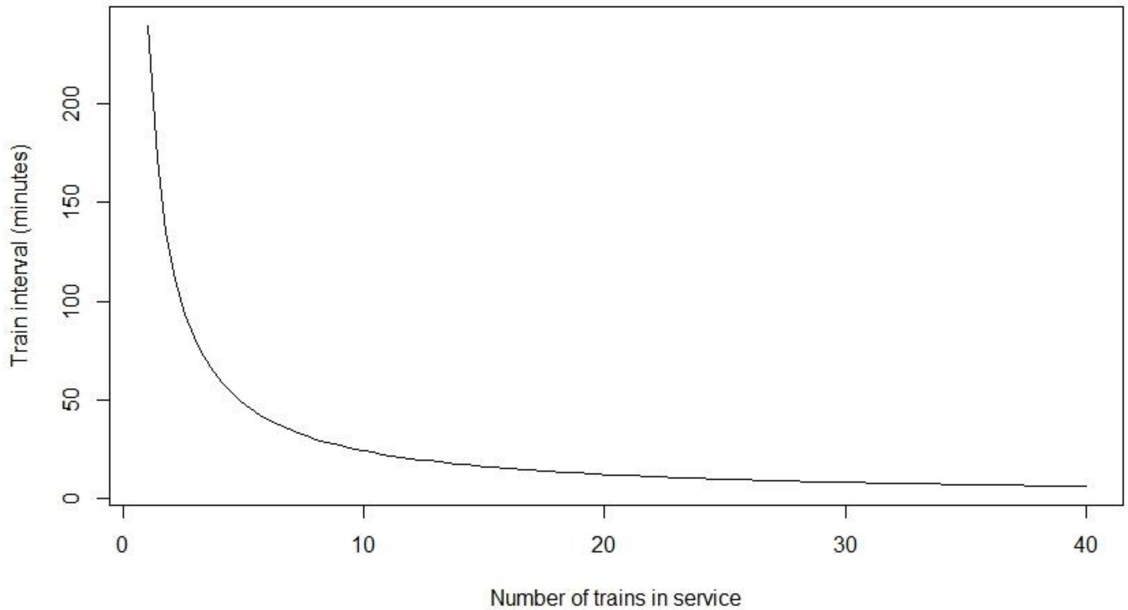


Figure 3. Relationship between Number of trains in service and Arrival interval (simply a reciprocal function, the natural relationship between time and frequency)

Moreover, there is a limit on the number of passengers that can board a train. Therefore, if the amount of passengers is twice as the limit, then each passenger has equal chance of boarding the first or the second train. In other words, half of passengers will have a longer wait time, and how bad that wait time is depends on train frequency. This is to model the fact that if there are lots of passengers and few trains, then the experience will be much worse than if there are fewer passengers and more frequent trains. This set up is based on almost ten years of my personal experience of taking subways in Beijing, where the limit on the number of passengers each train can take consistently caused "congestion" of public transit. However, it doesn't need to be one passenger on each train. So, in this model, the number of passengers on each train will not affect how much each passenger likes trains as long as the number of passengers is below the limit. The total time of taking trains to work for person j, $t_{tr,j}$, is given by:

$$t_{tr,j} = T_{tr} + \frac{\varepsilon_j - 1}{f}$$

On the right hand side, $T_{tr}$is the time needed for the trip, and the second part is the waiting time, where ε_j is a random integer denoting which train the person got onto (equal probability), ranging from 1 to the integer part of $\frac{\mathrm{N_{tr}}}{\mathrm{L}}$, the total number of trains required. $\mathrm{N_{tr}}$ is the number of people taking trains, and L is the limit on number of people per train.

So the average commute time of taking trains is given by

$$\frac{\sum_{i=tr,j=1}^{n} t_{i,j}}{n} = T_{tr} + \frac{int\left(\frac{N_{tr}}{L}\right) - 1}{2f}$$

The left hand side is the average train passenger's time used, because n is the total population, which equals 300 in this study.

The idea is that more frequent services is able to accommodate more passengers, while infrequent services can only serve a much smaller passenger load for whom the train service happens to be convenient. Big and small values of the limit will both be simulated.

Overall, the number of people driving will be given by:

$$N_c = \sum_{j=1}^{n} [1 - I(i_j)] = n - N_{tr}$$

where $I(i_j) = \begin{cases} 1\ if\, i_j = train \\ 0\ if\ i_j = driving \end{cases}$

And the number of people taking trains will be given by:

$$N_{tr} = \sum_{j=1}^{n} I(i_j) = n - N_c$$

The Train Service Provider

The train service provider will maximize ridership subject to maintaining positive profit by deciding how many trains to put into service. Maximizing ridership is a fairly common mandate for public transportation, especially for the rush hour service that is being simulated here, since it is similar in nature to utility services. For example, the transit plan of Houston Metro proposed to devote 80% of its resource to maximizing ridership[24]. The ticket price is assumed to be fixed, because the price is really stable in reality[25], and this model is more interested in the effect of service frequency.

Since increasing frequency will in general reduce wait time, it should boost ridership and thus also revenue. However, we also assumed costs associated with having more trains in service. Whether increasing service frequency would raise or reduce profit depends on the situation. Obviously, increasing train frequency from 0 to 10 per 4 hours is going to have a much bigger effect than increasing frequency from 100 to 110 per 4 hours. There are also a limited amount of people. So eventually, the train authority will suffer net loss from having more trains. It also has the mandate of maximizing ridership as long as profit remains positive. Therefore, it will try to find the point where the profit is slightly positive or zero and putting one more train into service will turn that profit into negative.

[24] Walker, Jarrett. "Houston: Transit, Reimagined." Human Transit, 09 May 2014. Web. 11 Sept. 2015.

[25] For example, base fare in Chicago was $0.45 in 1970, or $2.49 in 2009 adjusted for inflation, while the base fare in 2009 was $2/$2.25. For bus fare see http://www.chicagobus.org/history. inflation was calculated by http://www.usinflationcalculator.com/

However, the train service provider is assumed to have bounded rationality. It does not know where the global maximal ridership will be. Therefore, if it turns out that there are multiple local maxima (see fig. 1), which one will be chose depends on the starting point.

Agent based modeling (ABM)[26]

Based on the modeling needs, the most appropriate modelling approach would be agent-based modelling. The model that has been described above is an agent-based model. In agent-based models, a system is modeled as a collection of autonomous decision-making entities called agents and their common residing environment. Each agent individually assesses its situation and makes decisions on the basis of a set of rules. Agents may interact with other agents directly, or by reacting to changes in the system environment caused by actions of other agents. Through the interactive and adaptive behaviors of agents, complex phenomena may emerge from the ABM system. ABM is especially useful when agents could have nonlinear actions (either drive or take trains, nothing in between), when agents have memory and adaptation, and when agent interactions are heterogeneous (a diverse preference distribution) and can generate network effects (i.e. traffic congestion, train becomes too crowded).

Simulation

Calibration

There are many variables that can be calibrated, but they can be grouped so that the relative magnitudes within one group are more important than their absolute values.

[26] Bonabeau, Eric. "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Sciences 99.suppl 3 (2002): 7280-7287.

Population, limit on the number of people in each train, and how fast congestion time grows as the number of cars increase are in one group. If the population ($n = N_{tr} + N_c$) doubles, the limit (L) doubles, and the congestion time growth is halved (“a” is doubled), then the dynamics will not change[27].

$$ave(t_{tr}) = \frac{\sum_{i=tr,j=1}^{n} t_{i,j}}{n} = T_{tr} + \frac{int\left(\frac{N_{tr}}{L}\right) - 1}{2f}$$

$$ave(t_c) = T_c + (\frac{N_c}{a})^b$$

$$N_{tr} = \sum_{j=1}^{n} I(i_j) = n - N_c$$

The time in which a car and a train can cover the whole distance without congestion, the frequency of trains, and how fast congestion time grows as number of cars increases are in one group. A doubling of train frequency can attract more people to use trains, but the relative impact of that can be reduced by a doubling of the time needed for cars and trains to cover the distance. A 20 minutes wait time is much more bearable if the trip is 100 minutes long than if the trip is 5 minutes long. If it is very easy to get completely stuck in congestion, then however fast a car can go without congestion wouldn’t matter.

$$T\_exp_{tr,j} \propto ave(t_{tr}) = T_{tr} + \frac{int\left(\frac{N_{tr}}{L}\right) - 1}{2f}$$

$$T\_exp_{c,j} \propto ave(t_c) = T_c + (\frac{N_c}{a})^b$$

[27] Suppose both N_{tr} and N_c doubled, but L and “a” are doubled as well, then neither $ave(t_{tr})$ nor $ave(t_c)$ changes, so the system will remain stable if no other variable changes.

In light of the difficulty in determining the correct relative magnitude and the lack of relevant literature, the model will be set up with plausible parameter values and later tested against other possible relative setups.

Method and dynamics

The simulation was carried out through the computer programming language Python.

At the start of the simulation, the computer program first assumed a fixed train frequency. Then, 300 people were created with random preference values drawn from a distribution based on a prior assumption about the shape of the overall preference distribution of the population. In the beginning, they do not know anything about potential congestion or waiting time, so they make decisions without considering them. Then, there will be a number of cars on the road and a number of passengers waiting for trains. Each individual's travel time will be determined in part by the actual situations on the road and at the train station but also to a smaller extent by random variations. Several days later, they gradually get better ideas about congestion and waiting time and update their decisions based on recent experiences.

Every day, all people make decisions based on their own memories and preference values without knowing other people's decisions. Since the train frequency is assumed to be the same, over time, the number of people choosing each travel mode will stabilize as shown later in actual simulations. After 10000 rounds, the average ridership will be calculated. It could be a good proxy for long run average revenue of a certain service frequency because, as will be shown later, ridership in most cases stabilizes in less than 100 rounds. Then, the same simulation is repeated 100 times to calculate the expected average ridership at that frequency.

Now the expected average ridership was calculated for one service frequency. Then I repeat the whole process for all frequencies in 1, 3, 5… 39 per 4-hou-period because that provides enough range of values of possible frequencies and also gradual changes to show details along the way. For example, for a one-hour period the number of trains can only change from 1 to 2, but for a four-hour period it can be any number from 4 to 8, which translates into 1, 1.25, 1.5, 1.75, and 2 trains per hour. Having 6 trains per 4 hours is the same as having 1.5 trains per hour; only it is more intuitive.

With average ridership values of each frequency, a plot would be created with its x axis as the service frequency and its y axis as the long run average ridership at each frequency. The ridership could also be viewed as the long run average revenue because the train fare of each passenger is assumed to be fixed. Then, as described in the model components section, a cost value would be fitted to the curve so that it could be an approximation of the attractiveness of the service. A high value means the service could still be worth doing even if it would cost a lot. Subtracting the cost from the revenue, we would have a profit curve of the same nature as figure 1, which is what we needed for the purpose of this study.

Results

The result section will be organized by hypotheses. For each hypothesis, I changed some of the assumptions about the model to test the effect it will have on the model and the profit curve. All the assumptions in each case are summarized in the following table, with the changed assumptions in bold.

To begin with, two base cases were simulated, and they happened to be potential examples of multiple equilibria. They both have a preference distribution based on the beta distribution shown in the model description, and a congestion function shown in figure 1 (a=25, b=2). They both have cars traveling slightly faster than trains without congestion, and a fixed standard deviation of actual congestion time (σ^2) of 5. The first case has a bigger limit on number of passengers per train than the second case. These numbers were tentative explorations, but they were used as the bench mark later since they displayed the phenomenon of interest. To signify its special status, the first case is indexed case 0.

In case 2 through 5, I experimented with train speed, the congestion function, and the train limit to get a better idea of the sensitivity of previous results to their assumptions. To my surprise, all four cases displayed multiple equilibria; some of them were even more prominent than the base cases. Although there are considerable variations in the manifestation of multiple equilibria that showed the impact of the assumptions tested, the overall result demonstrated a considerable amount of robustness to these assumptions.

Case 6 and case 7 explored the effect of the assumption on the preference distribution. Case 6 explored an assumption about the distribution that is more spread-out, thus the model has a more heterogeneous population. Case 7 explored the case where all individuals are indifferent between the two travel modes. In both cases, multiple equilibria do not exist. Case 7 exemplified the advantage of agent based models; we would never have had examples of multiple equilibria without the heterogeneous agents permitted by agent based modeling.

Case 8 through 10 explored the assumptions about the reliability of both trains and cars. Trains were assumed to be perfectly reliable in previous cases. In case 8, they were assumed to arrive in random intervals based on an exponential distribution with the mean interval still determined by train frequency. Reliability of train arrival interval turned out to have a gigantic impact on ridership. In case 9 and 10, the variability of car arrival time, rather than fixed, was assumed to be proportional to the severity of the congestion.

After that, I plotted profit curves of some previous cases differently to show the random variations that have nothing to do with all assumptions tested or train frequency. These plots showed that a public transportation authority, even with the exact same service provided, could run a net profit or a net loss depending on what kind of potential customers established the habit of taking trains.

In the end, case 11 and 12 explored the consequence of a lower car ownership rate among the population. This was done by assuming a huge preference value (which would force them to always choose public transit) among a portion of the population. Case 11 has a larger group of people (one third) with no car than case 12 (one fifth). It turned out that lower car ownership rate will undermine the conditions of multiple equilibria.

Below, all these cases will be discussed in greater detail.

Table 3. A summary of all cases

case	Distribution of preference	Mean of the distribution	Standard deviation of the distribution	Train travel time (minutes, no waiting)	car travel time (minutes, no congestion)	Train limit (person)	Congestion function	Standard deviation of congestion (σ^2)	Train arrival schedule	Cost factor in graph
0	2*Beta(20,20)	1	0.156	25	20	30	a=25, b=2	5	fixed	7
1	2*Beta(20,20)	1	0.156	25	20	**100**	a=25, b=2	5	fixed	13.5
2	2*Beta(20,20)	1	0.156	**20**	20	30	a=25, b=2	5	Fixed	7

3	2*Beta(20,20)	1	0.156	25	20	30	**a=100, b=4**	5	fixed	5.5
4	2*Beta(20,20)	1	0.156	25	20	**100**	**a=100, b=4**	5	fixed	10
5	2*Beta(20,20)	1	0.156	25	20	**70**	**a=100, b=4**	5	fixed	8
6	**2 * Beta(3,3)**	1	0.378	25	20	30	a=25, b=2	5	fixed	7
7	**homogenous**	1	0	25	20	30	a=25, b=2	5	fixed	10
8	2*Beta(20,20)	1	0.156	25	20	30	a=25, b=2	5	**Exponential**	3.5
9	2*Beta(20,20)	1	0.156	25	20	30	a=25, b=2	**Proportional to congestion**	Fixed	15
10	2*Beta(20,20)	1	0.156	25	20	30	**a=100, b=4**	**Proportional to congestion**	fixed	9

Hypothesis 1: there could be multiple equilibria

Case 0 is the base case for testing, and case 1 changed the limit on the number of people in each train from 30 to 100 for the purpose of sensitivity test. The preference distribution for both cases is the default distribution, which is 2 times beta (20, 20) (Fig. 10). Trains are assumed to be slightly slower than cars when there is no congestion. Congestion is modelled by a quadratic function which was shown by figure 1. It is assumed that when there is congestion, on average all cars slow down, but some cars will be affected more than others. The standard deviation of car arrival time is assumed to be 5. Trains are assumed to arrive on a fixed schedule. These assumptions are summarized in tables and the results will be interpreted later in the discussion.

