

# Measuring Service Quality in Scheduled Bus Services

Paola Prioni and David A. Hensher  
University of Sydney

## Abstract

*Transit operators are increasingly being asked to provide services more cost efficiently. To do this, operators must find ways to reduce the costs of delivering a given level of service (in contrast to simply reducing cost at the expense of lower service levels). There is growing concern in many countries that operators are not focusing enough attention on identifying whether passengers are satisfied with existing service levels and what might be done to increase or at least preserve current service quality.*

*This article develops a stated preference model of service quality choice that provides the set of indicators required to represent a user-based measure of service quality. The service quality index (SQI) provides an operationally appealing measure of service effectiveness to assist regulators in administering and monitoring a performance assessment regime and operators in improving customer service. SQI has been readily accepted by many bus operators in New South Wales (Australia) as the preferred way of establishing and monitoring the effectiveness of service levels, in contrast to traditional stand-alone satisfaction scores based on independent assessment of each attribute in isolation from the entire service package that passengers actually experience.*

## Introduction

Although transportation researchers recognize output in all transportation industries as being multiproduct and multidimensional, empirical cost and demand studies mainly restrict the specification of output to simple physical

measures. Output in the bus industry has been measured in vehicle-kilometers, and alternatively in passenger-trips (Berechman 1987) or in bus-miles (Viton 1986). Other empirical studies have recognized the importance of the shape of the network (Windle 1988; Filippini et al. 1992). Examples of network descriptions are the length of the network (e.g., Caves et al. 1985) or the number of stops (or stations) (Filippini et al. 1992). More sophisticated descriptions consider the structure of the network in terms of the graph of a transport network (e.g., Filippini 1991). Using the notions of nodes (stops) and arcs (connections) of a transport network, graph theory enables the derivation of a network indicator that contains information about length and structure of the network.

Several authors (e.g., De Borger 1991) have pursued another approach based on hedonic output aggregator functions, specifying the cost function as a function of some output function.<sup>1</sup> De Borger, for example, specifies two Cobb Douglas-type functions for  $F$  (freight-kilometers) and  $R$  (passenger-kilometers) for Belgian railroads. The cost function is estimated jointly with the aggregator function. This procedure is usually less parameter intensive and allows the use of flexible functional forms (Jara-Diaz 1982).

Despite attempts to include service quality variables in cost and demand function estimation, and the large amount of literature on the underlying dimensions of service quality in user preferences surveys (e.g., Pullen 1993; Swanson et al. 1997; Cunningham et al. 1997; Kittleson and Associates 1996), previous cost and demand studies simplify the concept of service quality and measure it without input from service users. This article focuses on the development of an SQI, which was derived from a revealed preference discrete choice model, as a representation of the quality level currently on offer. The index, enriched by a stated preference experiment, shows that changes in service levels have a direct influence on costs as quality represents an input that has costs and an indirect influence via SQI and its parameterization. This, in turn, influences the demand for bus travel, which itself influences costs in a model in which physical output is defined by final demand.

Data for the empirical model system are sourced from a sample of private urban bus operators and their passengers in New South Wales.<sup>2</sup> These operators provide services under a contractual arrangement to the state government that

gives them sole authority to provide local route services within their operating area, subject to compliance with minimum service levels and maximum average age of the fleet (for full details of the contractual conditions, see Brewer and Hensher 1997). Operators compete with rail and car, but not with other bus operations servicing the scheduled route market. They do compete in the markets for tour/charter and dedicated school contracts. The results have important consequences for both bus operators and regulators. For bus operators, the results provide an improved understanding of customers' perceptions and use of different service levels as well as the marginal cost of providing best-practice service quality. For transport regulators, the results help establish rules for competitive tendering and definitions of a performance assessment regime.

This article introduces the approach to measuring service quality. It provides estimation results derived from the model and concludes with comments on particular issues.

### **Service Quality**

The concept of service quality includes aspects of transportation service that are not always well defined and easily measured. In this article, quality of service is defined in terms of a set of attributes that each user perceives to be the sources of utility or satisfaction in bus use. The dimensions of quality, viewed from a customer's perspective, are complex. Consumers might consider comfort at the bus stop and the time to get a seat, or only the comfort of the seats. Modal choice surveys have identified a large number of influences on the use of buses in contrast to other private and public modes. Service quality can be divided into six broad classes of effects, each containing different quality dimensions (Hensher 1991). The left column in Table 1 summarizes one possible classification of the main influences on the demand side.

Some demand-side measures can be translated (or mapped) into a set of supply-side equivalences (resources that the operator has partial or total control of) such as timetable, fleet age, and/or buses that are air-conditioned; the number of vehicles that are wheelchair accessible; the number of hours spent cleaning vehicles; and the money spent on driver training.

