WORKING PAPER
ITS-WP-99-14

The Missing Link in Contract Performance Assessment: The Integration of a Service Quality Index into a Competitive Tendering Regime

By

David A. Hensher
& Paola Prioni

July,1999

ISSN 1440-3501

Established and supported under the Australian Research Council's Key Centre Program.

INSTITUTE OF TRANSPORT STUDIES
The Australian Key Centre in Transport Management

The University of Sydney and Monash University

NUMBER: Working Paper ITS-WP-99-14

TITLE: The Missing Link in Contract Performance Assessment: The Integration of a Service Quality Index into a Competitive Tendering Regime

ABSTRACT: Over the last two decades the bus industry in many countries has been involved in a process of economic deregulation, competitive regulation and privatisation. Among the different policy practices designed to increase competition, competitive tendering represents a widespread policy intervention. Although there is extensive acceptance of competitive tendering, the focus has been on cost efficiency and cost effectiveness designed to identify the mix of inputs used to produce a given level of output at the lowest cost, where output is produced services (eg vehicle kilometres) on the efficiency measure and consumed services (eg passenger kilometres) on the effectiveness measure.

Regulators have been singularly unsuccessful in developing a robust specification of service quality levels, and have come into criticism that the focus of economic reform has concentrated too much on saving money at the expense of preservation and enhancement of service levels. The definition of service level has tended to ignore the quality of service, limiting the specification of a predetermined level of service to simple physical measures such as vehicle kilometres and passengers carried. In this paper we develop a method of filling in the missing link in the specification of contract performance - service effectiveness - which measures the effectiveness of a service in satisfying passengers.

AUTHORS: David A. Hensher &
Paola Prioni*

CONTACT: Institute of Transport Studies (Sydney & Monash)
The Australian Key Centre in Transport Management
C37, The University of Sydney NSW 2006
Australia

Telephone: +61 2 9351 0071
Facsimile: +61 2 9351 0088
E-mail: itsinfo@its.usyd.edu.au
Internet: http://www.its.usyd.edu.au

DATE: July, 1999

* *This paper was written during Dr Paola Prioni's time at ITS on a Swiss Government sponsored Post Doctoral scholarship.*

Introduction

Suggestions that we need to be more vigilant in the way that service quality is handled in competitive tenders has often led to concerns about adding complexity to contract design that would both discourage bidders and add unacceptable administrative costs to the evaluation and monitoring process. The extent to which required service quality targets would discourage bidders and/or add administrative costs will be dependent on how complex the service quality formula becomes and the extent to which it adds to the incentive (in)compatibility of the tendering process.

While cost efficiency gains are important in the establishment of a successful competitive bid, the definition preconditions the outcome on a given level of service. Such a (minimum) service level is typically defined by the amount of service kilometres and hours delivered over a network subject to a conditions on access distance to the network. This definition of service levels does not take into account what really influences a users perception of the effectiveness of a service.

This paper proposes a way of measuring service quality that results in a very simple and intuitively appealing formula that is transparent, is incentive compatible, easy to administer and monitor and which can be integrated into the specification of a competitive tender. In developing such a service quality index (SQI) we integrate the rich literatures on stated choice methods and performance measurement, redefining service effectiveness as a two-dimensional construct in which perceived service quality is a major component. The empirical study illustrates how a discrete choice model framework in which alternative bus service packages are evaluated provides an appropriate way of identifying the underlying service attributes that represent service quality and their relative importance from the passengers' perspective.

The inclusion of service quality opens up an opportunity to review the way that competitive tenders are structured to take into account improved service quality in line with benchmarked best practice. In addition to requiring the delivery of a specific level of service, we might include a requirement to provide this level of service to comply with a target service quality as specified by SQI. One possible strategy is to inform bidders of the current level of service quality on each of the dimensions of the SQI (together with the weights for each attribute) and to require that the successful bidder move the index up to a new level by adjusting the levels of one or more of the attributes in the index. For example moving the SQI from 1.3 to 1.6. The operator can determine how to achieve the target level and what it might cost and build this cost into the price of the bid. In this way we are encouraging improvements in service quality under incentive compatible tenders that are an improvement over the traditional cost only contracts. This SQI is the preferred indicator of passenger service quality in a competitive tendering and/or performance assessment regime.

This paper is organised as follows. We begin with a review of approaches to specify an indicator of service quality. This is followed by the justification for the stated preference paradigm with a focus on evaluating packages of service attributes. The empirical context and survey instrument are presented followed by the analysis results from a multinomial logit model and the construction of the service quality index. The final section preceding the conclusion suggest a schema for integrating SQI into the specification of a competitive tender, including the determination of targets and conditions of review and renewal.

The Search for an Operational Indicator of Service Quality

The literature on measuring the cost efficiency and cost effectiveness of bus services and operations is extensive (eg Hensher and Daniels 1995, Fielding et al 1985). A major data input is the level of service output, typically measured on the demand side by annual passenger trips or passenger kilometres and on the supply side by vehicle kilometres. As aggregate indicators of total output, these measures implicitly assume homogeneity in respect of service quality. Passengers however evaluate services in many ways, which may not be systematically associated with the amount of use of the service; indeed it is unclear whether differences in passenger satisfaction across the segments served by buses can be proxied by the preferred demand side indicator, aggregate passenger kilometres.

