

# **SURVEYS TO EVALUATE BRITISH RAIL'S ROLLING STOCK**

*George Terzis and Geoff Copley  
The MVA Consultancy, Woking, England*

## **1. Introduction**

The aim of this paper is to describe the context in which The MVA Consultancy applies the Blaise software to quantitative surveys. In particular, we discuss in some detail a recent application carried out for InterCity to assess demand for rail travel in relation to different types of rolling stock.

InterCity (IC) is one of British Rail's (BR) four passenger businesses mainly catering for medium and long distance travel in England and Wales. It operates to a commercial, profit making remit, with a turnover of around £830 million and in 1990/91 a trading profit of £50 million.

The IC Mark IV stock has been introduced on the East Coast Main Line since 1989. In 1991 the service was established on this route and most passengers had experienced both Mark IV and Mark III (which is also in service on this route) stock. IC required a comprehensive study to understand the perceived differences between the two types of carriages.

The MVA Consultancy was commissioned to carry out research with the following objectives:

- to obtain quantitative evidence of the differences between the two types of carriages as perceived by passengers;
- to isolate the stock effects from journey time improvements and to give overall and detailed valuations of the carriage aspects (appearance, seating, ride quality, etc).

The tight timescale of the study (three months to report) and the decision to customise the interview to suit each respondent's circumstances and perceptions required the use of Computer Assisted Personal Interviews (CAPI). In general in the field of transport research, the advantages of CAPI relate to speed of data collection but also most importantly to increased accuracy of the results and enhanced ability to elicit complex information.

## **2. Methodology**

Usually such research is preceded by a qualitative phase with in-depth interviews, either in group discussions or personal semi-structured interviews on trains, in order to understand the areas of interest and define the relevant attributes. In this case this was not necessary because of extensive past experience and a number of qualitative and quantitative surveys about BR's rolling stock. In particular, The MVA Consultancy had undertaken a similar study in 1985, with a substantial qualitative stage, to assess the differences between the Mark III carriage and its predecessor Mark II (at that time pen and paper interview (PAPI) techniques were used for the main fieldwork). In addition, a number of studies in the UK and in the Netherlands provide enough empirical evidence about the most important attributes of carriage quality.

It was decided to include the following seven aspects of carriage quality in the assessment:

- luggage;
- appearance and cleanliness;
- ride quality;
- seating;
- access/egress;
- toilet facilities;
- environment control.

A questionnaire was designed to collect information in the following areas in broad terms:

## *Surveys to evaluate British Rail's rolling stock*

- passenger and journey details to customise interviews;
- the ranked order of importance of the seven carriage quality aspects;
- the perceived quality levels for each aspect of the carriage where the interview took place;
- the satisfaction level for each aspect;
- the existence of noticeable differences between Mark III and Mark IV stock;
- the value of these noticeable differences using stated preference techniques.

The fieldwork was conducted on BR trains with lap top computers and more than 600 passengers were interviewed. They were recruited on either Mark III or Mark IV carriages on the East Coast Main Line. The computerised interview was facilitated by other visual aids including artist's impressions of the different levels of aspects of carriages. These artist's impressions were instrumental in introducing the subject of the research. They also helped the interviewee's concentration in carrying out ranking exercises, about attitudes and satisfactions of stock, customised to her/his perceptions. The respondents were asked to compare the carriage they were travelling in with the other type of stock. In addition to the artist's impressions of the stock details, photographs of each carriage were also shown.

A series of questions was incorporated concerning the quality of the carriage. The artist's impressions were used to help the respondent indicate the level which best described the quality of each aspect. The degree of satisfaction each respondent felt for the perceived level of quality was recorded on a five point scale, defined as: very good, good, adequate, poor and very poor.

Perceived differences between the two types of stock were noted on each quality aspect. Given that a difference on one of the aspects had been acknowledged, the questionnaire explored further differences at a more detailed level. For example, if a respondent indicated that there was a seating difference between Mark III and Mark IV, then the respondent was asked if the difference was in the following:

## *Surveys to evaluate British Rail's rolling stock*

- comfort of the seat;
- adjustment of the seat;
- legroom;
- the seating layout.

In some cases more than one difference was perceived. All seven main aspects of carriage quality were broken down into their important component parts. Each difference was recorded using the five point scale below to show which type of carriage was preferred:

- definitely prefer \$current carriage;
- slightly prefer \$current carriage;
- no difference;
- slightly prefer \$other carriage;
- definitely prefer \$other carriage.

where \$current and \$other were defined as MarkIII and MarkIV as appropriate.

