

Using differential pricing to mitigate the impact of congestion on metro: a big data study base on Shanghai metro

Haisen Kang

Shanghai Pinghe High School Shanghai, China

Abstract—Metro, which is one of the most popular way of public transportation, has shown inability to withstand the high intensity of congestion presented in the first two decades of the new millennia. Despite the effort made by government which includes adding multiple metro line, the situation is still grievous. Unlike other form of transportation such as train or airplane, the ridership of metro wasn't staggered and can be manipulated. In this paper, we employ differential pricing to alleviate traffic pressure on metro during the peak hours. As people have different elasticity of demand for metro transportation in different time interval, we can reduce the number of passengers who relatively treat metro transportation as unnecessary in specific time interval and place by setting up different price. Base on the data of shanghai metro, we show different aspects in our model in which we use system clustering, optimization model of social welfare and calculate an acceptable range of price using Ramsey pricing model. We validate our solution, using the agent simulation model, to be considerably capable at easing the traffic pressure of metro, whether comparably or statistically.

1 INTRODUCTION

A. Background

Metro is one of the most significant forms of public transportation for the general public. Since its invention, more and more countries have adopted this system to ameliorate their public transportation. This topic is not a regional consideration but more of global concern. Each year, billions of people take the metro for work or travel within the range of the city. In 2019, for shanghai alone, the maximum value of passengers reached 12 million people-times per day which equals one-third of the city's population is bearing on underground transportation. As more and more additional lines were built, the amount of people who use the underground is also proliferating. It has readily turned into a common sense that accompanied by the growing trend of demand is the increment of congestion in the world.

Currently, there is almost a universal agreement on the severity of china's metro congestion. Albeit the government effort of adding multiple new lines in the past ten years, the result did not make the public to be sanguine at all. Astonishingly, according to the data provided by the China Association of Metro, Guangzhou metro's line three experience ridership as much as 600000 per hour during its rush hours. Moreover, these 600000 people are not dispersed evenly along the entire line 3, but most of them are concentrated in some significant stations. These values had long exceeded the maximum volume that some station can withstand. This conundrum could cause significant

impacts such as danger during rush hours and uncomfortable experiences. However, most nefariously, it could induce substantial economic losses as passengers might miss several trains due to immense ridership at certain time intervals. All the time wasted by passengers could mean the increase of time cost, which if quantified, could mean billions of economic losses for the entire society.

To tackle these problems, the government had provided a variety of solutions in the past two decades which included adding more line, shortening time interval for each train's arrival and adding more stations in each line. However, as beholding, no trend of decrease in congestion intensity had evaded. These exorbitant solutions failed to match the skyrocket of ridership which caused a plethoric amount of fee wasted. It not only failed to solve the puzzle but even added insult to injury. The society urgently advocates for a new solution which shall be inexpensive, feasible, and viable.

B. Pricing Policy around the globe

The first concept about differential pricing is brought out by Vickrey [1], who is a representative economist on cost pricing in the US. Vickrey (1955) realized that pricing method that only depends on Marginal cost is not precise enough. He then started to expand the definition of cost to a broader range which includes fixed cost and other types of cost. However, this was still not a refined calculation as the difference in passenger volume during peak hours and ordinary time is neglected. [2] Glaister & Lewis (1978) this deficiency and started to construct a pricing model base on social behavioral patterns which include peak hour, sound

binbinsen@163.com

pollution and congestion. This Glaister-Lewis model can be counted as the foundation of differential pricing on reducing congestion and related social patterns.

In china, however, development of pricing models started relatively recently. These study on pricing policy generally treated the maximization of corporation’s profit as their major objective, considerations of social welfare were generally neglected.

C. Purpose of the study

As discussed in 1.1, current solutions to reduce congestion is generally over costly and inefficient. Thus, the primary purpose of this study is to determine an uncostly mechanism for reducing congestion. However, reduction of congestion does not require much sophistication as if the only purpose is to reduce congestion, raising the price to an unaffordable value achieve this goal, but it is certainly unacceptable. Thus, the purpose of this study is to determine a method to reduce congestion while ensuring the maximization of social welfare. This assigns the policy of percentage price addition with adequacy and reasonability.

D. General modeling

In this study, we adopted the use of agent simulation model, which is a model to simulate passenger behavior after price addition. Then, we aimed to calculate the total social welfare change by subtracting the time cost people reduce as congestion decrease and deadweight loss due to a price increase. Next, we iterate this process for percentage price addition from 1% to 100% and determine the optimal value of social welfare increase to be our final mechanism. Furthermore, we testify the adequacy of this result dually by employing the use of Ramsey pricing model to calculate an acceptable range for percentage price addition. If this value lies within the range of acceptable percentage price addition, we then conclude our model to be reliable and robust.

E. Significance

We might experience congestion on a daily casual so that we do not treat it as a serious matter. However, if we take a view collectively at the social value. It could mean billions of working hours wasted, productivity largely destroyed, immense loss in the economy. Moreover, congestion can also reduce people’s productivity not by reducing their work time but devastating their mood throughout the entire day. This can be validated if we imagine how are we going to feel in a day if it takes us 3 hours to get to work each day. Furthermore, our study possesses unique significance as we aim to determine the value of percentage price addition that maximizes the social welfare increase while reducing congestion extensively. This means that while reducing congestion, society could be still better off economically.

2 GENERAL DATA TREATMENT AND PROCESSING

We separate the time of a day into 5 specific time intervals: Morning non-rush hour (MNRH): before 7:00, Morning rush hour (MRH): 7:00-9:00, Noon non-rush hour (NNRH):

9:00-17:00, Evening rush hour (ERH): 17:00-19:30 and Evening non-rush hour (ENRH): after 19:30.

Within these specific intervals, we try to discern between the passenger volume who enter the station and leaves the station by giving a E after the name of time intervals for passengers who enter the station. For example: MRHE (morning rush hour enter). Conversely, we add the letter L which stands for leave after the passenger volume who exit the station such as MRHL (morning rush hour leave).

The data set used in this working paper is adopted from the dataset of 2018 SODA [3] competition. The original data set only shows five variables: passenger ID, enter time, enter station, exit time and exit station. Using pandas, we can add up the total passenger volume of each station in every time interval respectively. Shown in table 1:

TABLE I. PASSENGER VOLUME OF THE FIRST FIVE SITES IN EACH TIME INTERVAL

Code	MNRHE	MRHE	NNRHE	ERHE	ENRHE
1	235	1157	10304	11995	6550
2	280	1567	5445	3805	1459
3	705	4859	6892	3535	1526
4	1300	8503	4765	1031	591
5	729	3532	40753	39497	35696
	MRHL	MNRHL	NNRHL	ERHL	ENRHL
	517	8836	17031	4279	1597
	202	3366	5962	1863	1239
	293	2880	5755	4395	3322
	62	761	3034	5170	4675
	2880	29593	60260	23859	8102

A. Standardization

We then standardize the passenger volume to a percentage and make sure the entering and leaving percentage in each time intervals add up to 100%.

