

Equity Analysis of Urban Rail Fare Policy and Passenger Overload Delay: An International Comparison and the Case of Metro Manila MRT-3

Andra Charis MIJARES ^a, Mio SUZUKI ^b, Tetsuo YAI ^c

^{a,b,c} *Graduate School of Science and Engineering, Tokyo Institute of Technology, Tokyo, 226-8503, Japan*

^a *E-mail: mijares.a.aa@m.titech.ac.jp*

^b *E-mail: mios@enveng.titech.ac.jp*

^c *E-mail: tyai@enveng.titech.ac.jp*

Abstract: This study focused on the intra-modal equity of urban rail fare policy and passenger overload delay. A framework for evaluating equity of a fare policy is proposed, and used to make a macroscopic comparison among several urban rail fare policies. It was found that there is a trade-off between horizontal and vertical equity. A new passenger overload delay equity index is also proposed based on the theory of relative deprivation. The study also examined the case in Metro Manila where urban rail fares have been kept constant since 2000 amidst inflation and increase of non-rail public transport fares. This entailed estimations of station O-D matrix using gravity model and passenger waiting time using queuing theory. It was found that although it is commendable in terms of fare affordability, it appears that the very low rail fares have increased rail demand beyond capacity, thus bringing about inequity in passenger overload delay.

Keywords: Equity, Urban Rail, Fare Policy, Passenger Overload Delay

1. INTRODUCTION

Equity is defined as fairness in the distribution of goods and services among the people in an economy. It is one of the core values in society – it is intrinsic for humans to possess a sense of justice and equality – yet, it almost always gets second-rate treatment.

In the context of urban rail fare policy, there are two aspects of equity: intra-modal and intermodal. The first aspect refers to equity among urban rail passengers, while the latter is concerned with equity between rail and non-rail public transport modes. This study focuses solely on intra-modal equity. Furthermore, there are two types of transportation equity: horizontal and vertical. Horizontal equity demands within-group equity, that is, equal treatment among equal groups, while vertical equity denotes between-group equity and imposes special consideration towards transportation disadvantaged people such as the poor, students, elderly, and disabled people.

Meanwhile, fare policy has four components. The fare level stands for the price for making a trip and affects fare affordability – too expensive fares may price poor people off transit. Fare structure refers to the spatial structure that supports the fare system, that is, how fare levels change according to distance, zone, or time of day. Common fare structure types are flat, distance-based, and zone-based. Ticketing type refers to the payment medium such as single ticket, monthly pass and stored value cards, which may allow discounts. Concessions are the discounts offered to transportation disadvantaged people since they are typically on limited income and may have mobility difficulties.

Moreover, urban rail fare policy is linked to the level of service because fare levels can

trigger behavioral changes among riders and thus affect demand. Insufficient capacity with respect to demand leads to overcrowding and passenger overload delay, and an equity issue arises because delay is notably higher in some stations than others. In this context, the concept of equity can be extended to include the theory of relative deprivation, which occurs when individuals or groups subjectively perceive themselves as unfairly disadvantaged over others who are perceived as having similar attributes and deserving similar rewards. Thus, delayed passengers may compare themselves against all passengers and view themselves as deprived.

The specific case of Metro Manila is of interest because of its unusual situation: urban rail fares have been kept constant since 2000 amidst inflation and increase of non-rail public transport fares. Normally, rail fares are set higher due to its higher quality of service. However, the fares are kept artificially low through increasing subsidies sourced from the national government. This has made rail travel relatively cheaper, and has increased rail demand beyond capacity, inducing equity issues regarding user costs and level of service. As an illustration, the Metro Manila MRT-3 is designed to carry about 22,500 passengers per peak hour per direction (PPHD) but demand in 2008 has increased to 26,500 PPHD.

This paper is organized as follows. The next section contains a brief review of related literature. Section 3 proposes a framework for equity analysis of urban fare policies, and then uses this framework to compare several international fare policies. In Section 4, a passenger overload delay equity index based on the concept of equity and the theory of relative deprivation is proposed. It also examines the specific case in Metro Manila MRT-3, including O-D and waiting time estimation and delay equity analysis. Finally, conclusions are drawn and future work is discussed in Section 5.

2. PREVIOUS STUDIES

Equity is a subjective concept, but previous studies have used conventional inequality indices to quantify it. In transportation, Levinson (2007) has used the Gini coefficient to assess the equity of delay and mobility in ramp metering. Ramjerdi (2006) emphasized on the importance of using several inequality measures in analyzing equity. He utilized the mean, range, variance, coefficient of variation, relative mean deviation, logarithmic variance, variance of logarithms, Theil, Atkinson, and Kolm indices, and Gini coefficient in analyzing the change in equity of welfare after the application of a specific policy.

Another approach for measuring equity in the context of transit fare policy is through the distributional effects using transit demographics in combination with basic service consumption data, such as transit costs to draw conclusions regarding transit service equity. The costs of the operator are not considered explicitly, and transit subsidies are exogenous to the transit users. In other words, the analysis of equity is done from the passengers' viewpoint and not that of the operator's. Studies that use this approach include Leutze & Ugolik (1978), Pucher (1981, 1983), and Martinelli & Medellin (2007). This implies that passengers should pay according to the cost that they impose on the system, and this has been the subject of some studies by Cervero (1981, 1990). He pointed out that flat fare is inequitable since short-distance passengers end up subsidizing long-distance ones. In a report on U.S. fare policies (TCRP, 1994) that market-based pricing schemes such as monthly passes raise equity concerns as it requires an initial payment that is usually much higher than the cash fare.

Moreover, the most common assumption in waiting time measurement in transport studies is merely taking half the headway as the average waiting time, which is the result for a perfectly regular service with Poisson arrivals and sufficient capacity. De Cea and Fernandez (1993) argue that the oversimplification of waiting time assumption is justifiable as it is

impractical to go into a more complicated formulation for most purposes. However, in congested networks, waiting time increases as the discrepancy between demand and capacity increases. It is therefore useful to forego the “sufficient capacity, constant headway” assumption and include the probability of being refused in the estimation of waiting time.

Several studies such as that of Bell (1995) focused on the delay arising from residual queues that form when capacity is exceeded. In the context of transit, Lam *et al* (1999) called this type of delay as ‘passenger overload delay’ to refer to the time penalty that passengers will wait for the next coming vehicle or transfer to the alternative routes when they cannot board the first coming vehicle because of insufficient capacity of in-vehicle links. They also established a mathematical model for estimating it. This term has been used in subsequent studies including Lam *et al* (2001, 2003, 2009), Wu and Lam (2003), Wahba and Shalaby (2005), Yang and Lam (2006), Zhang *et al* (2010), and Szeto *et al* (2011). Refused passengers who have higher delay tend to be concentrated on bottleneck stations, thus resulting in equity problems. The equity of passenger overload delay was examined implicitly by Shimamura *et al* (2005) by incorporating the failure-to-board probability in their transit assignment problem. They defined a concept called of connectivity reliability as the probability of arriving at the destination without failing to board at any station, and thus measures congestion level. The Gini coefficient was then used as an equity measure of the connectivity reliability (and not of waiting time per se), and was stipulated as one of the objective functions in the bi-level programming problem for optimization with equilibrium constraints. Equity in waiting time due to queuing is also a topic of interest in the fields of telecommunications and computer systems (e.g. Avi-Itzhak and Levy, 2004) and consumer service (e.g. Goodwin *et al*, 1991).

Another prospective approach in measuring equity is through the use of distributional poverty gap measures, which are based on inequality indices and the theory of relative deprivation (Clark *et al*, 1981). So far, these measures are used to measure poverty, but could be extended to the concept of passenger overload delay.

3. INTERNATIONAL RAIL FARE COMPARISON

3.1 Framework for Intra-modal Equity Analysis

Figure 1 shows the relationship between the equity aspects, equity types, and fare policy components.

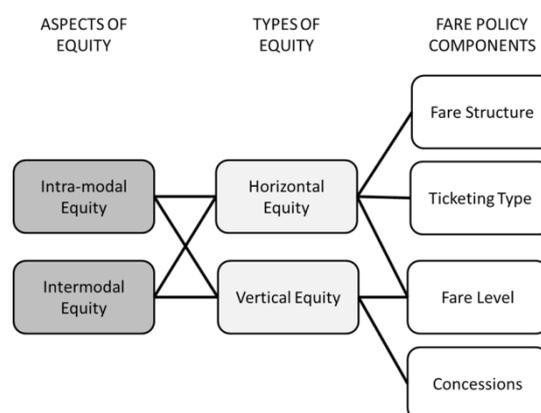


Figure 1. Theoretical Framework

Intra-modal equity is the focus of this study, so intermodal equity was not considered. The original theoretical framework of was used to conduct an intra-modal equity analysis and

macroscopic comparison of international rail fare policies. Table 1 shows the indicators used. The rationales for choosing the indicators are given in the subsequent sections.