**Table 1**  
**Demand-Side Effects and Their Equivalence on the Supply Side**

<i>Demand-Side Effects</i>	<i>Supply-Side Equivalence</i>
<p><i>Getting-to-the-bus-stop quality</i></p> <ul style="list-style-type: none"> <li>• ease, safety, time (distance), knowing where the bus stop is</li> </ul>	<p><i>Getting-to-the-bus-stop quality</i></p> <ul style="list-style-type: none"> <li>• frequency, availability of bus shelter and seats</li> </ul>
<p><i>Wait quality</i></p> <ul style="list-style-type: none"> <li>• wait time at stop, punctuality of bus</li> <li>• wait comfort, wait safety</li> </ul>	<p><i>Wait quality</i></p> <ul style="list-style-type: none"> <li>• frequency</li> <li>• availability of bus shelter and seats</li> </ul>
<p><i>Trip quality</i></p> <ul style="list-style-type: none"> <li>• time to board bus</li> <li>• time to get seat</li> <li>• moving to your seat</li> <li>• travel time</li> <li>• trip cost</li> </ul>	<p><i>Trip quality</i></p> <ul style="list-style-type: none"> <li>• frequency, percent of low-floor buses</li> <li>• number of seats available</li> <li>• average speed, network shape</li> <li>• travel time</li> <li>• fare</li> </ul>
<p><i>Vehicle quality</i></p> <ul style="list-style-type: none"> <li>• cleanliness</li> <li>• comfort of seats (types), spaciousness</li> <li>• temperature control (ventilation)</li> <li>• noise</li> <li>• safety</li> <li>• modernity</li> <li>• ease of use for those with disabilities</li> </ul>	<p><i>Vehicle quality</i></p> <ul style="list-style-type: none"> <li>• hours of vehicle cleaning/vehicle</li> <li>• percent of buses with cloth seats</li> <li>• percent of buses with air-conditioning</li> <li>• visual surveillance</li> <li>• average age of the fleet</li> <li>• wheelchair access (yes/no)</li> </ul>
<p><i>Driver quality</i></p> <ul style="list-style-type: none"> <li>• appearance</li> <li>• helpfulness</li> </ul>	<p><i>Driver quality</i></p> <ul style="list-style-type: none"> <li>• years of driving experience, money spent on drivers' training</li> </ul>
<p><i>Information quality</i></p> <ul style="list-style-type: none"> <li>• pretrip information</li> </ul>	<p><i>Information Quality</i></p> <ul style="list-style-type: none"> <li>• availability of timetable/destination</li> </ul>

Source: Brewer and Hensher 1997; Hensher 1991; Swanson et al., 1997.

Supply-side attributes are, in contrast to the quality attributes in the left column, to varying degrees, observable and under the direct control of the bus operator.<sup>3</sup> For example a change in the average fleet size will have a direct influence on the time to get a seat. On the other side, the supplied level of service quality is expected to be a function of consumer preferences. If the supplied quality level is a response to customer preferences, and not only to some regulatory restrictions, quality exogeneity cannot be assumed. In this circumstance, a capability to represent the quality of service as determined by users is needed. The discrete choice approach is an appealing framework.

## **Deriving a User-Based SQI**

Service quality could be achieved by analyzing customer preferences for different levels of bus service quality and deriving the utility of the actual supplied quality level. To this extent, bus travelers' preferences must be identified and quantified. This analysis is restricted to actual bus users; it does recognize that nonusers also provide useful information on levels of service offered by bus operators. Within a performance regime based on the acceptability of service levels to actual users and with a focus on the service quality that influences operator costs, the emphasis on users is appropriate.

In order to determine user preferences for service quality, information is needed on the behavioral responses to a wide range of levels of service quality defined on an extended set of attributes such as those given in Table 1. Revealed preference (RP) data are typically restrictive in their variance properties, but are an important input into the assessment. The preferred approach is a stated preference (SP) experiment combined with existing levels of service. A sampled passenger would evaluate a number of alternative service levels together with the level experienced and choose the most preferred. Systematically varying the levels of the attributes in repeated experiments provides a profile of each passenger's preferences for bus services. The data are analyzed as a discrete choice model in which the SP and RP data are combined to obtain estimated parameters for each attribute. The simple multinomial logit (MNL) model is estimated in which all random components are independently and identically distributed (IID).<sup>4</sup>

An SQI for each bus firm can be derived from the application of the parameter estimates to the current RP levels that each operator-specific passenger sample currently experiences. This index is not a probability (of choice) weighted indicator that is typically derived from a choice model (and referred to as the inclusive value or expected maximum utility index); rather it is an indicator based solely on the levels of service currently on offer. The SP-RP model's role is to provide a rich set of parameter estimates to weight each attribute of service quality.

Such a measure is useful as a stand-alone index of passenger satisfaction for a performance assessment regime (PAR). The role of service quality is not

only as an input, which influences operational costs, but it also represents an important determinant of passenger travel demand. One of the principal difficulties in passenger travel demand studies is the specification of the relevant set of service attributes. Previous empirical studies on passenger demand restricted service quality specification to some measurable characteristics of the supplied service, which are normally selected from a limited set of observable variables. The SQI derived here is based on user perception for different quality levels and not on some set of ad hoc plausible quality attributes. It is passenger perception of quality changes that has an impact on final demand and not the physical changes in some output characteristics.

Given its users-character, SQI can enter as a theoretically valid argument in a travel demand function. A general form of passenger demand can, therefore, be written as:

$$y = d(y_s, sqi_1, \dots, sqi_k, c, m, r) \quad (1)$$

Or alternatively:

$$y = d(y_s, SQI, c, m, r) \quad (2)$$

The first specification assumes passenger travel demand as a function of the physical output  $y_s$ , the cost of the competing mode  $c$ , the income  $m$ , some socio-economic variables  $r$ , and the parameterized quality attributes  $sqi_1, \dots, sqi_k$ . The  $sqi$ 's are the components of the SQI and represent the weighted quality attributes levels resulting from the discrete choice model estimation. The alternative specification (2) considers service quality in its weighted aggregate form (SQI).