This is the base scenario so that we can change the assumptions later to test for other hypotheses.

Case 0 (Base case)

Distribution of preference	2*Beta(20,20)
Mean of the distribution	1
Standard deviation of the distribution	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	5
Train arrival schedule	fixed
Cost factor in graph	7

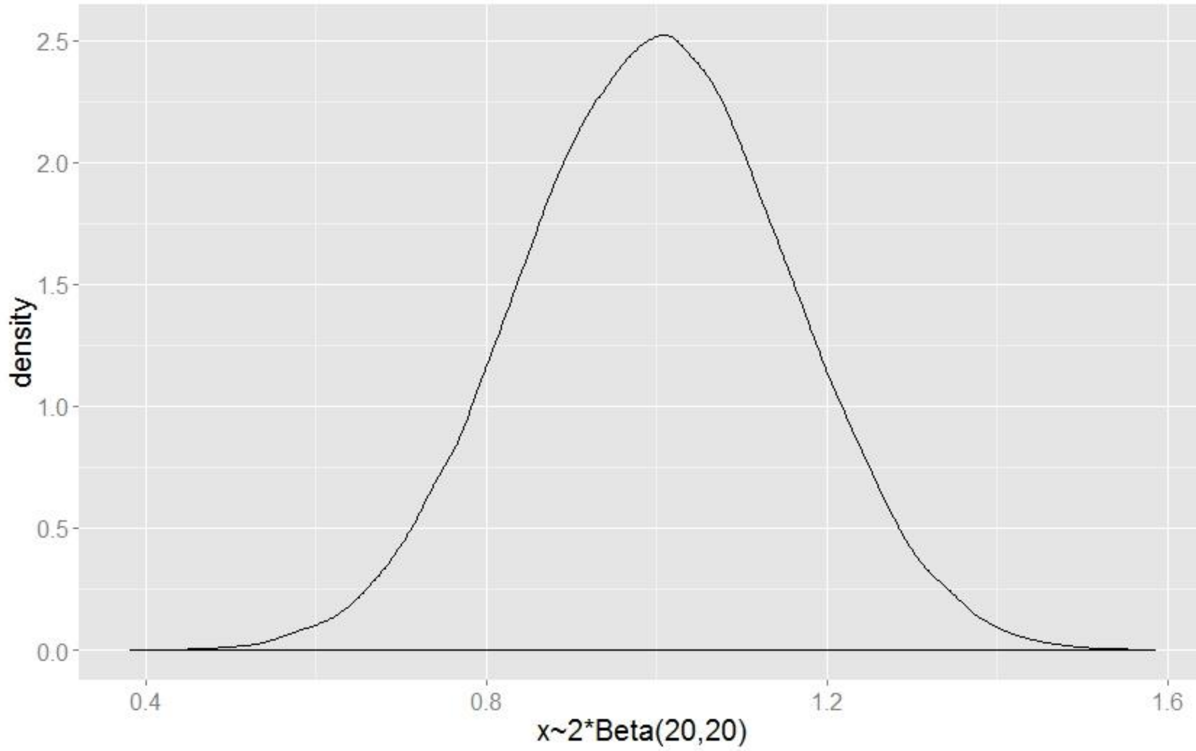


Figure 4. Density plot of 100000 samples from 2*Beta (20, 20)

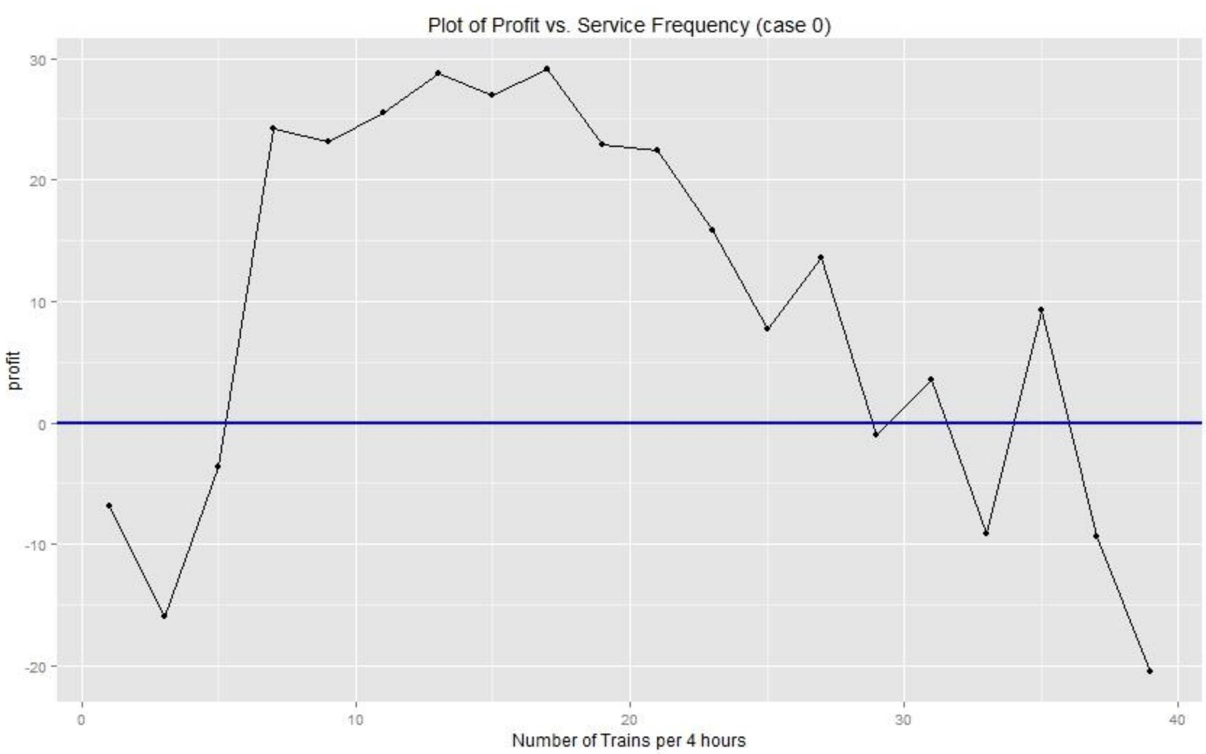


Figure 5. Plot of Profit vs. Service Frequency (case 0)

Case 1

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	100
Congestion function	Quadratic
Standard deviation of congestion	5
Train arrival schedule	fixed
Cost factor in graph	13.5

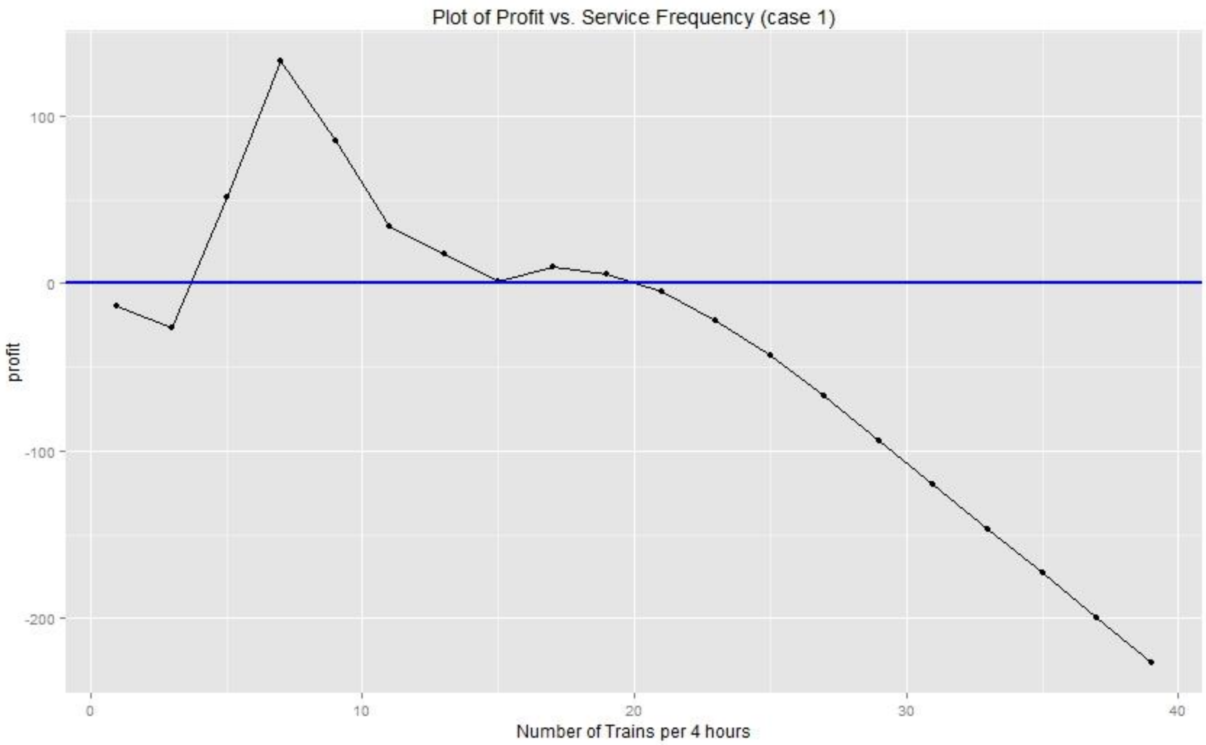


Figure 6. Plot of Profit vs. Service Frequency (case 1)

Discussion

1. The tradeoff between profit and ridership

   As shown in fig. 11 and fig. 12, due to the limited number of residents the train station serves, there clearly is a decreasing marginal return of higher frequency, which will result in plummeting profits eventually. But a frequency that is too low also means a lower ridership and usually incurs a negative profit unless there is a group of people that absolutely depend on the transit no matter how long it takes. For the segment with positive profit, there can be equilibrium at each frequency and the transportation authority faces a tradeoff between profit and ridership. Since the mandate is to maximize ridership under a budget, a well-functioning transit agency will choose the point where the profit curve crosses the x axis from above.

2. The possibility of multiple equilibria

   However, there could be more than one point where the profit curve crosses the x axis from above. For example, in case 0 the three such points are around 29, 32, and 36, and in case 1 there are two such points, 15 and 20. Unfortunately, the transit agency is not omniscient and it may feel satisfied at the point 15 instead of experimenting with higher frequency. Because the profit kept falling as the agency ramped up frequency, the agency has every reason to assume that this trend will continue and it will lose more money by increasing frequency furthermore. Although there is a hint of multiple equilibria, it does not seem to be a very strong one.

Hypothesis 2: a small increase in train speed will not significantly increase ridership, especially at low train frequencies.

Although according to the survey mentioned in the introduction, total travel time is one of the most important factors in determining ridership, at lower frequency the impact of speed should be trumped by the supposedly long waiting time. To test this hypothesis, the travel time needed for trains in case 2 is reduced by 5 to 20. All other assumptions of case 2 are the same as case 0. The cost factor in the result should be paid special attention because that number signifies how attractive trains are to the public in general.

Case 2

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	20
car travel time (minutes, no congestion)	20

Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	5
Train arrival schedule	Fixed
Cost factor in graph	7

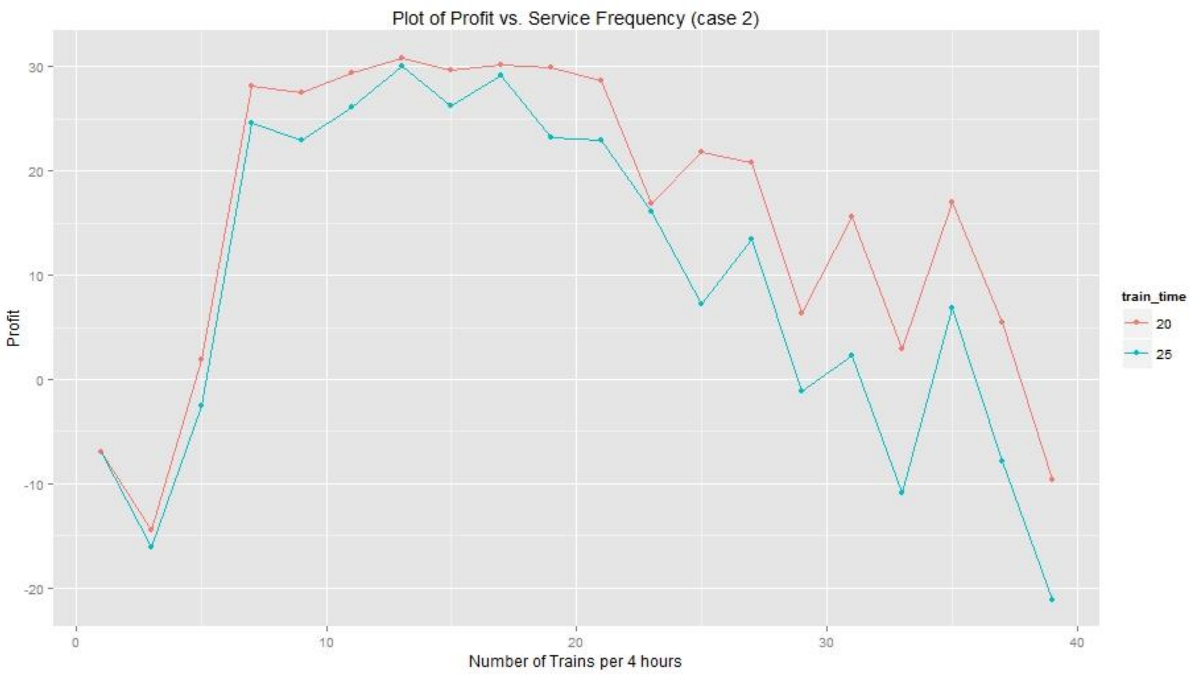


Figure 7. Plot of Profit vs. Service Frequency (comparison of case 2 and case 0)

There seems to be a relationship between train frequency and the difference in ridership between the two cases, so I ran a simple linear regression. I hypothesized that it matters more to have fast er trains when the train frequency is higher. The response variable is the difference, and the expla natory variable is the train frequency. Their values were calculated by subtracting the ridership at each frequency in the two cases shown in figure 7. The data size is 20. I assumed no intercept be cause when the train frequency is zero (no service), the ridership in both cases should be zero and thus there should no difference in ridership either.

$$\text{difference} = \beta * \text{frequency} + \varepsilon$$

I was interested in whether the coefficient of frequency is significantly difference from zero and how good is the overall fit. A significantly positive coefficient would be consistent with my hypothesis.