Table 1. Intra-modal Equity Indicators for International Rail Fare Comparison

Type of Equity	Indicator
Horizontal Equity	Distance-based equity: Equity of fare per kilometer traveled regardless of trip length
	Frequency-based equity: Equity of fare paid per trip regardless of trip frequency
Vertical Equity	Fare affordability: Ratio of two trips relative to minimum daily wage
	Level of concessions for students, the elderly, and disabled persons

3.2 Description of Data

Data regarding fare levels, fare structures, ticket types, concessions, as well as minimum daily wage, were collected for 136 urban rail fare policies. Rail lines that operate under the same pricing system (e.g. all Tokyo Metro lines) are considered as one policy to avoid redundancy. Majority of these urban rail fare policies are from Japan (30) and the USA (47) due to data availability, with other policies from Australia, Brazil, Canada, China, Denmark, Egypt, Finland, France, Greece, Hong Kong, Hungary, India, Italy, Malaysia, Mexico, Netherlands, Norway, Philippines, Russia, Singapore, South Korea, Sweden, Taiwan, Thailand, United Arab Emirates, and United Kingdom. The average line length is 31.6 km, and 39 lines are light rail and the rest are heavy rail.

Table 2. Classification by Fare Structure, Discount Ticketing Medium, and Location

Fare Structure/ Location	Discounted Ticketing Medium						Total
	MP	MT	PT	SC	SV	none	
Distance-based	45			14	4		63
East Asia	38			9	1		48
Europe	2						2
Middle East				1			1
North America	1			2			3
Southeast Asia	4			2	3		9
Flat Fare	42	2	2	6		11	63
East Asia	1			2		1	4
Europe	6	2	2				10
Latin America						4	4
Middle East						1	1
North America	35			4		5	44
Zone-based	4	1		4		1	10
East Asia					1		1
Europe					1		1
Middle East					1		1
North America	3					1	4
Oceania	1	1		1			3
Total	91	3	2	24	4	12	136

MP=monthly pass, MT=multiple tickets, PT=prepaid ticket, SC=smart card, SV=stored value card

The most predominant fare structure in North America and Latin America is the flat fare, while distance-based is more prominent in Asia. European and Middle Eastern countries

operate a variety of fare structures. Meanwhile for ticketing types, there is a variety of discounted ticketing media aside from the basic single journey ticket that are used by many transit agencies. Several fare policies differentiate the fare depending on the ticketing medium used and the frequency of use. The most commonly used is the monthly (or period) pass, which provides unlimited trips within a month and is priced at a certain number of breakeven trips to provide a discount to frequent riders, although some agencies cap the maximum number of trips to avoid potential abuse. Multiple tickets are pre-purchased at a discounted rate. Stored value cards have automatically loaded money on to card off-system and relevant fare automatically deducted. Prepaid tickets refer to those purchased off-site. Smart cards are integrated circuit card that serve mainly as a public transport payment card and an electronic wallet. A barrier for the use of discounted ticketing media is the affordability of the initial cash-out associated with it. This is a main issue especially for low-income users who receive their wage on a daily basis, and thus cannot afford to shell out a huge amount at once. The classification according to ticketing type, fare structure and area is tabulated in Table 2.

3.3 Assumed Trip Length Distribution

Since we do not have information regarding the actual O-D distribution for each fare policy, there is a need to assume an identical trip length distribution for the population that is applicable to all fare policies, and then subject these to equity analysis. In this case, four types of distributions were assumed (i.e. Gamma, Log-normal, Normal, and Flat Uniform) and the parameters for each distribution were made to vary ten times each to assess the sensitivity of the equity measure as well as the rankings of the fare policies according to the travel patterns. The average trip length was made to vary between 7-20 km and the standard deviation between 2-12 km. 20,000 samples were drawn for each distribution-parameter combination, and the frequency distribution was discretized into 1-km intervals for simplification. Thus, this is merely a simulation study and results should be interpreted as such.

3.4 Horizontal Equity

3.4.1 Distance-based equity

This indicator was measured using the Gini coefficient for fare paid per kilometer assuming different trip length distributions and different trip distances. It generally aims to measure the degree to which the fare structure differentiates against distance, and is independent of the actual fare level thus allowing comparison. Figure 2 shows the Lorenz curve for distance-based equity, in which the cumulative shares of fare per km against the cumulative percentage of the population.

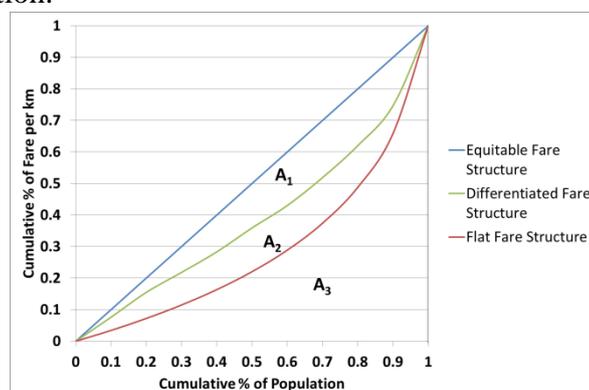


Figure 2. Lorenz Curve for Distance-based Equity

The Gini coefficients are expressed as:

$$G_{flat} = \frac{A_1 + A_2}{A_1 + A_2 + A_3} \tag{1a}$$

$$G_{diff} = \frac{A_1}{A_1 + A_2 + A_3} \tag{1b}$$

where A_1 , A_2 and A_3 correspond to the areas identified in the graph. The Gini coefficient is equivalent to the average difference between all possible pairs of resources expressed as a proportion of total resources.

The range of trip lengths was also varied from 0 to 15 km trips, 0 to 30 km trips, and 0 to 45 km trips to represent different city sizes. Since some fare policies only support a certain trip length (e.g. Metro Manila LRT-1 is only around 20 km long), fare policies had to be dropped as the trip length considered was increased. Distance-based equity is based on the level of differentiation of the fare structure with respect to distance, so all flat fare structures will have the same index, and thus all flat fare policies (63 out of 133 policies) were labeled as “Flat Fare” while differentiated structures were treated individually. It was observed that not all distance-based structures are alike – some differentiate fare level with respect to distance more strongly than others. Some employ a staggered fare structure that increases for every km or so (e.g. rail lines in the Netherlands and Japan), while some use a structure that only increases the fare once for the entire length (e.g. Shanghai Metro and Busan Metro), practically a flat fare structure. Some fare policies differentiate fare structure according to ticket types – charging a more differentiated fare (i.e. more equitable) for card than single ticket users (e.g. Singapore MRT and Metro Manila LRT-1) for operational purposes. Figure 3 shows the ranking of the policies according to trip length range and distribution type.