## Empirical Approach

The choice model requires data from a sample of bus users on their preferences for different bus service levels. An appealing paradigm, consistent with the economic theory, to obtain robust models of consumer behavior is the combination of RP and SP data. The merging of these two data sources has been successfully applied in several empirical studies (e.g., Hensher, Louviere, and

Swait 1999; Louviere, Hensher, and Swait 2000; Hensher 1994). The major advantage of SP data compared with RP data is that they exploit a more extensive attributes space. RP (or market) data vary only within the frontier of the existing alternatives, restricting the variability of the quality attributes. Even within the technology frontier, the variability of service attribute levels in real markets is typically limited. Increasing the attributes' range through SP:

- improves the identification of the willingness to pay for a particular service attribute,
- reduces the risk of confounding correlation between attributes, and
- produces more robust parameters in the discrete choice model estimation.

RP data have still to be considered because they represent both an anchor around which the SP attributes levels are systematically varied and an important input into the evaluation of the expected maximum satisfaction associated with a specific level of service quality (i.e., the SQI of a particular bus operator).

The following sections outline the major phases in choice modeling. They discuss the choice of the relevant attributes and their levels in the SP experiment and provide details on the survey and the experimental design.

### ***Attribute Selection and Levels for the Experimental Design***

To assist in the selection of attributes for the stated preference experiment, a broad-based pilot survey instrument was designed that listed the full set of service attributes (around 40) that the literature suggests are potential influences on bus users' preferences. Thirty-nine private operators in Sydney were faxed the instrument and asked to evaluate each attribute on a 0-to-100 scale, using their own experience in assessing the relative importance to their passengers.

Eighteen operators responded to the pilot survey. The attributes' rankings were analyzed using descriptive statistics, frequency tables, and factor analysis. To keep the experiment to a manageable size, the number of attributes was restricted to 13. The selection of the key 13 service attributes for the SP design was based on findings from the pilot survey, a literature research, and interviews with bus industry specialists. Table 2 presents the final set of the 13 attributes and their levels.

Each attribute has three levels. The levels are defined to be realistic with values varying in a broad interval to ensure clear trade-offs between attributes. The number in parentheses corresponds to the associated code in the SP experiment.<sup>5</sup>

### ***The Survey***

The sample selection is restricted to scheduled<sup>6</sup> bus users of 25 private urban bus operators in New South Wales. Since this study focuses on assessing traveler preferences or satisfaction for service quality for a specific bus service, the exclusion of other modes is defensible. The survey was conducted as a simple “paper-and-pencil” questionnaire.<sup>7</sup> In the revealed preference part of the study, scheduled bus users reported details of their current trip in addition to some socioeconomic characteristics. Survey forms were distributed and collected during April and May 1999.

### ***Statistical Design of the SP Instrument***

Through a formal statistical design, the attribute levels are combined into bus packages before being translated into a survey form. The full factorial design (i.e., all possible bus packages) consists of  $3^{13}$  combinations of the 13 attributes in each of three levels. To produce a practicable and understandable design for the respondents, the number of combinations was restricted to 81 (i.e., 81 choice sets) using a fractional design. Fractional designs permit the reduction in the number of combinations (i.e., the number of bus packages) without losing important statistical information (see Louviere, Hensher, and Swait 2000).

The pretest of the survey showed that respondents were able to evaluate consistently three choice sets (i.e., different scenarios of bus packages). The researchers reduced the number of different survey forms to 27. Each of the 32 bus operators received 8 survey sets each consisting of 27 different survey forms or  $27 \times 8 = 216$  survey forms and instructions on how to organize the survey. An example of a choice set consisting of three paired comparisons is given in Table 3.

Each of the three coded lines in Table 3 corresponds to a choice set or a scenario. The values 0, 1, and 2 correspond to the levels assigned to an attribute



**Table 2**  
**The Set of Attributes and Attributes Levels in the SP Experiment**

<i>Attribute</i>	<i>Levels</i>	<i>Interpretation of Levels</i>
Reliability	3	<ul style="list-style-type: none"> <li>• On time (2)</li> <li>• 5 minutes late (1)</li> <li>• 10 minutes late (0)</li> </ul>
Fare	3	<ul style="list-style-type: none"> <li>• 25% more than the current one-way fare (2)</li> <li>• Same as now (1)</li> <li>• 25% less than the current one-way fare (0)</li> </ul>
Walking distance to the bus stop (in minutes)	3	<ul style="list-style-type: none"> <li>• Same as now (0)</li> <li>• 5 minutes more (1)</li> <li>• 10 minutes more (2)</li> </ul>
Waiting safety	3	<ul style="list-style-type: none"> <li>• Very safe (2)</li> <li>• Reasonably safe (1)</li> <li>• Reasonably unsafe (0)</li> </ul>
Travel time	3	<ul style="list-style-type: none"> <li>• 25% quicker than the current travel time (0)</li> <li>• Same as now (1)</li> <li>• 25% longer than the current travel time (2)</li> </ul>
Bus stop facilities	3	<ul style="list-style-type: none"> <li>• Bus shelter with seats (2)</li> <li>• Seats only (1)</li> <li>• No shelter or seats at all (0)</li> </ul>
Air-conditioning	3	<ul style="list-style-type: none"> <li>• Available with no surcharge (2)</li> <li>• Available with a surcharge of 20% on existing one-way fare (1)</li> <li>• Not available (0)</li> </ul>
Information at the bus stop	3	<ul style="list-style-type: none"> <li>• Timetable and map (2)</li> <li>• Timetable but no map (1)</li> <li>• None (0)</li> </ul>
Frequency	3	<ul style="list-style-type: none"> <li>• Every 15 minutes (2)</li> <li>• Every 30 minutes (1)</li> <li>• Every 60 minutes (0)</li> </ul>
Safety on board	3	<ul style="list-style-type: none"> <li>• The ride is very smooth with no sudden braking (2)</li> <li>• The ride is generally smooth with rare sudden braking (1)</li> <li>• The ride is jerky; sudden braking occurs often (0)</li> </ul>
Cleanliness of seats	3	<ul style="list-style-type: none"> <li>• Very clean (2)</li> <li>• Clean enough (1)</li> <li>• Not clean enough (0)</li> </ul>
Access to the bus	3	<ul style="list-style-type: none"> <li>• Wide entry with no steps (2)</li> <li>• Wide entry with 2 steps (1)</li> <li>• Narrow entry with 4 steps (0)</li> </ul>
Driver attitude	3	<ul style="list-style-type: none"> <li>• Very friendly (2)</li> <li>• Friendly enough (1)</li> <li>• Very unfriendly (0)</li> </ul>