In the 1970's British rail argued that maximisation of passenger kilometres was a good proxy indicator of social welfare maximisation and accessibility offered to passengers. At the time, however, there had not been any substantive investigation of how passengers perceive the level and quality of services in their determination of passenger satisfaction. A number of studies have since refocused on the measurement of service quality, investigating the role of trade-off methods such as stated preference (eg Hensher 1991, Swanson et al 1997) and univariate procedures that rate individual service items on a satisfaction scale (eg Cunningham et al 1997).

Although specific aspects of service quality may be particularly positive or negative in a passenger's perception of (and satisfaction with) a service, we make the assumption that the overall level of passenger satisfaction is best measured by how an individual evaluates the total package of services on offer. Appropriate weights attached to each service dimension will reveal the strength of positive and negative sources of overall satisfaction. The stated preference (SP) paradigm enables us to develop preference formulae for a large number of service level scenarios, which can be implemented at the bus business level to establish operator-specific indicators of service delivery quality and effectiveness. The resulting satisfaction (or utility) indicators emanating from the estimation of the stated preference experiments measure the expected utility that a passenger obtains from the current levels of service and how this might change under alternative service level regimes[1].

---

[1] Given the heterogeneity of the population of bus passengers, segment-specific service quality indicators can be identified.

The Stated Preference Paradigm

The task is to develop an approach to incorporating a service quality indicator into a performance assessment regime that is a meaningful measure of service effectiveness from a passenger perspective. In addition, such an index should have the ability to be decomposed into its constituent sources of passenger satisfaction, as well as mapping into an aggregate demand-side indicator of passenger output to establish the role of the latter as a practical approximation of the social welfare significance of the bus service levels.

The starting position is a recognition that passengers purchase a package of service attributes when travelling on a bus and thus the contribution of each underlying elemental attribute must be assessed in the context of the overall quality of service on offer. The attributes are not necessarily independent of each other and hence the use of univariate methods such as satisfaction scales associated with each attribute are quite misleading as a basis if inferring the role of each attribute in the measurement of overall consumer satisfaction.

With a complex disaggregation of service quality, revealed preference data (RP) is inappropriate. There is too much confoundment in RP data, best described as ‘dirty’ from the point of view of statistical estimation of the individual influences on choice. Furthermore some attributes such as air conditioning do not exist today on many urban buses so we are unable to establish their influence.

An alternative data paradigm emerged in the late 1970s (Louviere and Hensher 1982, 1983, Wordworth and Louviere 1983). As a redefinition of the SC approach of earlier years, the focus was not on a different set of measures of trading between service attributes but a re-specification of the way in which the choice outcome and the attribute levels were defined. In contrast to RP data in which the choice outcome was exactly known (by observation) and the attributes of each alternative were measured with error (due the reporting process), SC data were defined by a set of attributes with precise levels and a choice outcome that was reported with error. Although the initial SC models were based on a rating or ranking exercise, which gave them limited credibility within the economic paradigm the turning point of acceptance came when it was shown that SC data with a choice response is identical to RP data except for its measurement and specification properties (Louviere and Hensher 1983, Woodworth and Louviere 1983, Hensher 1994, Hensher Louviere and Swait 1999).

Stated preference (SP) methods provided the richness required for the service quality index (SQI). It involves a stated choice experiment in which we systematically vary combinations of levels of each attribute to reveal new opportunities relative to the existing service levels on offer. Through the experimental design paradigm we observe a sample of travelers making choices between the current trip attribute level bundle and other attribute level bundles. This approach is capable of separating out the independent contributions of each service component and hence is capable of providing an SQI that is a rich representation of the sources of service (dis)utility.

Defining the Empirical Setting and the SP Experiment

To assist in the selection of attributes for the SQI, we undertook an extensive review of the literature as well as a survey of bus operators who have a wealth of experience on what customers look for in a good service (see Prioni and Hensher 1999). We found that thirteen attributes describe the major dimensions of service quality from a user's perspective. The range of levels of each attribute in Table 1 provided us with a mechanism for establishing the weights that signal the contribution of each attribute to the overall SQI.

Table 1. The Set of Attributes and Attributes Levels in the SP Experiment

Attribute	*Interpretation of levels*	*Attribute*	*Interpretation of levels*
Reliability	-on time -5 minutes late -10 minutes late	**Info at the bus stop**	-on time -5 minutes late -10 minutes late
Frequency	-every 15 minutes -every 30 minutes -every 60 minutes	**Travel Time**	-25% quicker than the current travel time -same as now -25% longer than the current travel time
Walking distance to the bus stop	-now -5 minutes more -10 minutes more	**Bus stop facilities**	-Bus shelter with seats -Seats only -No shelter or seats at all
Waiting safety	-very safe -reasonably safe -reasonably unsafe	**Fare**	-25% more than the current one-way fare -same as now -25% less than the current one-way fare
Access to the bus	-Wide entry with no steps -Wide entry with 2 steps -Narrow entry with 4 steps	**Driver attitude**	-Very friendly -friendly enough -Very unfriendly
Air conditioning	-Available with no surcharge -Available with a surcharge of 20% on existing one-way fare -Not available	**Safety on board**	-The ride is very smooth with no sudden braking -The ride is generally smooth with rare sudden braking -The ride is jerky; sudden braking occurs often
Cleanliness of seats	-Very clean -Clean enough -Not clean enough		

