Improving Bus Service Reliability: The Singapore Experience

Waiyan Leong[a*], Karen Goh[b], Stephane Hess[c], Paul Murphy[d]
[a] *Land Transport Authority, No. 1 Hampshire Road Singapore 219428.*
[b] *Ministry of Communications and Information, 140 Hill Street, Singapore 179369.*
[c] *Institute of Transport Studies & Choice Modelling Centre, University of Leeds, 34-40 University Road, Leeds LS2 9JT, UK.*
[d] *AECOM, AECOM House, 179 Moss Lane, Altrincham, WA15 8FH, UK.*

* *Corresponding author: Land Transport Authority, No. 1 Hampshire Road Singapore 219428*
Phone: +65 6295 7726. Email: leong_wai_yan@lta.gov.sg

Keywords:
Bus Service Reliability, Singapore, Excess Wait Time, Land Transport Authority, Bus Contracting Model.

Classification codes:
R480, L920

ABSTRACT

In February 2014, Singapore embarked on a 2-year trial of a Bus Service Reliability Framework (BSRF) to improve en-route bus regularity and reduce instances of bus bunching and prolonged waiting times. Based on London's Quality Incentive Contract, the Singapore model also imposes penalties or provides incentives to operators for increases/reductions of Excess Wait Time (EWT) beyond a certain route-specific baseline.

Drawing on insights derived from research on performance-based contracts, this paper describes some key considerations surrounding this particular innovation in Singapore's overall bus regulatory framework. We also discuss an important advancement in our understanding of how bus users value reliability improvements through estimates obtained from stated preference data. At the same time, early indications from the trial have been encouraging.

1. Introduction

For many public transport users, service reliability is a key attribute of the travel experience. The importance of reliability is amply demonstrated by the multitude of papers concluding that variability in travel time impacts well-being more negatively than the actual journey time itself (see review by Carrion & Levinson, 2012). With regards to the bus in particular, an unreliable bus service, characterised by unequal headways or bus bunching for high frequency services, can lead to longer waiting time and travel time for bus passengers. Moreover, in cases where passenger loading on a particular bus route is already high, unequal arrival times can mean severe crowding on the first bus that arrives after a long headway. The generally unpleasant in-vehicle experience adds another layer of frustration to passengers who have already endured a longer than expected wait, if they are not denied boarding in the first place. Unreliability begets further unreliability as dwell times increase at bus stops to cater for the higher passenger movements on and off the bus.

This paper describes Singapore's experience with improving bus service reliability. Section 2 provides a review of reliability measures and how reliability is achieved in various jurisdictions worldwide. Section 3 briefly describes Singapore's bus industry before discussing Singapore's trial of its Bus Service Reliability Framework (BSRF). Section 4 is largely an empirical section which discusses commuter awareness of the

BSRF, stated preference strategies to measure improvement in reliability and crowding, and outcomes to date. Section 5 concludes.

2. Literature Review

2.1 Overview of Reliability Measures Used Worldwide

Recognising the centrality of this aspect of service quality in passenger experience, industry regulators around the world have introduced various service reliability frameworks in their performance monitoring regimes. TriMet in Portland, Oregon, uses the Bus Dispatching System (BDS) to monitor public transport reliability (Feng & Figliozzi, 2012). The BDS combines Automatic Vehicle Location and Automatic Passenger Counters data to provide detailed information on bus service performance. The two performance measures are headway deviation and actual headway spatial distribution. Headway deviation looks at the difference between actual headway and scheduled headway. Actual headway spatial distribution depicts the proportion of actual headways deviating from scheduled headway against different stops along the route. Spikes and dips on the distribution would suggest congestion or chokepoints that require improvements. Strathman et al. (2000) report that the BDS has decreased bus service running time by 3 percent after implementation.

In Shanghai and Jiangyin City, a normalised average headway index is used to determine the actual headway deviation from the scheduled headway (Guo, Luo, Lin & Feng, 2011). An index below 100% indicates that the bus is earlier than scheduled, while an index above 100% indicates that the bus is behind schedule (Eq. (1)).

$$\overline{H_i} = \frac{\sum_{2}^{n} H_{\delta i}/H_{\delta}}{n-1} \times 100\% \qquad (1)$$

where: $\overline{H_i}$ = average headway of normalised *i*-th bus stop;
H_δ = the departure interval at departure station of two specific adjacent bus deployments;
$H_{\delta i}$ = headway of the two specific adjacent bus deployments at *i*-th bus stop;
n = number of bus deployments.

In Changzhou’s Bus Rapid Transit, four measures are used as indicators of reliability (Huo, Zhao, Li, & Hu, 2014). One statistic used is the coefficient of variation of headway, which is the standard deviation of headway divided by its mean. It indicates service reliability from the operator’s perspective. On the other hand, potential waiting time, equivalent waiting time and reliability buffer time indicate service reliability from the users’ perspective. Potential waiting time refers to the difference between 95^{th} percentile waiting time and mean waiting time. Equivalent waiting time is the weighted sum of mean and potential waiting time. Reliability buffer time is the extra time that commuters need to provision beyond typical journey time to ensure on time arrival at destination with 95% probability.

In Chicago, the Automatic Vehicle Location data is used to determine running time adherence and headway regularity (Lin, Wang & Barnum, 2008). Running time adherence measures the average difference between actual and scheduled run times, while headway regularity measures the average difference between actual and scheduled headways. A high metric value for these two indicators will indicate irregular bus services and poor reliability (Eq. (2)).

$$Running\ Time\ Adherence = \frac{\sum_m \left|\frac{Actual\ Run\ Time - Scheduled\ Run\ Time}{Scheduled\ Run\ Time}\right|}{m} \times 100\%$$

$$Headway\ Regularity = \frac{\sum_n \left|\frac{Actual\ Headway - Scheduled\ Headway}{Scheduled\ Headway}\right|}{n} \times 100\% \qquad (2)$$

In Sydney, Transport for New South Wales uses Key Performance Indicators to monitor bus services performance. It measures punctuality of buses at the commencement of trip, mid-point of trip and at the last transit stop, requiring at least 95% of the trips to be between 2 min early and 6 min late.