**Table 3**  
**Example of Three Choice sets**

<i>Bus Package of Bus Company A</i>	<i>Bus Package of Bus Company B</i>
0 1 1 0 2 0 1 1 2 0 1 2 1	2 1 1 1 0 1 1 1 1 0 1 1 2
0 1 1 0 1 0 2 2 1 0 2 1 2	1 2 2 2 0 2 2 2 2 0 2 2 1
0 1 1 1 1 1 0 1 1 1 1 1 1	1 1 0 1 1 0 1 1 0 1 0 1 1

(see Table 2). For example, row 1 begins with a 0, which is the low level (10 minutes late) of the attribute “reliability” for the bus package of Bus Company A in the first scenario. In the SP experiment, passengers were asked to evaluate two other bus packages in addition to the service level experienced on their current trip and indicate which of the three bus packages they preferred. An example of a SP question is shown in Figure 1 with a response question.

Service Feature	Bus Package of Bus Company A	Bus Package of Bus Company B
Reliability	On time	10 minutes late
One-way fare	25% less than the current one-way fare	Same as now
Walking distance to the bus stop	10 minutes more than now	Same as now
Waiting safety	Reasonably unsafe	Reasonably unsafe
Travel time	Same as now	25% longer than the current travel time
Bus stop facilities	Seats only	Seats only
Air-conditioning	Available with a surcharge of 20% on existing one-way fare	Not available
Information at the bus stop	Timetable and map	None
Frequency	Every 60 minutes	Every 15 minutes
Safety on board	The ride is very smooth with no sudden braking	The ride is jerky; sudden braking occurs often
Cleanliness of seats	Very clean	Clean enough
Ease of access to the bus	Wide entry with no steps inside the bus	Wide entry with 2 steps inside the bus
Driver attitude	Friendly enough	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 traveling on

**Figure 1. Example of the SP instrument for one choice set**

### Empirical Findings

This section examines the results of the user preference model. It also discusses the SQI for each operator.

#### Results of the User Preference Model

Twenty-five operators returned data from the onboard RP–SP user survey. Each operator was invited to collect the data over an eight-week period in April–May 1999. A total of 3,849 useable observations (out of 4,334 returns)

was incorporated in the estimation of the discrete choice model. An 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 unlabeled alternatives. Consequently all design attributes were given the same parameter estimates (or weights) across the three alternatives. In addition, the current trip alternative considered alternative-specific characteristics of the passenger (income, gender, age, and car availability) and of the operator, together with a number of other potential influences on relative utility such as treatment effect, trip purpose, and access mode.

The final user attribute choice model is summarized in Table 4. The model includes the attributes of the SP experiment, operator-specific dummy (i.e., 1, 0) variables and three user characteristics. The overall goodness of fit (adjusted pseudo  $R_2$ ) of the model is 0.324. The service attributes provide very important information on the contribution of each service dimension to overall service quality (equating service quality from a user perspective to the derived utility from bus use). The 13 service attributes have been specified as either continuous, where they are ratio scaled, or as dummy variables on each attribute level relative to a base level. Reliability, fare, access time, and bus travel time are ratio scaled and enter each utility expression as stand-alone attributes. The other 9 attributes, each of three levels, are represented by 18 variables in the choice model. Frequency of service, although a potentially continuous variable, has been specified as two dummy variables for reasons given below.

The great majority of the SP attributes are statistically significant. Service reliability (i.e., the extent to which buses arrive on time), fares, access time, and travel time are all highly significant with the expected negative sign. The value of bus travel time savings implied by the ratio of the parameter estimate of travel time to fare is \$4.01 per person-hour; and the value of access (to bus) travel time savings is \$5.39 per person-hour. This is impressive, lining up closely with the evidence from other studies. This adds much credence to the empirical outputs. When the dummy attributes are considered, the researchers systematically found plausible results. Relative to “reasonably unsafe,” they found a positive (almost) significant parameter estimate for “reasonably safe” (0.1510) and for “very safe” (0.1889). The higher estimate for “very safe” in