TABLE II. PERCENTAGE PASSENGER VOLUME OF THE FIRST FIVE SITES IN EACH TIME INTERVAL

Code	MNRHE	MRHE	NNRHE	ERHE	ENRHE
1	0.777	3.826	34.07	39.665	21.659
2	2.230	12.480	43.366	30.304	11.620
3	4.025	27.739	39.345	20.180	8.712
4	8.030	52.520	29.432	6.368	3.650
5	0.606	2.938	33.902	32.857	29.695
	MRHL	MNRHL	NNRHL	ERHL	ENRHL
	1.603	27.390	52.793	13.264	4.950
	1.599	26.647	47.198	14.748	9.808
	1.760	17.302	34.575	26.404	19.958
	0.452	5.554	22.143	37.732	34.119
	2.310	23.733	48.326	19.134	6.498

This process is necessary because k-means primarily cluster data according to its distance. If not standardize, cluster would only be determined by the number of passenger volume of each station, where big station would be in the same cluster and contrarily small stations in one cluster. this mean that we will fail to observed the pattern

of passenger volume discern, which this paper would discuss later.

3 MODELING

A. System clustering

While every station has its pattern of passenger volume, providing all stations with a universal pricing policy would be unduly onerous. For instance, stations that experienced minor passenger amount during the morning rush hours should not have the same differential pricing policy as those that experience prodigal ridership. However, it is also fallacious and overcomplicating to provide different pricing policy for every station, respectively. Instead, clustering all the stations into different groups would be a more eclectic remedy that satisfies the different ridership pattern exhibited in each station respectively and remains in a reachable range. In this process, we adopt the k-means model, which is one of the unsupervised models dedicated to clustering tasks.

K-means basically separates the data sample into k clusters by the distance between each individual sample. Suppose we cluster the sample into (C_1, C_2, \dots, C_k) , and we define E as the square error of the sample, then the major task for k-means is to minimize the E of the total sample, illustrated in equation 1:

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 \quad (1)$$

Where μ_i is the mean value vector of C_i , in other word, the centroid of each cluster, expressed as:

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \quad (2)$$

As it is over complicated to calculate the above equation with normal mathematical method, we adopt the usage of computer iteration. In k-means, we separate the sample set $D = \{x_1, x_2, \dots, x_m\}$ to k clusters $C = \{C_1, C_2, \dots, C_k\}$, where k here is 3. This leads to an important aspect of k-means which the value of μ_i is defined by our prior knowledge on the sample that we want to cluster. Based on the discrepancy between the ridership that enter and exit, we hypothesize that the clustering result would resemble the function of the region located by the stations such as business zone, transition zone and residential allotment.

However, we need to clarify that this x is a vector of 8 dimension which includes: MNRHE, MRHE, NNRHE, ERHE, MNRHL, MRHL, NNRHL, ERHEL. ENRHE and ENRHL were not included because the enter station value and exit station value is turn into a percentage which is added up to 100. Determining 4 value out of 5 can automatically determine the fifth as their total value is fixed.

We first select k centroids which are $\{\mu_1, \mu_2, \mu_3\}$, than we calculate the distance between x_i and each centroids $\{\mu_1, \mu_2, \mu_3\}$ expressed in equation 3:

$$d_{ij} = \|x_i - \mu_j\|_2 \quad (3)$$

For each x_i , we chose the smallest d_{ij} and cluster it into its according type λ_i and then update $C_{\lambda_i} = C_{\lambda_i} \cup \{x_i\}$. After this being done, for every $j = 1, 2, 3$ we recalculate the new centroid which is show as the following:

$$\mu_j = \frac{1}{|C_j|} \sum_{x \in C_j} x \quad (4)$$

If any of the k centroid vectors changes, we iterate the above process. If no changes are observed, we output cluster $C = \{C_1, C_2, C_3\}$.

1) Clustering Results

As the discrepancy between the passenger volume who enters and those who exit a station is meniscal and do not show significant patterns during the non-rush-hours time interval, it is essential to describe these clustering result by its ridership patterns exhibited during the rush-hour time intervals.

a) Cluster 1:

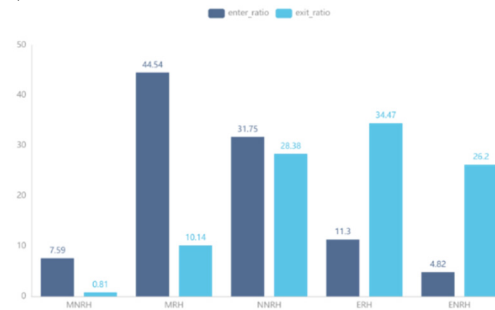


Figure 1. Passenger volume in each time interval of cluster one

Figure 1 indicates an important aspect of this cluster which we can see that the passenger volume that enters the station during the morning is much higher than the passenger volume that exit. However, we can behold a different circumstance during the night rush hour, which was illustrated transparently in the figure that passenger who exits far exceeds those who enter.



Figure 2. Passenger volume per minute of cluster one

Following figure 2, which shows the passenger volume that enters the station, we can see that the people who enter the station during the morning surpass the people who enter at night's rush hours. According to these traits, we can easily associate this cluster to a real scenario where people tend to enter the station in the morning and leave it at night. While people need to go to work during the morning, they usually enter the stations near their residents. Conversely, they leave the station near their residents when they come back whether from work or school. As the traits of the

station presented in cluster one highly over-lapse the characteristics of those near people’s residents, it is reasonable to associate these two as the same type of station which in other words: cluster one resembles those stations that presented in the residential allotment. Evidence in figure 2, also imposed a validation on this inference. For instance, YUAN SHENG stadium station is one station in the cluster, and unsurprisingly located in the residential allotment.

b) Cluster 2:

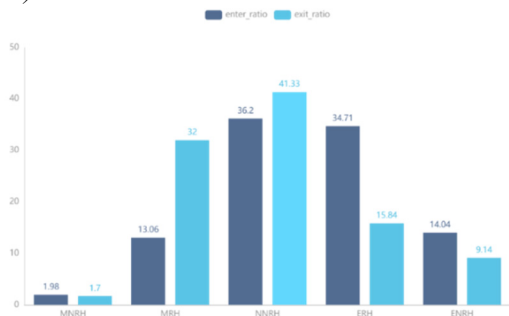


Figure 3. Passenger volume in each time interval of cluster two

This cluster turns out to show a completely different passenger volume pattern. According to figure 3, the passenger who enters during the morning is more than those who leave during the morning. And a reverse entirely contrastive relationship was observed at night. People frequently leave the stations adjacent to their working location and enter them at night when work is over. These stations exhibit the same traits as cluster 2 had shown, thus, reasonable to associate them.