Through a formal statistical design the attribute levels are combined into *bus packages* before being translated into a survey form. The full factorial design (ie all possible bus packages) consists of 3^{13} combinations of the 13 attributes each of three levels. To produce a practicable and understandable design for the respondents, we restricted the number of combinations to 81 (ie 81 choice sets) using a fractional design. Fractional designs permit the reduction in the number of combinations (ie the number of bus packages) without losing important statistical information (see Louviere, Hensher and Swait in press).

A pre-test of the survey showed that respondents were able to evaluate consistently three choice sets (ie different scenarios of bus packages), resulting in 27 different survey forms. To allow for a rich variation in the combinations of attribute levels to be evaluated as service packages in the SP experiment, each bus operator received 8 sets of 27 different survey forms (ie 216 forms) and instructions on how to organise the survey. An example of an SP question is shown in Table 2.

Table 2. A Typical Stated Preference Exercise

SERVICE FEATURE	BUS PACKAGE OF THE BUS COMPANY A	BUS PACKAGE OF THE BUS COMPANY B	BUS PACKAGE OF THE CURRENT BUS
Reliability	**10 minutes late**	**on time**	**7 minutes late**
One-way fare	**same as now**	**same as now**	**2 dollars**
Walking distance to the bus stop	**5 minutes more than now**	**5 minutes more than now**	**5 minutes**
Personal Safety at the bus stop	**reasonably unsafe**	**reasonably safe**	**very safe**
Travel Time	**25% longer than the current travel time**	**25% quicker than the current travel time**	**30 minutes**
Bus stop facilities	**No shelter or seats at all**	**Seats only**	**Seats only**
Air conditioning	**Not available**	**Available with no surcharge**	**Not available**
Information at the bus stop	**Timetable but no map**	**Timetable but no map**	**Timetable and a map**
Frequency	**Every 15 minutes**	**Every 30 minutes**	**Every 60 minutes**
Safety on board	**The ride is jerky; sudden braking occurs often**	**The ride is jerky; sudden braking occurs often**	**The ride is jerky; sudden braking occurs often**
Cleanliness of seats	**Clean enough**	**Clean enough**	**Very clean**
Ease of access to the bus	**Wide entry with no steps inside the bus**	**Wide entry with 2 steps inside the bus**	**Wide entry with 2 steps inside the bus**
Driver behaviour	**Friendly enough**	**Very friendly**	**Very friendly**

If BUS A and BUS B were available today, which bus service would you choose?

☐ BUS A ☐ BUS B ☐ The bus you are travelling on.

Results of the User Preference Model

Scheduled[2] bus users of 25 private bus operators in NSW participated. Survey forms were distributed and collected during April and May 1999. A total of 3,849 useable observations (out of 4,334 returns) were incorporated in the estimation of the discrete choice model. A multinomial logit (MNL) specification was selected. This is appropriate for a model form in which the utility expressions associated with the current trip and two attribute packages are unlabelled (or unranked) alternatives. Consequently all design attributes were generic across the three alternatives. In addition in the current trip alternative we considered alternative-specific characteristics of the passenger (income, gender, age and car availability) and of the operator together with a number of other

[2] School children were excluded from the sample, as they are captive users and might have a biased perception towards the attributes.

potential influences on relative utility such a treatment effect, trip purpose and access mode.

The user attribute choice model is summarised in Table 3. The model includes the attributes of the SP experiment, operator-specific dummy variables and three user characteristics. The overall goodness of fit (adjusted pseudo-R^2) of the model is 0.324. The great majority of the design attributes are statistically significant. Service reliability (ie the extent to which buses arrive on time), fares, access time and travel time are all highly significant with the expected negative sign. Relative to 'reasonably unsafe', we find a positive (almost) significant parameter estimate for 'reasonably safe' (0.1510) and for 'very safe' (0.1889). The higher estimate for 'very safe' in contrast to 'reasonably safe' is plausible. The infrastructure at the bus stop appears not to be a major influence on service quality with both 'seats only' and 'bus shelter with seats' not being statistically significant relative to 'no shelter or seats'. If reproducible in further studies this has important policy implications as to priorities in service improvement. The availability of air conditioning is another interesting result. We find that 'air conditioning without a fare surcharge' is not statistically significant relative to no air conditioning. In contrast the provision of air conditioning with a 20% surcharge on existing fares is statistically significant with a negative sign suggesting that users would sooner not have air conditioning if it means paying higher fares.