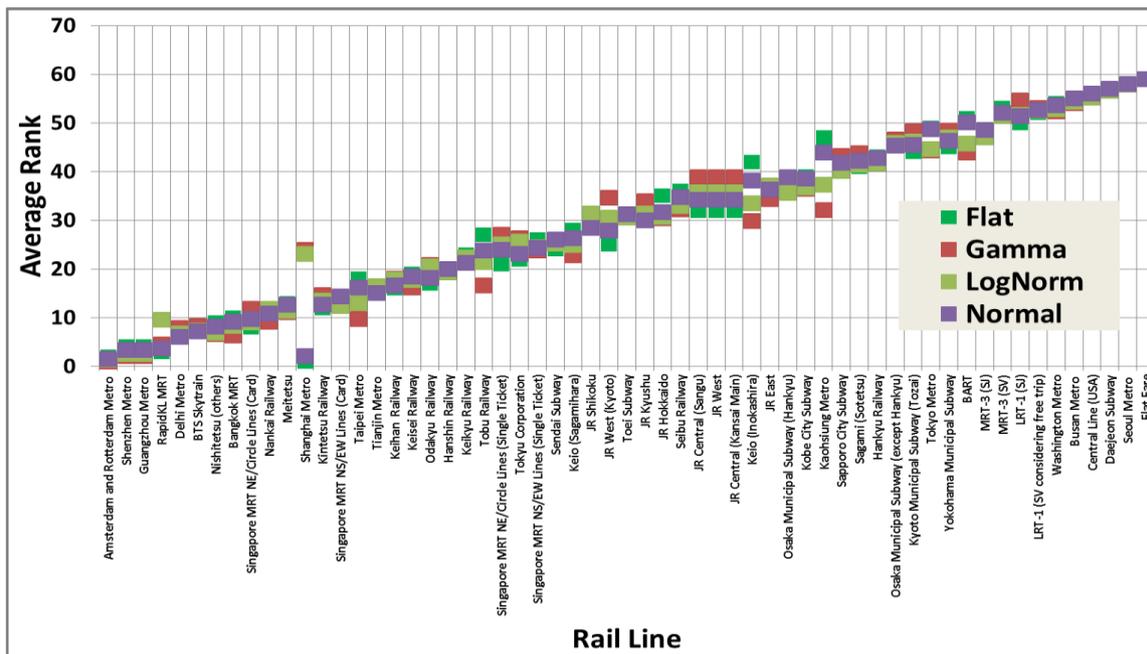


Figure 3. Average Rank of Distance-based Equity for Trip Length Range of 0-15 km

Assuming a flat distribution, the analysis of the distance-based equity of fare policies for 15-km, 30-km and 45-km trip lengths reveals two types of fare policies, as divided by the yellow line, as seen in Figure 4. On the left hand side, fare structures become more equitable as trip length increases, indicating strong fare differentiation even for longer lengths and

whose ranks are generally high. On the right hand side, the opposite trend is observed – the fare structures turn out to be more equitable for short trip lengths and become increasingly inequitable as trip length is increased. Fare policies that belong to this group have widely varying or consistently low ranks.

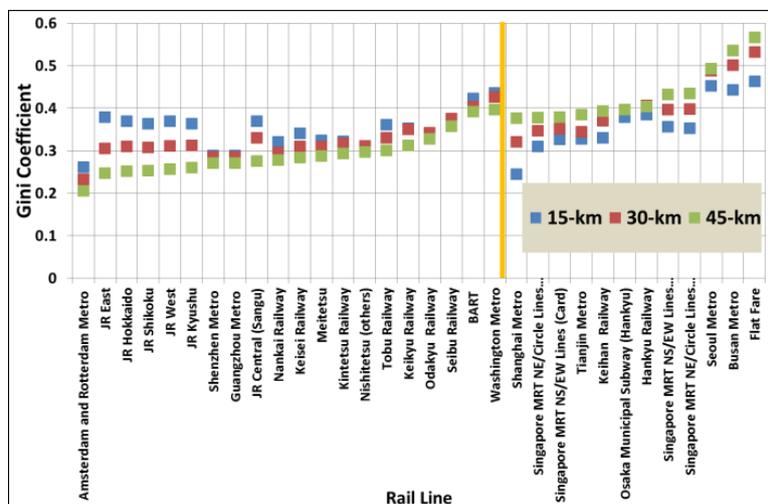


Figure 4. Grouping of Fare Policies According to Distance-Based Equity (Flat distribution)

3.4.2 Frequency-based equity

This indicator aims to measure the discrepancy of fare paid according to the ticket type used by employing the Gini coefficient. To enable the analysis, five distributions of frequency of trips within a month were assumed as well as the change of demand with respect to the affordability of the initial cash-out. It was also assumed that a certain user will purchase a monthly pass if his trip frequency is greater than or equal to the number of trips to break even, provided that he can afford it. For other discounted media (i.e. stored value card, smart card, multiple tickets and prepaid tickets), it was assumed that users who make at least ten trips per month would consider purchasing such media.

It was found that monthly passes with low breakeven trips but are priced high with respect to the minimum daily wage (e.g. Copenhagen Metro) are generally less equitable because it allows higher discounts for people who travel more frequently. For other discounted media, there is higher inequity for policies that offer greater discounts as opposed to single ticket users (e.g. London Underground) and those that require high initial cash-out. The most equitable fare policies from the perspective of frequency-based equity are those policies that do not provide any differentiation according to trip frequency, with $Gini=0$.

3.5 Vertical Equity

Fare affordability is defined as the ratio of a round-trip fare and the minimum daily wage. The minimum daily wage was chosen as the parameter because it targets working commuters with the lowest income as to whether or not they can afford the use of travel via urban rail. This is consistent with the World Bank’s definition of affordability as the extent to which the financial cost of journeys put an individual or household in the position of having to make sacrifices to travel or the extent to which they can afford to travel when they want to (Carruthers *et al*, 2005). In their case, an affordability index was defined as bus fares for 60 monthly trips as a percent of average per capita income for the poorest 20 percent (quintile) of

population. Moreover, in 1988, the Economic Commission for Latin America and the Caribbean constructed an affordability index according to how much would be needed to make fifty trips per month, expressed as a percentage of the minimum wage, for ten cities in Latin America.

It was found that a round-trip base fare costs an average of 4.0% of the minimum daily wage, while that for two 10-km trips takes up an average of 7.4%. A 10-km trip is just used as a benchmark since we do not have information on the actual trip length for each line. It was observed that distance-based structures tend to be cheaper for short trips, but flat fare structures tend to be cheaper for longer trips, as seen in Figure 5.

Moreover, using the assumed trip length distributions, the Gini coefficient for fare affordability regardless of trip length was computed. Plotting it against the Gini coefficient for distance-based equity for rail lines with trip length range of 0 to 45 km, it was found that there is a trade-off between the two, as seen in Figure 6. Linear regression shows this trade-off relationship. This indicates that some fare policies prioritize fare affordability regardless of trip length, such as the flat fare policy. On the other hand, strongly differentiated fare policies such as the JR lines and Amsterdam Metro emphasize on charging passengers the cost they impose on the system.

Additionally, concessions for transportation disadvantaged people are provided by almost all fare policies, with several policies giving free trips to elderly and disabled persons (e.g. several lines in Australia, South Korea and China). Student discounts are generally lower, with only a few policies providing free trips under certain conditions (e.g. Chicago 'L' Transit for full-time college students from participating schools during the school term, and Brasilia Metro during weekdays).

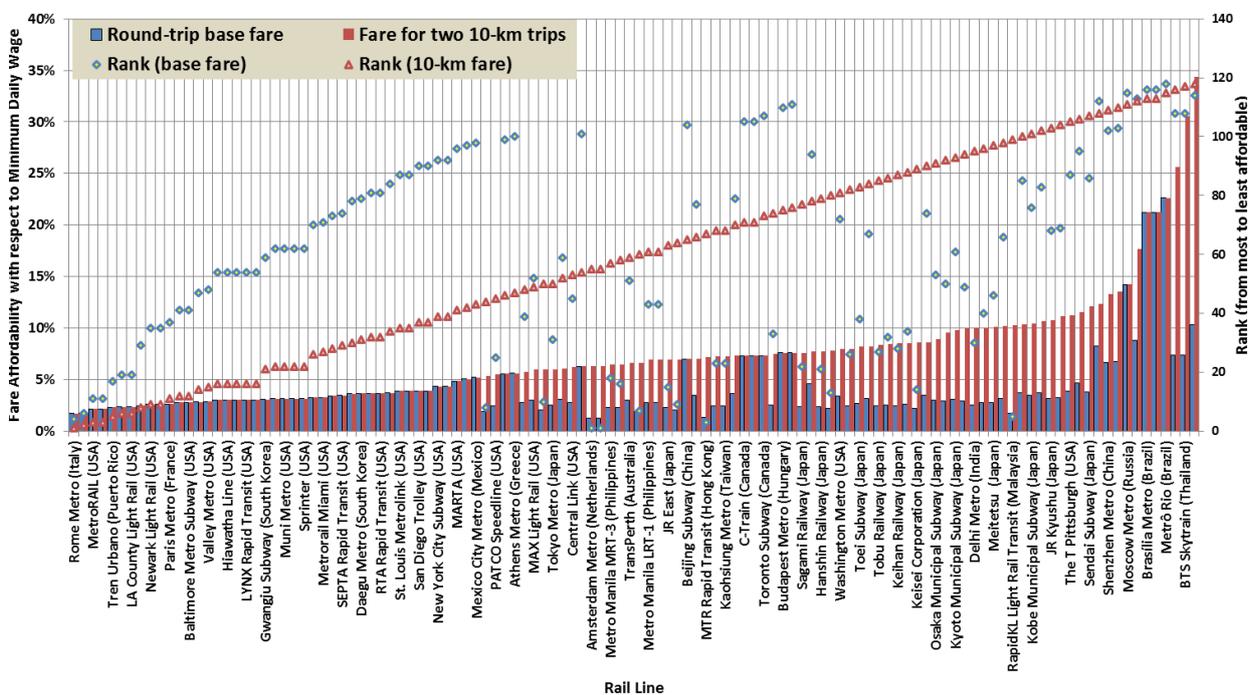


Figure 5. Fare Affordability of Round-trip Base Fare and two 10-km trips (arranged from most affordable to least affordable)