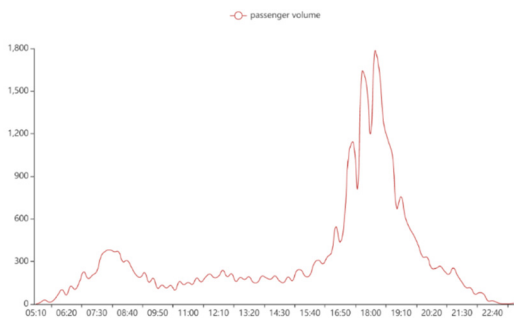


Figure 4. Passenger volume per minute of cluster two

In figure 4, we can also see that passenger volume who enter the station during the morning is relatively inferior to the passenger amount who enter at night. This piece of evidence furtherly supports the previous association. Thus, we can generally conclude that cluster two refers to the stations that locate at business zones where people majorly work.

c) Cluster 2:

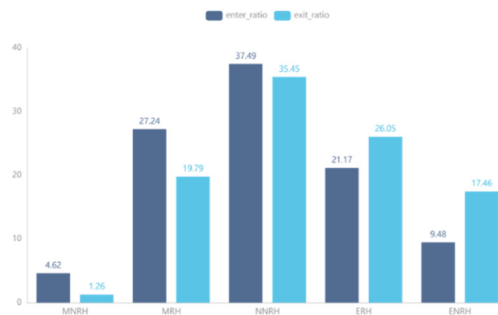


Figure 5. Passenger volume in each time interval of cluster three

Cluster 3, unlike the previous two, is primarily consisted of stations that show no significant difference between ridership that enters and leave. According to figure 5, enter ratio and exit ratio are only differed by 7.45 percent during MRH, 4.88 percent during ERH and 2.04 percent during NNRH. Comparing to the 30 percent difference that of cluster 1 during MRH, the difference exhibited in cluster 3 is rather diminutive.



Figure 6. Passenger volume per minute of cluster three

A more commanding view is observed in figure 6, which the passenger volume that enters the station during the morning show no significant difference from the passenger volume that enters the station at night. Specifically, the peak of passenger volume that enters the station during morning rush hour is around 650 whereas that of the night is 600. There is only a passenger volume difference of 50, which is relatively insignificant comparing to the substantial basis. These patterns typically present in the stations that are in transition zones between business allotment and residential allotment. To illustrate, imagine the passenger volume of a shopping mall, where the number of people who enter the mall shown no significant difference between the amount that leaves during a specific time interval. Thus, cluster 3 is consisted majorly by those station located in the transition zone that shows an equilibrium pattern between ridership entering the station and that leaves.

The above conclusion of the characteristic of these three clusters is even more explicit in figure 7 where the patterns of stations from 3 different cluster were combined in a single diagram.

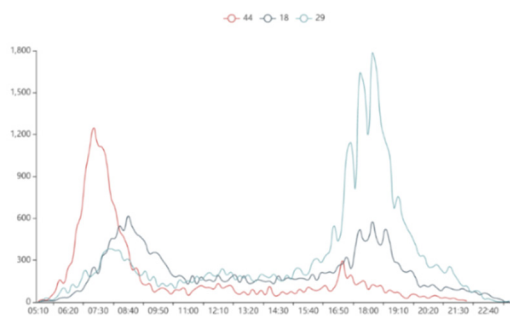


Figure 7. Passenger volume per minute of site 44, 18 and 29

Moreover, the associations of the clusters to their real location is also further validated in figure 7. We can substantiate cluster one as those stations located outside of the city center and in the residential area. Meantime, the stations in cluster 3 are also found in downtown where these places are majorly business area and commercial allotment. Stations in group 2 are also proven to present within the transition zone between the business area and residential distribution.

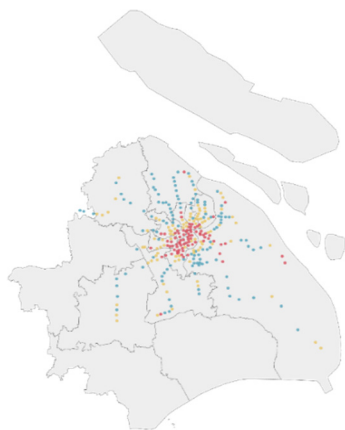


Figure 8. Distribution of different clusters of sites

2) Application of clustering result

The application of clustering is that we only offer price addition on those stations which belongs to cluster 2 and 3 as they possess a significant difference in passenger volume between the rush hours and non-rush hours. These stations during rush hour would exhibit a higher level of congestion. Even if stations in cluster 2, who do not show a significant difference in passenger volume between rush hours and non-rush hours, have a high total passenger volume, the intensity of congestion is not as severe as the other cluster because the passenger volume was distributed evenly in each time interval. The specific plan for price addition will be discussed in the following passage.

B. Agent simulation model

People typically want to measure the impact of a policy on society and the market before it is carried out. It would be meaningless to measure the impact of the policy after these impacts are already delivered. Thus, the construction of a simulation model is customarily needed to test an imaginary policy in an imaginary world. Specific to our circumstances, where every passenger is independent

individuals who make decisions based on their own interest and only on their interest, we adopted the use of agent-based simulation model. The use of agent-based simulation model is adequate as the researcher explicitly describes the decision processes of simulated actors at the micro-level. "Structures emerge at the macro level due to the actions of the agents, and their interactions with other agents and the environment." said professor from Newcastle University [4].

1) Assumptions

- We only place treatment on the congestion during morning rush hour and night rush hour. Although there might be congestion during time intervals apart from these two, those are only exceptions that do not deserve out-rated significance. Even from personal experience, we can observe that congestion during rush hours are those that deserve our attention due to both its intensity and universality.
- Waiting time for each passenger is the smaller one between 500 seconds and one-third of the interval between they enter one station and exit another station. As some of the data shown anomaly where the time between the passenger enters a station and leave another one is only 6 minutes at total. However, we recognize this as impossible due to the minimal time between two stations is around 3 minutes. Thus, we adopt five minutes as their waiting time in order not to let these outliers affect the final result of our simulation model.
- We separate the passenger into two major types which are the working population and non-working population. This separation is self-evident as a person is neither at work or not at work. This separation will aid in our further procession as different types of people have a different sense of neediness against metro at a particular time interval which in short, different elasticities of demand toward metro transportation. However, it is impossible to distinguish every single elasticity of different types of people, respectively. Thus, it is reasonable to separate the whole population into working, and non-working as these two types of people exhibit a relatively clear difference in their elasticities. In order to discern passengers' identity through primary data, working populations shall satisfy the following requirement: 1. Must take metro twice; 2. He or she shall enter an arbitrarily station before nine in the morning and leave the same station after four at noon.
- Working populations and non-working ones have different elasticity of demand toward the metro. We generally assume that the elasticity of the working population is 0.1 and those of the non-working ones are 0.5. This is according to a research done by professor Mayworm [5] in 1981, published in Journal for Transport Economics & Policy, which claim that the elasticity of travel demand in the city with population over 1 million is about -0.24. In this working paper, as the ratio of the working population to non-working ones is about 7:3, the weighted average for the elasticity

of metro demand in shanghai is about -0.22. Thus, assigning working population with an elasticity of 0.1 and non-working ones is adequate because the population of shanghai long exceed 1 million.

- The probability for a passenger to change its station equals to the product of his or her elasticity of demand toward metro and the percentage increase of price.
- The probability for a passenger to leave the metro transportation system is simply 0.1.
- The probability for the passengers, who are designated to change a station, is the reciprocal of the distance between his or her original station and 5 nearest station that have no price additions. People generally would choose the second-best alternative if the first one is unavailable, but the ignorance of information might cause irrational behavior.
- we assume that the time interval of a passenger from entering a station and leaving another one is reasonable if it is between 300 seconds and 10000. This is because it is impossible for a person to enter a station and leave another one within 5 minutes or greater than 10000 seconds. Any station not within this range will not be considered valuable and will be treated as outlier.