Residuals				
Min	1Q	Median	3Q	Max
-6.5784	-1.3037	-0.2388	1.5545	6.5815
Coefficients:				
	Estimate	Std. Error	t value	Pr(>|t|)
Frequency	0.32049	0.02855	11.22	7.95e-10 ***
Multiple R-squared:	Adjusted R-squared			
0.8689	0.862			

Significance code: *** p<0.01, ** p<0.05, * p<0.1

The effect of train frequency on the difference in ridership is significantly positive at the 1 percent level. The r-squared is surprisingly high, signifying a good fit. This result is consistent with my hypothesis.

Discussion

Case 2 has similar shaped profit curve as case 0, as well as very similar cost factors. This means it matters very little to have a faster public transit, especially when the service frequency is low. It could make more sense to improve train speed when the train frequency is higher. Moreover, the relative importance of frequency increases in a public transit network, because a higher frequency reduces waiting time more when people have to make transfers in a network. So the speed of trains may have an even smaller effect in a more connected network. Therefore, in general, policy makers should focus more on train frequency.

Hypothesis 3: bigger road capacity will make public transportation less attractive but would not affect the likelihood of having multiple equilibria.

The following three cases were simulated to study the effect of the shape of the congestion function. They were based on the base cases, only with a different congestion function. The old congestion function has a=25 and b=2. The new congestion function has a=100 and b=4. Both values for "b" were reasonable values from table 1 and both values for "a" were calibrated to give the congestion function a reasonable range of values. The new congestion function is compared to the old one in figure 13. It reduces congestion, thus reflects an increased road capacity. The new congestion function is expected to reduce ridership, because intuitively, more people will choose to drive if the road would become less congested. A reduced cost factor in the result would corroborate this hypothesis.

Case 3, case 4, and case 5 have train limits of 30, 100, and 70 respectively. All their other assumptions are the same.

Case 3

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=100, b=4
Standard deviation of congestion	5

Train arrival schedule	fixed
Cost factor in graph	5.5

Instead of the old congestion function:

$$t_c = T_c + \left(\frac{N_c}{25}\right)^2$$

The new car travel time is given by:

$$t_c = T_c + \left(\frac{N_c}{100}\right)^4$$

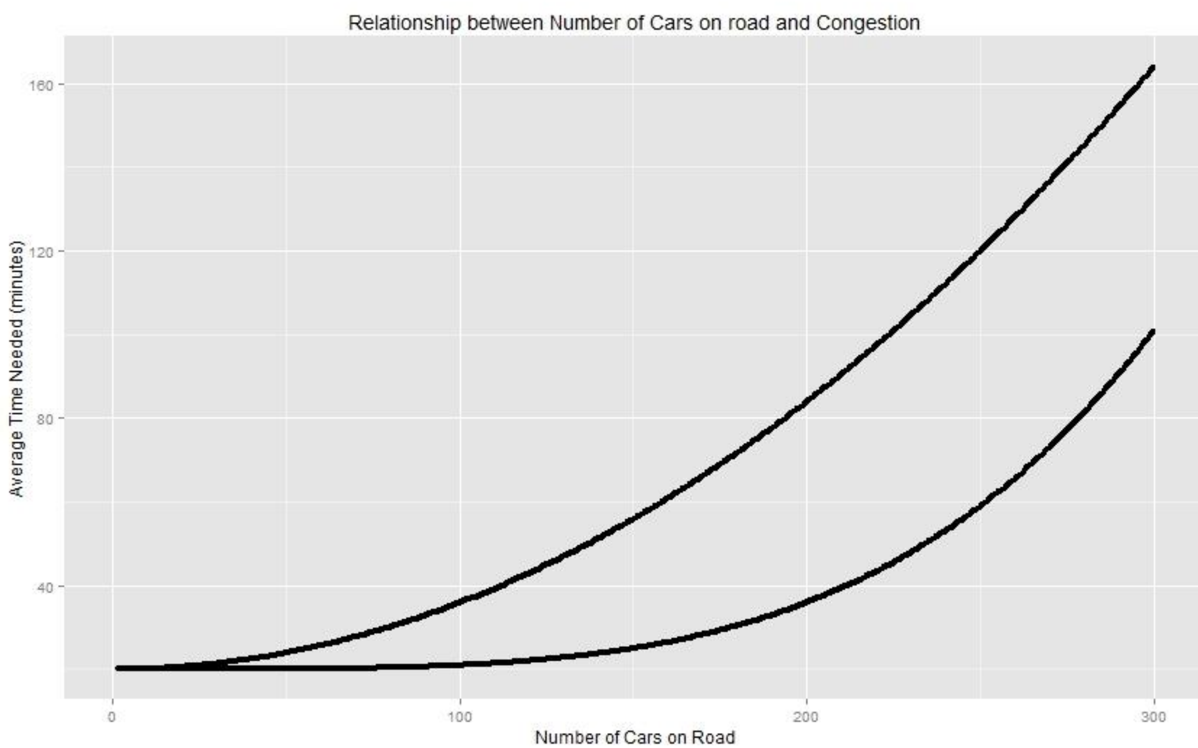


Figure 8. Comparison of the two congestion functions. The above one is the old one (quadratic), and the one below is the new one (quartic).

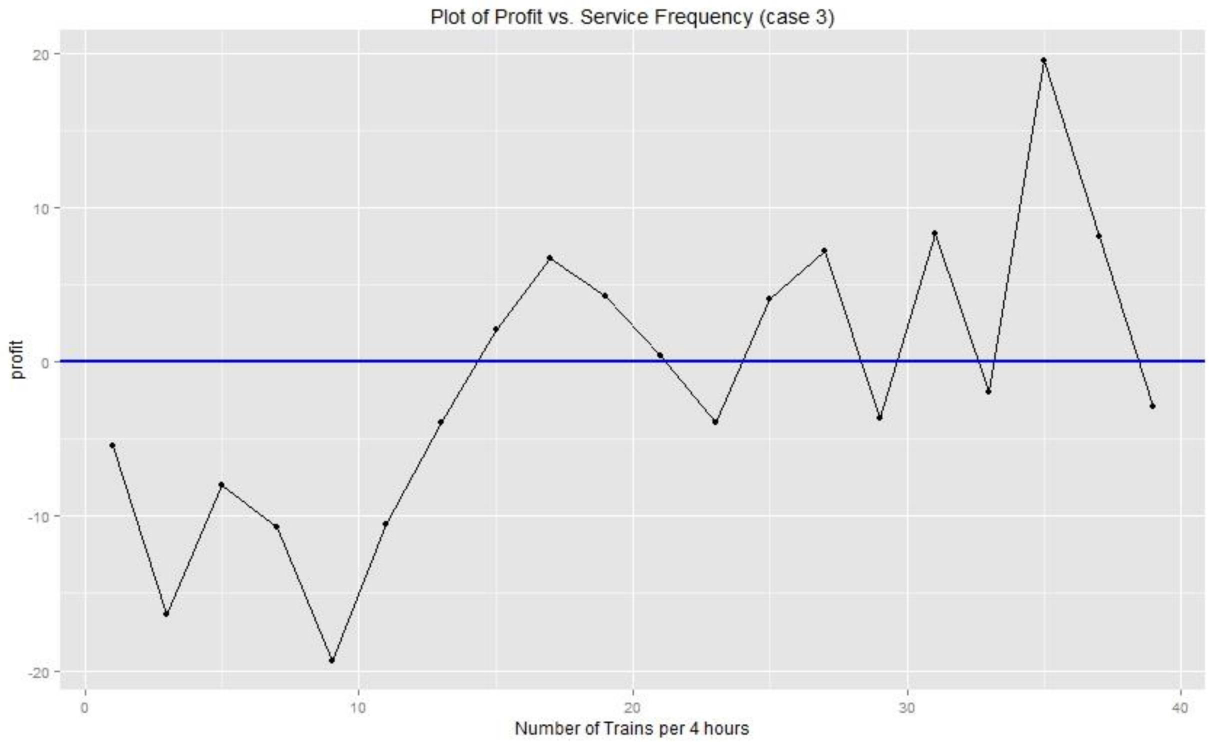


Figure 9. Plot of Profit vs. Service Frequency (case 3)

Case 4

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	100
Congestion function	a=100, b=4
Standard deviation of congestion	5
Train arrival schedule	fixed

Cost factor in graph	10

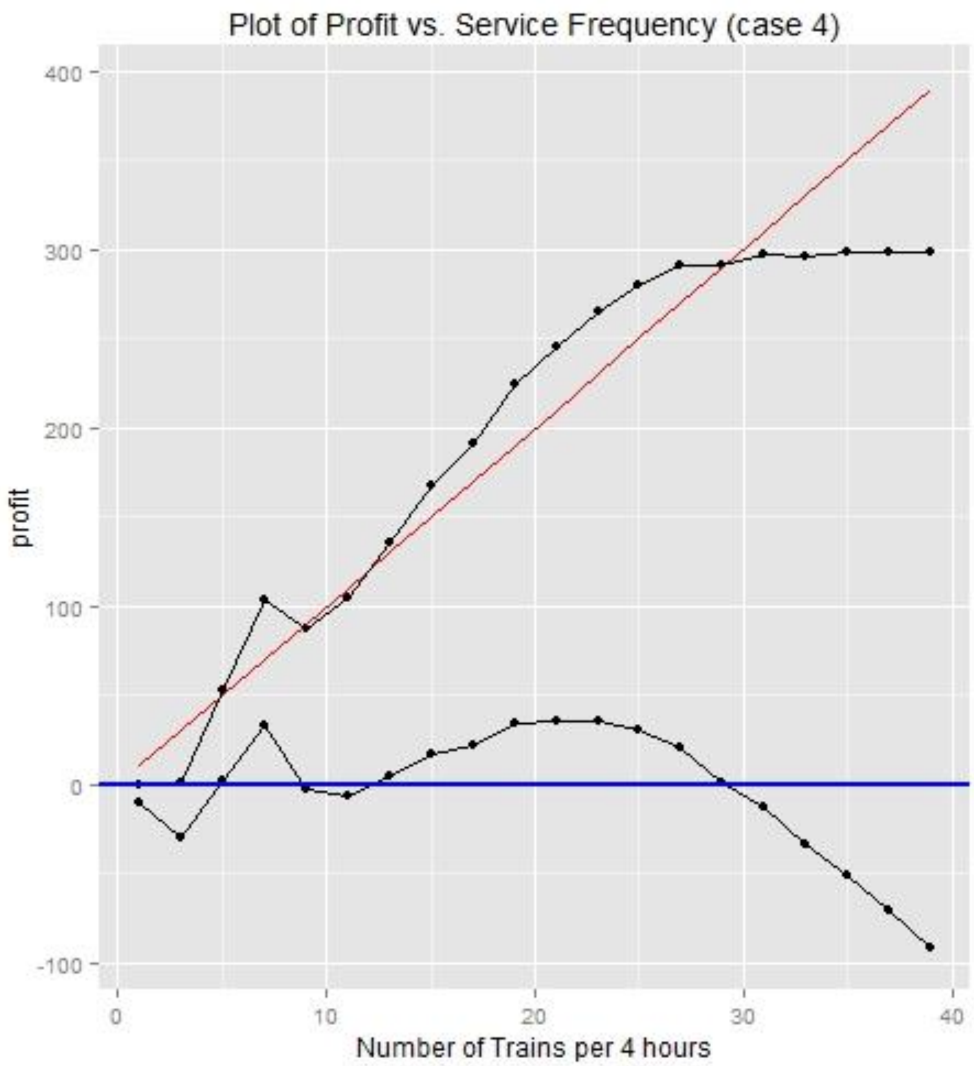


Figure 10. Plot of Profit vs. Service Frequency (case 4). The black curve below represents profit, the black curve above represents revenue, and the red line represents the "cost".

Case 5

Distribution of preference	2*Beta(20,20)

mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	70
Congestion function	a=100, b=4
Standard deviation of congestion	5
Train arrival schedule	fixed
Cost factor in graph	8

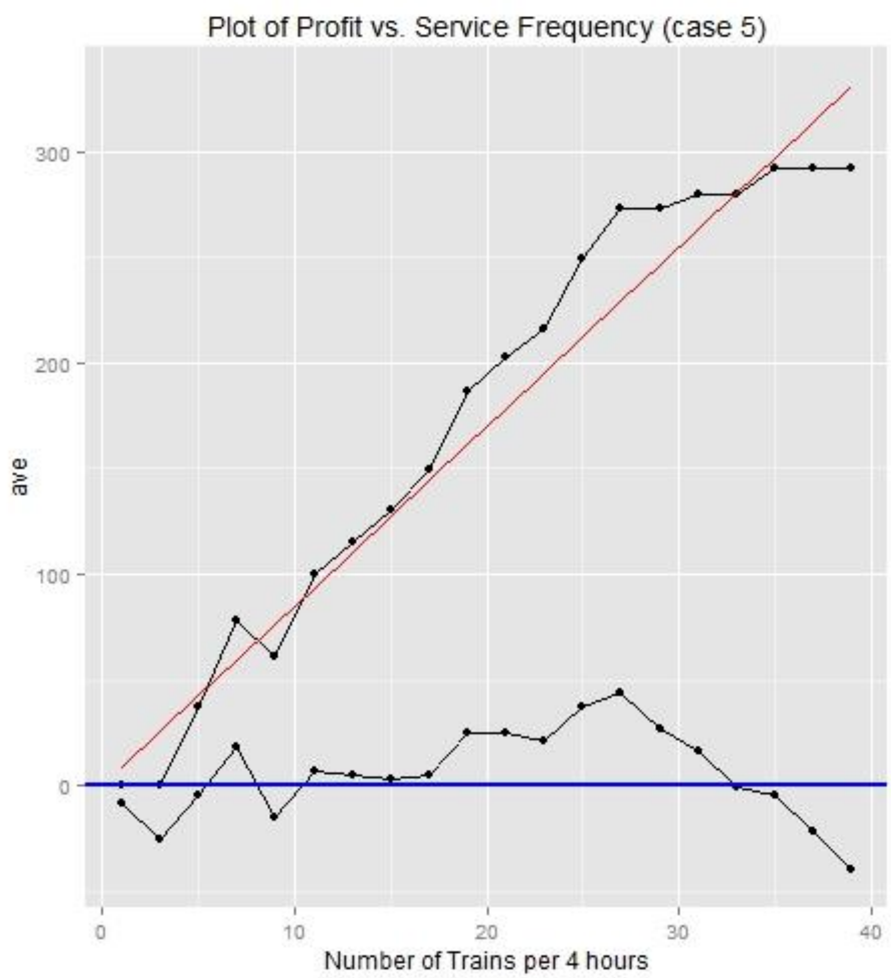


Figure 11. Plot of Profit vs. Service Frequency (case 5). The black curve below represents profit, the black curve above represents revenue, and the red line represents the “cost”.