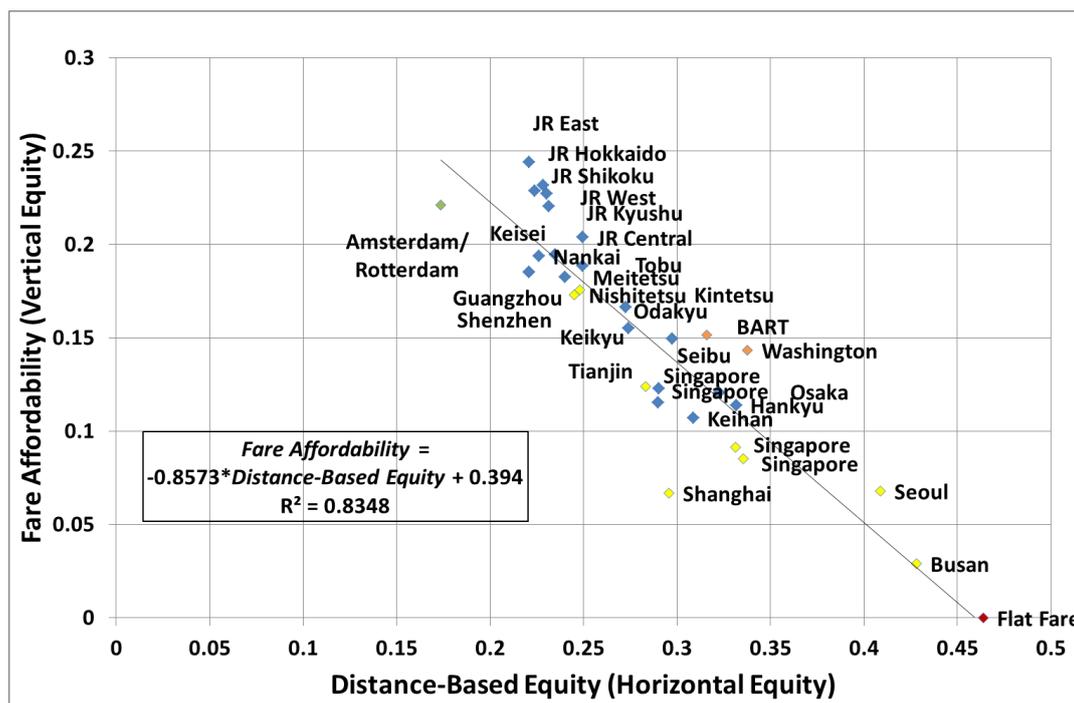


Figure 6. Plot of Distance-Based Equity against Fare Affordability (assuming Normal (15,10) and trip length range of 0 to 45 km)

3.6 Overall Equity Assessment

Overall equity assessment is done by considering all aspects and types of intra-modal equity. To enable comparison between fare policies, it is necessary to normalize, weight and aggregate the indicators. Min-max benchmarking is used, wherein X_{min} is the lowest value among all fare policies considered and X_{max} is the highest one. Thus, equity is scored on a relative rather than absolute basis.

Equations 2a and 2b show the normalized value I_{qp} for the raw indicator X_{qp} for an indicator that increases, and decreases in value as equity increases, respectively.

$$I_{qp} = \frac{X_{max} - X_{qp}}{X_{max} - X_{min}} \tag{2a}$$

$$I_{qp} = 1 - \frac{X_{qp} - X_{min}}{X_{max} - X_{min}} \tag{2b}$$

Equal weighting method was then used, wherein equal weights are applied to each component and equal weights are applied to each indicator within each component. Linear aggregation is then applied, with the formula given as:

$$CI_p = \sum_{q=1}^Q w_q I_{qp} \tag{3}$$

With $\sum_{q=1}^Q w_q = 1$ and $0 \leq w_q \leq 1$, for all $q=1, \dots, Q$ and $p=1, \dots, M$.

Where w_q is the weight of each component q . p refers to the policy, and I_{qp} refers to the indicator for each component of the policy considered. CI_p is the composite indicator of equity for policy p . Linear aggregation allows trade-off between indicators; a low score on one indicator can be compensated by a high score on another. The range of the equity score is

[0,1], wherein the rail policy with the highest score on *all* indicators will have a score of 1.

Truncated normal distribution with a mean of 15 km and standard deviation of 10 was used as the representative trip length distribution. A plot of horizontal equity (i.e. distance-based and frequency-based equity) and vertical equity (i.e. fare affordability and concessions) is shown in Figure 7. Shenzhen and Guangzhou Metro are seen as the most equitable due to their high scores on both vertical and horizontal equity. Moreover, Metro Manila rail lines were found to rank #60, #66 and #93 out of 116 policies considered. This result however, is limited by the assumptions made on trip length and frequency and on the definition of equity, and should be interpreted as such.

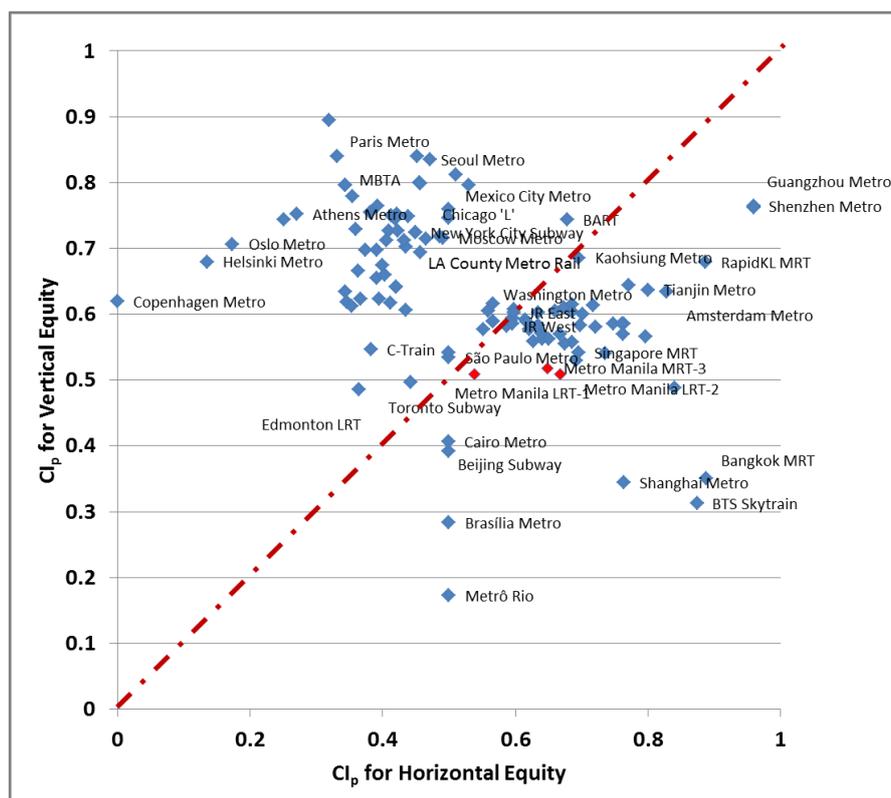


Figure 7. Plot of Relative Horizontal vs. Vertical Equity Index (Overall Equity Assessment)

4. CASE STUDY: METRO MANILA MRT-3

The section focuses on another part of the horizontal aspect of intra-modal equity. While the previous section analyzed fare policy components only, this section extends the analysis to the equity level of passenger overload delay, which is indirectly caused by the fare policy. Fares influence passenger behavior, thus affecting demand and O-D patterns, which in turn dictate the level of service of the system according to the available system capacity.

4.1. Proposed Passenger Overload Delay Equity Index

Following the definition by Lam *et al* (1999), we use the term passenger overload delay to refer to the time penalty that passengers will wait for the next coming vehicle when they cannot board the first coming vehicle because of insufficient capacity of in-vehicle links. In this case, other sources of delay are not considered (e.g. variation in headway, accidents) and the passengers do not have an option to transfer to alternative routes (i.e. no transit

assignment problem). In other words, when all trains arrive on time, passenger overload is the only source of delay. Since passenger overload delay is a time penalty and the maximum expected waiting time for a line with sufficient capacity is equal to the headway, it is then reasonable to define passenger overload delay as the difference between the actual waiting time and the headway. Thus, delayed passengers are those who are refused on the first train, and thus have a waiting time greater than the headway. It follows that undelayed passengers are those with waiting times less than or equal to the headway.