2) *Generating new agents*

The above assumptions can be elaborated in to algorithm by using basic principles of probability. The new agents generated in this working paper follows the following probability distribution:

- We generate the velocity of the passenger transferring to another station with uniform distribution. Illustrate as follow:

$$V_k \sim U[2, 4] \tag{5}$$

- We generate the time interval of new agents between entering one station and leaving another is with normal distribution. The mean and variance are based on history data. The expression is shown below:

$$\begin{cases} t_{ij} \sim N(\mu_{ij}, \sigma_{ij}^2) \\ \mu_{ij} = \frac{1}{n} \sum_{k=1}^n t_{ijk} \\ \sigma_{ij}^2 = \frac{1}{n-1} \sum_{k=1}^n (t_{ijk} - \mu_{ij})^2 \end{cases} \tag{6}$$

C. *Optimization model*

1) *Synopsis*

The purpose of this study is to achieve the diminution of congestion within metropolitan cities. While differential pricing is adopted base on the assumption that different population has different elasticity of demand toward metro, no people will take the metro if we dramatize the price. Differential pricing is most reasonable only if reducing congestion if the price is in an acceptable range. Thus, it is

required to find this acceptable range and optimal value for our pricing model. In this study, we measure society's total utility change by using the time cost that people reduce to minus the deadweight loss of the market by elevating the price to another level.

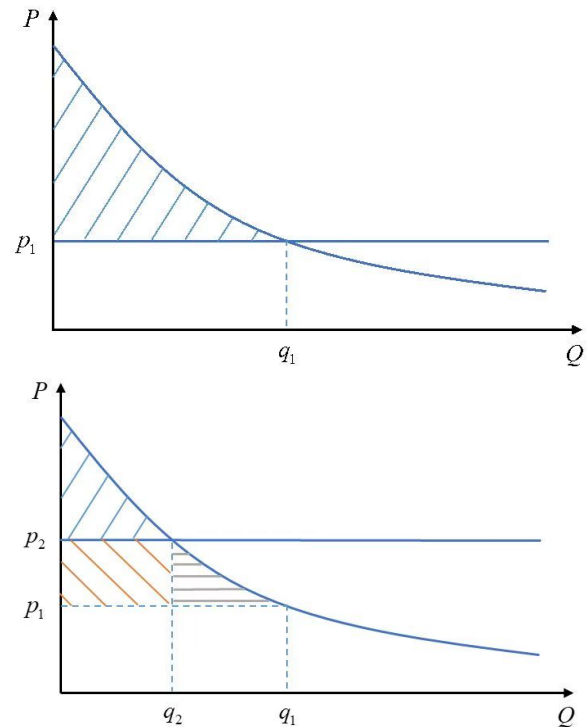


Figure 9. Deadweight loss

2) *Model Construction*

a) *Calculating deadweight loss*

The deadweight lost here simply account for the people who left the underground transportation due to a higher price. As shown in figure 9(b), the deadweight loss which equals to the grey shaded area with horizontal line could be express in the following equation:

$$DW = -\frac{1}{2} \Delta p \Delta q \tag{7}$$

We modify the equation by multiplying $\frac{pq}{pq}$

$$DW = -\frac{1}{2} pq \frac{\Delta p}{p} \frac{\Delta q}{q} \tag{8}$$

Where, ΔP is the change of price and Δq is the change of quantity demanded due to the offering of an additional price. We also notice that price elasticity of demand, PES, can be shown in the following equation.

$$\varepsilon = -\frac{\frac{\Delta q}{q}}{\frac{\Delta p}{p}} \tag{9}$$

After organizing equation 9, we get:

$$\frac{\Delta q}{q} = -\frac{\Delta p}{p} \varepsilon \quad (10)$$

Thus, we can substitute certain element in equation 8 with the result from equation 10 we get:

$$DW = -\frac{1}{2} \Delta p \Delta q = \frac{1}{2} p q \varepsilon \left(\frac{\Delta p}{p} \right)^2 \quad (11)$$

However, we notice that the deadweight lost mentioned here is the total loss of passengers who quit taking metro. These passengers incorporate different types of people which have different PED (price elasticity of demand) toward underground transportation. Thus, the final expression of the deadweight loss shall be the sum of all deadweight loss of different types of people. As we assume before that there are two types of population, working and non-working, we have the following expression of deadweight loss:

$$DW = \frac{1}{2} \sum_{i=1}^2 (p q)_i \varepsilon_i \left(\frac{\Delta p}{p} \right)^2 \quad (12)$$

The PED here are not the same elasticity for different types of passenger in the agent simulation model. Instead, it was the price elasticity of demand of those passengers who choose to quit underground transportation after a price addition. Thus, PES here for working population is 0.01 and 0.05 for non-working population.

b) Quantifying passengers' time cost

The major concept consistently throughout this working paper is that people's time can be converted into monetary value. To do so, a specific plan to quantify time is required. We first establish a threshold to distinguish the stations that require additional pricing. And then we calculate the time that passengers gain due to a decrease in waiting time base on their waiting time before the establishment of our policy and the decrease of passenger volume in each station respectively.

Major assumptions:

- Passengers' waiting time equals to the fraction of distance and velocity they travel. we only offer an additional price on those stations that have a value of passenger volume per exit or enter greater than 3600. This is because the result of data analysis show that these stations consist about 10 percent of the total data. It is also reasonable to choose only the top 10 percent of most congested station for an additional pricing because this would minimize the lost of public transportation corporation and maximize the effect of reducing congestion. Conversely, if we offer an additional price universally, 90 percent of stations which are not seriously congested will also receive a price addition. However, this price addition will not operate effusively as there are no congestion originally.
- It is ubiquitously recognized that the time people wasted due to congestion correlates inextricably with the intensity of the congestion. We generally

assume that the number of people per exit or entrance is an evident manifestation of the congestion intensity. In other word, the more the people there are per exit, the more time people waste. According to the law of diminishing utility, we know that the time wasted by the passenger do not increase after the congestion reach a certain level.

- The new waiting time after the pricing policy shall be based on the old waiting time. Meantime, it shall also be based on the fraction of the old passenger volume and the new one.
- According to the law of diminishing marginal utility, the change of waiting time for each passenger shall diminish as the price change increase.

Specific Model establishment

According to assumption 1, we can have the following equation:

$$t = \frac{s}{v} \quad (13)$$

Where t is the waiting time of passengers, s , the distance of passengers and v , the velocity of passengers. Base on assumption 3, we have equation 14:

$$v = v_0 \left(\frac{3600}{\max\{3600, num\}} \right)^2 \quad (14)$$

Where v_0 is the velocity of the passengers when there is no congestion presented in the stations. num here simply refers to the passenger volume per entrance in each station respectively. The above equation means we choose the bigger value between number and 3600. Moreover, this equation also shows that we only provide an additional pricing on those station whose passenger value surpass 3600. Thus, under this premises and base on assumption 1 and 3, we have the following equation

$$t_{old} = \frac{s}{v_0 \left(\frac{3600}{old} \right)^2} = 500 \quad (15)$$

In equation 15, old refers to the passenger volume per entrance before our pricing policy. t_{old} is the waiting time of each passenger before the ordination of our pricing policy. We generally assume that the waiting time of passengers in these stations with congestion are universally 500. Similarly, we have the waiting time for the passengers after our policy being implemented shown in the following equation:

$$t_{new} = \frac{s}{v_0 \left(\frac{3600}{new} \right)^2} \quad (16)$$

Where t_{new} is the waiting time of passengers after implementation of the pricing policy. new is the number of passenger volume per entrance after the policy. After a simple substitution, we have:

$$t_{new} = 500 \left(\frac{new}{old} \right)^2 \quad (17)$$

The change in waiting time can be calculated by subtracting the old waiting time with the new one:

$$\Delta t = t_{old} - t_{new} \quad (18)$$

By substituting the values in equation 15 and equation 17, we have equation 19:

$$\Delta t = 500 \left[1 - \left(\frac{new}{old} \right)^2 \right] \quad (19)$$

According to the equation of PED, we recognize that:

$$\Delta q_0 = -\varepsilon_0 q_0 \frac{\Delta p}{p} \quad (20)$$

Where Δq_0 is the change of quantity demanded of underground transportation system for the working populations, $-\varepsilon_0$ is the price elasticity of demand of the working populations and Δp is the addition of price before and after the implementation of the pricing policy. Similarly, we can resolve another equation for non-working populations:

$$\Delta q_1 = -\varepsilon_1 q_1 \frac{\Delta p}{p} \quad (21)$$

Δq_1 is the change of quantity demanded toward metro for non-working population, ε_1 is the PED of non-working population, q_1 is the quantity demanded toward metro of non-working populations before operation of our policy:

$$\begin{cases} q_0 = old \times ratio \\ q_1 = old - q_0 \end{cases} \quad (22)$$

Where, *ratio* is the ratio of working population in the total populations who take metro and the total change of quantity demanded before and after the pricing policy is the difference between the old value and the new one.

$$new - old = \Delta q_0 + \Delta q_1 = -(\varepsilon_0 q_0 + \varepsilon_1 q_1) \frac{\Delta p}{p} \quad (23)$$

We also can resolve the fraction of old value and new value by using equation 23:

$$\frac{new}{old} = 1 - \frac{\varepsilon_0 q_0 + \varepsilon_1 q_1}{old} \cdot \frac{\Delta p}{p} = 1 - C \frac{\Delta p}{p} \quad (C = \frac{\varepsilon_0 q_0 + \varepsilon_1 q_1}{old}) \quad (24)$$

We now can substitute equation 19 with equation 24, and we can get a function of P :

$$\Delta t = f \left(\frac{\Delta p}{p} \right) = 500 \left[1 - \left(1 - C \frac{\Delta p}{p} \right)^2 \right] \quad (25)$$

According to the law of diminishing utility, the function above has to be a convex function. To testify the suitability of this function, we aim to find its second derivative shown in the following:

$$f''(x) = -1000C^2 \quad (26)$$

$f''(x)$ is smaller than 0, meaning the function itself is concave. Thus, proving the above function to be adequate at simulating people's change of time. we can then express the social welfare change as the subtraction of deadweight loss and the reduced of time cost by the consumer, express as the following:

$$W = \sum \Delta t_i \cdot timevalue_i - DW \quad (27)$$

D. Ramsey pricing model

Ramsey pricing model is a pricing strategy appropriate to business that do not treat profit as their priority. It is especially adequate at the case of metro which is also an infrastructure build by the nation that do not see profit as its top priority. The use of Ramsey price model in metro transportation is even more justify as other alternative is unable to give eligible solution. Due to a high fixed cost for urban rail transportation, employing the use of average cost pricing will be exorbitant for the consumer who are the general-public. Similarly, the enterprise will be in a huge loss if marginal cost pricing is adopted [6].

Apart from most of the optimal pricing model, Ramsey pricing is a sub-optimal pricing model which allow the optimal allocation of resource and the business to profit. To reach this goal, the model manages to maximize the sum of consumer surplus (*cs*) and producer surplus (*ps*) under the premises which the business does not run into deficit. The basic idea of Ramsey pricing model can be written as the following:

$$\max(ps + cs) \quad (28)$$

The premises which is the business do not run into deficit can be express as the following equation:

$$p(q)q - C(q) = 0 \quad (29)$$

Where $p(q)$ is the price function, q is the quantity purchased and $C(q)$ is the total cost.

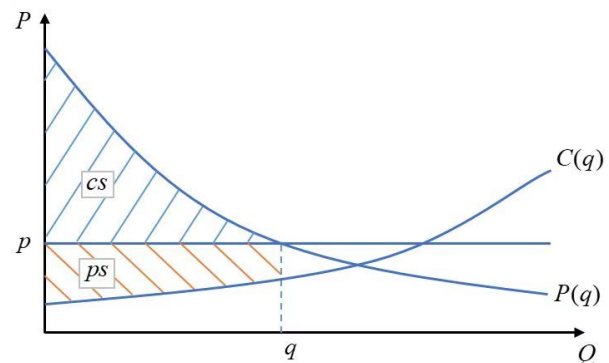


Figure 10. Consumer surplus and producer surplus

To determine the value of $\max(ps + cs)$, we first have to determine the expressions for ps and cs , as producer surplus is expressed as the difference between the total revenue of the business and the total cost for the company to produce a certain product. We can express ps in equation 30:

$$ps = p(q)q - C(q) \quad (30)$$

According to figure 10, cs is the area shaded blue in the figure. Thus, we can illustrate cs as the following:

$$cs = \int_0^q p(s)ds - qp(q) \quad (31)$$

The equation of maximizing $(cs + ps)$ is shown as the following:

$$\max(ps + cs) \quad (32)$$

To satisfy one major premise for Ramsey pricing model which is that producer do not run into deficit, we simply make sure the revenue of the producer is larger than its cost, shown in equation 33:

$$p(q)q - C(q) = 0 \quad (33)$$

As the goal of Ramsey pricing model is to determine the maximum value of equation 32, we need to search for a mathematical method to determine the maximum value. We adopted the use of Lagrange multiplier which is one of the prevalent way for solving conditional extremum. The equation is shown below where we simply added an additional part toward the original equation. While this additional part equals to zero, it doesn't affect the result of the conditional extremum:

$$L(q, \lambda) = \int_0^q p(s)ds - p(q)q + \lambda [C(q) - p(q)q] \quad (34)$$

We organize the equation and get the following version:

$$L(q, \lambda) = \int_0^q p(s)ds - (1 + \lambda)p(q)q + \lambda C(q) \quad (35)$$

The conditional extremum is obtained when making the partial derivative of the equation 35, shown in equation 36:

$$\frac{\partial L}{\partial q} = p(q) - (1 + \lambda) \left[\frac{dp}{dq} q + p \right] + \lambda MC = 0 \quad (36)$$

The price elasticity of demand is shown in the following equation 37:

$$\varepsilon = - \frac{\frac{dp}{dq} q}{p} \quad (37)$$

After substituting equivalent value, we can resolve a new formula with the elasticity of demand, shown in equation:

$$p - (1 + \lambda) \left(p - \frac{p}{\varepsilon} \right) + \lambda MC = 0 \quad (38)$$

Transform the equation, we have:

$$\lambda MC = p \left(- \frac{1 + \lambda}{\varepsilon} + \lambda \right) \quad (39)$$

Move p to the right side alone and we can resolve a new expression of p :

$$p = \frac{\lambda MC}{- \frac{1 + \lambda}{\varepsilon} + \lambda} \quad (40)$$

We define $R = \frac{1 + \lambda}{\lambda}$ where R is the Ramsey coefficient. As a result, we can obtain an expression of p related to marginal cost, elasticity of demand and R , we can observe this in equation 42:

$$p = \frac{MC}{1 - \frac{R}{\varepsilon}} \quad (42)$$

Here, p is negatively related with the price elasticity of demand and proportional to marginal cost. Equation 42 is the final expression of Ramsey pricing model.