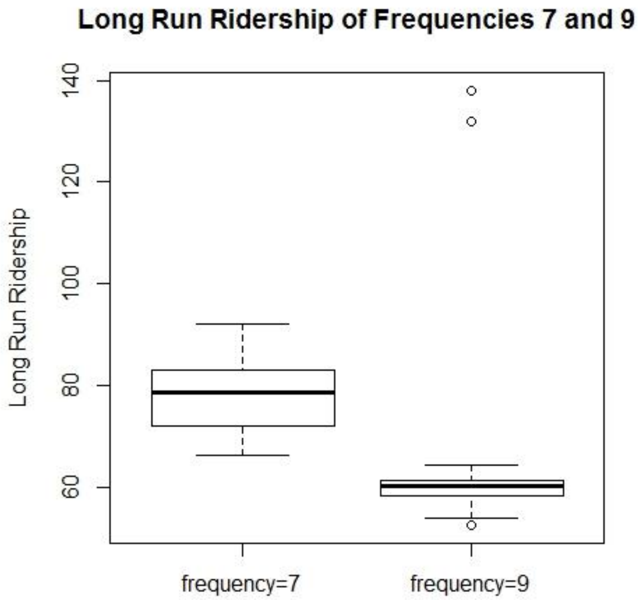


Figure 12. boxplot of long run ridership at frequency 7 and 9 for case 5. (t=13.004, df = 164.71, p-value < 2.2e-16)

Discussion

Indeed, we can see that with a bigger road capacity comes a smaller cost factor, which means the transit is less appealing to riders and it can bear less cost. The phenomenon of multiple equilibria also seems to be more prominent in these cases than in the base cases. Unfortunately, there does not seem to be an obvious explanation or a test that can show it is indeed more prominent. More interestingly, and quite counterintuitively, the total ridership (thus revenue) drops significantly at frequency 9 for the train limit of 70 and 100 cases (see figure 10 and 11). The decrease in ridership is statistically significant (p<2.2e-16). In previous cases, the total ridership (not shown) could plateau despite increasing frequency, which caused profit to plummet, and then start increasing again, which caused profit to become positive again. Never has total ridership dropped due to higher frequency. However, in case 4 and 5, there is a drop in total ridership due to higher frequency for unknown reasons.

Overall, increasing the road capacity indeed makes public transportation less attractive. However, it could have complicated effects on the existence and the prominence of multiple equilibria, and could make the optimal frequency more difficult to predict. This also showed the importance of making more reliable assumptions on the relationship between traffic load and congestion as it could fundamentally change the result. Practical models for the purpose of informing real world decisions could calibrate the congestion function based on measurements of the particular road under consideration.

Hypothesis 4: more heterogeneous passenger preference will make multiple equilibria more likely to occur.

One of the most influential assumptions of the model is the preference distribution. This hypothesis is to explore the effect of preference distribution on the existence of multiple equilibria. To test this hypothesis, two other preference distributions were run, one was more heterogeneous than the previous one (beta (3, 3)), the other one was homogeneous.

Case 6

Distribution of preference	**2 * Beta(3,3)**
mean	1

Standard deviation	0.378
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	5
Train arrival schedule	fixed
Cost factor in graph	7

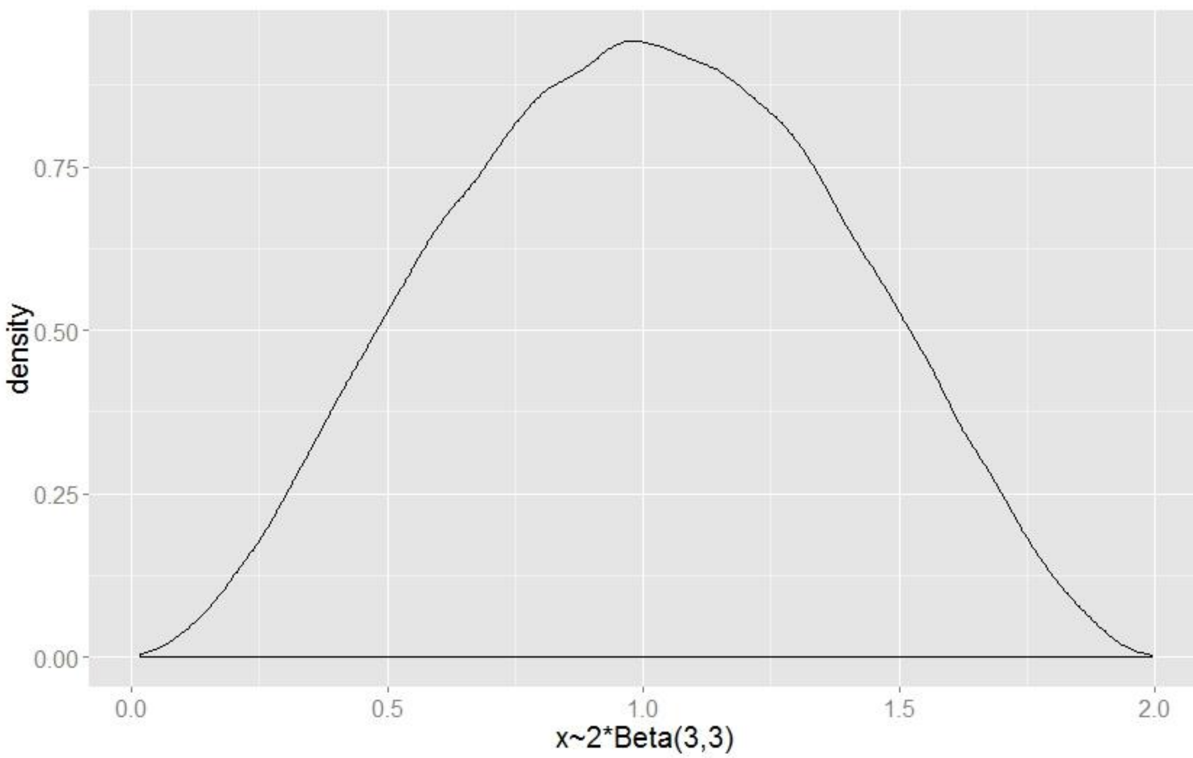


Figure 13. Density of 100000 samples of 2*Beta (3, 3)

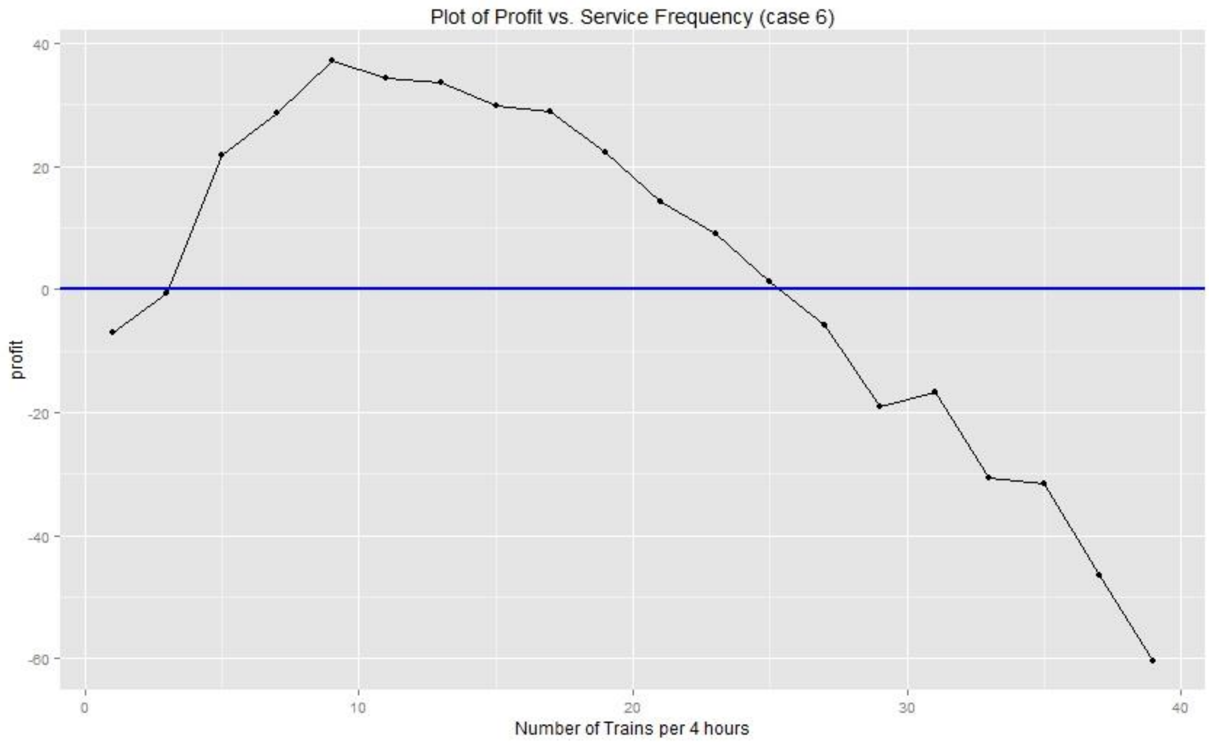


Figure 14. Plot of Profit vs. Service Frequency (case 4)

Discussion

Compared the previous cases, the one listed above have populations with a more variable preference. This is clear from the more-than-doubled standard deviation, as well as from comparing figure 4 and 13. Interestingly, the profit curve is much smoother and no multiple equilibria have been found however I twist the cost factor. This curve resembles the standard textbook profit curve of a firm. Cases with even more heterogeneous populations were simulated and their results were similar to that of case 6 (not shown to conserve space). In general, a similar shaped but more heterogeneous population does not lead to multiple equilibria. A possible explanation is that the existence of multiple equilibria depends on there being a large group of people who are more or less indifferent between the two transportation modes so that they would only switch to using public transit after huge improvements in service quality have been made. That group would need to be big enough in number to overturn the previously negative profit into positive. There also need to be few people who are going to take public transportation with little regards to the service quality so that a lackluster improvement in service quality would not suffice. With a combination of these two factors, terrible services and terrific services would both be profitable, but mediocre services would not. A more heterogeneous population like the one in case 6 increases the group of people who are reliable on public transit and reduces the group of people who are more demanding. Therefore, a more heterogeneous population would undermine the conditions for the existence of multiple equilibria.

Case 7

Distribution of preference	**homogenous**
mean	1
Standard deviation	0
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	5
Train arrival schedule	fixed
Cost factor in graph	10

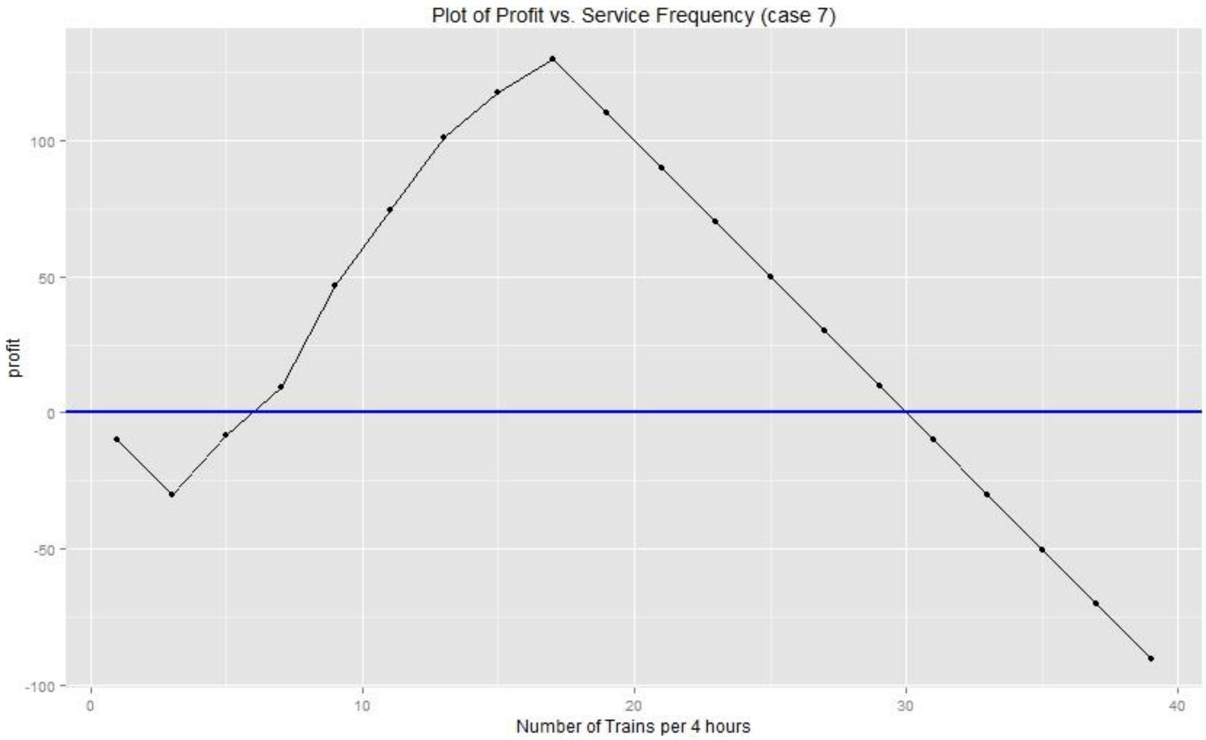


Figure 15. Plot of Profit vs. Service Frequency (case 7)

Discussion

On the other extreme, a population with the least variable preference values would be a homogenous population. In this case, all the people are assumed to have a preference value of 1, which means they are indifferent between the two transportation modes and will choose the faster one. The cost factor is much higher than that of case 0, which means that people are much more attracted to public transportation in this case. One should note that this case is a lot less realistic because such a homogenous group never exists. No multiple equilibria exist in this case either. The ridership shoots to the maximum and falls back all in a steady manner.