Transport for London (TfL) characterises London bus services depending on whether they are high or low frequency. The reliability of high frequency services, defined as those with headways of less than 15 minutes, is assessed based on average excess wait time (EWT) experienced by commuters. Unreliable bus services, as evidenced by irregular spacing of buses, will result in high EWT. On the other hand, low frequency services are assessed on percentage of buses departing on time according to bus schedules.

2.2 Achieving Reliability through Performance Based Contracts

How to meaningfully measure bus service reliability is one, but certainly not the only consideration that regulators need to address. Another important question involves mechanism design – how one might achieve even better bus service reliability performance. Fortunately, on this latter issue, the existing literature offers substantial guidance, particularly through the use of performance-based contracts (PBCs). Based on the extensive research, PBCs are now used across manufacturing and service industries, in public and private domains (Selviaridis and Wynstra, 2014).

In the realm of bus service provision in particular, Hensher and Houghton (2002) proposed a system that takes into account various external costs such as costs of congestion, among many others, with social surplus maximisation as the underlying motivation in order to ensure that bus operators deliver the optimal service level that is consistent with the needs of stakeholders, especially the government. Working along similar lines, Hensher and Stanley (2003) highlighted the importance of PBCs as a crucial factor that aligns commercial objectives with social objectives by rewarding operators for achieving a minimum level of service (MSL) and for an increase in ridership. Selviaridis and Wynstra (2014) also highlighted a form of PBCs where negative or positive incentives are given although there can also be ‘dead zones’ for acceptable performance levels for which there is neither penalty nor extra rewards.

A notable example of an implementation of PBCs in the provision of public bus service is the Hordaland framework (Larsen, 2001). In the Hordaland framework, social benefits such as reduced waiting time, reduced number of transfers and transfers of riders from car to public transport is internalised into the operators' remuneration contracts based on revenue kilometres and passengers The framework attempts to induce operators to deliver the socially optimal level of services through performance-based subsidies as part of the total payment per passenger received by operators (Hensher and Stanley, 2003).

Numerous studies have also shown that reliability remains a crucial component of bus service quality. dell'Olio, Ibeas and Cecín (2010) found that bus reliability is one of the most important attribute alongside waiting time. They highlighted the need for transport companies to place greater emphasis on punctuality and headways between buses. Not only is bus reliability an important factor in determining people's behaviour choices (Disney, 1999; König & Axhausen, 2002), it is also found to be a main reason behind users' dissatisfaction towards public bus services (Edvardsson, 1998). In addition, bus reliability may be of a decisive factor in determining people's tolerance level in other areas such as comfort and ventilation (Disney, 1999). Similarly, a key finding uncovered by Bates, Polak, Jones, and Cook (2001) shows that as travellers, users highly value punctuality. In order to improve service regularity, Cats (2014) highlighted the need to incorporate reliability measures such as regularity indicators into an incentive scheme.

The incorporation of bus service reliability into a PBC regime for operators has been widely implemented around the world, such as in New Zealand and in London (Transport for London, 2015; Ian Wallis Associates Ltd & The TAS Partnership, 2013; Vincent, 2008). In London's case, an incentive mechanism embedded in its Quality Incentive Contract is used to encourage bus operators to provide reliable service, with a bonus of up to 15 percent or deduction of up to 10 percent of the contract price relative to the required standards. These standards are known as the Minimum Performance Standards (MPS) and are a crucial part of London's scheme as they act as a reliability benchmark for operators. Operators are not paid for any mileage not operated for reasons within the operator's control, such as staff shortages or mechanical issues. Setting MPS properly is a very important part of the contracting process, as failure to could so could mean that bus operators would pass the risk of poor performance to the public authority by putting the anticipated penalty in the contract bid.

Since the implementation of the Quality Incentive Contracts in 2001, EWT of high frequency services has fallen from 2.2 minutes to 1.0 minutes and the percentage of on-time low frequency services has risen from 68 percent to 83 percent in 2014 (Transport for London, 2014). With greater service reliability, passengers are also able to time their arrival time to coincide with bus arrival time to reduce waiting time. TfL reports that service improvements are estimated to have accounted for some 30 percent of the growth in demand for bus services from 1997 to 2012 (Transport for London, 2014).

3. The Bus Service Reliability Framework (BSRF) in Singapore

3.1 Overview of Singapore's Bus Industry

Bus services in Singapore are primarily provided by two major bus operators, SBS Transit (SBST) and SMRT Corporation (SMRT). As of June 2015, both companies operated over 350 bus services, with an average total daily vehicle-kilometres travelled of 856,600 bus-km. Each bus operator has been assigned separate areas of responsibility that correspond to satellite residential towns developed by government planning agencies. They operate a mixture of short intra-town routes (called “feeder services”) and longer routes (called “trunk routes”). Having been awarded the first bus package of public bus services under the new Bus Contracting Model (see Goh, Swee, & Low, 2015), Tower Transit became Singapore’s third major bus operator as of May 2016. SBS Transit operates 247 bus routes, with a total route network of 5054 km, while SMRT Corporation operates 111 bus routes with a total route network of 2326 km.

Before it was required to undertake a new advisory role on public transport matters, the Public Transport Council (PTC) oversaw regulation on, among other things, bus services and bus service operators. As part of its regulatory framework, the PTC established Quality of Service (QoS) standards, in particular Operating Performance Standards which measure minimum daily or monthly operational deliverables, covering aspects of bus reliability, loading and safety. On bus reliability standards, the PTC required operators to ensure that, on a daily basis, each service had at least 85% of its trips depart the bus interchanges and terminals not more than 5 min from its scheduled headway.

3.2 Overview of the Bus Service Reliability Framework (BSRF)

It has often been pointed out that the PTC’s regulatory stance failed to sufficiently address commuters’ concerns about reliability as the only key performance indicator measures reliability at origin and not **en-route** regularity. In 2014, the Land Transport Authority (LTA) decided to make en-route reliability an explicit service requirement by introducing the Bus Service Reliability Framework (BSRF) initially as a two-year trial for the two major bus operators (Land Transport Authority, 2014b, 2014a). In the first phase of the trial described in this paper, 22 bus services are covered, comprising a mixture of trunk and feeder services. The intention of the trial is to improve en-route bus regularity, reduce instances of bus bunching and reduce prolonged waiting times for bus users.