**Table 4**  
**Final User Preference Model**

<i>Variable</i>	<i>Units</i>	<i>Acronym</i>	<i>Parameter</i>	<i>t-value</i>
Reliability	mins	RELI	-0.05821	-8.411
Bus fare	\$	TARIF	-0.4780	-6.406
Access time	mins	ACCESST	-0.04317	-5.311
Bus time	mins	TRATIM	-0.03200	-5.435
Very safe	1,0	VSAFE	0.18895	2.255
Reasonably safe	1,0	RSAFE	0.15108	1.820
Seats only at bus stop	1,0	SEATS	-0.03411	-0.510
Seat plus shelter	1,0	SEATSHEL	0.09040	1.503
Air-conditioning free	1,0	AVALFREE	0.07131	1.112
Air-conditioning at 20% extra fare	1,0	AVALPAY	-0.17432	-2.207
Ride generally smooth	1,0	GSBRAKE	0.20788	2.963
Ride very smooth	1,0	VSNBRAKE	0.35232	4.904
Clean enough	1,0	CENOUGH	0.13867	1.830
Very clean	1,0	VCLEAN	0.20446	2.713
Wide entry with 2 steps	1,0	WIDE2STP	0.09589	1.499
Wide entry with no steps	1,0	WIDENSTP	-0.10319	-1.372
Driver friendly enough	1,0	FRIENDEN	0.19798	2.572
Driver very friendly	1,0	VFRIEND	0.42287	5.564
Timetable only	1,0	TIMNOMAP	0.29609	4.745
Timetable and map	1,0	TIMWMAP	0.19720	3.021
Frequency/every 60 minutes	1,0	FREQ60	-0.58595	-6.902
Frequency/every 30 minutes	1,0	FREQ30	-0.12221	-1.640
Female	1,0	FEMALE	0.09986	1.198
Personal income	\$'000s	PINCO	0.00905	3.817
Age of passenger	years	AGES	0.01379	5.787
Operator 1	1,0	Op1	0.37358	1.671
Operator 2	1,0	Op2	0.19642	0.654
Operator 3	1,0	Op3	-0.94098	-5.497
Operator 4	1,0	Op4	-0.17726	-1.080
Operator 5	1,0	Op5	-0.12964	-0.653
Operator 6	1,0	Op6	0.97267	1.937
Operator 7	1,0	Op7	-0.18127	-0.982
Operator 8	1,0	Op8	0.35723	1.294
Operator 9	1,0	Op9	-0.26210	-1.215
Operator 10	1,0	Op10	-0.56626	-1.845
Operator 11	1,0	Op11	-1.2555	-4.850
Operator 12	1,0	Op12	-0.22189	-0.842
Operator 13	1,0	Op13	-0.47366	-1.210
Operator 14	1,0	Op14	0.01784	0.072
Operator 15	1,0	Op15	0.06911	0.084
Operator 16	1,0	Op16	-0.37973	-1.685
Operator 17	1,0	Op17	0.06878	0.292
Operator 18	1,0	Op18	-0.36574	-0.825
Operator 19	1,0	Op19	1.1207	4.218
Operator 20	1,0	Op20	0.10014	0.488
Operator 21	1,0	Op21	0.11275	0.546
Operator 22	1,0	Op22	0.32239	0.781
Operator 23	1,0	Op23	-0.53292	-1.845
Operator 24	1,0	Op24	0.08878	0.161
Log-likelihood	2839.25			
Pseudo $R^2$ (adjusted)	0.324			

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. “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 percent surcharge on existing fares is statistically significant with a negative sign, suggesting that users would prefer not to have air-conditioning if it meant paying higher fares.

Onboard safety, defined by smoothness of the ride, is a statistically strong attribute. Relative to “the ride is jerky with sudden braking occurring often,” researchers found 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 for “very clean” relative to “not clean enough.” The nonstatistical (1.830) significance of “clean enough” suggests 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 enough to not be concerned with the configuration of steps and entry widths. Driver attitude is a statistically strong influence on a user’s perception of service quality. Indeed, relative to “very unfriendly” a significant increase might be expected in the mean parameter estimate when going from “friendly enough” to “very friendly.” This is the most nonlinear 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-minute 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. Generally, older individuals and those with higher incomes preferred the levels of service offered by the existing trip than by the alternative packages. This suggests that as individuals age and their incomes rise, they see existing service quality as increasingly satisfying their requirements. Alternatively, younger users and those on lower incomes perceive a greater need for improved service quality. Car availability was not statistically significant.

The researchers investigated the potential for systematic bias due to the sequence in which the SP treatments were given on the survey instrument. There was no evidence of bias in selection from the current and two alternative service packages. The researchers also analyzed possible effects of the survey administration since a range of data collection procedures were implemented across the 25 operators. For example, both drivers and inspectors were involved in the forms' distribution. A series of dummy variables were introduced to distinguish distribution and collection by (a) the driver, (b) an inspector who stayed on board and explained the survey, and (c) an inspector handing out forms with a reply post-paid envelope to return the forms at a later date. The distribution and collection procedure was not a statistically significant influence on the choices made by respondents, despite the suggestion from some bus operators that the responses would be systematically biased (in favor of current service) by an approach that may appear to be coercing passengers to participate.

Trip purpose, with the exception of commuting, did not statistically impact choice, while commuting was marginally significant. Commuting was excluded from the final model in order to limit the amount of data needed to construct an SQI for operators who were not participating in the SP survey. As a consequence, the commuter effect is absorbed into the operator-specific dummy variables (assigned to the existing trip alternative). With 25 bus opera-

tors, the researchers had 24 operator-specific effects. These effects account for other influences on choice that are unique to each operator. A negative sign on the parameter estimate implies that a bus operator is perceived by users as delivering a quality of service that is, relative to the base operator, worse. By comparing the absolute magnitude of the parameter estimate, one can see the extent to which an operator is delivering a service that is worse than other operators after allowing for the attributes explicitly taken into account from the SP experiment. Table 4 shows that Operators 3 and 11 have the highest negative operator-specific parameter estimates while Operators 1, 6, and 19 have the highest positive operator-specific estimates.