A possible explanation for the absence of multiple equilibria could be the following. The indifferent group of people are more sensitive to the service quality of public transportation, thus could easily be crowded out by the group of people who badly needed public transit. The group relying on public transit makes ridership abnormally high at very low service frequency, thus makes a lackluster improvement seemingly pointless. Without such a group, the indifferent group would switch to using public transit gradually, resulting in a more steady change in ridership.

The previous two cases showed the importance of the assumption of heterogeneous preferences in modelling transit demand as even a subtle difference on that can make a difference on the existence of multiple equilibria.

Hypothesis 5: less reliable arrival time will severely reduce ridership.

In previous cases, we assumed the train arrives on a fixed schedule. However, the reality may not be as perfect. Moreover, according to the survey by the Transit Center, travel time reliability is among the top two considerations when deciding travel mode. Therefore, it would be interesting to test the effect of a less reliable train arrival time on ridership. In case 8, trains were assumed to arrive based on an exponential distribution whose mean value (i.e. expected arrival time) is the same as that in the fixed arrival case. Although the expected train travel time would stay the same, this set up would introduce considerable variability in actual travel times.

Firstly, the real time fluctuation of ridership during the first 10000 rounds will be plotted for both the base case and the new case. Then, the usual profit versus service frequency plot will be discussed.

Comparison of fixed and exponential arrival time in single simulations

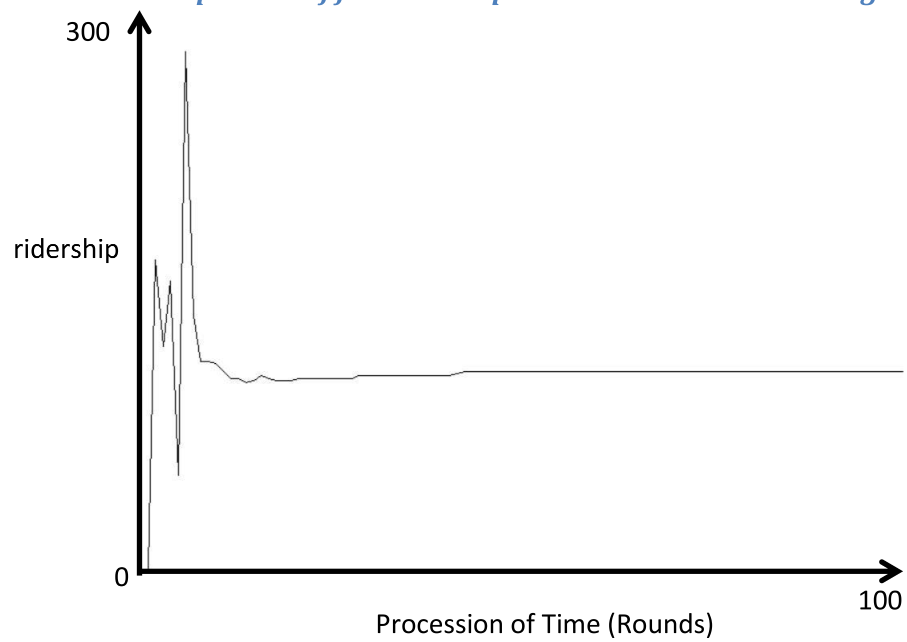


Fig. 16. Ridership in a single simulation (x=time, y=ridership) – base model, 100 rounds

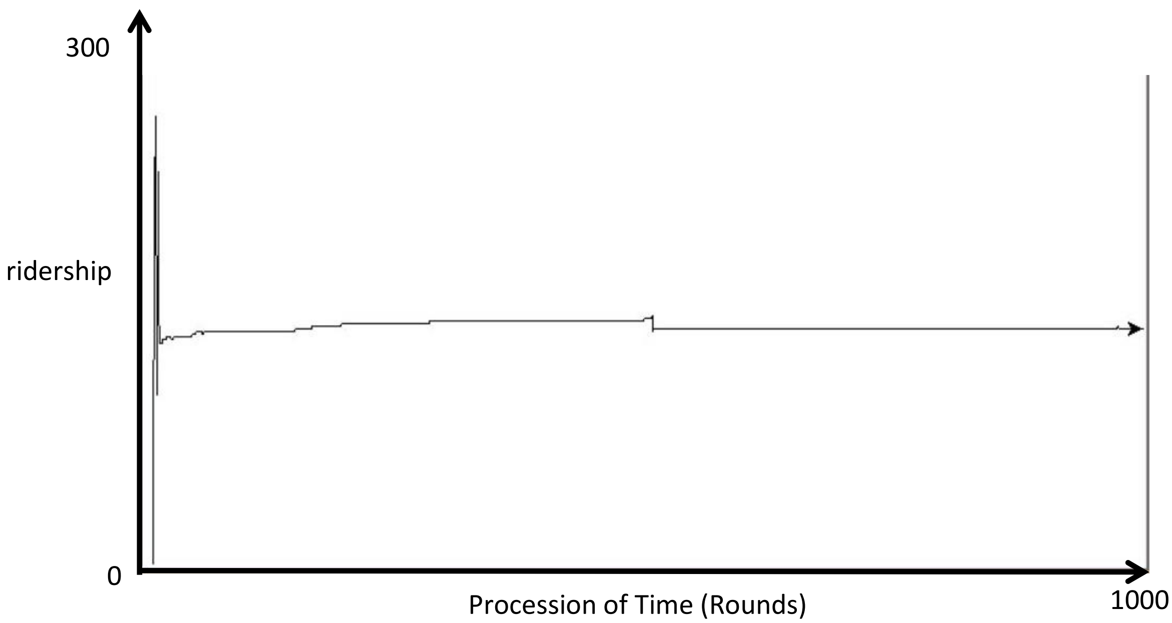


Fig. 17. Ridership in a single simulation (x=time, y=ridership) – base model, 1000 rounds

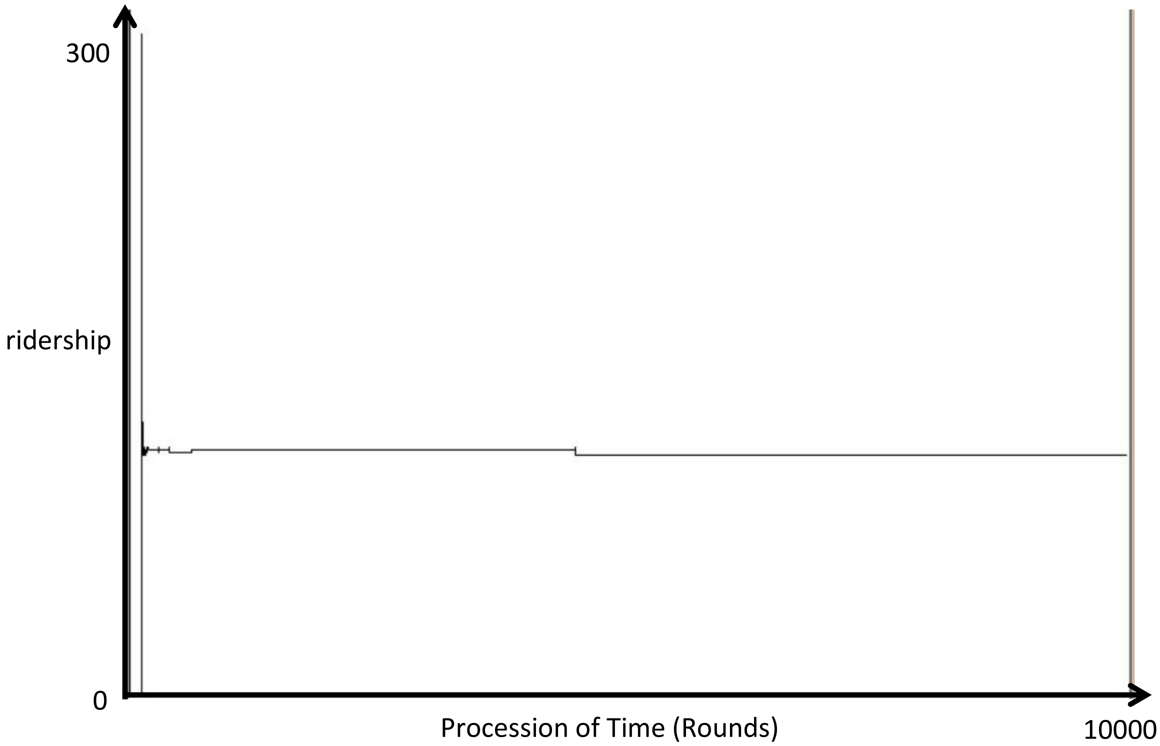


Fig. 18. Ridership in a single simulation (x=time, y=ridership) – base model, 10000 rounds

Figure 16 to 18 are graphical representations of train ridership under fixed train frequency (which is 10 in the above graphs). The x axis represents the passage of time, and the y axis represents ridership (0-300). Clearly, the ridership becomes stable very fast. In these cases, trains are assumed to arrive on fixed schedule, as shown by the equation:

$$T_{tr\_interval} = \frac{1}{f} = \frac{T_{total}}{N_{trains}}$$

$$t_{tr,j} = T_{tr} + \frac{\varepsilon_j - 1}{f}$$

Therefore, stable ridership means passengers will have the same experience in the future, thus the number of people driving and taking trains will remain fairly stable as well.

In the case shown by the following two graphs, trains are assumed to arrive in intervals which are drawn from the exponential distribution that has its mean equal to the mean expected interval determined by the train frequency. First, all passengers wait for the first train to come. Then, people are randomly selected to fill the first train until it is full, and the rest wait for the next train. The interval between the (k-1)th train and the kth train, T_k, is given by

$$T_k \sim Exponential(\frac{1}{f})$$

So that

$$t_{tr,j} = T_{tr} + \sum_{k=1}^{\varepsilon_j} T_k$$

This set up creates a lot more uncertainty in the time needed for taking trains, so we would expect the ridership to be less stable than the base case. This is exactly what happened.

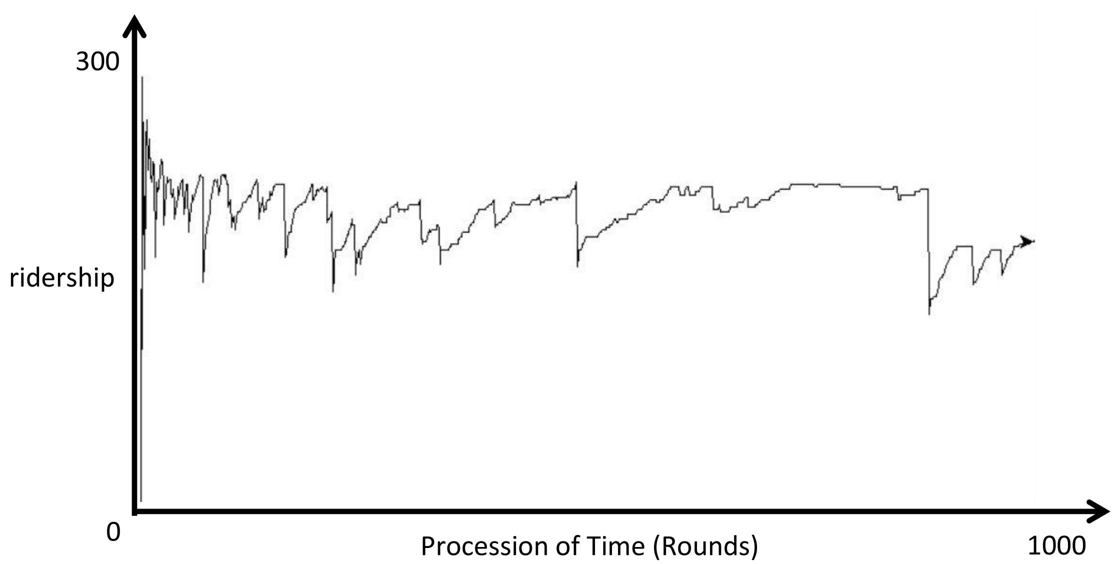


Fig. 19. Ridership in a single simulation (x=time, y=ridership) – exp model, 1000 rounds

Since the ridership did not seem to have stabilized, more rounds (10000) were simulated.

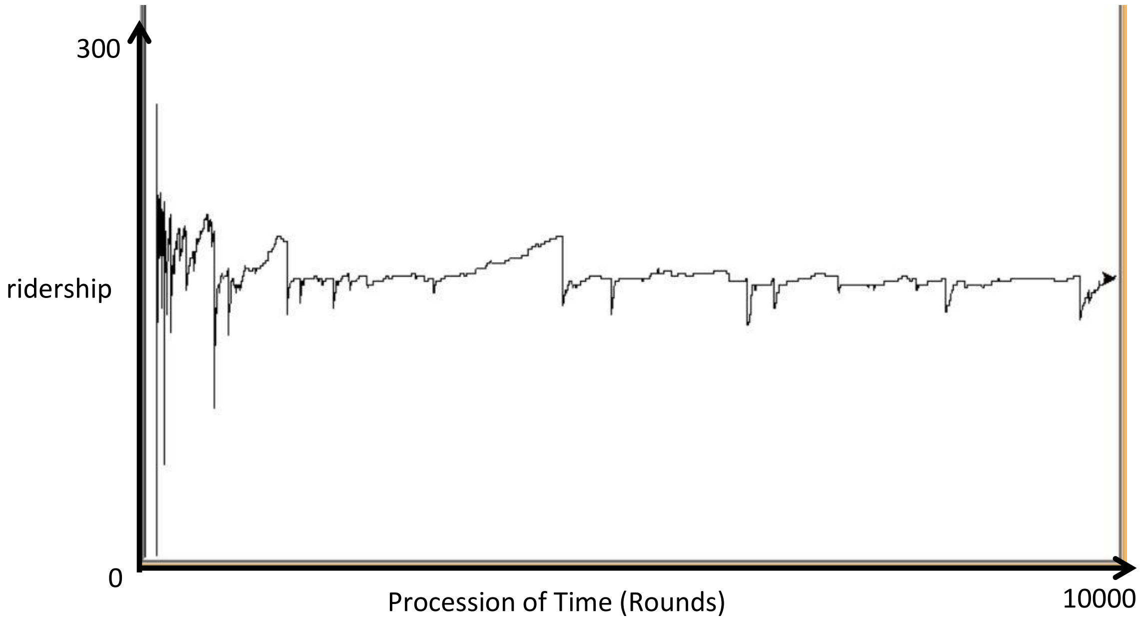
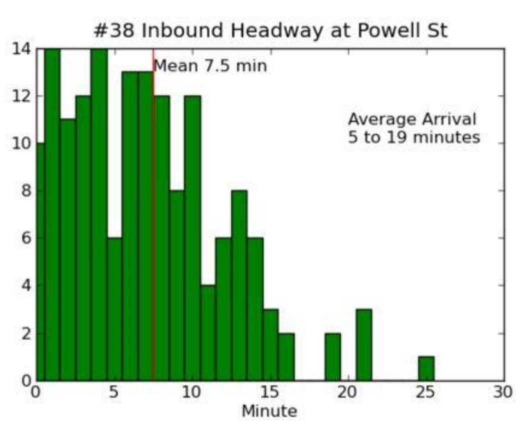


Fig. 20. Ridership in a single simulation (x=time, y=ridership) – exp model, 10000 rounds

In the base model, ridership seems to stabilize after about only 10 rounds (Fig. 16). In comparison, it seems the exponential model will never be as stable; ridership fluctuates around a certain value even after 10000 rounds. All cases that are studied later have fixed arrival schedules unless explicitly specified.