Although the BSRF trial only covers a modest number of bus services, nevertheless, the LTA hopes that the trial will provide a better understanding of bus reliability improvements that could be made on different types of routes, as well as the BSRF’s effectiveness to get bus operators to improve service reliability. A successful roll-out of the BSRF would complement other regulatory measures such as more strictly enforcing bus lanes and enhancing bus priority schemes and measures. With the Bus Contracting Model that is being introduced, the lessons learnt from the trial will also be valuable to LTA as it is now responsible for setting regulatory standards for the quality of bus services in the bus packages that are tendered out.

Like London, Singapore also categorises bus services by high and low frequency. Low frequency services are those with the majority of headways being more than 15 minutes where punctuality of bus arrivals is more important. High frequency services are buses that arrive at frequencies of 15 minutes or less where commuters can “turn up and go” and generally do not refer to the timetable. The vast majority of basic bus services in Singapore are ‘high frequency’ services.

Following the TfL model, the BSRF assesses the regularity of a bus service using the concept of Excess Wait Time (EWT). EWT is the average additional wait time actually experienced by commuters at bus stops, compared to the expected wait time if the buses arrived at regular intervals. It is defined as the difference between the Actual Wait Time (AWT) and the Scheduled Wait Time (SWT), that is, EWT = AWT – SWT, with AWT and SWT defined in Eq. (3).

$$AWT = \frac{\sum_n actual\ \ headway_n^2}{2 \times \sum_n actual\ \ headway_n}\ ;\ SWT = \frac{\sum_n scheduled\ \ headway_n^2}{2 \times \sum_n scheduled\ \ headway_n} \qquad (3)$$

The EWT methodology assumes uniform arrival rate of passengers and increases if there is bus bunching which results in prolonged waits for the subsequent bus. Conversely, it goes down if bus arrivals at each bus stop become more regular. As a result, commuters would experience greater ease in boarding as the passenger load is spread more evenly across the various bus trips. If a bus service arrives perfectly regularly, the EWT will be 0 min.

The EWT explicitly quantifies the commuting experience using the excess time experienced by commuters due to variability of the bus services. This allows useful and direct comparisons across different bus routes and bus stops along the line. The EWT is an indicator of reliability from the commuter's perspective, rather than the supply side perspective such as the operator's ability to dispatch buses on time (Oort, 2014). Moreover, the EWT can be used a consistent indicator across different countries even if there are different definitions and range of standards for punctuality.

In Singapore, for the trial, the EWT is measured during peak and off-peak hours from Mondays to Fridays, excluding Public Holidays. The AM peak is defined as actual arrivals between 6.30am and 8.29am; AM off-peak as between 8.30am and 4.59pm; PM peak as between 5.00pm and 6.59pm and PM off peak as between 7.00pm and 10.59pm. EWT is also measured across all the trips for a single bus service, and at several critical bus stops or "intermediate timing points" (ITPs). The number of ITPs varies by route length, with more ITPs for longer routes. EWT are also weighted heavier for the peak periods compared to the off-peak. Finally, EWT are averaged across the calendar month by direction before summation across both directions to obtain the overall EWT for the service for the month. 22 bus services were chosen for the BSRF trial because of public feedback of poor reliability. These include a mix of long and short trunk services as well as feeder services[1].

Each bus service will have its own existing "baseline" EWT depending on the current performance and the characteristics of the route. Typically, a long trunk route will have a higher existing baseline EWT. The operators' performance is benchmarked to historical performance in order to ensure the reasonableness of the standards that operators need to comply with. A holistic measurement that assesses individual bus services as a whole over the entire month, and not individual bus trips at specific times of the day, is used in order to ensure that improvements in reliability are in fact sustained over the longer term. Table 1 lists the baseline EWT scores for the bus services covered in the initial rollout of the BSRF.

Although the baseline EWT benchmarks are conceptually similar to the London MPS, unlike London, operators in Singapore are not allowed to curtail service, such as

[1] Seven SMRT bus services - 176, 184, 188, 302, 858, 901 and 911 - were placed on the BSRF from February 2014. SBS Transit Services 17, 52, 228 and 242 were placed on the BSRF from 28 February 2014 and Services 3, 39, 241 and 325 from 24 March 2014. From 23 June 2014, another seven bus services - SBS Transit Services 51, 154, 292, and 354, as well as SMRT Services 189, 853 and 962 - were implemented under BSRF.

short turning without penalty when there is traffic. As the operating environment is different in Singapore, such curtailments are not expected to be necessary and operators are aware of this requirement. Under the Bus Contracting Model, the government specifies frequency requirements for the route when calling for a tender. Bidders are to submit their optimised required number of vehicles for operation as part of their tender proposal.

Table 1: Baseline EWT scores

Fleet size category	Operator	Bus Service	Trunk/ Feeder	Length of Direction 1 (km)	Length of Direction 2 (km)	EWT Baseline (min)
Category 3 (≥20 buses)	SBST	3	Trunk	19.5	20.3	1.5
		39	Trunk	26.2	26.4	1.2
		51	Trunk	37.7	36.3	2.2
		154	Trunk	32.5	34.4	1.9
	SMRT	176	Trunk	23.5	24.4	1.6
		188	Trunk	21.6	22.1	1.4
		858[a]	Trunk	73.4	0	2.1
Category 2 (10 to <20 buses)	SBST	17[a]	Trunk	25	0	1.8
		52	Trunk	25.9	25.2	2.0
	SMRT	184[a]	Trunk	22.4	0	1.3
		189[a]	Trunk	19.8	0	1.4
		302	Feeder	7.5	0	1.0
		853	Trunk	31.1	32.3	1.5
		911	Feeder	11.8	0	1.3
		962[a]	Trunk	16.9	0	1.4
Category 1 (< 10 buses)	SBST	228	Feeder	10.7	0	1.1
		241	Feeder	6.2	0	1.2
		242	Feeder	5.7	0	0.8
		292	Feeder	5.9	0	0.9
		325	Feeder	12.3	0	0.8
		354	Feeder	4.5	0	1.1
	SMRT	901	Feeder	11.8	0	0.9

[a] These trunk services are loop services and therefore have no Direction 2.