However, in a situation where the service reliability is imperfect, train operations is an additional source of delay. A comparison is given below in Figure 8. In this study, we consider perfect train service reliability and focus on the equity of passenger overload delay.

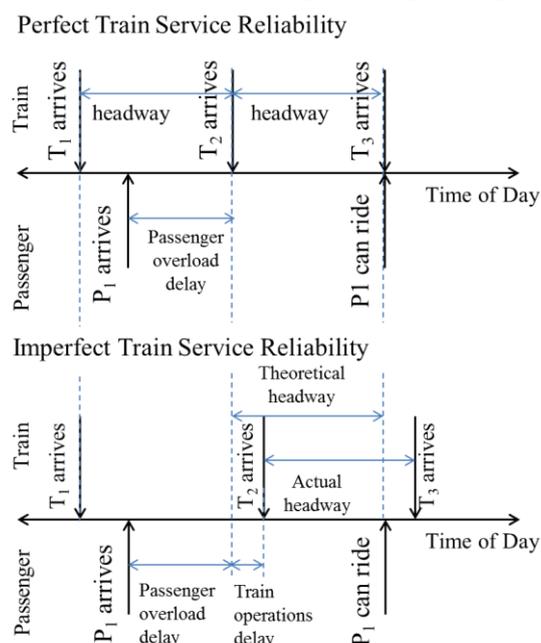


Figure 8. Definition of Passenger Overload Delay

A passenger overload delay equity index that captures the concepts of equity and relative deprivation among passengers is proposed. When delayed passengers compare themselves with less delayed and undelayed passengers, they would feel relative deprivation. Conversely, when undelayed passengers compare themselves to delayed passengers, they would feel relative gratification. Using the headway as the reference point implies that the passenger was not able to board the first train. Let a “time interval” be the period between just after the previous train left up to until the next train arrives. A passenger who arrives at a certain time interval expects to ride on the train that arrives at the end of the interval, and if he does so, he is undelayed.

The proposed delay equity index entails the comparison of waiting times among all passengers under a censored waiting time distribution rather than the original waiting time distribution. This means that all undelayed passengers are considered to be equal to the headway. It would also reflect the improvement in social welfare due to a decrease of the number of delayed passengers. This is similar to a distributional poverty gap measure called Takayama Index, which measures poverty under a censored income distribution wherein all non-poor people have income equal to the poverty line (Clark *et al.*, 1981). To apply this in the context of passenger overload delay, the following equations are defined.

Head count ratio, H , refers to the portion of delayed passengers q among all rail passengers n during the morning peak period.

To formulate the social welfare-theoretic delay equity index, it is necessary to establish the relationship between the frequency distribution of waiting time y , $f(y)$, and the frequency distribution of the social (dis)utility $-w$, $g(-w)$. For any individual with waiting time y_i , there is an associated disutility level $d(y_i)$, which is the deprivation function for an individual due to waiting time.

$$d(y_i) = \frac{1}{\alpha} [\max(z, y_i)]^\alpha \tag{3}$$

Waiting time is used in the deprivation function instead of passenger overload delay because the use of the latter would imply that only the relative deprivation of delayed passengers among themselves is considered, and not against the undelayed passengers. It is reasonable to assume that delayed passengers would compare their situation with undelayed passengers as well. In effect, a censored waiting time distribution is used because the Poisson arrival is considered. Some degree of inequity is already inherent for a Poisson arrival process, thus, if a censored waiting time distribution is used, the contribution of the assumed randomness of the arrival process itself is reduced and the inequity would largely be attributed to insufficient capacity.

The social welfare function is increasing and additive, and is given as the sum of all deprivation functions. It represents the total disutility felt by society due to the waiting time.

$$-w(y_p, z, \alpha) = \sum_i d(y_i), \quad i = 1, \dots, n \tag{4a}$$

$$= \left(\frac{1}{\alpha}\right) \sum_i y_i^\alpha + \left[\frac{n-q}{\alpha}\right] z^\alpha, \quad i = 1, \dots, q \tag{4b}$$

The inequality aversion parameter is $\alpha \geq 1$ for concavity in waiting time, and it represents the importance given to passengers with higher delay. The above equation means that undelayed passengers are considered to have a waiting time equal to headway, given that the actual value of waiting time for undelayed passengers is entirely due to the randomness of arrival. Since a smaller value of y_i is desirable, as when $\alpha > 1$, more weight is placed on large waiting times in determining $-w(y_p, z, \alpha)$ and in the limit, as $\alpha \rightarrow \infty$, only the largest waiting time matters and $-w(y_p, z, \alpha)$ becomes maximin with respect to waiting time.

From here, we then define an equally distributed equivalent waiting time, y^* , for all passengers, which is the value of waiting time that if shared by all passengers yields the same level of social welfare as the censored waiting time distribution. The equally distributed equivalent waiting time, y^*_p , is for delayed passengers only.

$$-w(y_p, z, \alpha) = \left(\frac{n}{\alpha}\right) y^{*\alpha} = \left(\frac{q}{\alpha}\right) y_p^{*\alpha} + \left[\frac{n-q}{\alpha}\right] z^\alpha \tag{5}$$

A situation of no passenger delay would mean that all passengers in the censored waiting time distribution have a waiting time equal to the headway z (i.e. all passengers can ride on the first train assuming that they arrived at the start of the period). The equally distributed equivalent waiting time is always greater than or equal to z . The resulting delay equity index is then:

$$P = \frac{-w^0 - (-w)}{-w^0} = \frac{y^*}{z} - 1 \tag{6}$$

where $-w^0$ is the social welfare level for a situation of no delay.

The significance of this delay equity index is that it is effectively the ratio between the

delay (difference between the equally distributed equivalent waiting time and headway). This can also be interpreted as the percentage increase in social welfare level (i.e. worsening in social disutility) from case of no passenger delay to the current situation. The delay equity index ranges from zero (i.e. the case when everyone is undelayed; equality) up to infinity. The index satisfies the monotonicity axiom, that is, the reduction of waiting time of a delayed passenger must improve equity. It also satisfies the transfer axiom, that is, a pure transfer of waiting time from a delayed passenger to another passenger with lower delay must improve equity if the difference between their delays are less than in the initial case, *ceteris paribus*.

4.2 Description of Data

The proposed delay equity index is applied to the case of MRT-3, which is the most crowded among the three urban rail lines in Metro Manila. The 16.7-km line runs parallel to EDSA, and handles about 500,000 riders daily, which is beyond its 350,000 capacity. It operates a distance-based staggered fare structure that ranges from PhP10 to PhP15.

Hourly station entry and exit ridership data for MRT-3 on July 7, 2005 (Thursday) was used as the marginal data. The major origins are Stations #1, 2, 4 and 13, while the major destinations are Stations #7, 10, 11 and 13. Moreover, the study used a stated preference survey by DOTC in 2009 for the station O-D estimation for the sample data. It had 1,661 respondents of which 59.8% were male, 52.1% were aged 21-30, 43.5% had a monthly income less than PhP10,000, 69.1% were employees, 53.8% started work or school between 8AM-10AM, and 87.1% did not own a car. One question in the survey regarding their usual station O-D in the morning peak, of which 1,363 passengers responded, was specifically used. Figure 9 shows a comparison on the sampling rate between the sample and marginal data. Pearson's Chi-square test was done to test whether there is a significant difference between the marginal and sampling distributions to test for sampling bias. Results indicate that they are significantly different, thus the sampling rate was adjusted.

