27th Australasian Transport Research Forum, Adelaide, 29 September – 1 October 2004

Paper Abstract

Paper title: **Designing customer focussed public transport services on a constrained budget**

Author(s) name(s): Andrew G. Parsons[1], Anne Stewart[1], and Claire McAlpine[2]

Organisation(s): (1) Department of Marketiing, University of Auckland Business School and (2) Transport Department, Auckland Regional Council

Contact details:

Postal address: Department of Marketing, University of Auckland Business School, The University of Auckland, Private Bag 92019, Auckland 1, New Zealand

Telephone: +64 9 373 7599 ext 87575 (A Parsons)

Facsimile: +64 9 373 7444

email: a.stewart@auckland.ac.nz
a.parsons@auckland.ac.nz

Abstract (200 words):
A key issue in transport provision is how to determine where investment in a network's service can be most effective, given a budget that does not allow all the investments desired. Methodologically this can also be an issue, as often there are a myriad of options available – too many for the customer to easily trade-off, and too many for traditional approaches such as conjoint analysis. Using both a non-parametric approach and the indices approach advocated by Swanson, Ampt and Jones (1997), 600 rail passengers are examined across 15 alternative scenarios, to derive customer preferences for proposed changes to the Auckland Rail network through evaluation of their affect (liking) for the scenarios and likely resultant behaviour. Fisher's (1974) environmental quality scale is used to assess customer affect.

Introduction

In the situation where a public transport funder cannot support all desired service options because of budget constraints it is important for the funder to consult with passengers as part of efficient governance of public monies (Pina and Torres, 2001). The problem with such an approach to the public users is that often there are too many available options for the respondent to easily grasp, trade-offs are difficult to ascertain because differences make it like comparing apples with oranges, and stated preferences are often not good indicators of actual liking or behaviour.

The regional council responsible for public transport funding / provision across a major urban area needed to identify priorities for rail investment. The council had designed a future station and service delivery that was going to be the standard across the network in the near future. However, there were fourteen service options that were being debated that could be added to the design (the base scenario). The council did not have the budget to include all fourteen however.

This study reports the findings of research that compared two alternative methods for assessing multiple options in a public transport setting, and used the more discriminating to assess options on the basis of affect (liking), stated future behaviour, and projected future behaviour. Competing methods is considered a sound scientific approach when testing unknowns like this (Armstrong, Brodie and Parsons, 2001).

The study

Fourteen alternative scenarios were constructed around a base scenario that described, using text and pictures, the concept of a standard service for the rail network (see figure 1). Each alternative incorporated an optional investment into the service (see table 1 for a list of the options – each was described fully with additional text, and where appropriate, additional drawings). The approach of the research was to get passengers to evaluate a scenario (one of the fourteen or the base) in terms of affect (the liking they had for what was described) and likely resultant behaviour.

Six hundred rail passengers viewed one of fifteen alternative scenarios (n = 40 per scenario). It is important to note that passengers were randomly selected from set populations (i.e. those waiting to catch a train) rather than a convenience intercept sample. On this basis, the assumption is made that random assignment results in each group of 40 having similar distributions of characteristics, and are thus directly comparable.

Affect (consumer liking for something) was measured using Fisher's (1974) environmental scale, made up of thirteen items. These thirteen items, employing 7-point semantic differential scales, were averaged to give a score for the individual's affect. The nearer to one, the more an individual passenger liked the scenario presented to them, the nearer to 7 the less they liked it.

The passengers were also asked to rate the scenario on likely future behaviour by stating how much more (or less) likely they were to use the new service than the current (to assess current passenger patronage changes); and the prototypicality (reflective of the respondent's self-image of what a suitable service is) of the scenario (to project new passenger patronage).

The suburban station you use is an open platform with bright lighting on the platform and approaches during evening hours. A large timetable is displayed at the station, along with a station sign visible from the approaches and from the train carriages. Basic maintenance of the station is checked daily, with repairs undertaken in a timely manner (e.g. potholes repaired, facilities repaired in the reasonable time it takes to undertake such work). There is safety fencing around the station environment, preventing crossing of open tracks. The train arrives within five minutes of the scheduled time, though up to 15% of trains may be delayed. If a train is delayed by more than 12-15 minutes, text messages are sent to the cell-phones of passengers who have listed with the service.

Carriages are fully cleaned internally on a daily basis, with litter removed regularly during the day. Externally the carriages are cleaned twice per week

Figure 1 Base scenario

Recall that we did not wish to simply know preference but also to identify possible trade-offs that would allow the council to make more informed governance of public money under a constrained budget. Whilst the common approach for such a trade-off is something like conjoint analysis, this was not appropriate in this case because research parameters required the study to be done with existing passengers as they were waiting for a train. Even by using a factorial approach the number of alternatives that would need to be presented to passengers would have been too time consuming.