With the respondent thinking in terms of the differences in rolling stock quality, the two Stated Preference (SP) experiments were introduced. SP techniques are used in transport modelling as an extension of 'revealed preference' (RP) methods. Traditionally, demand analysis is based on RP data; that is choices and decisions that have actually been made in the marketplace. With theoretical advances in econometrics, there has been great expansion in this field. But there are practical limitations to the RP approach, largely connected with survey costs and the difficulty of distinguishing the effects of attributes such as quality or convenience. In addition, for obvious reasons, models based on RP cannot handle new alternatives that might be introduced into the market.

As a consequence, a practice has developed of basing demand estimates on an analysis of responses to hypothetical choices. The alternatives are presented to respondents in terms of their component attributes described at a number of levels. The design of these SP techniques is based on the principles of experimental statistics. Traditionally the presentation of the alternative choices in SP surveys has been carried out using cards

and paper. The recent dramatic changes in the micro-computer hardware and software market have facilitated new developments in SP techniques. By using computers, choices can be customised to suit the respondent's circumstances.

In this study we presented two SP experiments to the respondents to assess their valuations of rolling stock, overall, and then discriminate between the various components of the quality of rolling stock. In the first experiment, the respondents were asked to rank a number of different BR services which included trade-offs between return fare, one-way journey time and type of rolling stock. The latter was defined at two levels, Mark III and Mark IV. Both return fare and journey time had three levels, the middle level representing the characteristics of the current journey and the other two a reduction and an increase in fare and journey time.

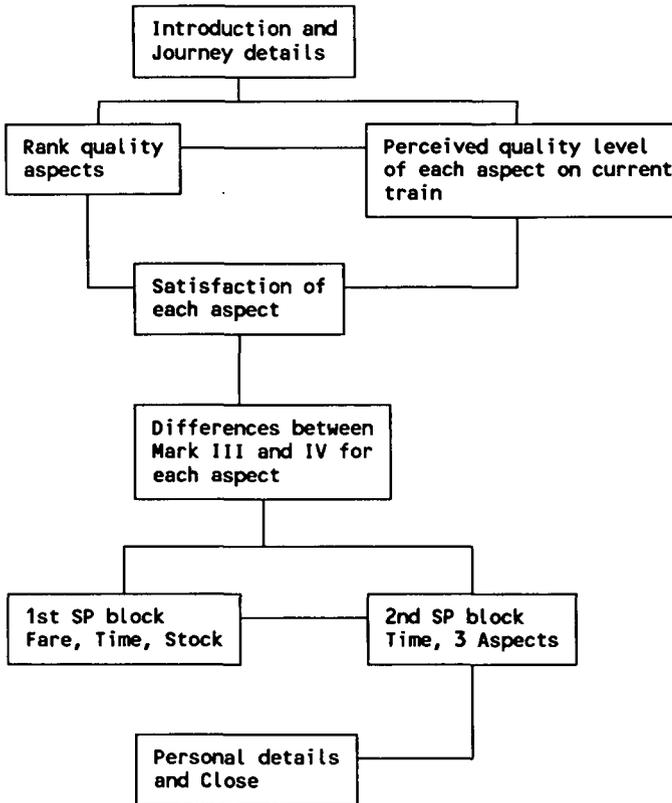
The full factorial design (all attributes presented at all levels in all combinations) for the above variables and levels includes  $(3 \times 3 \times 2 =)$  18 options. It was decided to present a total of 16 options to all respondents, but in blocks. Since four options can be clearly presented on each screen, it was decided to divide the design into four sets of four options to trade-off.

The second experiment presented trade-offs between services described in terms of journey time and three specific attributes of carriages. These qualitative attributes were selected for each interviewee according to his importance ratings and perceptions of change. In total the full factorial of the second SP includes  $(3 \times 2 \times 2 \times 2 =)$  24 options, of which 12 were presented to the respondents in three sets of four options. The respondents were asked to rank the four options within each set in order of importance. Both the SP experiments were customised to suit current individual circumstances so that realistic questions could be asked.

### **3. Application**

We usually start questionnaire design by drawing a block diagram of questions and groups of questions to be included in the research (see figure 1). After the decision is taken about the basic contents of the research and its feasibility is assessed, the questionnaire is directly designed in the Blaise language and the specification is stored in the computer. In practice, substantial parts of a questionnaire are readily drawn from a library of questions and blocks which we have built up in three years' experience with Blaise.

Figure 1. Structure of the questionnaire



Blaise 2.22 was used in this application. The questionnaire included a main section with questions about journey and personal details. Rankings and satisfactions of the quality of carriages and the SP experiments were included in separate blocks. In the main section individual details were established which were used to customise later questions. It was important to establish the journey time and the purpose of travel because the SP designs were customised on the basis of this information.