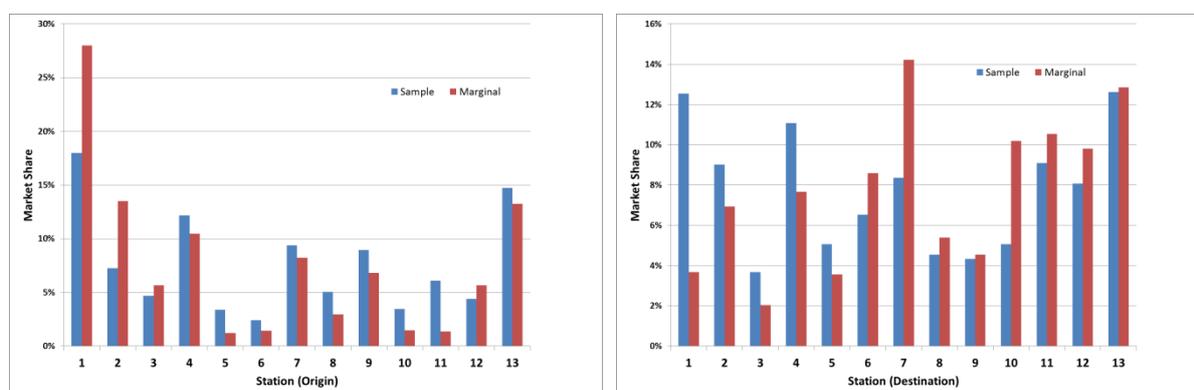


Figure 9. Comparison of Hourly Data (Marginal) and SP Survey (Sample)

Data regarding MRT-3 specifications and operational characteristics (e.g. station and train timetables) and ridership trends was collected from its official website. In the morning, off-peak period headway is 5 minutes, while the peak period headway is 3 minutes. Marginal data was increased according to the annual ridership rate from 2005 to 2011 by simple growth rate method using the total annual ridership trend as it was found that the station annual ridership trend does not vary yearly. It was assumed that passenger delay is due to waiting time only (i.e. no train operations delay), that the crush capacity is 1,302 passengers per train (9 passengers per sqm), and that everyone is willing to wait (i.e. no reneging).

4.3 MRT-3 Station O-D and Waiting Time Estimation

The procedure for the estimation of station O-D and waiting time is shown in Figure 10.

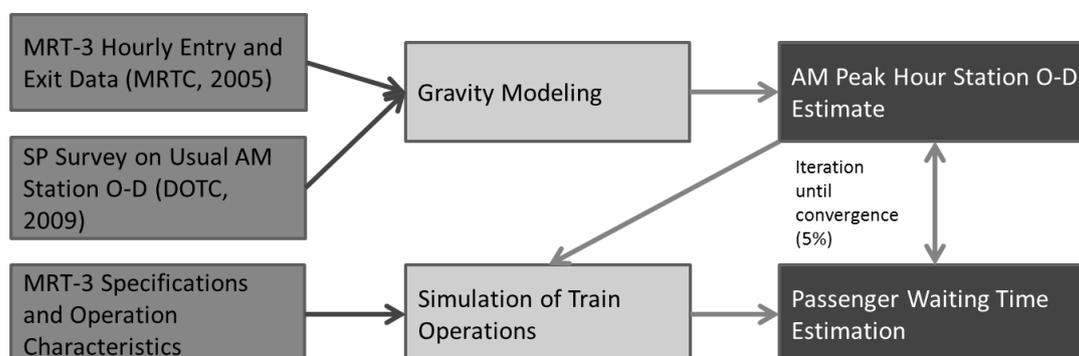


Figure 10. Station O-D and Waiting Time Estimation Procedure

A gravity model with the following form was used for the station O-D estimation:

$$T_{ij} = kT_i^\alpha T_j^\beta e^{-\gamma C_{ij}} \quad (7)$$

where T_{ij} is the number of trips from station i to station j , T_i is the number of trips produced at i , T_j is the number of trips attracted by j , and C_{ij} is the cost matrix.

Two modifications to the classic gravity model were introduced. First is the replacement of zero cells by “1” in the sample data prior to enlargement and model estimation. The purpose of this is to see whether similar model results are obtained with and without replacement. Second is the replacement of the conventional cost matrix that considers the generalized cost for MRT-3 alone by a cost matrix that considers the generalized cost difference between MRT-3 and bus for trips between i and j . T_j is adjusted according to the waiting time results, while T_i is considered constant. C_{ij} is the generalized cost difference between MRT and bus rather than just the generalized cost of MRT.

The model fit was seen to significantly improve with the modified C_{ij} , most likely because public transport users compare the cost of using MRT with that of the bus service that runs parallel to it. In equation form, the modified C_{ij} is:

$$C_{ij} = f_{ij} + \beta_1 v_{ij} + \beta_2 y_{ij} + \beta_3 a_{ij} \quad (8)$$

Where f_{ij} is the fare difference, v_{ij} is the in-vehicle travel time difference, y_{ij} is the waiting time difference, a_{ij} is the access time difference, and β_1 , β_2 and β_3 represent the values of each travel time component. The value of time is 50% of average hourly wage rate (Global Environment Facility, 2001) and the values of waiting and access time were taken from Fillone (2005).

It can be observed from the Figure 11 that the model greatly improves its fit with the replacement of the conventional cost matrix with the modified one, which is the (translated) difference between generalized costs of MRT-3 and bus. In addition, the use of the conventional matrix in estimation involves disregarding sample trips lower than 14 in the estimation in order to obtain the correct parameter signs.

The unique point of this waiting time estimation is its dynamic aspect in space and time, as passenger queues are allowed to accumulate on the platform and priority is always given to passengers that arrive first, and the relationship between waiting time to O-D is considered.

Various waiting time models based on queuing theory were studied, and it was decided that the appropriate way of modeling waiting time is through an $M/D^{[K]}/1/FIFO$ model, that is, a single-server advanced Markovian queuing model with Poisson arrival rate and deterministic service rate with bulk servicing. Following the approach by Petersen *et al* (1995) in modeling airport congestion using an $M/D^{[K]}/c$ model, a time horizon (peak period) is divided into time intervals (headway period), wherein each time interval is characterized by its own arrival rate and service rate. A period in this study is considered as one headway interval ($t = 0$ to $t = h$; where t stands for time and h refers to headway). Bulk service with maximum capacity K is considered at the end of the period. K is equal to the excess capacity at that interval and is variable according to the demand and O-D pattern in the previous stations. Refused passengers are added to the queue of the next time interval. Mode choice is ignored because we only want to describe the waiting time phenomenon and provide a general expression for passenger overload delay in terms of headway, upstream boarding and alighting demand, station demand and train capacity.

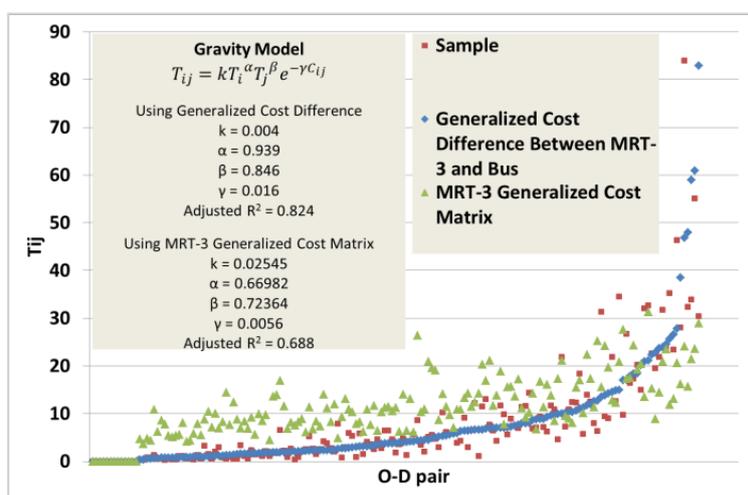


Figure 11. Effect of Varying the Cost Matrix

There are two kinds of queuing model dynamics: spatial and temporal. The spatial model refers to the same train as it travels to all stations, while the temporal model denotes the same station across different time periods. Waiting time for each passenger is estimated one train and one station at a time. If there are r trains within the period and j stations, the model will then be run $r \times j$ times for each direction, with each run dependent on the previous station and time period. The maximum number of passengers that may board the train at any station i at time interval t is

$$B_{it} = \min((Q_{ln} - f_{bit(\cdot)}), f_{wit(\cdot)}) \quad (9)$$

where Q_{ln} is the train capacity, $f_{bit(\cdot)}$ is the user flow inside the train and $f_{wit(\cdot)}$ is the user flow willing to board. The first expression inside the \min function represents K in the $M/D^{[K]}/1$ model, or the available capacity in the train. If the available capacity is less than the user flow willing to board, then there will be refused passengers.