On-board safety defined by the smoothness of the ride is a statistically strong attribute. Relative to 'the ride is jerky with sudden braking occurring often', we find that 'the ride is generally smooth with rare sudden braking' and 'the ride is smooth with no sudden braking' are both very important positive attributes of service quality. This suggests both policy initiatives in driver skill as well as vehicle quality. Cleanliness of the bus is statistically significant when 'very clean' relative to 'not clean enough'. The non-statistical (1.830) significance of 'clean enough' suggests that we really have a dichotomy between very clean and not very clean. Ease of access to a bus, closely linked to the issue of accessible transport turns out to be not so important overall, presumably because the majority of users (including many aging users) are sufficiently healthy to not be concerned with the configuration of steps and entry widths. The attitude of the driver is a statistically strong influence on a user's perception of service quality. Indeed, relative to 'very unfriendly' we might expect a significant increase in the mean parameter estimate when we go from 'friendly enough' to 'very friendly'. This is the most non-linear effect on utility of all the attributes of service quality. Finally, the availability of information at the bus stop (timetable and map) is statistically important compared to 'no information', although surprisingly the key information item is a timetable, with a map being a liability (possibly because of experience with vandalism?).

Finally, bus frequency defined as 15, 30 and 60 minutes was found to be significant when treated as a dummy variable distinguishing 60 minutes from 15 and 30 minutes. There is a strong negative sign for the 60 minute dummy variable suggesting that a 60 minute service reduces relative utility significantly compared with a service frequency of every 15 or 30 minutes. Not statistically significant is the 30 minutes dummy variable, defined equal to one for frequencies equal to 30 minutes.

The socioeconomic characteristics sought from bus users were limited to personal income, age, gender and car availability. We found that individuals on higher incomes and of more years were more likely to prefer the levels of service offered by the existing trip than by the alternative packages. What this suggests is that as individuals age and increase their income, they see existing service quality as increasingly satisfying their requirements for service quality. Alternatively it is the younger and those on lower incomes that see a greater need for improved service quality. Car availability was not statistically significant. Further details are given in Prioni and Hensher (1999).

Table 3. Final User Preference Model

Variable	*Units*	*Acronym*	*Parameter*	*t-value*
Reliability mins		RELI	-.05821	-8.411
Bus fare	$	TARIF	-.4780	-6.406
Access time mins		ACCESST	-.04317	-5.311
Bus time	mins	TRATIM	-.03200	-5.435
Very safe	1,0	VSAFE	.18895	2.255
Reasonably safe	1,0	RSAFE	.15108	1.820
Seats only at bus stop	1,0	SEATS	-.03411	-.510
Seat plus shelter	1,0	SEATSHEL	.09040	1.503
Air conditioning free	1,0	AVALFREE	.07131	1.112
AC at 20% extra fare	1,0	AVALPAY	-.17432	-2.207
Ride-generally smooth	1,0	GSBRAKE	.20788	2.963
Ride-very smooth	1,0	VSNBRAKE	.35232	4.904
Clean enough	1,0	CENOUGH	.13867	1.830
Very clean	1,0	VCLEAN	.20446	2.713
Wide entry/2 steps	1,0	WIDE2STP	.09589	1.499
Wide entry/no steps	1,0	WIDENSTP	-.10319	-1.372
Driver friendly enough	1,0	FRIENDEN	.19798	2.572
Driver very friendly	1,0	VFRIEND	.42287	5.564
Timetable only	1,0	TIMNOMAP	.29609	4.745
Timetable and map	1,0	TIMWMAP	.19720	3.021
Frequency/every 60 mins	1,0	FREQ60	-.58595	-6.902
Frequency/every 30 mins	1,0	FREQ30	-.12221	-1.640
Female	1,0	FEMALE	.09986	1.198
Personal income	$'000s	PINCO	.00905	3.817
Age of passenger	years	AGES	.01379	5.787
Operator 1	1,0	Op1	.37358	1.671
Operator 2	1,0	Op2	.19642	.654
Operator 3	1,0	Op3	-.94098	-5.497
Operator 4	1,0	Op4	-.17726	-1.080
Operator 5	1,0	Op5	-.12964	-.653
Operator 6	1,0	Op6	.97267	1.937
Operator 7	1,0	Op7	-.18127	-0.982
Operator 8	1,0	Op8	.35723	1.294
Operator 9	1,0	Op9	-.26210	-1.215
Operator 10	1,0	Op10	-.56626	-1.845
Operator 11	1,0	Op11	-1.2555	-4.850
Operator 12	1,0	Op12	-.22189	-0.842
Operator 13	1,0	Op13	-.47366	-1.210
Operator 14	1,0	Op14	.01784	.072
Operator 15	1,0	Op15	.06911	.084
Operator 16	1,0	Op16	-.37973	-1.685
Operator 17	1,0	Op17	.06878	.292
Operator 18	1,0	Op18	-.36574	-0.825
Operator 19	1,0	Op19	1.1207	4.218
Operator 20	1,0	Op20	.10014	.488
Operator 21	1,0	Op21	.11275	.546
Operator 22	1,0	Op22	.32239	.781
Operator 23	1,0	Op23	-.53292	-1.845
Operator 24	1,0	Op24	.08878	.161
Log-likelihood	-2839.25			
Pseudo R^2 (adjusted)	0.324			