### ***The Service Quality Index***

The SQI for each operator can be calculated by the application of the utility expression in Table 5 and the levels of each of the attributes associated with the current trip experience of each sampled passenger. In this initial study, the researchers estimated a single set of utility weights across the sample of 3,849 passengers using the services of 25 operators. They investigated possibilities of differences in weights between segments of operators (e.g., 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 2 of the mean and standard deviation of each attribute for each operator). Rather, it shows that the marginal utility of each attribute (i.e., 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 summarized in Table 5 and graphed in Figure 2 at its mean for each operator. The researchers normalized the SQI in Figure 1 to a base of 0 for the operator with the lowest relative SQI. The range is from 0 to 2.70. This estimate is the SQI indicator imported into the passenger-demand model.

From the parameter estimates, other interesting results can be derived. Figure 3 shows the contribution (in terms of utility) of each single, quality

**Table 5**  
**Summary Statistics of Service Quality Index**

<i>Operator</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Sample Size</i>
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

Note: The starred sample sizes are too small to be able to infer any substantive evidence from these specific operators. Ongoing survey research has boosted these numbers to enable us to investigate the performance of these operators in more detail.

attribute over the entire sample.<sup>9</sup> Tariff (UTARIF), travel time (UTRATIM), and access time (UACCESST) have the highest impact on service quality. On the positive side of SQI, the major influence is given by the friendliness of the driver (UVFRIEND) and the smoothness of the ride (UVSNBRAK).

### **Service Effectiveness and Contracts<sup>10</sup>**

Table 6 gives an example on how to integrate SQI targets into a contracting process. It assumes that from a survey of a sample of existing users, the user-defined quality of current service of three operators has been identified.



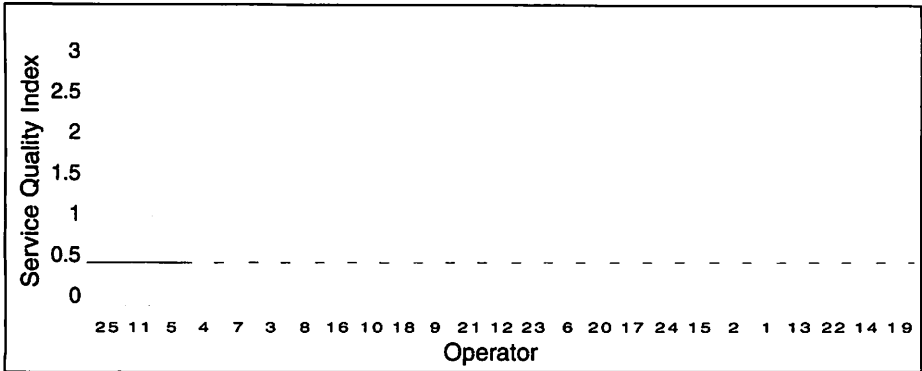


Figure 2. The benchmarking of mean SQI across operators (operator 25 set =0)

Operator 1 achieved an SQI of 1.4 by providing a service that is on average two minutes late, clean enough for 60 percent of the sampled users, costs on average \$2.1, etc. Operators 2 and 3 have SQIs, respectively, of 1.3 and 2.0. Assuming that these operators are comparable, Operator 3 exemplifies best practice.

In many countries, monopolistic bus operating conditions exist in urban areas. Urban areas also have vastly different characteristics with respect to key factors such as travel distances, topography, urban structure, etc. Often there exists a situation of few and protected operators, each operating under usually vastly different conditions. In light of this, it may be difficult to determine best practice by comparing current operators. Under these conditions, a slight variation in

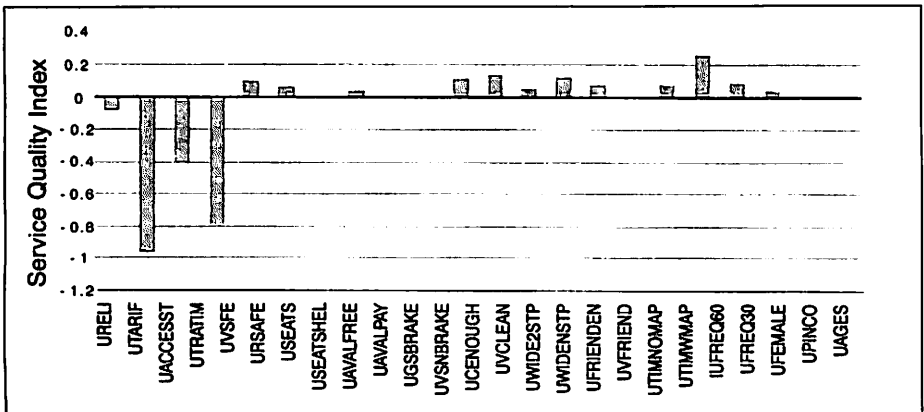


Figure 3. SQI composition

approach may be to establish an existing SQI for the current operator in each area and use this as the starting point from which service delivery must be improved.

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 operator must be at best practice, this may discourage bidders and so a target level is set that is recognized as achievable by potential bidders. The level should be incentive compatible.