1) Defining cost function

The Ramsey pricing model is now used to determine the fare level of urban metro system. However, we first need to establish a methodology to determine the value of MC (marginal cost), R and ε . To illustrate, MC could be described by a function of MC about the value of passenger volume. We generally assume that the value of passenger volume or ridership could affect the total operational cost of a city's metro system. This is because as the passenger volume or ridership increase, it inevitably brings damage and abrasion to the train and cause an additional demand for more labor force to ensure the security of metro station. Function relationship between passenger volume and total operational cost is presented in equation 43:

$$C(Q) = aQ^b \quad (43)$$

Where, $C(Q)$ is the total operational cost of metro system, a and b is constant and Q is the passenger volume. From this equation, we can resolve that as passenger flow increase in a certain value, total operational cost also increase. Equation 43 can be transformed to linear form which as shown in the following equation 44:

$$\ln C = \ln a + b \ln Q \quad (44)$$

To determine the value of a and b , and to testify the suitability of the equation at describing the relationship between passenger volume and total operational cost, a regression will be carry out in the following part.

2) Regression

To construct a cost function related to passenger volume, it is required to analyze the relationship between these two variables. Thus, a dataset of passenger volume and total operational cost are necessary for the analysis. In this study, we use data provided by Larry little field who

organized the data from national transit database. [7] The dataset includes the following information: Annual Passenger volume of New York city from 1991-2015, Annual Total operational cost of New York metro system from 1991-2015, All the statistic related to monetary value were adjusted after considering the effect of inflation.

3) *Quality of the data*

The credibility of the data is self-evident as it was adopted from national transit database [8] and New York university research center [9] which is the most authoritative database in the US. Meanwhile, the dataset possesses high reliability as we also consider the effect of inflation which might cause the value of MC to be substantially higher.

4) *Result of the regression*

To test our hypothesis about the relationship of passenger volume and total operational cost and to test equation 44, a regression model had been set up which is derived from the above data set. And the results are shown in the following table:

TABLE III. REGRESSION RESULT OF EQUATION 44

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9.351772	2.267696	4.123909	0.0004
LOG_DAY_VOLUME	0.442908	0.149985	2.953023	0.0071
R-squared	0.274913	Mean dependent var		16.04769
Adjusted R-squared	0.243388	S.D. dependent var		0.180819
S.E. of regression	0.157283	Akaike info criterion		-0.784928
Sum squared resid	0.568969	Schwarz criterion		-0.687417
Log likelihood	11.81159	Hannan-Quinn criter.		-0.757882
F-statistic	8.720347	Durbin-Watson stat		0.187366
Prob(F-statistic)	0.007134			

According to table 3, the p-value of the coefficient before the explanatory variable log_day_volume is 0.0071, which shows the linear relationship is significant. Then, equation 44 can be rewritten as the following:

$$\ln C = 9.3518 + 0.4429 \ln Q \quad (45)$$

the equation can be transformed to:

$$C = e^{9.352} Q^{0.443} \quad (46)$$

As we know that MC is the derivative of equation 46, MC could be express as the following:

$$MC = \frac{dC}{dQ} \quad (47)$$

$$MC = 0.443 \frac{e^{9.352}}{Q^{0.557}} \quad (48)$$

Equation 48 is the final expression of MC in relation with Q, which is the passenger volume per day. As this paper wants to solve the congestion within shanghai, we will adopt the use of Shanghai's passenger volume per day which is 12000000. Specific equation is expressed as the follow:

$$MC|_{Q=12000000} = 0.581 \quad (49)$$

Thus, the volume of MC for metro system in Shanghai is 0.581 dollars and the value of MC in equation 49 is also 0.581 dollars.

5) *Range of possible R*

The value of ε in total can be obtain by calculating the weighted average of the elasticities that the two different type of passenger possess. It can be illustrated in the following expression:

$$\varepsilon = \frac{\frac{\Delta q}{q} = \frac{\Delta q_0}{q_0} \cdot \frac{q_0}{q} + \frac{\Delta q_1}{q_1} \cdot \frac{q_1}{q}}{\frac{\Delta p}{p}} = work_ratio \cdot \varepsilon_0 + (1 - work_ratio) \cdot \varepsilon_1 = 0.216 \quad (50)$$

as the value of MC is determined by the above portion, we can now express P in R and R only:

$$P(R) = \frac{MC}{1 - \frac{R}{\varepsilon}} = \frac{0.581}{1 - \frac{R}{0.216}} \quad (51)$$

the figure of price as a function of R can be drawn in the following figure:

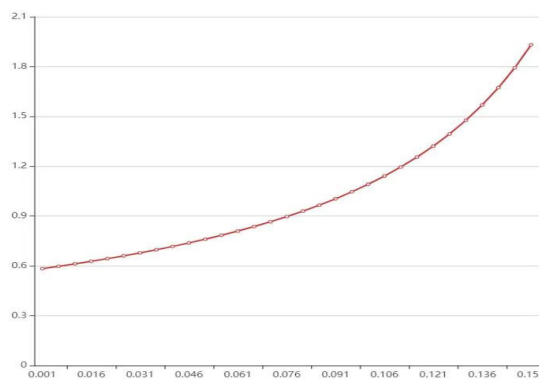


Figure 11. Function of price in relation with R. (R on horizontal axis and P on vertical.)

The purpose of Ramsey pricing model in this paper is to distinguish an acceptable range of price addition. To achieve this, we must find an acceptable range of R. this can be done by finding the point with the highest value of curvature. As figure 11 shows a distinct change of the slope of the function P about R, we believe that the points after this changing point, which is the point with the highest curvature, is unreasonable as the price increase over rapidly in response to a change of R. conversely, the points too close to the y-axis are also unserviceable as the goal to reduce congestion cannot be attained. Thus, we believe points, which have values in the range between half the value of the point with the highest curvature and that point, are desirable. The curvature can be expressed in the following equation:

$$K = \frac{d\alpha}{ds} \quad (52)$$

Where da can be expressed as:

$$d\alpha = d(\arctan(\frac{dy}{dx})) = \frac{d(\frac{dy}{dx})}{(\frac{dy}{dx})^2 + 1} \quad (53)$$

And ds can be shown as:

$$ds = \sqrt{(dx)^2 + (dy)^2} = dx[(\frac{dy}{dx})^2 + 1]^{\frac{1}{2}} \quad (54)$$

Thus, K can be calculated as:

$$K = \frac{\frac{d}{dx}(\frac{dy}{dx})}{[(\frac{dy}{dx})^2 + 1]^{\frac{3}{2}}} = \frac{y''}{(y'^2 + 1)^{\frac{3}{2}}} \quad (55)$$