3.3 Trial of BSRF Incentive-Penalty Framework

It is widely acknowledged that improving reliability or reducing EWT is operationally challenging. Bus operators will have to put in additional resources, such as hiring more service controllers to manage bus services and having standby buses to inject mid-route if there are delays to buses which are already en-route. The operators may also choose to deploy another bus if the designated bus that is scheduled to turnaround has been delayed on the preceding trip due to congestion. By running such outside-schedule trips to maintain headway, the initiative taken by the operators is rewarded with a better EWT score.

Hence, should significant improvements in EWT be made, incentives are provided to allow operators to recoup costs. These incentives are calibrated in accordance with the efforts and operational costs involved in improving the reliability of the services. The incentive-penalty framework of the BSRF is modelled after London’s experience. Operators are rewarded only when they achieve improvements in bus service regularity, and are penalised if the service is not so.

Incentives and penalties are determined based on 6-month average performance of each service. No incentives or penalties apply in a neutral zone of ± 0.1 minutes on either side of the route-specific baseline EWT. Outside this neutral zone, the monthly incentive and penalty apply for every 0.1 minute improvement or deterioration in EWT score when compared to the baseline (Table 2). The ratio of incentives to penalties is approximately in the order of 3:2 (S$1 = US$0.75).

Table 2: BSRF incentive and penalty amounts

Fleet Size Category[a]	**1 (< 10 buses)**	**2 (10 to <20 buses)**	**3 (≥20 buses)**
Incentive Amount per 0.1 minute improvement per month	S$ 2000 (US$1500)	S$ 4000 (US$3000)	S$ 6000 (US$4500)
Penalty Amount per 0.1 minute deterioration per month	S$ 1300 (US$975)	S$ 2600 (US$1950)	S$ 4000 (US$3000)

[a] This is the scheduled fleet for each bus service

To allow bus operators to adjust to the new framework, for the trial, LTA granted both operators a transition period from 3 February until 31 May 2014 when EWTs will be monitored but no incentives or penalties will be applied. The transition period was designed to give the bus operators more time to train their bus drivers and service controllers and to fine tune their operational procedures to regulate bus arrival times.

4. Evidence and Outcomes

4.1 Stated Preference Surveys

To provide some evidence on how bus users in Singapore value different aspects of the bus service they use, a Stated Preference (SP) component focussing specifically on bus waiting times was included in the wider SP study conducted in Singapore in 2015 to elicit a number of economic measures relating to travel. The bus waiting time choice scenarios explicitly look at how bus users value patterns of waiting time at the bus stop, and hence EWT improvements and enhancements to bus service reliability.

As part of the SP study, a representative cross section of the Singapore population were asked to record the trips they had made in the past 2 days and were then asked about one of these journeys in detail. They were asked to record the times and costs involved in making their journey and for motorised modes they were asked to provide similar information about their alternative mode assuming they could not make the journey by their current mode. This information was then used as the basis of a number of different sets of SP scenarios, or games. Most of the SP games were within mode so respondents were presented with two options relating to their chosen mode. However some respondents were presented with a mode choice game where they chose between the different travel modes, such as car, bus, MRT, and taxi.

A final sample of 791 bus users received the Bus Waiting Time game. Below we provide the context behind this game, an example of how the options were presented plus a description of the SP design issues relating to each design. These are D-efficient designs and have been developed using Ngene. The SP presented each respondent with seven separate unlabelled binary choice tasks. Respondents were asked which of the two hypothetical options they would choose on the basis of the information presented.

In the Bus Waiting Time SP, respondents were asked to think about the journey they had provided information about but imagining a situation where they had the choice between two future hypothetical bus services. The scheduled arrival pattern of the bus was shown as a reference in all games (always arriving every 10 min, depicting the scenario of perfect reliability) and in each game, two sets of experimentally generated intervals between buses were shown as hypothetical alternatives A and B. The bus fare they would have to pay was also shown. It was assumed that buses arrived frequently enough that they “forget the timetable”; in other words, respondents were told to assume that their arrival time at the bus stop was completely arbitrary. Everything else about their journey remained unchanged. The bus fare was pivoted around the current single-trip fare reported by the respondent, with pivot levels of – 30%, – 15%, 0%, +15%, +30%.

Several SP studies (see Li, Hensher, & Rose, 2010 for a review) have attempted to estimate the value of service reliability using presentations that rely on probabilities of early or late arrivals of a bus (relative to a schedule) as key attributes of the alternatives. In this work however, because most bus services in Singapore run at headways of less than 15 min and therefore do not have a posted schedule at the bus stop, it was not very informative to describe a bus service as experiencing schedule delays or of being “early”, “on-time” or “late”. Instead, we decided to opt for a pictorial representation of reliability based on arrival intervals, such as in Fig. 1, which is also consistent with how LTA has been communicating the BSRF in public (LTA, 2014a). Moreover, the design of the Bus Waiting Time game also allowed for the underlying EWT associated with each of the options to be easily calculated and hence, the estimation of WTP measures for EWT improvements, which was a key objective of the study.

For the service reliability attribute, respondents were asked to consider bus arrival patterns over a time interval of 60 min. The scheduled arrival times, which depict the perfect scenario, do not vary across choice tasks and were based on a scheduled headway of 10 min, for a frequency of six buses per hour corresponding to six time intervals between arrivals. A scheduled headway of 10 min was deemed a reasonable assumption for the choice experiment as under the PTC QoS standards effective August 2009, operators are required to provide at least 80% of bus services at frequencies of not more than 10 minutes during weekday peak periods.

For the hypothetical scenarios A and B, the design contains six interval attributes, each between 4 and 16 min and with a condition that the total headways sum to 60 min in both options. The EWT in the two hypothetical alternatives ranges from 0.05 min to 1.35 min, with an average at 0.66 min.