In reality, the arrival schedule is less reliable than the fixed case, but not as variable as the exponential one. Transportation blogger Jarrett Walker displayed a distribution of real world arrival intervals of the San Francisco Muni collected by one of his readers.



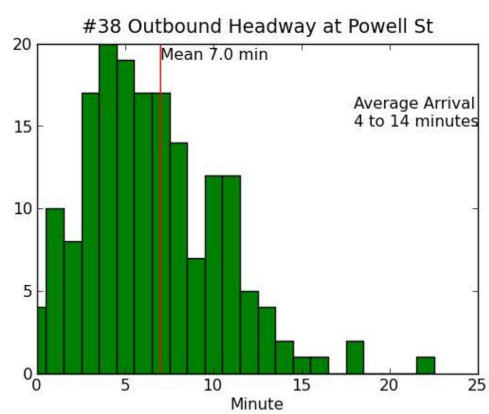


Figure 21. Distribution of Arrival Intervals at Powell St Station[28]

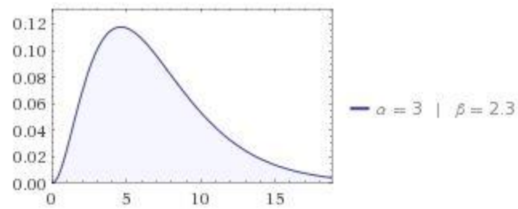


Figure 22. Probability Distribution Function of gamma (shape = 3, rate = 2.3)

The outbound distribution is very similar to a gamma distribution (3, 2.3), which has an average of 6.9 and a standard deviation of 3.98. On the other hand, an exponential distribution with a mean of 7 has a standard deviation of 7. Therefore, the actually data should be less variable than the exponential distribution, but certainly more variable than a fixed schedule, which has a standard deviation of 0. The reliability of public transportation in reality is in between the two cases shown above, so the variability of real world ridership may also be between that of the two cases. Exponential arrival was simulated in case 8. The gamma arrival, which better resembles the real world scenario, was also simulated, but would not be reported here, since it turned out to be quite similar to the exponential case.

A more reliable service not only makes passengers less hesitant in choosing public transportation, but also makes it easier for public transportation authorities to adjust service frequencies.

Case 8

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156

[28] Walker, Jarrett. "Now, Anyone Can Monitor Reliability." Human Transit, 30 June 2010. Web. 11 Sept. 2015.

Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	5
Train arrival schedule	Exponential
Cost factor in graph	3.75

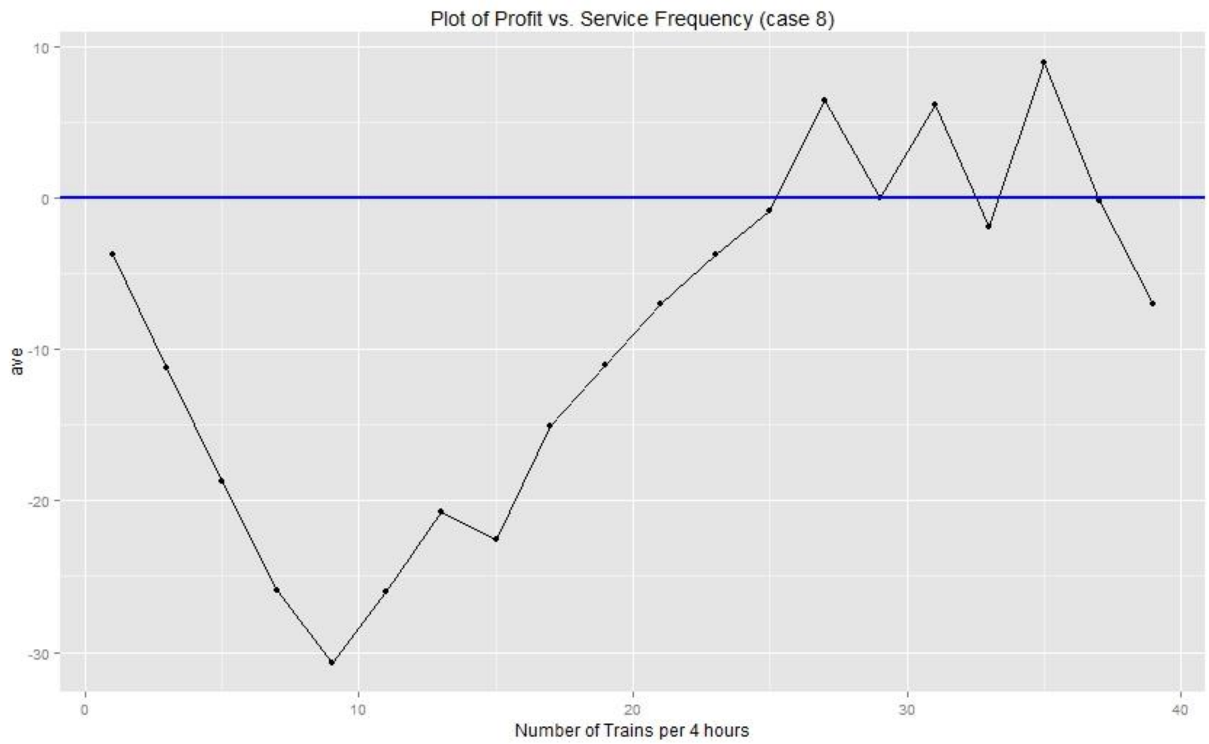


Figure 23. Plot of Profit vs. Service Frequency (case 8)

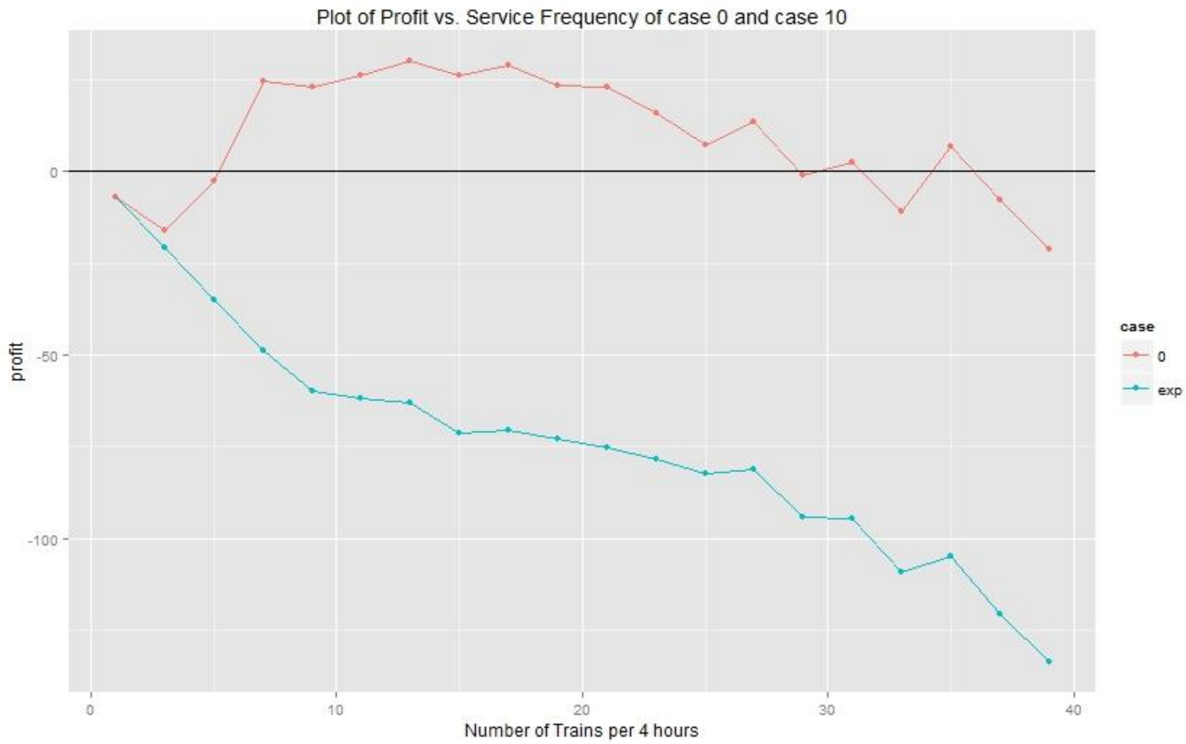


Figure 24. Comparison of Profit vs. Service Frequency of case 0 v s. case 10

Discussion

If the train arrival time is very unreliable and follows an exponential distribution, then the ridership will be severely reduced, especially at lower frequencies (fig. 23). This can also be shown by its lowest cost factor (3.5) among all cases. Its only difference with case 0 is the train arrival schedule. A comparison of the two cases is shown in figure 24 by assuming the same cost factor (7) for both cases. This is in accordance with Jarrett Walker's argument that reliability makes a huge influence on ridership.

On the opposite side, the time needed for driving can become less reliable as well. The next two cases simulated what would happen when driving time becomes less reliable.

Case 9

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=25, b=2
Standard deviation of congestion	**Total time – time if no congestion (20)**
Train arrival schedule	Fixed
Cost factor in graph	15

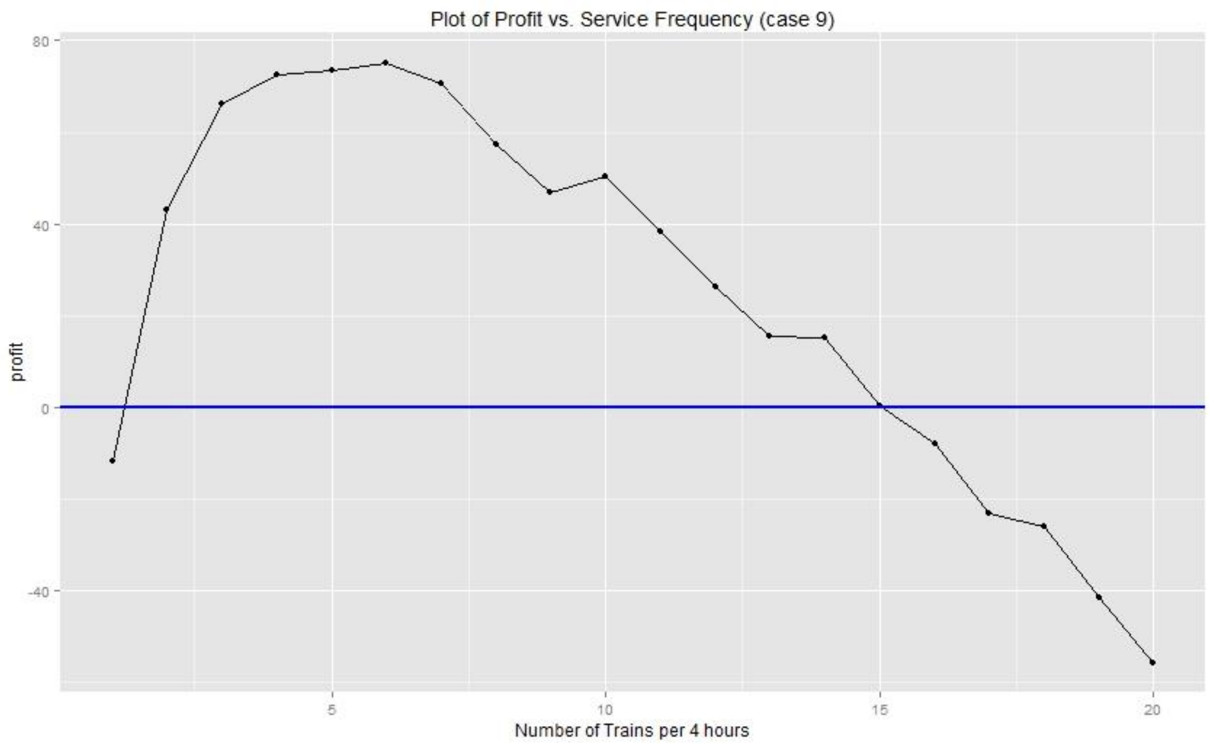


Figure 25. Plot of Profit vs. Service Frequency (case 9)

Case 10

Distribution of preference	2*Beta(20,20)
mean	1
Standard deviation	0.156
Train travel time (minutes, no waiting)	25
car travel time (minutes, no congestion)	20
Train limit (person)	30
Congestion function	a=100, b=4
Standard deviation of congestion	Total time – time if no congestion (20)
Train arrival schedule	fixed
Cost factor in graph	9