In this case, however, we do not search for the point with maximum curvature algebraically but geometrically. This is because the value of y axis which is the price of subway system is 10 time larger in scale than the x axis which is the value of R. We can find the geometrical or visual maximum curvature point by making the scale of two axis equal, in other words, multiplying the x axis by 10. As this being made clear and the above function about curvature being obtained, we now search for mechanism to find the point with greatest curvature. But before that, the function of P against R can be illustrated as the following:

$$f(R) = \frac{P(R)}{10} = \frac{k}{0.216 - R} \quad (56)$$

Where, $k = 0.216 \times 0.581 \times 0.1 = 0.0125496$. The derivative and second derivative needed for the calculation of curvature are shown below:

$$f'(R) = \frac{k}{(0.216 - R)^2} \quad (57)$$

$$f''(R) = \frac{2k}{(0.216 - R)^3} \quad (58)$$

We can resolve K which is the curvature of the points on the function in R as the following:

$$K(R) = \frac{f''(R)}{[1 + f'(R)^2]^{\frac{3}{2}}} \quad (59)$$

We can derive the maximum value of K by getting the value of square K:

$$K^2 = \frac{f''(R)^2}{[1 + f'(R)^2]^3} = \frac{4k^2}{[(0.216 - R)^2 + \frac{k^2}{(0.216 - R)^2}]^3} \quad (60)$$

When K^2 is at its maximum value we have $(0.216 - R)^2 = k$, where $R = 0.216 - \sqrt{k} = 0.104$. Thus, we conclude that $[0.052, 0.104]$ is the range of desirable R.

6) Applications

We now want to discover the desirable range for the price of the subway system in shanghai. By having the

optimal range of R being obtain, the value of adequate P can be ascertain by inserting the result into the function:

$$P(R)|_{R=0.052} = 0.765 \quad (61)$$

$$P(R)|_{R=0.104} = 1.121 \quad (62)$$

Thus, the adequate range for price of subway system will be from 0.765 dollars to 1.121 dollars. After transforming dollars into RMB and implanting the value of Shanghai's current pricing system, we found the percentage increase of price shall be the following:

$$[\frac{1}{1 - \frac{R}{\epsilon}} - 1]|_{R=0.052} = 31.7\% \quad (63)$$

$$[\frac{1}{1 - \frac{R}{\epsilon}} - 1]|_{R=0.104} = 92.9\% \quad (64)$$

Thus, the desirable range for percentage price increase of Shanghai's metro system is between 31.7% to 92.9%.

4 RESULTS

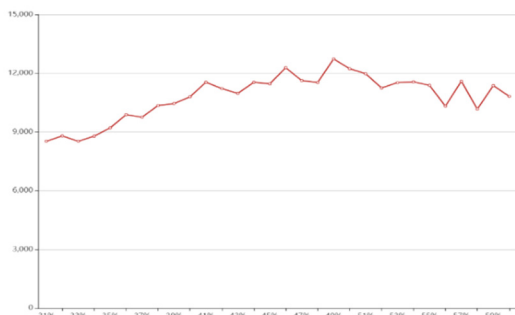
We derive the result of the data base on the agent simulation model, which relies on probability distribution. The price addition percentage is adopted by selecting the one with higher social welfare turnout while reducing congestion of the metro system. This social welfare or utility is calculated according to equation 27.

Moreover, the percentage price addition that leads to maximum social welfare needs to satisfy the desirable percentage price addition range provided by the Ramsey pricing model. If the result lies within this range, then we ascertain this result to be valid or beneficial.

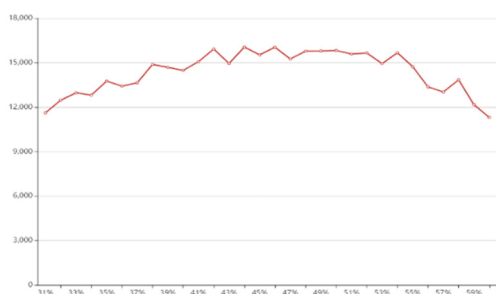
A. Percentage price addition

We first simulate the new passenger patterns after the price addition base on the agent simulation model. Agent simulation model relies on the elasticity of different types of people which includes working and non-working population. It also depends on passengers' patterns in the present day. These conditions are then elaborated using normal and uniform distribution.

Based on the new passenger behavior or patterns after the price addition generated by the agent simulation model, we then calculate the total social welfare change using the mechanism expressed in equation 27. This equation can be separated into two portions: one is the increase of deadweight loss after the price addition, and the other is the time cost gained by consumers as congestion is reduced. The time cost of consumers can be calculated by defining quantifying time into monetary value for different kind of passengers, respectively. The result of mean welfare change in relation with percentage price addition is shown in the following figure for MRH and ERH.



(a)



(b)

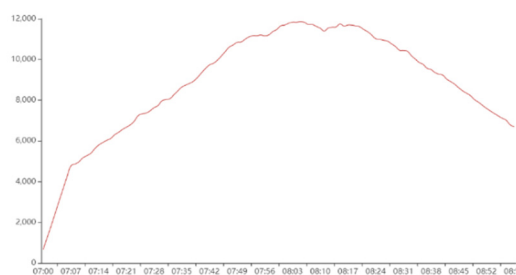
Figure 12. Increased social welfare versus percentage price addition

In both figures for MRH and ERH, we can observe a trend of first increase and then decrease. Thus, there must exist such a maximum value, which is the desired value of this study, during the middle of the increase and decrease of the social welfare change. This trend is indicated by equation 27, which is the expression of welfare change before and after the price addition. As the equation is proven convex, the value of social welfare change is not going to decrease if the general trend shows to be increasing. Similarly, it would not increase if the general trend exhibits to be decreasing. Thus, we only take a portion of the percentage price increase and social welfare figure in the range of 1-100% because the increasing trend can be observed near 1%-50% and decrease trend seen from 50%-100%. Thus, testing percentage price addition in a range of 1-100% is reasonable.

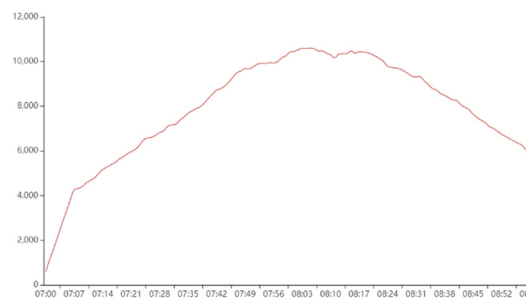
According to figure 12(a), the optimal level of social welfare is obtained when the percentage price addition in MRH (morning rush hour) is 49%. Similarly, according to figure 12(b), we obtain the optimal value of social welfare while the percentage price addition is 44% during ERH (evening rush hour).

B. Congestion reduction

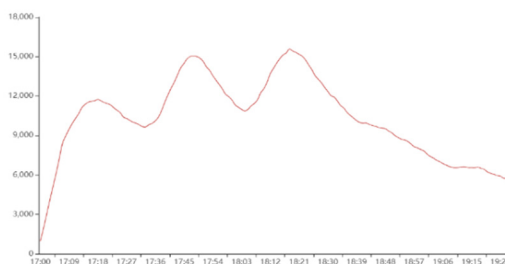
The primary purpose of this study is to reduce the congestion of significant stations during rush hours of a day. Calculating the total welfare of the society is only a mean to optimize the price addition and make it reasonable. We now examine the efficiency of our model at reducing congestion which the following figure shows the total amount of passenger volume before and after the price addition during morning rush hours. These values are acquired by calculating the total passenger volume of the stations that experience price addition, before the addition and after it.