Given the gap between an operator's SQI and that of best practice (e.g., 0.6

**Table 6**  
**Including SQI Targets in the Contact Design**

Operator	Current Service Description Attributes					SQI Target after		
	Reliability	Bus Fare	Clean Enough	Travel Time	Etc.	Realized	2.5 Years	5 Years
1	2 minutes late	2.1	60%	25 minutes	...	1.4	1.6	1.8
2	1 minute late	2.4	78%	26 minutes	...	1.3		
3	1 minute late	2.0	80%	21 minutes	...	2.0		

for Operator 3), the researchers propose a formulation  $SQI+z$  where  $z$  is the predesignated improvement over a period of time (e.g., 0.2 in both subperiods). The  $SQI+z$  formula provides a target in line with a predesignated increase in the service quality level. In the case of the service previously provided by incumbent Operator 1, authorities might 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).

The required service quality level can 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 internalized). Once successful in winning the contract, the operator has a strong incentive to meet the new levels of service.

### **Monitoring and Responses**

To ensure contract compliance, the supplier must be monitored during the contract period. Assuming a contract length of 5 years, a performance assessment should be imposed 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 transaction costs of the bid and include it in the bid price. Table 7 summarizes the four possible outcomes of a contractual process.

If the operator is compliant, it becomes a political decision whether the

<b>Table 7</b>		
<b>Possible Outcomes of a Tender</b>		
	<i>Renewal</i>	<i>Retender</i>
Compliant	End of the 5 years	End of the 5 years
Noncompliant	Retendered	Retender: Warning after 2.5 years

contract will be renewed or retendered at the end of the contract period. In the case of noncompliance after the first half of the contract period, the noncompliant operator should be warned about underperformance without losing 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 noncompliant operator, the tendering authority must determine if the reasons for noncompliance are internal to the contractor (i.e., under its control) or external (i.e., not under operator control). Only internal failure needs to be corrected through sanctions. In the case of external factors influencing the operator's service quality, the tendering authority should review the preagreed targets.

The extent of benefits from competitive tendering depend not only on the size of the targeted SQI, 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 are 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

This article has presented a new approach to measuring service quality for incorporation of a user-based SQI in a passenger-demand and operator-cost function. An SQI enables the regulator and bus operator to benchmark service effectiveness, adding this much-neglected dimension of performance assess-

ment. A gap in the literature is now filled on performance measurement, establishing a global measure of service effectiveness to parallel the global indicators used to measure cost efficiency and cost effectiveness (i.e., total factor productivity) (see Hensher 1992; Fielding et al. 1985). While this article has focused on existing users, the method can be implemented for potential users, although the costs of data collection will increase.<sup>11</sup>

The parameter estimates identified in the development of the SQI can be implemented in bus operation contexts, which are comparable to the range of service levels evaluated in the stated preference survey detailed above. Where an operator exhibits service levels noticeably different to this range, a new SP survey would need to be undertaken, using the exact method developed here. Regulators wishing to implement a performance assessment regime based on a subset of the attributes evaluated in this article can select the subset of parameter estimates and derive a partial SQI indicator, without the need to reestimate the user preference model.

### **Acknowledgments**

An early version of this article was presented at the sixth International Conference on Competition and Ownership of Land Passenger Transport, Cape Town, September 19–23, 1999. The authors are indebted to the Bus and Coach Association of New South Wales and the 25 bus operators who participated in the study. The SQI developed herein provides a way of implementing the performance assessment regime to be introduced in the state. The comments of three referees are much appreciated.

### **Endnotes**

1. Spady and Friedlaender (1978) used the hedonic approach first in the U.S. trucking industry.
2. The private bus industry has 39 percent of the share of the urban transit passenger task (the balance being 24% public operator, 36% train, and 1% ferry). Private bus operators service all of the suburban areas of the major cities with government operators servicing the central areas. The operators in this study receive no capital or operating subsidy but do receive substantial reimbursements from the state government under the school subsidy travel scheme based on the number of students with bus passes. State governments in Australia are responsible for con-

tracting out services to private operators and monitoring the operators' performance to ensure that they satisfy the terms of their contract. In contrast to the United States, state governments set policy and channel federal funds to the operators within their state. Any performance monitoring activities are generally tied to federal reporting requirements and not to the continued provision of funding. Culturally, the passenger market is typical of any major Western city in which the car dominates and is certainly comparable to cities in the United States, Canada, and Britain.

3. Examples of influences not under the control of the operator are pedestrian environment, transit priority provision, and traffic congestion.
4. The MNL and more advanced methods are discussed in detail in Louviere, Hensher, and Swait (2000). The MNL model is a simple choice model that evaluates choices based on the relative levels of the attributes of the alternatives in the choice set. Weights for each attribute that represent the relative importance of each attribute to the overall satisfaction of each alternative are established through statistical analysis.
5. The attribute "air-conditioning" was added in a second stage, as experts found that the availability of air-conditioning on buses influences operating cost and is an important policy issue in New South Wales. Similarly, the attribute "access to the bus" was introduced and defined by the number of steps and the width of the bus entrance as suggested by some specialists.
6. Children traveling to school were excluded from the sample.
7. The survey method is very cost effective. Each bus operator was given the survey forms at a briefing session and asked to distribute them according to a sampling scheme. A number of collection procedures were available. The university funded the survey design, sampling, data entry, and analysis and reporting. The cost of the study per operator if all costs are taken into account would have been no more than U.S. \$5,000, although on this occasion most costs were met by the Institute of Transport Studies.
8. Indeed, 0.324 is a very good fit for a nonlinear choice model. It cannot be compared to the  $R^2$  in a simple linear regression where much higher explanatory power is expected. Typically, a pseudo  $R^2$  between 0.2 and 0.4 in a nonlinear discrete choice model is equivalent to an  $R^2$  in a linear model of between 0.7 and 0.85 (see Louviere, Hensher, and Swait 2000).
9. See Table 4 for a complete list of the attribute definitions.