The Service Quality Indicator (SQI)

The Service Quality Index (SQI) for each operator is calculated by the application of the utility expression in Table 3 and the levels of each of the attributes associated with the current trip experience of each sampled passenger. In this study we have estimated a single set of utility weights across the sample of 3,849 passengers using the services of 25 operators. We investigated possibilities of differences in weights between segments of operators (eg Sydney metropolitan vs regional vs country towns) and found no statistically significant differences. This is most encouraging, suggesting a similar pattern of preferences of passengers across all operating environments. This does not mean however that the levels of service offered on each service attribute are the same (indeed there is substantial variation as shown in Figure 1 of the mean and standard deviation of each attribute for each operator). Rather what we are noting is that the marginal utilities of each attribute (ie the mean parameter estimate of part-worth weight) is well represented by a single mean estimate across all operators.

The SQI developed for each operator is summarised in Table 4 and graphed in Figure 1 at its mean for each operator. We have normalised SQI in Figure 1 to a base of zero for the operator with the lowest relative SQI. The range is from 0 to 2.70.

Table 4. Summary Statistics of Service Quality Index

Operator	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*	*Sample Size*
1	0.5311	0.788	-2.39	2.28	249
2	0.3900	0.894	-1.87	2.00	96
3	-.8178	1.248	-4.88	1.92	508
4	-1.098	0.927	-5.58	0.58	374
5	-1.2840	1.406	-5.46	0.84	196
6	-.8377	0.383	-.525	0.80	24 *
7	-.9263	1.297	**-6.74**	1.82	412
8	-.7113	0.566	-2.12	0.44	150
9	-.4597	0.685	-2.55	1.06	173
10	-.5805	0.904	-3.06	0.67	64 *
11	-1.628	0.979	-4.55	0.55	90
12	-.3923	1.000	-3.80	1.40	100
13	0.5435	0.483	-.434	1.28	41 *
14	0.7636	0.940	-2.28	2.61	180
15	0.2079	0.637	-.638	0.692	9 *
16	-.6345	0.958	-4.00	1.03	159
17	-.0649	1.089	-2.86	2.09	190
18	-.5687	1.206	-3.24	1.04	27 *
19	**1.0174**	0.947	-.990	**2.70**	203
20	-.0444	0.639	-1.43	1.55	224
21	-.4212	0.852	-3.45	1.17	227
22	0.6466	0.643	-.600	2.01	46 *
23	-.3076	1.034	-4.28	.808	65*
24	.1051	1.156	-2.17	1.42	22*
25	**-1.7579**	.875	-3.01	-.096	20*
All	-.4067	1.224	-6.74	2.70	3849

In developing the SQI indicator we have taken into account the differences in the socio-economic composition of the travelling public (eg age, income, car availability), the method of data collection (eg on board vs hand out and mail back) and location of operator. The contribution of each service quality attribute across all 25 operators in summarised in Figure 2 (and defined in Table 5).

The challenge for an operator is to compare themselves against best practice and to establish how best to improve overall service quality through implementing changes that reduce the magnitude of the attributes below the zero axis in Figure 2 and increase the magnitude of attributes above the zero axis.

Figure 1. The Service Quality Index

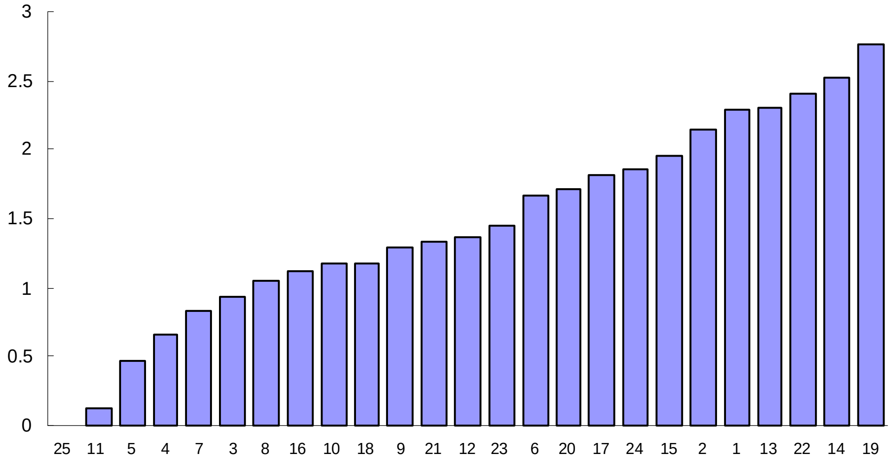


Figure 2. The composition of the service quality index (all operators in the sample*)*