We selected a ranking-based approach using a combination of Kruskal-Wallis and Mann Whitney U tests because the council wanted a ranking that showed absolute and relative rankings of the alternative service options. As a competing method however, Swanson, Ampt and Jones (1997) had approached a similar problem for bus passengers using an indexed approach. The two approaches are discussed next, followed by the results.

The ranking approach

The Kruskal-Wallis test (see Greensted, Jardine, and MacFarlane, 1986, for a clear explanation of this test) is a non-parametric (distribution free) alternative to the one-way analysis of variance (ANOVA) test, which looks for differences amongst three or more samples of data. Churchill and Iacobucci (2002) p650, illustrate diagrammatically when this test should be employed, which is when there are ordinal data for a single variable and the multiple samples can be considered independent. Taking the affect data from the Fisher (1974) scale we can group all the data and rank the scores regardless of the service option they originated from (so, in this case, we have 15 service options with 40 passengers sampled for each option, which gives us scores ranked from 1 to 600 – with ties accommodated). The Kruskal-Wallis test then re-classifies the rank numbers back into the service options, allowing rank sums to be calculated. These rank sums give us our rank order, and it is these that are also then used to test the null hypothesis that there is no significant difference. The chi-square distribution is used as a source of critical values (and permits unequal numbers in each class if required, though this was not necessary in this study), with degrees of freedom being one less than the number of classes (samples) – so 14 in this study, giving a critical value of 23.685.

The problem with leaving it at that is the similarity to the ANOVA. Just as in ANOVA, finding that there is (or is not) a significant difference does not in itself tell us much from a managerial point of view. In our study we want to rank the service options so that we might determine trade-offs. Across the fifteen the distribution may suggest no difference, but between the first ranked and the fifteenth ranked there may be enough difference to be significant. Thus, not only do we want relative position (1^{st}, 2^{nd} etc…), which is given by the Kruskal-Wallis rank sums, but we also want the absolute position – how far is first from second, and third and fourth etc… One way to think of it is that we want to turn something ordinal into something interval. The chi-square critical value only tells us that on average, the differences between first and second, and first and third, and second and third, etc… are significant (or not). From a trade-off managerial point of view we need to know which are the ones that have significant differences and which do not. We can do this by employing another of the non-parametric tests of association, the Mann-Whitney U, which is the equivalent of the parametric t-test (for a comparison of various correlation coefficients and non-parametric test see Babakus and Ferguson, 1988, Kendall and Gibbons, 1990), in a sequential manner to assess the trade-off ability between pairs of service options.

The index approach

Swanson, Ampt and Jones (1997) use an index approach based on mean ‘willingness to pay’ scores. Effectively what they do is estimate ‘monetary values’ for service offerings through the use of multiple linear regression. Dummy variables are employed to represent the presence (or not) of service options, fitting the following model;

$$R = \text{constant} + \sum b_i \delta_i + c.f$$

where R is the response scale (the stated preference), the δ_i are dummy variables representing the presence or absence of service options, and the b_i are parameters to be estimated. For their model the term f is the fare increases that Swanson, Ampt and Jones (1997) have included to ‘pay’ for the service options, while c is another parameter to be estimated. The monetary value for option i is the ratio b/c. Recognising that individuals’ reference points in existing experience will vary, the modellers then take the perception of the current service offering and use that as the start point for comparisons of additional benefit derived from service options. The approach means that constant equivalence between money and the preference scales for each service option does not have to be assumed, and the contextual situation of each individual is taken into account.

Whilst Swanson, Ampt and Jones (1997) include a fare term, by setting this to 1 to allow for the fact that in this case there is no monetary trade-off required by the passenger we simply turn the estimated parameter c into another constant (albeit unknown) which is then computationally subsumed into the constant term at the beginning of the equation. This allows us to then have a direct comparison of service options weighted as dummy variables versus the ranked Kendall’s W approach discussed earlier. Therefore the adapted Swanson, Ampt and Jones (1997) model used to compare with our suggested approach is;

$$R = \text{constant} + \sum b_i \delta_i$$

where the constant includes the $c.f$ term. The ‘willingness-to-pay’ weights used by Swanson, Ampt and Jones(1997) are all equal, and so the indices become simply the beta coefficients indexed against the largest one. Effectively, Swanson, Ampt, and Jones (1997) are employing dummy regression to assess the trade-offs between options based on largest influence, whilst we are employing a non-parametric statistic sequence identifying the absolute and then relative positions of options.