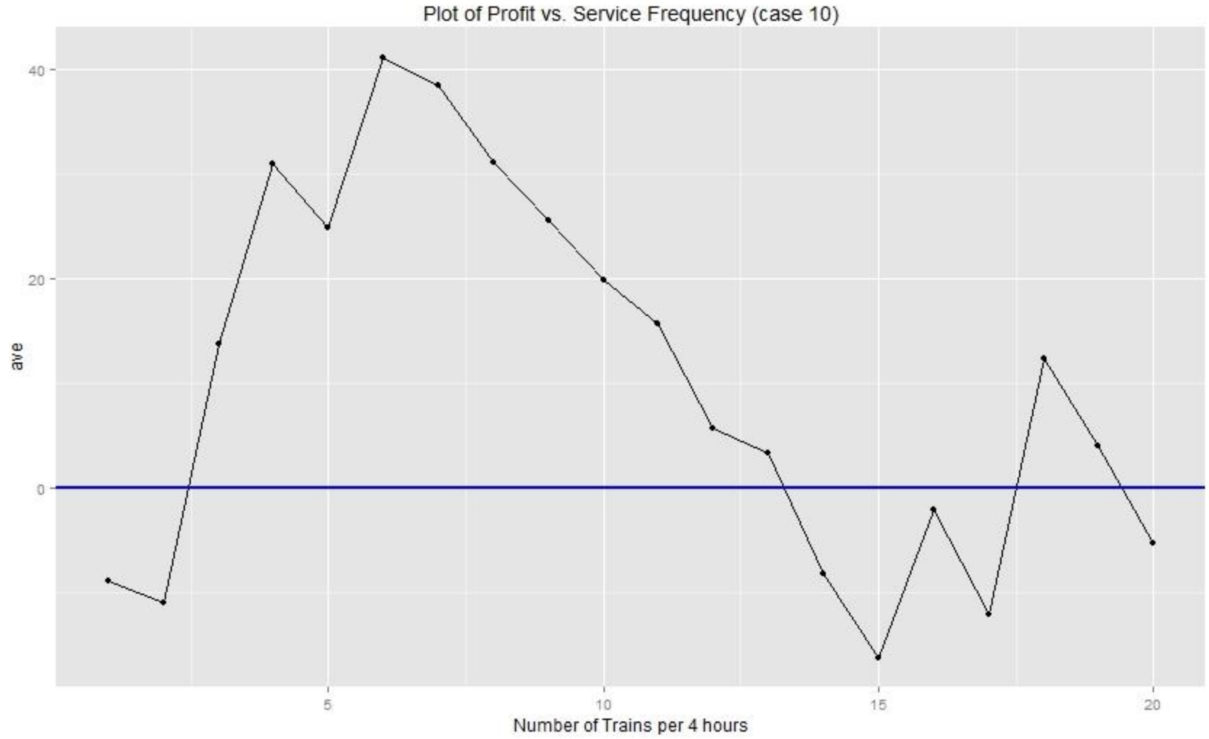


Figure 26. Plot of Profit vs. Service Frequency (case 10)

Discussion

In these cases, instead of assuming constant variance of the time each car would be stuck in congestion, I assumed that with heavier congestion, the congestion time would also become more variable. By comparing the cost factors of case 8 and 9 with those of case 0 and case 3, we do find that less reliable car time will result in more people choosing trains, but the overall shape of their profit curves do not seem to be affected as much as we would expect.

By comparing the cost factors of more variable congestion cases with those of their respective base cases, such as case 0 (cost factor=7) and case 9 (cost factor=15), or case 3 (cost factor=5.5) and case 10 (cost factor=9), we showed the substitution effect between the two transportation modes. Case 9 and case 10 have much bigger cost factors than case 0 and case 3, and their only difference is that the earlier two cases have less variable congestion time. We can think of it as an improvement of road conditions or an increase in road capacity, and that would reduce ridership of public transportation as well as profit level. This result reminds us the substitution effect between driving and taking public transportation. A city that wants to promote its public transportation should be careful with enlarging road capacity as that could make it harder to promote public transportation.

Hypothesis 6: there will be large random variations of long term average profit between different simulations (thus hypothetical worlds).

In previous cases, the ridership in all plots was actually the average from 100 simulation runs. But it would be interesting to see not only the average ridership, but also ridership of each simulation, because the reality corresponds to one single simulation rather than the average of them. Ridership values of simulations with the exact same set up are expected to have substantial variation. In this section, all data points were plotted for several selected previous cases to show whether there exist large variations.

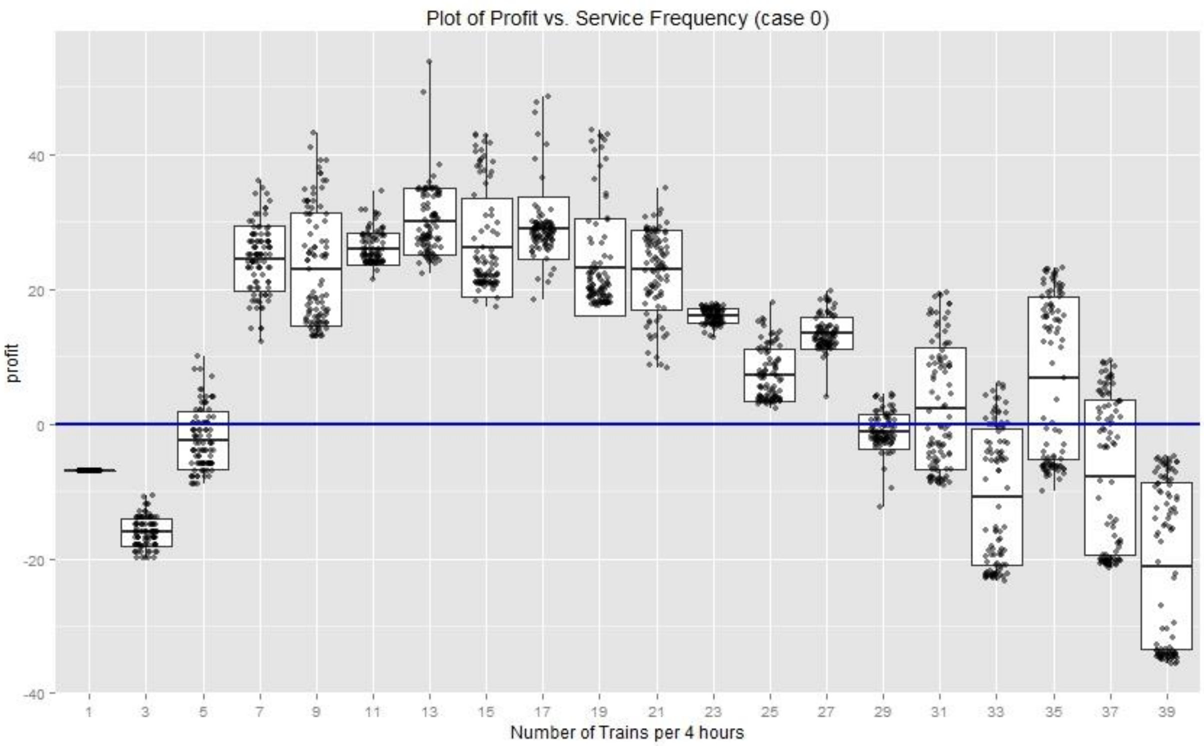


Figure 27. Box-plot of Profit vs. Service Frequency (case 0). Each dot represents the long run ridership at the corresponding frequency for a single simulation run (thus a single hypothetical world). The heights of the boxes represent two standard deviations.

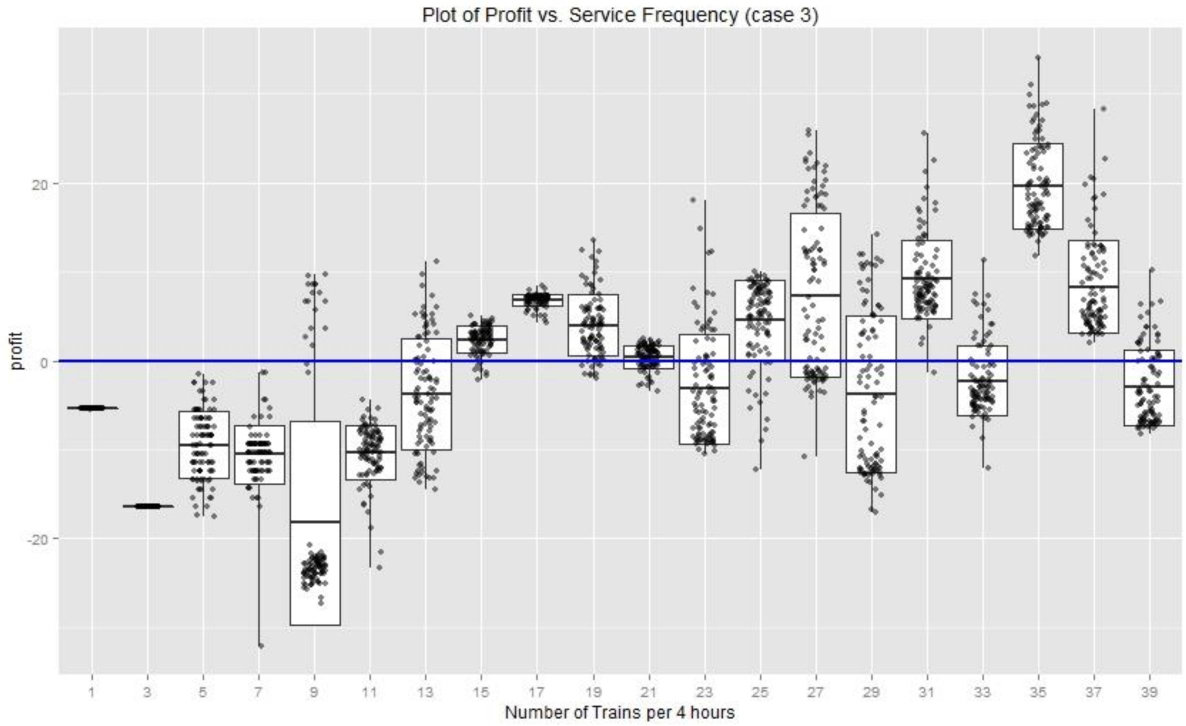


Figure 28. Box-plot of Profit vs. Service Frequency (case 3). Each dot represents the long run ridership at the corresponding frequency for a single simulation run (thus a single hypothetical world). The heights of the boxes represent two standard deviations.

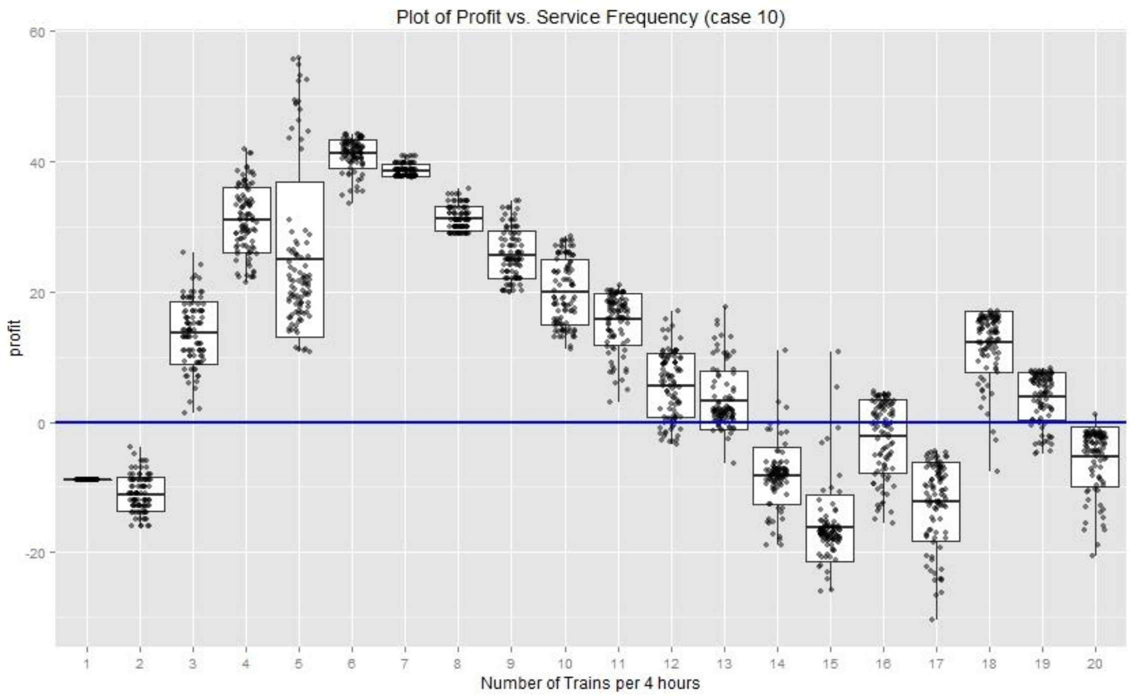


Figure 29. Box-plot of Profit vs. Service Frequency (case 10). Each dot represents the long run ridership at the corresponding frequency for a single simulation run (thus a single hypothetical world). The heights of the boxes represent two standard deviations.

Discussion

In the above graphs, each dot represents the long term (between 2,000 rounds and 20,000 rounds depending on how fast ridership stabilizes) average profit at the frequency corresponding to the x axis value in one simulation. Dozens of simulations were run at each frequency and plotted to show different possible outcomes that can arise even if everything stays the same. The variability is not day to day, but rather reflects difference in long term average profit between hypothetical worlds, because as shown in figure 16 - 18, ridership almost never changes after about the first 10 rounds. They show really big variability at some frequencies, but little variability at other frequencies. Also, the variability is big enough so that a profit could be made at frequencies that on average should result in a deficit, and vice versa. In the fortunate scenario, the public transit authority could make positive profits at almost every service frequency (fig. 28), thus would easily end up at a high service frequency and a high ridership if it truly cared to maximize ridership. On the other hand, the transit authority could also make negative profits at all frequencies between 19 and 33 (trains per four-hour period). In such an unfortunate scenario, it would be must harder to maximize ridership.

This result is caused by the randomness in whether people can get onto the first available train or have to wait for the next one. If the group of people who are really hesitant between driving and taking trains had terrible experiences waiting for trains, then they will choose driving and never switch back, because that unpleasant memory will be engraved in their memory and trains will never get a second chance. On the other hand, if they had great experiences using trains, they might just stick with it unless they have consistently bad experiences later on. Therefore, the long run ridership partly depends on which group of people established the habit of using the public transportation system. This really stressed the uncertainty in the real world situation, and the importance of convincing people to give public transit another try after improvements were made.

Hypothesis 7: lower car ownership will make multiple equilibria less likely to occur

In the following two cases, a portion of the population is assumed to have no car. They would have an infinite preference value, which means they could only choose to use the public transit however long it would take them. For practical purposes their preference values were set as 1000. Case 11 has one third of the total population with no car, and case 12 has one fifth. The rest of the population in both cases was assumed to have preference values drawn from the same distribution as the base case (fig. 4). All other assumptions were the same as those of the base case. The comparison of case 11 and 12 with the base case would reveal the effect of the demographical difference of the user population.