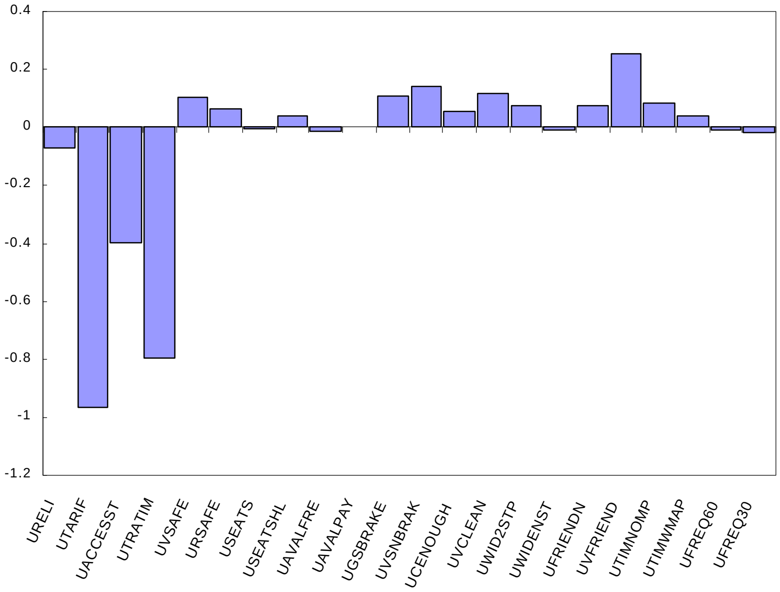


Table 5. Notation for Figure 2

URELI	Late minutes	UVSNBRAKE	Ride very smooth
UTARIF	Bus fare	UCENOUGH	Clean enough
UACCESST	**Access time**	UVCLEAN	Very clean
UTRATIM	Travel time	UWIDE2STP	Wide entry and 2 steps
UVSAFE	Very safe	UWIDENSTP	Wide entry no steps
URSAFE	Reasonably safe	UFRIENDN	Friendly drivers
USEATS	Seats only at bus stop	UVFRIEND	Drivers very friendly
USEATSHEL	Seats plus shelter at stop	UTIMWMAP	Timetable and map
UAVALFREE	Free Air conditioning	UTIMNOMAP	Timetable, no map
UAVALPAY	**Air conditioning at 20% extra fare**	UFREQ60	Frequency 60 minutes
UGSBRAKE	Smooth ride	UFREQ30	Frequency 30 minutes

Operationalising SQI as a Regulatory Tool

Integrating SQI targets in the specification of tenders

A growing criticism of competitive tendering is that economic regulators have failed to build into the specification of tender documents information on the quality of incumbent services from the users' perspective. This gap in the tendering process denies potential bidders the opportunity to prepare their bid offers with full knowledge of the effectiveness of existing service levels (Domberger et al 1995, Van de Velde and Sleuwaegen 1997).

SQI provides an appealing index to compute and operationalise service quality from an user perspective in an easy and scientific way. Because of its simplicity and its ability in capturing every important user-defined service quality component in a single index, SQI is a preferred operational tool in the specification of tendering contracts. SQI makes explicit through the revelation of information on current service quality the requirement to take into account the cost of maintaining and even enhancing service quality in bid offers, minimising the selection of low bids accompanied by low service quality delivery.

Traditional contracting makes no allowance for loss or gain of passengers. All that it requires is that the provider of the service satisfies conditions on delivery of minimum service levels specified by for example the total number of vehicle-kilometres, the network configuration, or the frequency of the service; but if a successful tender has for example a bad driver it might loose passengers and that feature of service is not accommodated in the specification of the contract. Our method takes this into account. An identification of SQI prior to tendering would allow the responsible authority to gain information on customers' satisfaction with the current levels of service quality and to include this information in the form of service quality targets in the contract specification.

Table 6 gives an example on how one might integrate SQI targets into the tender process. Let us assume that from a survey of a sample of existing users, we have identified the user-defined quality of current service of three operators. Operator 1 achieved an SQI of 1.4 by providing a service that is on average two minutes late, clean enough for 60% of the sampled users, costs on average $2.1, etc. Operators 2 and 3 have SQI's respectively of 1.3 and 2.0. Assuming that these operators are comparable, Operator 3 is best practice.

Regulators can use the SQI in the contract design to specify how much service improvement they require relative to the current levels as illustrated in the last two columns of Table 6. Although one might impose the requirement that each and every bus operator must be at best practice, this may discourage bidders and so we prefer to set a target level that is recognised as achievable by potential bidders. The level should be incentive compatible.

Table 6. Including SQI targets in the contact design

	Current Service description Attributes					***SQI***		
							Target after	
Operator	*Reliability*	*Bus fare*	*Clean enough*	*Travel time*	*etc…*	*realised*	**2.5 yrs**	**5 yrs.**
1	2 minutes late	**2.1**	**60%**	**25 minutes**	**…**	**1.4**	1.6	1.8
2	1 minutes late	**2.4**	**78%**	**26 minutes**	**…**	**1.3**		
3	1 minutes late	**2.0**	**80%**	**21 minutes**	**…**	**2.0**		

Given the gap between an operator's SQI and that of best practice (e.g. 0.6 for operator 3), we suggest a formulation $SQI+z$ where z is the predesignated improvement over a period of time (e.g. 0.2 in both sub periods). This is analogous to the CPI-x productivity formula used to regulate public utilities where the franchised operator raises its fare level every year in line with the general level of inflation (CPI) less a fixed amount (*x*) that reflects productivity improvement. The SQI+*z* formula provides a target in line with a pre-designated increase in the service quality level. In the case of the service previously provided by incumbent operator 1, authorities impose an SQI target of 1.6 after 2.5 years and a final SQI target of 1.8 at the end of the contract (5 years).