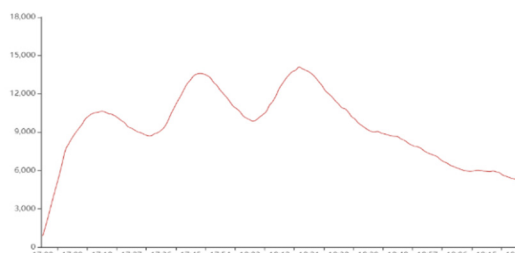
(a)



(b)



(c)



(d)

Figure 13. Passenger volume of the price-added sites before and after the addition

In figure 13(a)&(b), the maximum passenger volume during MRH decrease about 15 percent before and after the price addition. Similarly, the peak passenger volume of ERH also decreases by about 10 percent according to figure 13(c)&(d). These results are accompanied by great significance as congestion is vicious that decrease the overall social productivity. On the one hand, congestion means people's time is wasted, but on the other hand, it will also affect people's mood during a day and indirectly threaten the working population's productivity. Reducing congestion is not only about saving several minutes for individual, but more of what it does is that it saves several

minutes for every individual in every day, every month and every year and collectively, this means billions of hours for the society. Moreover, this congestion reduction is under the circumstances of our study is even more desirable as the total welfare of the society do not decrease but increase.

C. Result's adequacy according to Ramsey pricing model

To dually authenticate the reasonability of our results, we have to make sure that our results are allocated in the acceptable range of percentage price addition determined by the Ramsey pricing model. In Ramsey pricing model, we maximize the area of consumer surplus and producer surplus and resolve a function of price in relationship with R . we assume that the points of R half the distance before the maximum curvature point is acceptable as there exists an evident increase of the slope in the function. If we select R after this curvature

5 CONCLUSION

Congestion has been a nettlesome problem for hundreds of years. From the invention of public transportation to the development of cars, as long as there are public transportation and economic development, there exists congestion. Congestion is a relatively concerning issue as it not only wasted people's working hours but also interrupts people's mood of a day and destroy their incentive for providing quality work. For decades, government around the world had shown its effort at reducing congestion in public transportation. These actions include erecting more roads, building more metro lines, and even technological improvement were made in recent years. However, these methods all have a common problem which is they require colossal government expenditure that come from the tax-payers hand.

To tackle this issue, scholars around the world had provided solutions that do not consume significant government expenditure but control people's incentive from getting into the transportation stations. These solutions include differential pricing and intensely congested station express. However, the importance of differential pricing is explicitly not how it operates to reduce congestion because it is almost a common sense that while price increase, the passenger volume will decrease accordingly, the main focus is instead how to determine this level of price addition and past studies had given solutions like ladder pricing or even employing the usage of questionnaires. The problem, however, is that these solutions are over-simplistic and subjective that cannot represent the general behavior of the passengers. Thus, in this study, we search for a way of differential pricing based on big data and people's general behavior according to agreed economic principles such as elasticity. Also, by combining the mechanism of Ramsey pricing model and optimization model with a latest developed simulation model, we aimed to determine a reasonable policy of differential pricing that reduces the congestion while making sure the social welfare to be optimized.

In this study, we first generate new passenger patterns after the pricing policy with the use of agent simulation model. This simulation model is based on the behavior of passengers in reality and elaborated into results by specific

algorithms and probability distribution. After the results being generated, we calculate the social welfare change before and after the price addition using the subtraction of deadweight loss due to a price increase and again in time as the congestion decrease. The time is quantified into monetary value according to the time cost of different types of population. We iterate this process for different percentage price increases and then select the one with maximum social welfare increase as our final result. Furthermore, we determine an acceptable range of percentage price increase using Ramsey pricing model, which is a pricing mechanism based on the maximization of the social surplus. We finally illustrated that our results of optimal percentage price addition are in the acceptable range generated by the Ramsey pricing model. Thus, by testing the value dually using both optimization model and Ramsey pricing model, we are confident of claiming that our results are not only reasonable but desirable and robust.

We believe that our work can provide a useful reference for policymakers dedicated to reducing congestion as we make sure the increase, not decrease, in society's overall welfare or utility during the process of decreasing congestion. Reducing congestion is not only about saving several minutes, but it is about saving billions of work hours for the collective society and improve people's overall welfare.

We expect policy regarding differential pricing can be adopted in the future to reduce congestion as it would solve the problem efficiently by promising an increase in social welfare. Moreover, it can also save the policy makes billions of dollars as the costly efforts to reduce congestion by constructing new roads, railways, and systems can be eliminated.

In the future, we hope the continual development in theories of differential pricing to give a guideline for the policymakers. We also expect the development of other kinds of uncostly mechanisms at reducing congestion such as broad station express. Comparing to a society with intense congestion and traffic jam, a society with cozy transportation will not only have higher productivity that will cause a higher GDP, gain a comparative advantage at the international competition, but will also possess a relatively desirable social culture and paragon ambiance.

ACKNOWLEDGMENT

First, I would like to express my most genuine appreciation to my mentor Mr. Yang Liu Yong. We met during my study trip to Zhejiang University. Although you are a professor at Zhejiang University, and I am only a high-school student in Shanghai, you never show any condescension. Instead, it was your gratis constructive suggestion and examination that make this working paper possible.

Second, I shall express my gratitude to my further math tutor Mr. Cao. Without him, I wouldn't have had the mathematical basis for this working paper. It was also you who provide technical support when I have problems with my algorithm.

I would also like to give thanks to my parents. They are the ones who give me invaluable supervision and psychological supports.

Lastly, thanks to our school's dean of the student because it is your detention that inspired me for doing a research paper about reducing congestion. But moreover, it is the recognition of congestion's devastating impact that contributed to my willingness to choosing this topic.

REFERENCES

1. W. Vickrey, "Some implications of marginal cost pricing for public utilities," *The American economic review*, Vol. (45), No. 2: 605-620, 1955.
2. S. Glaister, D. Lewis, "An integrated fares policy for transport in London," *journal of public economics*, 1978(9): 341-355.
3. SODA2018 competition statistics, <http://soda.shdataic.org.cn/data/?cid=0&exclude=10,3,2,5,4,7,6,9,8>.
4. U. Aickelin, Introduction to Multi-agent Simulation, University of Nottingham, <https://arxiv.org/ftp/arxiv/papers/0803/0803.3905.pdf>, unpublished.
5. Mayworm, Logo and McEnroe, Transit Service Elasticities published in *Journal for Transport Economics & policy*, UK, 1981.
6. B. Nahata, A. Izyumov and V. Busygin : *J. Energy Economics*, Vol.29, p.105, 2007.
7. Saying the Unsaid in New York, 2015, <https://larrylittlefield.wordpress.com/2017/05/11/metro-ny-transit-costs-data-from-the-2015-national-transit-database/>.
8. "The National Transit Database (NTD)." Federal Transit Administration, United States Department of Transportation, 21 Aug. 2015, <https://www.transit.dot.gov/ntd>.
9. "Rudin Center Research." NYU Wagner, <https://wagner.nyu.edu/rudincenter/publications>.