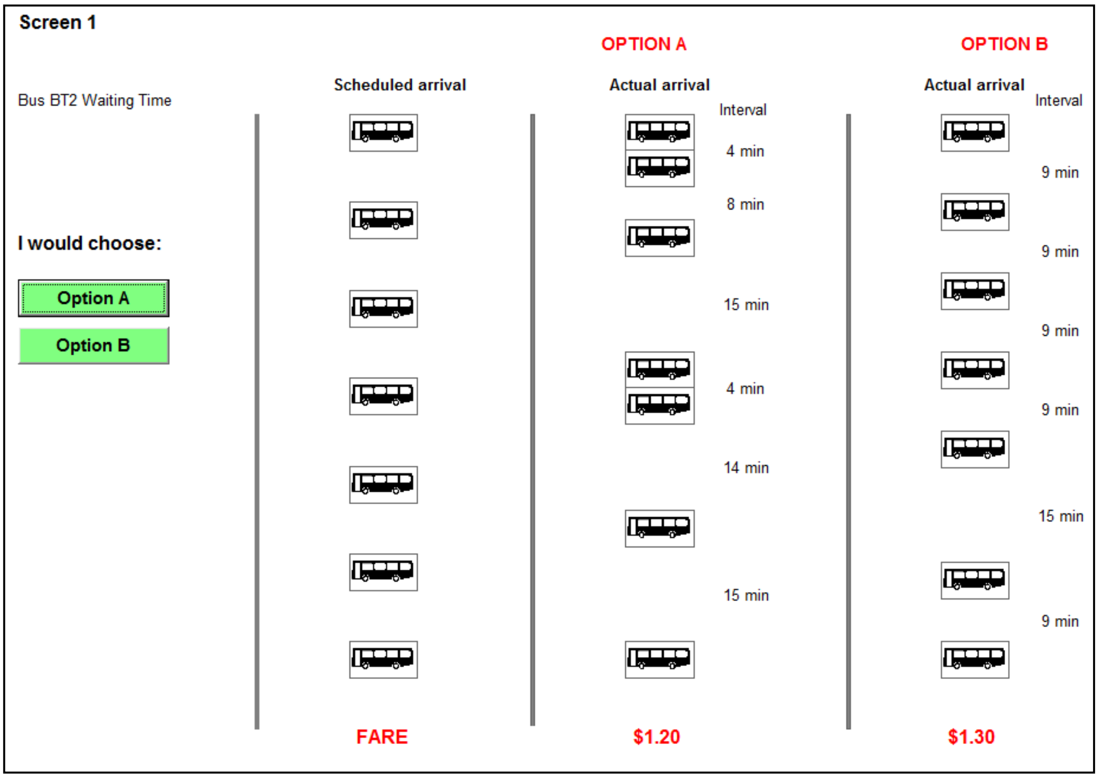

Figure 1: An example screenshot of a EWT choice experiment

The data collected from this SP survey was next analysed using advanced discrete choice models. We specified the models in valuation space, such that the utility for alternative *i* would be given by

$$V_i = \beta_{fare}(fare_i + VTTEWT \cdot EWT_i) + \varepsilon_i$$

where ε_i is a type I extreme value error term, $fare_i$ and EWT_i measure the fare and EWT of alternative *i*, β_{fare} is an estimated fare sensitivity and *VTTEWT* is the monetary valuation of a reduction in EWT. We worked with Mixed Logit models, in which we allow for random heterogeneity across individual decision makers. We similarly estimated the correlation between willingness to pay measures and cost components.

For distributional assumptions, we largely relied on lognormal distributions, with negative lognormal distributions for cost components, and positive lognormal distributions for willingness-to-pay measures. With the long tail of the lognormal distribution, a very small number of outlying values can lead to extreme valuations, and with this in mind, attempts at very minor censoring of the distribution were made, as discussed below.

The resulting model structure is highly advanced and flexible, with often large numbers of randomly distributed coefficients, with correlation between them. The degree of flexibility goes beyond what has been used in many previous studies, where especially the estimation of the full covariance matrix leads to substantial gains in model flexibility. These advantages were confirmed in our empirical work, with large further gains in model fit and more importantly also a greater ability to distinguish valuations across individual journey components.

However, this increase in flexibility also comes at the cost of increased model complexity. Classical estimation techniques for Mixed Logit were found to be incapable of estimating these models in a stable and timely manner and we instead turned to Bayesian estimation. The underlying model structure is still MMNL, it is simply the process used to obtain estimates that changes. For a detailed discussion of Bayesian techniques for Mixed Logit, see Train (2009, chapter 12). In our Bayesian estimation, we made use of 1,000,000 burn-in iterations for each model, with results then obtained by averaging over 50,000 post-burn-in iterations.

After model estimation, we also produced posterior estimates, thus generating for each individual the most likely value for the various willingness-to-pay measures as a function of the choices they were observed to have made (see Train, 2009, chapter 11). These individual-specific values are then used in posterior segmentation work to attempt to uncover further deterministic heterogeneity.

The model contains one cost component (fare) and one time component (EWT). We made use of a negative lognormal distribution for fare and specified the models in WTP space relative to fare, using a positive lognormal distribution for the valuation of EWT. This leads to two randomly distributed coefficients, where we estimate a full covariance matrix, meaning two diagonal terms (the variances of the individual utility components) and one off-diagonal term (the covariance between the two individual utility components). The estimated parameters relate to the underlying Normal distribution, i.e. the log of the absolute values of the coefficients (Table 3).