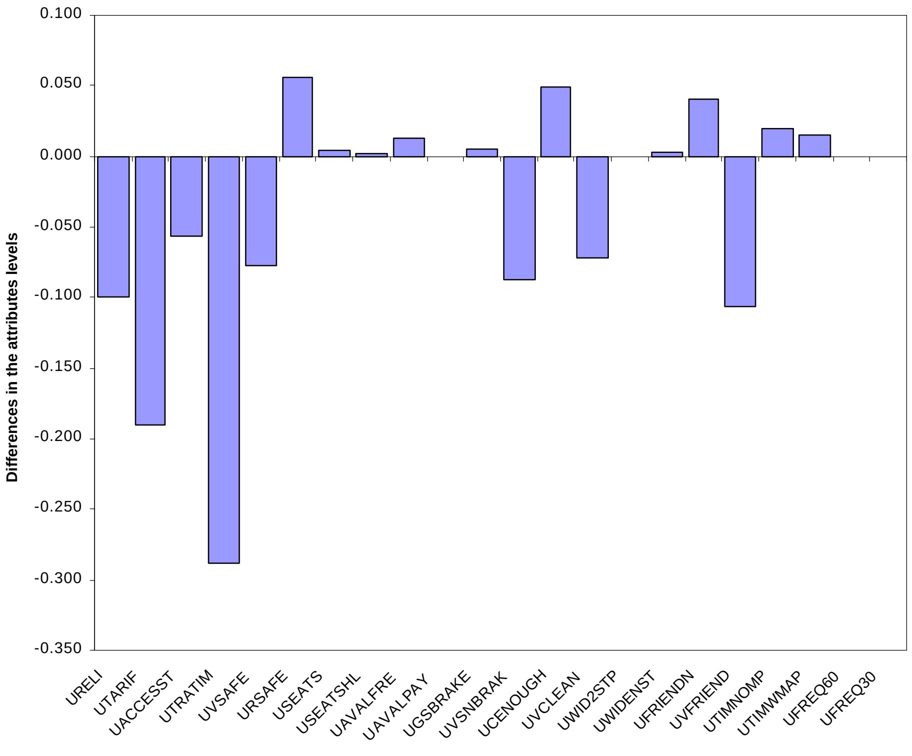
We must recognise that best practice will change over time and hence the target will be revised. Such a revision should be used to reset the value of *z* for the next 2.5 years and not backdated. In practice all potential entrants must be provided with the computational formula for SQI. According to their managerial and operational capability they will decide on how to decompose the index into the individual attribute components to achieve the targeted SQI. For example an operator might prefer to put more effort into the cleanliness of the vehicles and less into the reliability attribute (due to the difficult traffic conditions) but still comply with the targeted SQI.

The required service quality level will then be evaluated by bidders and added into the cost of providing the higher level of service to determine the bid price. The contract will be awarded to the lowest price offer (with the cost of service quality internalised). Once successful in winning the contract the operator has a strong incentive to meeting the new levels of service. Compared to the traditional tender contract specification, the inclusion of SQI in the contract secures improvements in cost efficiency while meeting the new levels of service effectiveness as prescribed by a user-defined service index.

Establishing the actual target

There are no specific arguments for establishing a particular SQI target other than to ensure an incentive compliant tendering process. The success of competitive tendering is determined among other things by the number of bidders: the greater the number of bidders the lower the bid price (see Glaister and Cox 1991, White and Tough 1995). For this reason the target SQI needs to be achievable and not necessarily set equal to best practice. In the incremental approach proposed, the value of *z* can be predetermined through negotiation between the regulator and industry although the size of *z* should not violate the conditions for an incentive compliant tendering process. As part of the process of establishing the value of *z*, one might look to existing evidence on the differences between best practice SQI and a specific operator's SQI. That is, best practice might be used as a mechanism for partitioning the targets over the life of the contract. Figure 3 illustrates such a difference.

Figure 3. A comparison of an operator against best practice



Authorities can define the targeted service quality to be based on all or a subset of the attributes reported herein. One can remove attributes that the regulator and/or the bus and coach industry might argue are not inclusions in an operational service quality indicator. For example, if it was argued that travel time and fares should be excluded this can easily be achieved with a new ordering of operators. SQI is therefore not only appealing for its simplicity (only one number) but also for its flexibility in accommodating changes in external factors (like changes in government policy or in the socioeconomic structure of the service area).

Moreover bus operators can be classified on a number of criteria agreed on between government and industry to arrive at operator membership of a segment. Benchmarking can then be undertaken within each segment (e.g. metropolitan/non metrpolitan area).

Monitoring and responses

To ensure contract compliance the supplier must be monitored during the contract period. This involves collecting and interpreting information that can be used to determine whether the specified bus services are achieving the new targeted SQI. Assuming a contract length of five years we propose a performance assessment at the midpoint. An operator would have to conduct a user survey after 2.5 years to establish compliance. To avoid any disputes on who should pay for the survey, it makes good sense to include the monitoring cost as part of transactions costs of the bid and included in the bid price. Table 7 summarises the four possible outcomes of a contractual process.