If b_{ij} refers to the number of passengers from station i to j that were able to board during a certain time interval ending at $t = L$ (i.e. L^{th} train), then $f_{b(\cdot)}$ at a certain station I is equal to:

$$f_{b(\cdot)} = \sum_{i=1}^{l-1} \sum_{j=l+1}^N b_{ij} = \sum_{i=1}^{l-1} B_i - \sum_{j=2}^l A_j \quad (10)$$

If B_i is the total number of boarding passengers at station i and A_j is the total number of alighting passengers at station j then $f_{b(.)}$ at station I can also be expressed as:

On the other hand, the user flow willing to board $f_{w(.)}$ at station i during a certain time interval from $t = L - h$ to $t = L$ (or the L^{th} train) is:

$$f_{w(.)} = \mu L - \sum_{t=0}^L B_{it} \quad (11)$$

Where μ is the average passenger arrival rate during the time interval and B_{it} refers to the number of passengers that can board at station i at time interval t .

According to Cascetta (2009), the average waiting time function for passengers in a scheduled service transportation system can be expressed in the equation below, provided that the available capacity is sufficient for the demand at any time interval.

$$TW_{ln} = \frac{\theta}{\varphi_{ln}} \quad (12a)$$

Where θ is equal to 0.5 if the headway is constant ($M/D^{[K]}/I$) and equal to 1 if the headways are distributed according to a negative exponential random variable.

He also specified a function relating the average waiting time to the flow of users staying on board and those waiting to board a single line, such that it would account for the refusal probability that arises from insufficient capacity. The expression is given as:

$$TW_{ln} = \frac{\theta}{\varphi_{ln(.)}} \left(\frac{f_{b(.)} + f_{w(.)}}{Q_{ln}} \right) \quad (12b)$$

Where $\varphi_{ln(.)}$ is the actual available frequency of line ln i.e. the average number of runs of the line for which there are available places, $f_{b(.)}$ is the user flow staying on board, $f_{w(.)}$ is the user flow willing to board, and Q_{ln} is the line capacity. This formula is only applicable for lines with insufficient capacity, that is, $f_{b(.)} + f_{w(.)} > Q_{ln}$. Since perfect service reliability is assumed, the outcome represents the best-case scenario, as rail operation delays would worsen passenger delay due to waiting time.

The model for station O-D and waiting time estimation was run, and after just two iterations, it converged under a convergence criterion of 5%. Two scenarios were considered according to the operation schemes observed in the MRT-3: (a) constant operations: one train every three minutes per direction; and, (b) “skip train” operations: an empty train skips the first two stations every 15 minutes. It is an actual countermeasure that is sometimes employed by MRT-3 to allow the accumulated passengers in the middle stations to ride the train during morning peak hours.

Figure 12 shows that the delay is concentrated on the fourth station only, with the maximum waiting time being around 31 minutes. In contrast, all other stations do not experience waiting times higher than the headway. This means that even with perfect train service reliability, passengers in Station 4 still experience passenger overload delay due to capacity constraints. It should be noted that other sources of delay such as late arrival of trains may worsen passenger waiting times and spread it to other stations.

The situation improves by employing the “skip train” operations countermeasure by spreading the delay to other stations and decreasing the maximum waiting time to around 17 minutes. Efforts to spread the delay by employing “skip train” operations results in lower delay for the third and fourth stations, but causes those at the first, second and third stations to experience delay as well.

These results are consistent with actual observations at the MRT-3 CCTV live streaming website (Metrostar Express) during the morning peak period (7-9AM) wherein many passengers in the first five stations heading towards the southbound direction were observed to wait for several trains before being able to board. The waiting time was observed to be most severe for the fourth station, which is similar to the results in this study.

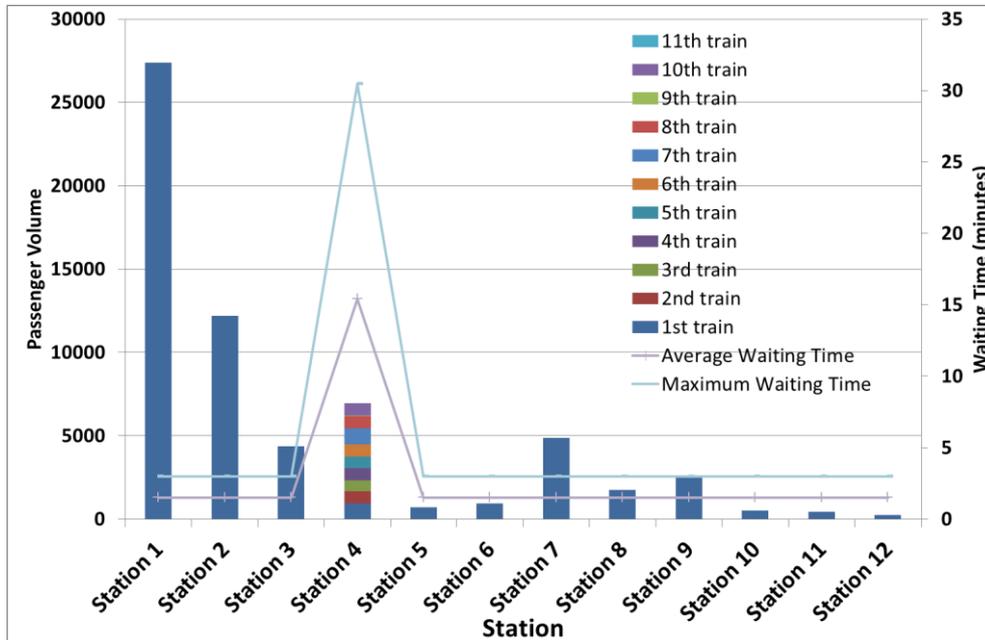


Figure 12. Waiting Time for Constant Operations

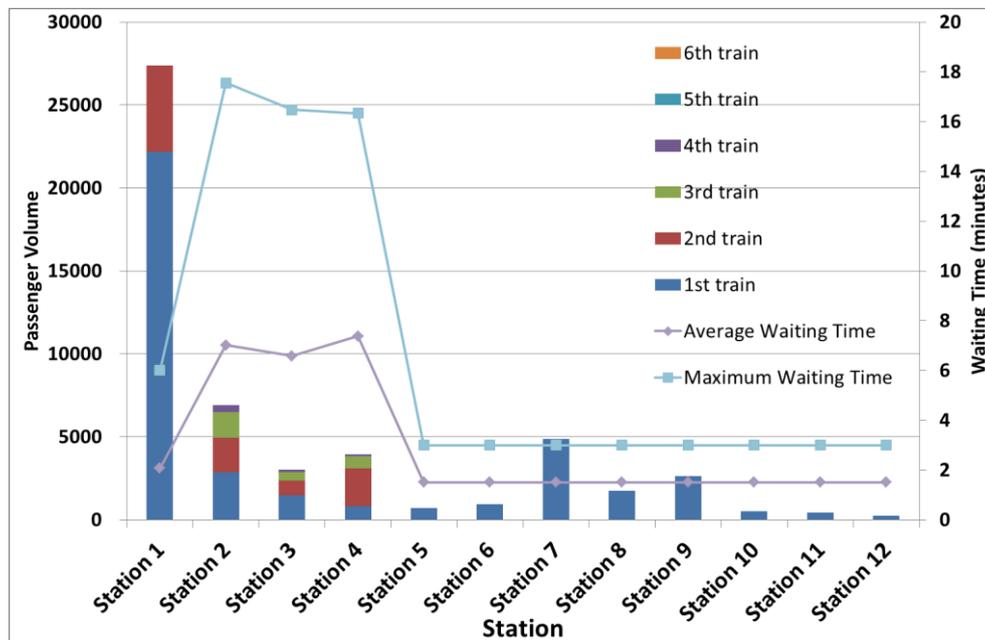


Figure 13. Waiting Time for "Skip Train" Operations

4.4 Passenger Overload Delay Equity Analysis

Several parameters, including total delay, maximum delay, head count ratio, Gini coefficient,

social welfare level and the proposed passenger overload delay equity index, were employed to assess the equity of the distribution of delay among passengers for the morning peak period in the southbound direction.

It can be observed from Table 3 that the Gini coefficient indicates an improvement of delay equity from constant operations to the “skip train” operations. It can be seen that according to Gini coefficient as well as for the proposed delay equity index, the constant operations scenario is less equitable than “skip train” operations. The same result is seen for the proposed delay equity index. However, when more weight is given for people with higher delay (as α increases), equity and social welfare levels are seen to worsen.

However, the total delay and number of delayed passengers are seen to increase, implying that there is a trade-off between equity and efficiency (i.e. minimization of total system delay).