Table 3: Estimation results for the EWT choice experiment

Parameters	Estimate	
Respondents	791	
Observations	5,537	
Estimated parameters	5	
Log-likelihood	-2,863.00	
adj. ρ^2	0.25	
	Est.	t-ratio
Fare (underlying Normal mean for log of negative of coeff)	0.99	7.85
VTT EWT (underlying Normal mean for log of coeff)	-1.89	-7.13
cov(1,1)	5.52	7.18
cov(1,2)	-8.83	-6.99
cov(2,2)	14.24	5.98

Before studying the results on the valuation of EWT in detail, two core points need to be made. Firstly, the ranges of EWT presented in the experiment were by definition very narrow, given the use of a scheduled headway of 10 min, and a maximum and minimum time in between bus arrival times of 4 and 16 min, respectively. This leads to a maximum EWT of just 1.35 min, with an average of 0.67 min. The resulting boundary EWT values, i.e. the trade-offs that respondents were faced with, ranged from 7.69c/min to 4800c/min (i.e. SGD48). This would be the valuation of EWT a respondent would need to have to choose the more expensive option in a given choice task (and hence the one with the lower EWT). The median accepted boundary (i.e. from choices where respondents accepted to pay more for lower EWT) was 75c/min, while the median rejected boundary (i.e. from choices where respondents refused to pay more for lower EWT) was 160c/min, with respective means of 114.87c/min and 374.05c/min for accepted and rejected boundaries. This thus directly explains the high estimated value of EWT we will now turn to, which also needs to be put into context by noting that in practice, a one minute change in EWT is a major difference.

For the valuation of EWT, we censored the lognormal distribution at the highest accepted boundary value, which was 800c/min, i.e. (SGD8). This is possible in this game where the choice in each scenario can be quantified by a single boundary value. This censoring led to a drop in log-likelihood to -2,921.65 units, i.e. a drop by 58.65 units. This is a non-trivial drop in fit, but this approach to censoring was needed in this game in order to obtain reasonable results. The result average valuation of EWT is 71.74c/min (Table 4). This is much higher than the valuations of in vehicle time from the same study, and exceeds the average wage rate by a factor of more than two. However, the value is a realistic estimate of the real valuation, being very close to the median accepted boundary value, and it again needs to be borne in mind that achieving a minute reduction in EWT is a far bigger step than a minute reduction in travel time.

Table 4: Implied EWT valuations

	Mean	Std. Dev.	Mean as % of wage rate
Value of EWT (c/min)	71.74	144.75	238
Wage rate from sample (SGD/hr)	18.11		

In our analysis of the posterior estimates, i.e. the most likely value for each component for each individual, we can make the following observations, where the small sizes for some of the subgroups need to be taken into consideration (Table 5):

- There is a strong indication of higher valuation of EWT reduction on school trips, potentially as a result of scheduling constraints.
- The EWT games show the strongest evidence of a meaningful income effect, with higher valuations for higher income respondents.
- Surprisingly, valuations are higher for students than for other respondents, a finding that is difficult to explain except on the basis of a higher household income for their families and a perception that it is not them who pay for bus fares.
- The valuation is higher for those travelling in a group.

Taken together, the results will help LTA focus its bus reliability improvements spatially and temporally across sections of the network where such improvements are most likely to be valued according to the profile of bus users.

Table 5: Posterior analysis of EWT valuations

	Sample size	Value of EWT (c/min)
Trip purpose		
No purpose	3	59.10
Home-based Other	229	60.44
Home-based School	101	90.07
Home-based Work	297	70.69
Not-home based	161	78.52
Income levels		
No income	312	69.98
Income < 2000	168	70.02
Income 2000-4000	142	78.74
Income 4000-7000	62	83.00
Income > 7000	23	90.04
Income missing or refused	84	56.57
Employment status		
Work full-time, part-time or self-employed	469	72.31
housewife	104	55.94
student	128	92.46
retired	66	48.49
unemployed or work NA	24	82.64
Travelling party size		
Travelled alone	646	70.45
Travelled with others	145	77.51

4.2 Commuter awareness of the BSRF

While it may be argued that bus users need not be burdened with the technical details of the BSRF as long as they perceive an improvement in bus service reliability, it is important from the Government public relations perspective that the public attributes the significant resources expended on the scheme to the efforts of the Authority/regulator. In this regard, it is thought that the most salient feature of the BSRF to the public would be the financial incentives/penalties for operators who exceed/fail to meet EWT baseline measures. However, an internal Government survey conducted in late 2014 found that the BSRF received just low to moderate awareness among bus users (Figure 2).When asked how aware they were of the BSRF on a scale of 1 – 5 (1 – Not at all aware, 2 – Slightly aware, 3 – Somewhat aware, 4 – Moderately aware, 5 – Extremely aware), just slightly over half of the bus users (56 percent) were at least somewhat aware of the framework, while close to a third were not at all aware (32 percent). It is possible that the survey results are simply a reflection of the limited scale of the trial at that time and that a similar survey to be conducted in a year's time may yield different results as the BSRF is ramped up across the island.

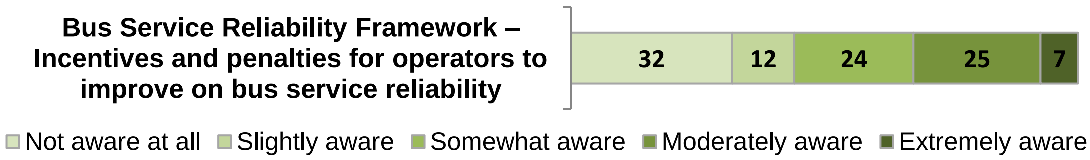

Figure 2: Proportion of respondents aware of the Bus Service Reliability Framework on varying degrees of awareness

4.3 Resultant Reliability Improvements

For the first assessment period of June to November 2014, the reliability of 20 out of the 22 bus services under trial has improved, as indicated in Table 6 (Land Transport Authority, 2015a). Of the bus services which improved, 18 earned incentives. The remaining four services had performances in the neutral zone and neither earned any incentive nor incurred any penalty. All trunk bus services saw improvements in EWT.