Table 7. Possible outcomes of a tender

	Renewal	*Retender*
Compliant	End of the 5 years	End of the 5 years
Non compliant	Re tendered	Retender: - (Warning after 2.5 years)

If the operator is compliant it becomes a political decision whether the contract will be renewed or retendered at the end of the contract period. In case of non-compliance after the first half of the contract period, the non-compliant operator should be warned about under performance without loosing a contract. If the operator is unable to achieve the target performance by the end of the contract period the contract should be retendered. In the case of a non-compliant operator, the tendering authority must determine if the reasons for non-compliance are internal to the contractor (ie under his control) or external (ie not under control of the operator). Only internal failure needs to be corrected through sanctions. In case of external factors influencing the operator’s service quality the tendering authority should review the pre-agreed targets.

The extent of benefits from competitive tendering depend not only on the size of the targeted SQI (see previous section) but also on other factors influencing the amount of competition. The size of irrecoverable costs, the length of the contract and the perceived probability of success will be critical factors in determining how many bidders come forward. The provision of information on existing service quality levels of the incumbent is essential to the success of the broadened specifications of competitive tenders if potential bidders are to be forthcoming.

Conclusions

We have developed a new approach to quantifying a service quality index (SQI) to enable the economic regulator and bus operators to benchmark service effectiveness, adding this much neglected dimension of performance assessment. The inclusion of service quality standards (ie SQI) in contract design avoids evaluation criteria exclusively based on a supply principle (cost efficiency). Competing offers can be judged according to their cost efficiency and service effectiveness. Indeed we have convincing evidence from ITS's international benchmarking program for the bus and coach industry that best practice operators on overall cost efficiency are seldom close to best practice on service quality.

We can now claim to have established a global measure of service effectiveness to parallel the global indicators used to measure cost efficiency and cost effectiveness (ie total factor productivity). Further details of the methods are presented in Prioni and Hensher (1999).

References

Cunningham, L.F., Young, C. and Lee, M. (1997) Developing customer-based measures of overall transportation service quality in Colorado: Quantitative and qualitative approaches, *Journal of Public Transportation*, 1 (4), 1-22.

Domberger, S., Hall, C., Ah Lik Li, E. (1995) The Determinants of Price and Quality in Competitively Tendered Contracts, *The Economic Journal*, 105, 1454-70.

Fielding, G.J., Babitsky, T.J. and Brenner, M.E. (1985) Performance evaluation for bus transit, in Hensher D.A. (Ed.) Competition and Ownership of Public Transit, A Special Issue of *Transportation Research*, 19 (1), 73-82.

Glaister, S., Cox, W. (1991) The Bidding Process (for the supply of bus services), *Transportation Planning and Technology*, 15 (2/4), 145-52.

Hensher, D.A. (1991) Hierarchical Stated Response Designs and Estimation in the Context of Bus Use Preferences, *Logistics and Transportation Reviews*, 26 (4), December, 299-323

Hensher, D.A. (1994) Stated preference analysis of travel choices: the state of the practice, *Transportation, 21 (2),* 107-33.

Hensher, D.A. and Daniels, R. (1995) Productivity Measurement in the Urban Bus Sector, *Transport Policy*, 2 (3), July, 179-94.

Hensher, D.A. Louviere, J.J. and Swait, J. (1999) Combining Sources of Preference Data, *Journal of Econometrics*, 89, 197-221.

Louviere, J.J. and Hensher, D.A. (1982) On the Design and Analysis of Simulated Choice or Allocation Experiments in Travel Choice Modelling. *Transportation Research Record* No. 890, 11-17.

Louviere J., and Woodworth G. (1983): Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data, *Journal of Marketing Research*, 20: pp350-36.

Louviere J., and Hensher D.A. (1983) Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event, *Journal of Consumer Research*, 10 (3), pp348-361.

Louviere, J.J., Hensher, D.A. and Swait, J. (in press) *Stated Choice Methods: Analysis and Applications in Marketing, Transportation and Environmental Valuation*, Cambridge University Press, Cambridge, 697pp.

Prioni, P. and Hensher, D.A. (1999) *Measuring Service Quality and Evaluating its Influence on Costs of Service Provision*, Institute of Transport Studies, Faculty of Economics, The University of Sydney, July.

Swanson, J., Ampt, L. and Jones, P. (1997) Measuring Bus Passenger Preferences, *Traffic Engineering and Control*, June, 330-36.

Van de Velde, D.M., Sleuwaegen, L.I.E. (1997) Public Transport Service Contracts: Searching for the Optimum, *International Journal of Transport Economics*, 24 (1), 53-74.

White, P., Tough, S. (1995) Alternative Tendering Systems and Deregulation in Britain, *Journal of Transport Economics and Policy*, 29 (3), 275-89.

ITS

INSTITUTE OF TRANSPORT STUDIES

The Australian Key Centre in Transport Management

The University of Sydney and Monash University