Results

The results of the index approach are given in table 1. The adjusted R-squared was only 0.003 with the constant the only significant coefficient. The constant in this case was the base scenario service option as this was deemed logically the best candidate for omission from the dummy variables (for a straightforward discussion of why this omission is computationally required see Parsons, 2001). It is unclear from Swanson, Ampt and Jones (1997) paper whether their constant, which must also have been one of the service options in dummy variable format, was included in the indexing. In our case, the b for the constant is nearly ten times that of the next largest service option b. Whilst the absolute positions obtained would not change, inclusion of the constant would make relative position assessment less clear. For this reason we decided to omit the constant and index on the next largest coefficient.

While we did not include the base scenario (which was the constant) in the indexing, we show the position in the index of the constant with respect to the change in direction of the *b* coefficients from the regression. Those service options that have a positive index score are ones whose *b* coefficients suggest they would add an incremental affect (liking) value to the base scenario. Those service options that have a negative index score are the ones whose *b* coefficients suggest they would subtract an incremental affect value from the base scenario. This also allows us a better visual comparison of the index results with the non-parametric results shown in table 2.

Table 1 Service options indexed

Service Option (the Scenarios presented)	Indexed Score	t*
Graffiti at station removed within 24 hours	100	-0.762
30 day passes from date of purchase	67	-0.511
Duress alarm/ help post a station	47	-0.358
Toilets accessible from station	34	-0.261
Seating at station (park bench style)	29	-0.216
Friday/Saturday service runs until midnight	21	-0.159
Base Scenario		
Shelter available at station	-15	0.115
Roving security patrols on trains / stations	-51	0.385
Ability to purchase tickets at convenient locations	-122	0.928
PA announcements of delays	-144	1.099
2, 4, or 6 stage passes	-149	1.137
Wayfaring signage	-152	1.160
Remote CCTV	-167	1.275
Pukekohe service on Saturday	-194	1.480

* note: the scale was a reverse scale in that the higher number suggested less liking, so sign direction is affected in the t

Table 2 shows the rank based on the Kruskal-Wallis approach, with the significant differences highlighted based on the Mann-Whitney U. An α of 0.05 was used to be consistent with the Swanson, Ampt and Jones (1997) approach.

Using the index approach, the results can be interpreted one of two ways; we can look at the statistical result and conclude that there is no real difference between any of the service options (the *t*'s translate into no significant *b* coefficients), and that putting in the base scenario will achieve the only significant effect. Alternatively we can look at the managerial result and say that all those positively indexed (above the base scenario in Table 1) contribute to enhancing the base scenario, and should be included if budget allows, whilst all those negatively indexed (below the base scenario) do not contribute so should not be included.

Table 2. Service options ranked

Service Option (the Scenarios presented)	Rank	Difference*
Graffiti at station removed within 24 hours	1	
30 day passes from date of purchase	2	
Seating at station (park bench style)	3	
Duress alarm/ help post a station	4	
Toilets accessible from station	5	
Base Scenario	6	
Friday/Saturday service runs until midnight	7	
Shelter available at station	8	
Roving security patrols on trains / stations	9	
Wayfaring signage	10	
Remote CCTV	11	
PA announcements of delays	12	
Ability to purchase tickets at convenient locations	13	
Pukekohe service on Saturday	14	
2, 4, or 6 stage passes	15	

* This is simply highlighting where the significant differences lie as discussed above

The managerial result is very similar to the managerial interpretation of the rankings in Table 2. Only the midnight service changes position. So, if budget allows, either approach would come to very similar conclusions as to what should be kept in and what should be left out. However, if budget does not allow for so many service options, the statistical analysis in the ranking approach provides a clearer picture. The Mann-Whitney U tests suggest that seven of the options are in that statistical 'no-mans land' where they do not significantly differ from anything and are hovering around the base scenario. These should be considered subject to budgetary whim – if they can be afforded, all well and good, but if not the base scenario satisfies passengers equally as well. Three of the service options however differ significantly from five others. Graffiti removal, 30 day passes, and seating are ranked above the base scenario (in 1st, 2nd, and 3rd respectively), so are considered to enhance the base scenario, and *are significantly different* from the bottom five service options in the case of graffiti removal, the bottom two in the case of 30 day passes, and the last ranked in the case of seating.

Statistically, the Swanson, Ampt and Jones (1997) approach suggests there is no difference between any of the service options, so from a budgeting point of view the decision maker is no better off in choosing which options to keep and which to discard. The ranking approach however suggests that if necessary, keep graffiti removal, 30 day passes, and seating, and we can discard CCTV, PA announcements, convenient ticket locations, a Saturday service for Pukekohe, and 2, 4, or 6 stage passes. This approach therefore gives the decision-maker much clearer choices when making their decision.