These results indicate that the existing operation strategies used in the MRT-3 are not enough in addressing passenger overload delay even with perfect service reliability assumption, indicating that it is the capacity constraint that is causing delay equity among passengers. The MRT-3 employs a scheduled skip train operations as well as crowd control procedures whenever necessary yet excessive waiting time is still observed (Metrostar Express). With delay at this level, it is possible that passengers at the stations that experience delay would be deprived of the opportunity to ride the rail, further aggravating equity problems.

Table 3. Delay Equity in Metro Manila MRT-3

Parameter		Constant operations	“Skip train” operations
Total delay (minutes)		200004.32	405650.14
Average delay (minutes)		3.14	6.38
Maximum delay (minutes)		30.47	17.55
Standard deviation		6.59	4.05
Coefficient of variation		2.10	0.64
Gini coefficient		0.61	0.50
Proposed Delay Equity Index	$(\alpha=1)$	0.50	0.56
	$(\alpha=1.5)$	1.45	1.25
	$(\alpha=2)$	4.03	2.56
	$(\alpha=3)$	31.31	10.15
Social welfare level	$(\alpha=1)$	-2.86E+05	-2.97E+05
	$(\alpha=1.5)$	-5.39E+05	-4.95E+05
	$(\alpha=2)$	-1.44E+06	-1.02E+06
	$(\alpha=3)$	-1.85E+07	-6.38E+06
Delayed passengers		6,581	24,246
Undelayed passengers		57,040	39,375
Head count ratio		0.10	0.38

5. CONCLUSIONS AND FUTURE WORK

This study examined two equity issues: urban rail fare policy, and passenger overload delay.

First, an original framework for equity analysis of rail fare policies was developed by considering equity aspects, equity types and fare policy components, and use of inequality indices. Distance-based equity is dependent on the city size (range of commuter trip length) as

well as the type of distribution and its parameters. There is a trade-off between distance-based equity and fare affordability, and weakly-differentiated fare structures strike a balance to address both equity objectives. A ranking of fare policies according to overall equity was made, but this result is limited to the hypothetical assumptions on trip length and frequency, and the definition of equity. To make these results more meaningful, the actual trip length and frequency distributions for each city should be considered in the future.

This study has also shown that an $M/D^{(K)}/I$ queuing model can be used to estimate passenger overload delay due to capacity constraints. This model largely depends on the O-D demand, available capacity, and headway for each interval. Moreover, a delay equity index that is based on the theory of relative deprivation, rather than common inequality indices such as Gini, is seen as more appropriate in describing the equity level of passenger overload delay for all O-D pairs. Considering the censored waiting time distribution rather than just passenger overload delay would let passengers compare their deprivation functions among all passengers, regardless of whether they are delayed or not.

For MRT-3 station O-D estimation, the fit greatly improves by using a modified cost matrix that considers the generalized cost difference between MRT-3 and bus, rather than a conventional matrix.

Intra-modal equity analysis shows that passenger overload delay is concentrated on just one station, indicating inequity. The employment of “skip train” operations as a countermeasure improves equity and social welfare level, but passenger overload delay is still substantial. This indicates that capacity of MRT-3 is simply not adequate, even with perfect train service reliability and implementation of a countermeasure. The relationship of this phenomenon with public transport fare levels and private transport costs will be the subject of future study. Another direction that the researchers intend to take is the inclusion of the proposed delay equity index as well as other transport modes in the fare optimization problem. It would also be interesting to incorporate the equity-efficiency and horizontal-vertical equity trade-offs in designing fare policies.

REFERENCES

- Avi-Itzhak, B., Levy, H. (2004). On the Fairness of Queues. *Advanced Applied Probability*, 36, 919-936.
- Bell, M.H. (1995). Stochastic User Equilibrium Assignment in Networks with Queues. *Transportation Research Board B*, 29B (2), 125-137.
- Carruthers, R., Dick, M. Saurkar, A. (2005). Affordability of Public Transport in Developing Countries. The World Bank Transport Sector Group TP-3. Retrieved from: <http://www.worldbank.org/transport/>
- Cascetta, E. (2009). *Transportation Systems Analysis: Models and Applications*. 2nd ed. New York, NY: Springer.
- Cervero, R. (1981). Flat vs. Differentiated Transit Pricing: What's a Fair Fare? Department of City and Regional Planning, University of California, Berkeley. *Transportation*, 10 211-232
- Cervero, R. (1990). Transit Pricing Research: A Review and Synthesis. *Transportation*, 17:117-139.
- Chicago Transit Authority. (n.d.). Retrieved from: <http://www.transitchicago.com/upass/>
- Clark, S., Hemming, R., Ulph, D. (1981). On the Indices for the Measurement of Poverty. *The Economic Journal.*, 91 (362), 515-526.
- Companhia do Metropolitano do Distrito Federal – Metrô. (n.d.). Retrieved from: <http://www.metro.df.gov.br/estacoes/tarifas.html>.

Cowell, F. (2009). *Measuring Inequality*. New York, NY: Oxford University Press. Retrieved from: http://darp.lse.ac.uk/papersDB/Cowell_measuringinequality3.pdf

De Cea, J., Fernandez, E/ (1993) Transit Assignment for congested public transport systems: An equilibrium model. *Transportation Science*, 27, 133-1147.

DOTC-Metrostar Express (2009). MRT 3 Passenger Stated Preference Survey Report. Part of the Pre-feasibility Study for MRT 3 West Extension to Bay City.

Fillone, A. (2005). *Discrete Choice Modeling of Work Trips in Metro Manila and Urban Transport Policy Application*. PhD Dissertation. University of the Philippines School of Urban and Regional Planning.

Goodwin, C., Smith, K., Verhage, B. (1991). An Equity Model of Consumer Responses to Waiting Time. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 129-138.

Global Environment Facility (2001). Project Appraisal Document for the Philippines (MMURTRIP), Report No: 20767-PH.

Lam, W.H.K., Gao, Z.Y., Chan, K.S., Yang, H. (1999). A stochastic user equilibrium assignment model for congested transit networks. *Transportation Research Part B*, 33, 351-368.

Leutze, C.B.; Ugolik, W.R. (1978). *Who Pays the Highest and the Lowest Per-Mile Transit Fares?* Albany, NY. New York State Department of Transportation

Levinson, D. (2002). Identifying Winners and Losers in Transportation. *Transportation Research Record*, 1812, 179-185.

Litman, T. (2011). *Evaluating Transportation Equity: Guidance for Incorporating Distributional Impacts in Transportation Planning*. Victoria Transport Planning Institute. Retrieved from: <http://www.vtpi.org/equity.pdf>

Martinelli, D., Medellin, L. (2007). *Assessment of Bus Transit Equity in Two Metropolitan Areas*. Mid-Atlantic Universities Transportation Center. West Virginia University.

Metrostar Express. (n.d.). Retrieved from <http://dotcmrt3.gov.ph/>

OECD, European Commission, Joint Research Centre (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*, by Nardo, M. M. Saisana, A. Saltelli and S. Tarantola (EC/JRC), A. Hoffman and E. Giovannini (OECD), OECD publication Code: 302008251E1.

Peterson M., Bertsimas D., Odoni A. (1995). Models and Algorithms for Transient Queuing Congestion at Airports. *Management Science*, 41 (8), 1279 – 1295.

Proost, S., Van Dender, K. (2008). Optimal Urban Transport Pricing in the Presence of Congestion, Economies of Density and Costly Public Funds. *Transportation Research A*, 42, 1220-1230.

Pucher, J. (1981). Equity in transit finance: Distribution of transit subsidy benefits and costs among income classes. *American Planning Association Journal*, 47, 387-407.

Pucher, J. (1983). Who Benefits from Transit Subsidies? Recent Evidence from Six Metropolitan Areas. *Transportation Research A*, 17A (1), January, 39-50

Shimamoto, H., Iida, Y., Kurauchi, F., Schmocker, J., Bell, M.G.H. (2005). Evaluating Public Transit Congestion Mitigation Measures Using a Passenger Assignment Model. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 2076-2091.

Szeto, W.Y., Jiang, Y., Wong, K.I., Solayappan, M. (2011). Reliability-based stochastic transit assignment with capacity constraints: Formulation and solution method. *Transportation Research Part C* (Article in Press).

Wahba, M., Shalaby, A. (2005). Multiagent Learning-Based Approach to Transit Assignment Problem. *Transportation Research Record*, 1926, 96–105.

Zhang, Y., Lam, W.H.K., Sumalee, A., Lo, H.K., Tong, C.O. (2010). The multi-class

schedule-based transit assignment model under network uncertainties. *Public Transportation*, 2, 69-86.