*Surveys to evaluate British Rail's rolling stock*

Extensive use of array questions and array variables was made in the rankings and satisfactions block. An important consideration was to show the respondents their own rankings and allow them to reconsider. In the SP exercises extensive use of substitution enabled the design to be customised for the different types of travellers. As an example we show in figure 2 one set of four options that respondents were asked to choose from in the first SP. The Blaise application is shown at the end of the paper.

*Figure 2. Presentation of the first Stated Preference design*

<p>SERVICE A          Journey time \$sameasnow          Return fare \$£1.00\$more          Carriages \$current</p>	3	2	<p>SERVICE B          Journey time \$sameasnow          Return fare \$sameasnow          Carriages \$other</p>
<p>SERVICE C          Journey time \$10mins\$more          Return fare \$£2.00\$less          Carriages \$current</p>	4	1	<p>SERVICE D          Journey time \$sameasnow          Return fare \$£1.50\$less          Carriages \$other</p>

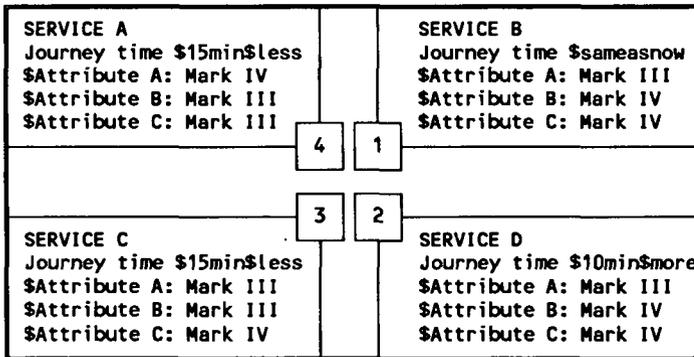
In the application of the second SP experiment the following criteria were used to decide which of the seven attributes should be included in the trade-offs:

- if the respondent recognised three or more differences between the two types of stock, then the three ranked most important were incorporated into the SP;
- if the respondent recognised only two different aspects, then these two variables were joined in the SP by the highest ranked of the other five quality aspects;
- if the respondent recognised only one different aspect of carriage quality, then the two highest ranked of the other six were also incorporated into the SP;

- if no aspect was acknowledged to be different between the two types of stock, then the respondent was asked to trade-off the three highest ranked carriage quality aspects.

These three variables were presented as in Mark III and Mark IV carriages and the fourth variable to trade-off was always one-way journey time, at three levels. It was customised to suit the circumstances of each passenger's journey (e.g. see figure 3). The above criteria were implemented into Blaise by means of a series of loops and conditional statements in the *Route* paragraph. Finally, the three variables of quality of carriages shown to each respondent were stored as hidden questions.

Figure 3. Presentation of the second Stated Preference design



An important aspect of our Blaise applications is the ability to check each interview for consistency interactively. For example routing is different depending on the type of passenger being interviewed (eg business and leisure). Most of the common type of interviewers' errors can be avoided (eg answer out of range, origin of journey same as destination, etc). In addition, calculations are carried out by the computer to reduce the interviewers' (and interviewees') burden. In the particular application described in this paper, the computer guided the interviewer to ask and enter the correct departure and arrival time of the journey. Then journey time was calculated and shown to the interviewee together with all the other particular details of his journey.

In the application of the SP experiments the checking facilities were particularly useful. It is a common error in PAPI approaches to rank the same option twice. In this case, this was prevented with a few lines of code (conditional and loop check) in the *Check* paragraph. Another common cause of attrition in the final sample for SP experiments is irrational ranking of options. In order to avoid correlation in the design some cases of dominance are left in the SP design. Thus, some options are always worse, or better than others. If the respondent fails to understand this, the recorded information is useless at the analysis stage. Our approach here is to use either a soft check (*Signal* paragraph), or warnings in the *Route* paragraph to notify both interviewer and interviewee about the error. They are encouraged to go back and examine the rationality of their answers and correct their rankings. It is generally acknowledged that these types of ranking errors can lead to an attrition of 10% to 15% at the analysis stage. Our experience shows that by using the above CAPI techniques attrition can be eliminated or reduced to less than 5%.

#### 4. Discussion

Towards the end of the 1980s The MVA Consultancy evaluated a number of CAPI systems. The decision was taken at the end of 1988 to use Blaise which is at the forefront of CAPI software.

Since then a number of surveys have been written and administered in Blaise by the MVA Consultancy in the fields of market research and passenger and freight transport demand assessment, with considerable success. Beyond doubt we are, in general, satisfied by the software in its present form. However, we are also eagerly looking to further developments to accommodate the evolving nature of our research. We welcome the opportunity to discuss different aspects of Blaise with other practitioners and academics and present our views on current limitations and areas for further development, examples of which are shown below.