Following the analysis for affect, we applied the ranking approach to the two behavioural measures. The results are shown in tables 3 and 4.

Table 3 Behaviour-based ranking

Scenario Ranked on Behaviour (*The higher the ranking, the more likely a passenger will use the service more often than the current service*)	**Rank (in parentheses is the corresponding rank for affect from table 2)**
Purchase tickets at convenient location	1 (13)
Duress / help post	2 (4)
30 day passes valid from date of purchase	3 (2)
BASE	4 (6)
Graffiti removed within 24 hours	5 (1)
Remote CCTV	6 (11)
Pukekohe service on Saturday	7 (14)
Park bench seating	8 (3)
Wayfaring signage	9 (10)
Friday and Saturday service until midnight	10 (7)
Shelter available	11 (8)
Toilets accessible	12 (5)
Roving security patrols	13 (9)
PA announcements of delays	14 (12)
2, 4, or 6 stage passes	15 (15)

Table 4. Prototypicality-based ranking

Scenario Ranked on Prototypicality (*The higher the ranking, the more a passenger sees the service as something people like them would use*)	**Rank (in parentheses is the corresponding rank for affect from table 2)**
Graffiti removed within 24 hours	1 (1)
PA announcements of delays	2 (12)
Friday and Saturday service until midnight	3 (7)
Roving security patrols	4 (9)
BASE	5 (6)
Purchase tickets at convenient location	6 (13)
Toilets accessible	7 (5) *
Duress / help post	7 (4) *
Pukekohe service on Saturday	9 (14)
30 day passes valid from date of purchase	10 (2)
Wayfaring signage	11 (10)
Park bench seating	12 (3)
Shelter available	13 (8)
2, 4, or 6 stage passes	14 (15)
Remote CCTV	15 (11)

* tied

We can see that as the passenger moves from what they like (affect) to what will change their behaviour, there are some important shifts in priorities that the index approach of Swanson, Ampt and Jones (1997) would not identify. These are discussed next in the conclusions.

Conclusions

Table 2 shows a distinctive ranking of the scenarios in terms of what is liked. The things that are liked more are likely to give greater satisfaction scores. So, if it is the aim of the public transport funder to have greater satisfaction amongst current customers, they could look at graffiti removal within 24 hours, introducing 30 day passes, and placing park benches on station platforms, and, if they have the additional budget, having a duress/alarm post at stations, and having toilets accessible from the station for passengers. The ranking-based approach would also suggest that in having these service options, the funder could definitely omit the five lowest ranked options.

If the funder is looking at increasing current passenger use, then having convenient ticket locations should also be considered, if necessary at the expense of the seating and toilets (remembering that if the trade-off is necessary then satisfaction will slip). If the funder is looking at increasing the number of current passengers by attracting similar types of passengers then timetable issues can be addressed in terms of Friday/Saturday night services and announcements regarding delays. This latter measure is useful if a provider only wants to look at incremental improvements to an existing service (as in this case) rather than considering wholesale changes that may attract a different market segment.

On the question of which analytical approach to use, it is clear that the rank-based approach is superior to an index-based approach when trade-offs need to be justified by a public transport funder.

References

Armstrong, J Scott, B, Roderick J and Parsons, A G (2001) Hypotheses in Marketing Science: literature review and publication audit *Marketing Letters 12* (2), 171-187

Babakus, E and Ferguson, C E Jr (1988) On choosing the appropriate measure of association when analysing rating scale data *Journal of the Academy of Marketing Science 16* (Spring), 95-102

Churchill, G A Jr and Iacobucci, D (2002) *Marketing Research: Methodological Foundations* (8th ed) Fort Worth: Harcourt Inc

Fisher, J D (1974) Situation-specific variables as determinants of perceived environmental aesthetic quality and perceived crowdedness *Journal of Research in Personality 8*, 177-188

Kendall, M G and Gibbons, J D (1990) *Rank Correlation Methods* (5th ed) New York: Oxford University Press

Parsons, A G (2001) The association between daily weather and daily shopping patterns *Australasian Marketing Journal 9* (2), 78-84

Parsons, A G (2003) Assessing the effectiveness of shopping mall promotions *International Journal of Retail and Distribution Management 31* (2), 74-79

Pina, V and Torres, L (2001) Analysis of the efficiency of local government services delivery. an application to urban public transport *Transportation Research A 35*, 929-944

Swanson, J, Ampt, E, and Jones, P (1997) Measuring bus passenger preferences *Traffic Engineering and Control*, June, 330-336