10. See Hensher and Prioni (2000) for further details.
11. A nonuser survey can be undertaken as face-to-face or a mail/fax out and mail/fax back. Since the user survey was a self-completion survey, it is highly likely that a self-administration data collection strategy is feasible. While the formula will be the same, the weights might be expected to differ, providing useful information for a marketing campaign.

## References

- Berechman, J. 1987. Cost structure and production technology in transit: An application to the Israeli bus transit sector. *Regional Science and Urban Economics* 17 (4): 519–534.
- Brewer, A., and D. A. Hensher. 1997. *Operating a bus and coach business: Insights and practice*. Sydney: Allen & Unwin.
- Caves, D. W., L. R. Christensen, M. W. Tretheway, and R. J. Windle. 1985. Network effects and the measurement of economies of scale and density for U.S. railroads. In Daughety, A. F., ed. *Analytical studies in transport economics*. Cambridge: Cambridge University Press, 97–120.
- Cunningham, L. F., C. Young, and M. Lee. 1997. Developing customer-based measures of overall transportation service quality in Colorado: Quantitative and qualitative approaches. *Journal of Public Transportation* 1(4): 1–22.
- De Borger, B. L. 1991. Hedonic versus homogeneous output specifications of railroad technology: Belgian railroads 1950–1986. *Transportation Research A* 25A(4): 227–238.
- Fielding, G. J., T. J. Babitsky, and M. E. Brenner. 1985. Performance evaluation for bus transit. *Transportation Research* 19(1): 73–82.
- Filippini, M. 1991. *La struttura dei costi delle ferrovie private Svizzere secondo la teoria del duale: Implicazioni per una politica delle fusioni*. Zuercher Dissertation, Zentralstelle der Studentenschaft, Zuerich.
- Filippini, M., R. Maggi, and P. Prioni. 1992. Inefficiency in a regulated industry—The case of the Swiss regional bus companies. *Annals of Public and Cooperative Economics* 63(3): 437–455.
- Hensher, D. A. 1991. Hierarchical stated response designs and estimation in the context of bus use preferences. *Logistics and Transportation Reviews* 26 (4): 299–323.
- . 1992. Total factor productivity growth and endogenous demand: Establishing a

- benchmark index for the selection of operational performance measures in public bus firms. *Transportation Research* (Special Issue on Performance Measurement) 26B (5): 435–448.
- . 1994. Stated preference analysis of travel choices: The state of the practice. *Transportation* 21 (2): 107–133.
- Hensher, D. A., J. J. Louviere, and J. Swait. 1999. Combining sources of preference data. *Journal of Econometrics* 89: 197–221.
- Hensher, D. A., and P. Prioni. March 2000. *The missing link in contract performance assessment: The integration of a service quality index into a competitive tendering regime*. Sydney: Institute of Transport Studies, University of Sydney.
- Jara-Diaz, S. 1982. The estimation of transport cost functions: A methodological review. *Transport Reviews* 2: 257–278.
- Kittleston and Associates. 1996. Development of transit capacity and quality of service principles, practice, and procedures. *Transit Cooperative Research Program Project A-15*. Washington, DC: Transportation Research Board.
- Louviere, J. J., D. A. Hensher, and J. Swait. 2000. *Stated choice methods, analysis, and applications in marketing, transportation, and environmental valuation*. Cambridge: Cambridge University Press.
- Pullen, W. T. 1993. Definition and measurement of quality of service for local public transport management. *Transport Reviews* 13(3): 247–264.
- Spady, R. H., and A. Friedlaender. 1978. Hedonic cost functions for the regulated trucking industry. *Bell Journal of Economics* 9: 159–179.
- Swanson, J., L. Ampt, and P. Jones. 1997. Measuring bus passenger preferences. *Traffic Engineering and Control* (June): 330–336.
- Viton, P. A. 1986. The question of efficiency in urban bus transportation. *Journal of Regional Science* 26(3): 499–513.
- Windle, R. J. 1988. Transit policy and the cost structure of urban bus transportation. In Dodgson, J. S. and N. Topham, eds. *Bus deregulation and privatization: An international perspective*. Avebury: Aldershot, 119–140.

### About the Authors

DAVID HENSHER (Davidh@its.usyd.edu.au) is professor of management and director of the Institute of Transport Studies at the University of Sydney. He is also affiliated with the Department of Civil Engineering at Monash University. Dr. Hensher is a fellow of the Academy of Social Science in Australia, immediate past president of the International Association of

Travel Behaviour Research and a vice-chair of the International Scientific Committee of the World Conference of Transport Research.

Dr. Hensher is on the editorial boards of 10 of the leading transport journals and area editor of *Transport Reviews*. He has recently been appointed by Elsevier/Pergamon Press as series and volume editor of a new handbook series *Handbooks in Transport*. He has published extensively in the leading international transport journals and key journals in economics, and has written five books.

**PAOLA PRIONI** (Paolap@its.usyd.edu.au) is currently a management consultant in Switzerland. This article was written while Dr. Prioni was a Swiss government sponsored post-doctoral fellow at the Institute of Transport Studies. Her research interests are in public transport costing and service quality measurement. She completed her Ph.D. at the University of Zurich on cost efficiency of urban bus operations in Switzerland.