Table 6: Summary of EWT improvements in first year of BSRF trial

Fleet size category	Operator	Bus Service	EWT Baseline (min)	EWT for Jun to Nov 2014 (min)	ΔEWT (min)
Category 3 (≥20 buses)	SBST	3	1.5	1.1	0.4
		39[a]	1.2	1.1	0.1
		51	2.2	1.6	0.6
		154	1.9	1.4	0.5
	SMRT	176	1.6	1.4	0.2
		188	1.4	1.1	0.3
		858	2.1	1.8	0.3
Category 2 (10 to <20 buses)	SBST	17	1.8	1.4	0.4
		52	2.0	1.4	0.6
	SMRT	184	1.3	1.0	0.3
		189	1.4	1.2	0.2
		853	1.5	1.0	0.5
		911[a]	1.3	1.4	-0.1
		962[a]	1.4	1.3	0.1
Category 1 (< 10 buses)	SBST	228	1.1	0.7	0.4
		241	1.2	1.0	0.2
		242	0.8	0.6	0.2
		292	0.9	0.7	0.2
		325	0.8	0.6	0.2
		354	1.1	0.6	0.5
	SMRT	302[a]	1.0	1.0	0.0
		901	0.9	0.7	0.2

[a] *These services remained in the neutral zone and did not qualify for any incentive or penalty.*

Based on its performance over 6 months, SBST earned about S$700,000 (US$525,000) for reliability improvements to 11 services and SMRT earned S$345,000 (US$259,000) for improvements to seven services. In comparison, total bus fare revenue by SBST and SMRT for the whole of 2014 was S$754 million (US$566 million) and S$218 million (US$163 million) respectively, which works out to be an average of S$2.7 million (US$2 million) per route. There were no penalties deducted from both operators for this assessment period as none of their services had deteriorated more than 0.1 min to fall into the penalty zone. The incentives will help to offset the costs incurred by the PTOs to hire the additional service controllers to support the BSRF.

To further investigate the impact of the BSRF, we compared the daily ITP-specific EWT performance of a BSRF trunk route (SMRT 853) against the daily ITP-specific EWT of SMRT 852, a non-BSRF trunk route. SMRT 852 was chosen on the basis of a significant route overlap with SMRT 853 from a bus interchange in the northern part of the island, so that factors external to the operator, such as road, traffic and weather conditions, can be largely controlled for when comparing EWTs at the common ITP. Route diagrams of SMRT 853 and SMRT 852 with the approximate location of the ITP that the EWTs are based on are shown in Appendix 1.

A before and after comparison relative to the start of the incentive-penalty regime in June 2014 shows that SMRT 853's EWT has decreased by about 0.1 minutes, while SMRT 852's EWT has increased by the same magnitude (Table 7). Table 7 also reports a simple difference-in-difference analysis of the EWT of SMRT 852 relative to the EWT of SMRT 853 on a monthly basis from January 2014, suggesting that BSRF has helped to improve the reliability of SMRT 853, after controlling for changes in operating conditions that are proxied by the EWT scores of SMRT 852. Improvements were detectable from April 2014 onwards, such that by November of the same year, the EWT difference for SMRT 852 and SMRT 853 has widened by 0.7 minutes from the January baseline. These are encouraging indications for the BSRF trial in general.

The success of the BSRF trial in its first year may be attributed to operators injecting additional resources required to monitor and improve en-route reliability, such as having more standby buses and drivers, as well as employing more service controllers to communicate with drivers and manage bus movement. The surpassing of standards in most of the assessed bus services does not indicate low baseline standards but rather the concerted efforts of operators to improve reliability of services under BSRF. Likewise, there were services that did not improve despite operators putting in their best efforts to manage these services. With these and subsequent findings from the trial, LTA will continue to fine-tune the framework over time.

Table 7: Before-and-after EWT comparisons of a selected pair of BSRF/non-BSRF bus routes

	SMRT 853	**SMRT 852**	Dependent variable:	**$(EWT_{852} - EWT_{853})_t$**
Average EWT from Jan to May 2014 **before** Incentive/Penalty Start Date	0.81	0.89	Constant	–0.25*** (0.094)
			(0,1) variable for Feb	0.18 (0.14)
			(0,1) variable for Mar	0.32 (0.17)
			(0,1) variable for Apr	0.56*** (0.12)
			(0,1) variable for May	0.52*** (0.12)
Average EWT from Jul to Nov 2014 **after** incentive/Penalty Start Date	0.71	1.01	(0,1) variable for Jul	0.57*** (0.14)
			(0,1) variable for Aug	0.74*** (0.14)
			(0,1) variable for Sep	0.42*** (0.14)
			(0,1) variable for Oct	0.35** (0.14)
			(0,1) variable for Nov	0.70*** (0.149)
Difference in EWT (After – Before)	–0.10** (0.045)	0.12* (0.063)	Adj R^2	0.15
			No. of observations	202

Standard errors are reported in parentheses.
*, **, *** indicates significance at the 90%, 95% and 99% level, respectively.

5. Conclusions

This paper provides an overview of various mechanisms used by regulators to measure and improve bus service reliability standards, with a particular emphasis on the approach that Singapore has taken. From an empirical and methodological perspective, another significant contribution of this paper is the estimation of willingness-to-pay measures for EWT improvements, the results of which will be useful from a policy perspective in helping to determine a commensurate level of effort and resources to be expended on improving bus reliability.

Overall, the BSRF trial in Singapore has yielded promising results to date with a majority of bus services under the trial notching improvements in their EWT scores. In June 2015, LTA announced that the number of bus services on the BSRF trial would be doubled to 45, and that the trial would be extended from February to August 2016 (Land Transport Authority, 2015b). LTA also announced that a new indicator, on-time adherence[2], which is used to measure the punctuality of bus services, will be trialled for two low frequency services. Eventually, LTA intends to bring all bus services under the BSRF through the Bus Contracting Model. During the tendering process, EWT baseline standards for each bus service are clearly

[2] On-time adherence (OTA) is the percentage of arrivals at selected bus stops that are within a -2 min/+5 min range of the scheduled arrival time. The OTA target is currently set at 85%.

stipulated in contractual documents, which tenderers have to consider when planning their timetables.

With the start of operations by Tower Transit, Singapore's third bus operator, on 29 May 2016, Singapore's new Bus Contracting Model provides avenues for further research into how contract specifications and the injection of greater competition advance the improvement and implementation of reliability standards across the bus network. It would be interesting for example to study if there are varying degrees of success among the various operators (incumbents and new entrants) in terms of running more reliable bus services, and the factors that distinguish operators from each other. Amid rising commuter expectations and with a major push by the Government to provide better service and make public transport a choice mode, the emphasis on improving reliability through the introduction of new measures and regulatory frameworks could not come at a more opportune time.