Case 11 & 12

Distribution of preference	2*Beta(20,20) for $\frac{2}{3}N$ people, 1000 for $\frac{1}{3}N$ people	2*Beta(20,20) for $\frac{4}{5}N$ people, 1000 for $\frac{1}{5}N$ people
mean	1	1
Standard deviation	0.156	0.156
Train travel time (minutes, no waiting)	25	25
car travel time (minutes, no congestion)	20	20
Train limit (person)	30	30
Congestion function	a=25, b=2	a=25, b=2
Standard deviation of congestion	5	5
Train arrival schedule	Fixed	Fixed
Cost factor in graph	8	8

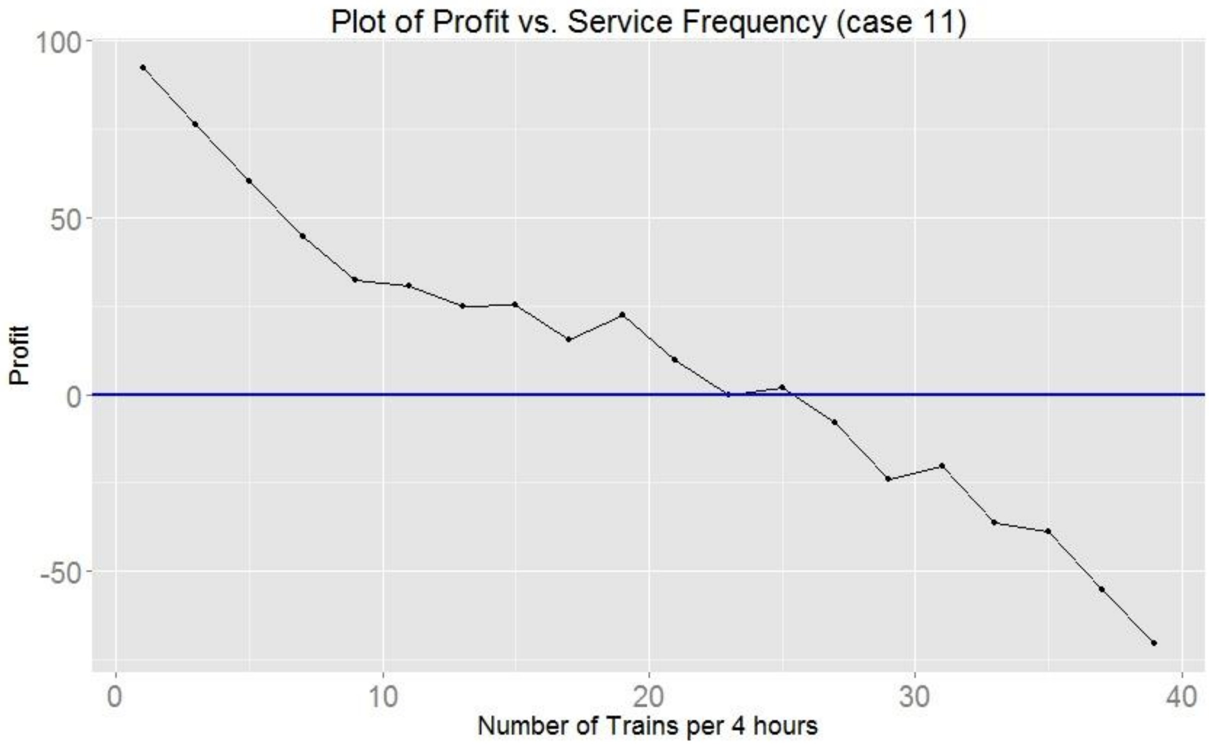


Figure 30. Plot of Profit vs. Service Frequency (case 11)

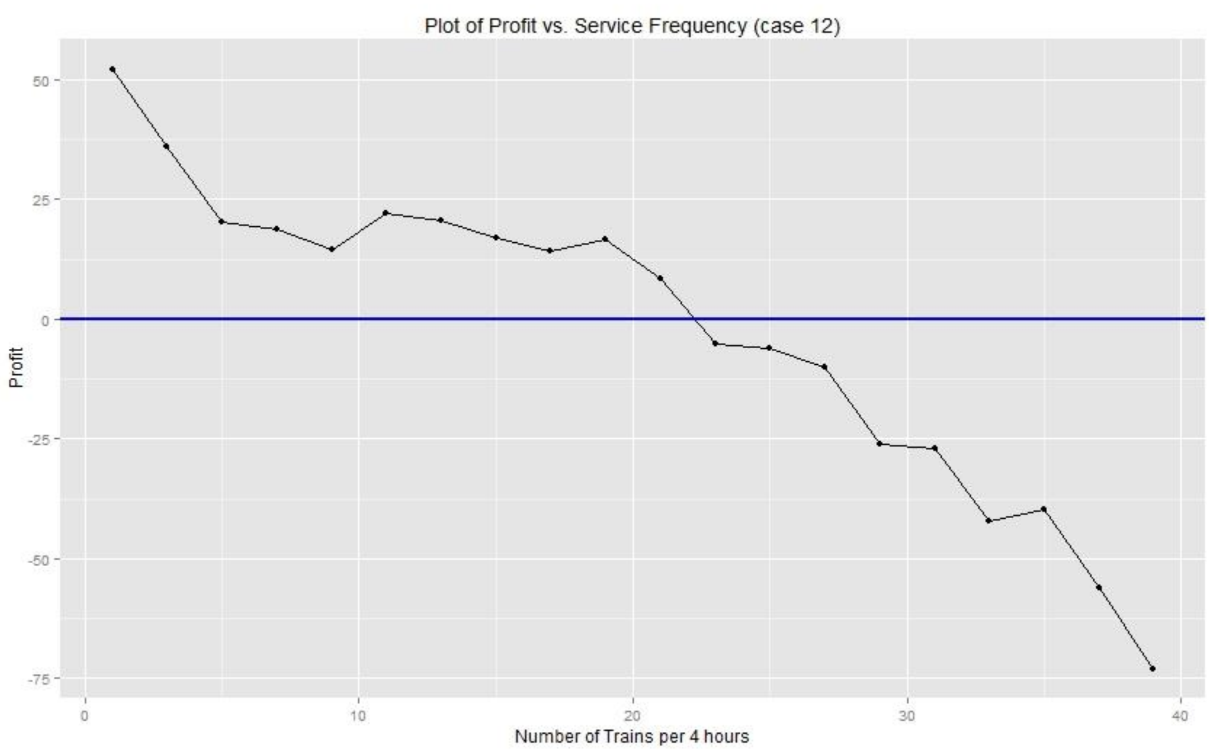


Figure 31. Plot of Profit vs. Service Frequency (case 12)

Discussion

The profit curve of case 11 is almost a simple downward sloping line. The implication of it is that the public transit agency would face a trickier trader-off between ridership and profit. In this case, there is almost always a considerable cost in increasing ridership, whereas in previous cases it is possible to increase ridership without requiring a bigger budget, because the increased revenue might be able to compensate for the cost. The profit curve of case 12 is a downward sloping line with a hint of multiple equilibria. It is thus in between case 11 and the base case.

These cases showed the difference between coverage lines and ridership lines. Coverage lines serve places where people badly need or feel entitled to public transportation[29]. They are often created for underprivileged communities to provide mobility and social inclusion[30]. Frequent lines, on the other hand, are created to maximize ridership so as to reduce congestion and emission caused by private cars. Due to their different purposes, ridership lines try to appeal to people who drive, while coverage lines serve areas where a portion of the population may not have cars. Therefore, models of ridership lines should have populations with higher car ownership rates than those in models of coverage lines. Cases 11 and 12, consequently, would be more suited to coverage lines, and previous cases would be more suited to ridership lines.

As case 11 and case 12 showed, coverage lines may face trickier trade-off between frequency and profit, thus tend to have lower frequency. Their ridership at low service frequencies would be higher than that of ridership lines at the same frequency. Also, since coverage lines are much less likely to have multiple equilibria, transit planners should not focus on frequency, but rather should focus on strategies such as better zoning or better design of transit networks.

[29] Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.
[30] Walker, Jarrett. "Purpose-driven public transport: creating a clear conversation about public transport goals." Journal of transport geography 16.6 (2008): 436-442.

Conclusion, Limitations, and Future Directions

Conclusion

This study used an agent-based model to model individual's choice between taking public transit and driving and the resulting public transit ridership. Agent based modeling allowed easier incorporation of assumptions such as adaptive expectations, simple decision heuristics, and a diverse population, as opposed to the classical assumptions of rational expectations, rational choice with utility maximization, and representative agents.

Simulation results showed the possibility of multiple local optimal points where the public transportation agency may locally maximize ridership under a budget constraint. The existence of multiple equilibria is moderately robust to changes in model parameters as multiple equilibria showed up in various cases.

The model then explored the effect of changes in various model components. Train speed was shown to be of little impact on ridership, especially at low service frequencies. Road capacity was shown to have a significant impact on the prominence and the form of multiple equilibria, though the specific impact could be complicated and hard to predict. The distribution of passenger preferences for public transit could determine the existence of multiple equilibria. Neither an entirely homogeneous population nor a really heterogeneous population could result in multiple equilibria. It is thus important to be able to model an arbitrarily diverse population since a subtle difference could be consequential. Reliability of train arrival schedule was shown to be of huge impact on ridership, especially at lower service frequencies. Multiple equilibria could still exist when train arrival schedule is unreliable. A reasonable amount of variation in car congestion time would not have as big an impact. A lower car ownership within the population would undermine the conditions for multiple equilibria as well as make the trade-off between profit and service frequency more difficult.

Simulation results also showed significant variation in long run average ridership of models with the same set up. It is thus possible to have a higher or lower ridership in each hypothetical world (or the real world) depending on which group of the population established the habit of using public transportation. This result could not have happened had we assumed rational expectations.

Limitations and Future Directions

- Public Transit Network

In this model, a simplistic case where there are only two train stations was studied. Real world public transportation networks are much more complicated and the real world population is much more spatially distributed. A more ambitious model could model all of that with the help of a Geographic Information System and a larger scale computer program. That extension would confer a lot more functionality to the model. Specifically, the more complete model would be able to study transfers within the transit network, optimal dispatch schedule of trains and buses, the effect of urban sprawl on people's mobility, to name a few.

- Model Calibration

In this study, almost all the parameters were not calibrated but rather experimented with to determine the reasonable range of values. This could be fine for an explorative study. But if a similar model were ever to be used to aid real world decisions, calibration of parameters would be not only necessary but difficult. However, this is the case for almost all models that aspire to inform real world decisions. Conversely, theoretical models used to explore certain phenomena often make assumptions based on intuition without calibrating them based on real world data. The real problem with this model is that it would be practically impossible to explore all possible combinations of different assumptions as there are a lot of assumptions with large ranges of possible values. The number of cases would be exponentially big.

The calibration of this model could be straightforward. Train travel time, frequency, and reliability can be monitored by passengers or the service agency. Car travel time, congestion function, and variability of congestion delay could be measured by mobile apps on the smartphones of a sample of drivers. The preference distribution of the population could be documented by surveys or observed by tracking real world decisions, but this could take more time than other calibrations.

- No Hypothesis Testing

Although agent based models enabled me to have more flexible assumptions, they do not provide closed form analytical solutions. Neither is it easy to do hypothesis testing under this frame work. Therefore, it is difficult to determine which factor contributed the most toward the emergence or disappearance of a certain phenomenon or to test hypotheses on that. On the other hand, the emergence of disappearance of a certain phenomenon could be the coordinated effect of many factors involved, thus the customary convention of trying to find a single cause of the phenomenon could be misguided. Nonetheless, the habit of looking for the main causes of a phenomenon still has its merit as it could help people better understand the phenomenon. There is a trade-off to be made between the flexibility of a model and the ease to interpret it. Agent based models provide more flexibility of modelling at the expense of interpretability.

Bibliography

Bonabeau, Eric. "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Sciences 99.suppl 3 (2002): 7280-7287.

Brown, Justine. "Can Technology Help Improve Mass Transit Use?" Government Technology, 14 May 2015. Web. 11 Sept. 2015.

Granovetter, Mark. "Threshold models of collective behavior." American journal of sociology (1978): 1420-1443.

Hommes, Cars H. "Heterogeneous agent models in economics and finance." Handbook of computational economics 2 (2006): 1109-1186.

Moore II, James E., and Thomas A. Rubin. "Better Transportation Alternatives for Los Angeles." Reason Foundation, 1 Sept. 1997. Web. 11 Sept. 2015.

Morris, Eric A. "Los Angeles Transportation Facts and Fiction: Transit." Freakonomics. N.p., 03 Mar. 2009. Web. 11 Sept. 2015.

O'Hara, Mary. "When Poor People Can’t Get on Due to Lack of Public Transport." The Guardian, 26 May 2015. Web. 11 Sept. 2015.

Ramey, Corinne. "How America’s Transportation System Discriminates Against the Most Vulnerable." America’s Unfair Rules of the Road. Slate Magazine, 27 Feb. 2015. Web. 11 Sept. 2015.

Saberi Kalaee, Meead. Investigating freeway speed-flow relationships for traffic assignment applications. Portland State University, 2010.

US. Federal Highway Administration. Office of Environment, Planning, and Realty. Travel Model Validation and Reasonableness Checking Manual. Federal Highway Administration, 24 Sept. 2010. Web. 1 Oct. 2015.

Walker, Chris. "Why Don't White People in L.A. Take the Bus?" L.A. Weekly. L.A. Weekly, 23 June 2014. Web. 11 Sept. 2015.

Walker, Chris. "Why Don't White People in L.A. Take the Bus?" L.A. Weekly, 23 June 2014. Web. 11 Sept. 2015.

Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.

Walker, Jarrett. "Explainer: The Transit Ridership Recipe." Human Transit, 15 July 2015. Web. 11 Sept. 2015.

Walker, Jarrett. "Houston: Transit, Reimagined." Human Transit, 09 May 2014. Web. 11 Sept. 2015.

Walker, Jarrett. "Los Angeles: The next Great Transit Metropolis?" Human Transit, 30 Mar. 2010. Web. 11 Sept. 2015.

Walker, Jarrett. "Now, Anyone Can Monitor Reliability." Human Transit, 30 June 2010. Web. 11 Sept. 2015.

Walker, Jarrett. "Purpose-driven public transport: creating a clear conversation about public transport goals." Journal of transport geography 16.6 (2008): 436-442.

Walker, Jarrett. "What Motivates Mode Choices for Urban Residents?" Human Transit, 2 Oct. 2014. Web. 11 Sept. 2015.

Resource Systems Group, Inc. "Who's on Board: Mobility Attitudes Survey." (2014): 1-27. The Transit Center, 2014. Web. 1 Sept. 2015.