Acknowledgements

The authors wish to acknowledge the invaluable research assistance rendered by Teck-Sing Chng, Zhi-Rong Chng, Jeffrey Jen and Zhi-Zhan Lo in the preparation of this paper. Comments from an anonymous reviewer are much appreciated.

The views expressed in this paper are those of the authors and do not necessarily reflect the official position of the Land Transport Authority or of the Government of Singapore. Any remaining errors are ours alone.

References

Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. Transportation Research Part E: Logistics and Transportation Review, 37(2), 191-229.

Carrion, C., and Levinson, D., 2012. Value of travel time reliability: a review of current evidence. Transportation Research Part A: Policy and Practice, 46(4), 720-741.

Cats, O., 2014. Regularity-driven bus operation: Principles, implementation and business models. Transport Policy, 36, 223 – 230.

dell'Olio, L., Ibeas, A., and Cecín, P., 2010. Modelling user perception of bus transit quality. Transport Policy, 17(6), 388 – 397.

Disney, J., 1999. Customer satisfaction and loyalty: The critical elements of service quality. Total Quality Management, 10(4-5), 491-497.

Edvardsson, B., 1998. Causes of customer dissatisfaction - studies of public transport by the critical-incident method, Managing Service Quality: An International Journal, 8(3), 189 – 197.

Feng, W., and Figliozzi, M., 2012. Developing a bus service reliability evaluation and visualization framework using archived AVL/APC data. Paper presented at the Western ITE Conference, Santa Barbara, California, USA.

Goh, P., Swee, A., and Low, J., 2015. Transition of Singapore's public transport industry. Paper presented at the International Conference Series on Competition and Ownership in Land Passenger Transport, Santiago, Chile, 30 August – 3 September 2015.

Guo, G., Luo, H., Lin, X., and Feng, C., 2011. Headway-based evaluation of bus service reliability. Paper presented at the 14th International IEEE Conference on Intelligent Transportation Systems, Washington, DC, USA.

Hensher, D.A., and Houghton, E., 2002. Performance-Based Quality Contracts for the Bus Sector: Delivering Social and Commercial Value for Money. Transportation Research Part B: Methodological, 38(2), 123 – 146.

Hensher, D.A., and Stanley, J., 2003. Performance-based quality contracts in bus service provision. Transportation Research Part A: Policy and Practice, 37(6), 519 – 538.

Huo, Y., Zhao, J., Li, W., and Hu, X., 2014. Measuring bus service reliability: An example of bus rapid transit in Changzhou. Journal of Public Transportation, 17(2), 113-133.

Ian Wallis Associates Ltd, and The TAS Partnership, 2013. Improving bus service reliability. NZ Transport Agency research report, 527. [online] Available from https://www.nzta.govt.nz/assets/resources/research/reports/527/docs/527.pdf [Accessed: 15 May, 2016]

König, A., and Axhausen, K.W., 2002. The reliability of the transportation system and its influence on the choice behaviour. Paper presented at the Swiss Transport Research Conference (STRC), Monte Verità/Ascona, March 20 – 22.

Land Transport Authority, 2014a. To improve bus regularity, reduce bus bunching, 22 bus services under BSRF starting from February 2014. [online] Available from:

http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=5e6512c1-d964-4229-ba34-92c539a7c6b8 [Accessed: 29 June 2015]

Land Transport Authority, 2014b. Early signs that BSRF is reducing instances of bus bunching and prolonged waiting times. [online] Available from: http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=d95b67ae-7659-4ff0-8ec2-00957f2becc1 [Accessed: 29 June 2015]

Land Transport Authority, 2015a. Bus service reliability framework (update for first assessment period. [online] Available from http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=edab460e-1915-43e2-a668-2e2d33792663 [Accessed: 29 June 2015]

Land Transport Authority, 2015b. Expanding the bus service reliability trial. [online] Available from http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=76c8984b-93ee-41a2-a06d-3623c094450d [Accessed: 29 June 2015]

Larsen, O. I., 2001. Designing incentive schemes for public transport in Hordaland county, Norway. Paper presented at the 7th International Conference on Competition and Ownership of Land Passenger Transport, Molde, Norway.

Li, Z., Hensher, D.A., and Rose, J.M., 2010. Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence. Transportation Research Part E: Logistics and Transportation Review, 46(3), 384-403.

Lin, J., Wang, M. L., and Barnum, D. T., 2008. A quality control framework for bus schedule reliability. Transportation Research Part E: Logistics and Transportation Review, 44(6), 1086-1098.

Oort, N., 2014. Incorporating service reliability in public transport design and performance requirements: international survey results and recommendations. Research in Transportation Economics, 48, 92-100.

Selviaridis, K., and Wynstra, F., 2014. Performance-based contracting: a literature review and future research directions. International Journal of Production Research, 53(12), 3505-3540.

Strathman, J., Kimpel, T., Gerhart, R., Turner, K., Taylor, P., Callas, S., and Griffin, D., 2000. Service reliability impacts of computer-aided dispatching and automatic vehicle location technology: A Tri-Met case study. Transportation Quarterly, 54(3), 85-102.

Train, K., 2009. Discrete choice methods with simulation. Cambridge UK: Cambridge University Press.

Transport for London, 2014. Travel in London: Report 7. [online] Available from https://www.tfl.gov.uk/cdn/static/cms/documents/travel-in-london-report-7.pdf [Accessed: 11 June 2016]

Transport for London, 2015. London's Bus Contracting and Tendering Process. [online] Available from http://content.tfl.gov.uk/uploads/forms/lbsl-tendering-and-contracting.pdf [Accessed: 15 May 2016]

Vincent, M., 2008. Measurement valuation of public transport reliability. Land Transport New Zealand Research Report, 339. [online] Available from https://www.nzta.govt.nz/assets/resources/research/reports/339/docs/339.pdf [Accessed 15 May 2016]

Table A.2 Route diagrams of SMRT 853 and 852