## *Surveys to evaluate British Rail's rolling stock*

- ability to call from Blaise program modules written by users in PASCAL;
- format and screen layout facilities;
- compile big CAPI questionnaires;
- facility to change order of questions and enumerated answers at random;
- more algorithmic and mathematical facilities (nested do loops in route paragraph, additional functions etc).

## Surveys to evaluate British Rail's rolling stock

12

### APPENDIX. Example of a Stated Preference design in Blaise

Block Spblock "Sp Cards" (Subfile);

Var

```
Niter, Oflag1, Oflag2, Oflag3, Iter, Flag: Integer;
Catch1, Catch2: Integer;
CardOrder: Array[1..4] of Integer;
Ch, Last1Rank, Last2Rank, Last3Rank, Last4Rank: Integer;
Lasx1Rank, Lasx2Rank, Lasx3Rank: Integer;
RanCh: Array[1..4] of String[6];
RanFil: String[6];
jti1, jti3, jtx1, jtx3: Integer;
jf1, jf2, jf3: Real;
far, frc, jtime: Real;
rou: Integer;
fact1, fact3: Real;
```

Quest

SpIntro

```
/////In this part of this questionnaire we want you, as a rail
traveller, to consider a series of imaginary choices.
//// In each case, please imagine that you were given a choice
//// between FOUR different types of services, for the part of
//// your journey that you are actually making on THIS train
//// today//": Pseudo (Empty, Screen);
```

BRack

```
///// PLEASE NOTE
/////The services are entirely imaginary and only set up for research
purposes to assess preferences for the Mark III and Mark IV
carriages.//":
Pseudo (Empty, Screen);
```

DesShow

```
///The different services which we will ask you to choose between are
each described in terms of
```

```
////RETURN FARE which is either ONE WAY JOURNEY TIME which is
/// * 'same as now' * 'same as now'
// OR * a reduction as shown OR * a reduction as shown
// OR * an increase as shown OR * an increase as shown
```

```
////And finally CARRIAGES which are either
```

```
//// * Mark III OR
// * Mark IV
//": Pseudo (Empty, Screen);
```

## Surveys to evaluate British Rail's rolling stock

### SameOthers

```
"/////Apart from return fare, journey time and carriages, you can
assume that the journey is exactly the same as the one you are making
today//": Pseudo (Empty, Screen);
```

### ChoosePlease

```
"//// Suppose you could choose between different services
//// for the part of your journey that you are making on THIS
//// train today. Please indicate which of the four services
//// (A, B, C and D) you would most prefer, which would be
//// your second and third choices, etc//":
Pseudo (Empty, Screen);
```

### Des1

```
"/////Which of the following is your $RanFil choice?//": Array[1..4]
of (A "      Service A
//return fare:      £ $jf1 less
//journey time:     $jt1 mins less
//rolling stock:    Mark III
//                * $CardOrder[1] *
//-----",
B "      Service B
//return fare:      £ $jf1 less
//journey time:     same as now
//rolling stock:    Mark IV
//                * $CardOrder[2] *
//-----",
C "      Service C
//return fare:      same as now
//journey time:     $jt1 mins less
//rolling stock:    Mark IV
//                * $CardOrder[3] *
//-----",
D "      Service D
//return fare:      £ $jf3 more
//journey time:     $jt1 mins less
//rolling stock:    Mark IV
//                * $CardOrder[4] *
//-----");
```

### Alarm

```
"///// ARE YOU SURE ?
/////PLEASE GO BACK AND CHANGE RANKS
//// Service D SHOULD NOT be better than service C
/////They are similar in journey time and rolling stock but Service D is
more expensive//":
(Warn "Please use <Up Arrow> to go back and change") (Empty,Screen);
```

## Surveys to evaluate British Rail's rolling stock

```
Des1Sure
"//These are your rankings for the first set of services
//                                     ARE YOU SURE ?//":
(A "      Service A
//return fare:      £ $jf1 less
//journey time:     $jti1 mins less
//rolling stock:    Mark III
//  RANK * $CardOrder[1] *
//-----",
B "      Service B
//return fare:      £ $jf1 less
//journey time:     same as now
//rolling stock:    Mark IV
//  RANK * $CardOrder[2] *
//-----",
C "      Service C
//return fare:      same as now
//journey time:     $jti1 mins less
//rolling stock:    Mark IV
//  RANK * $CardOrder[3] *
//-----",
D "      Service D
//return fare:      £ $jf3 more
//journey time:     $jti1 mins less
//rolling stock:    Mark IV
//  RANK * $CardOrder[4] *
//-----") (Empty, Screen);

.....

define imputed questions for store )
llr1: 1..4 (Hidden);
llr2: 1..4 (Hidden);
llr3: 1..4 (Hidden);
llr4: 1..4 (Hidden);

jjff1: 0.00..50.00 (Hidden);
jjff3: 0.00..50.00 (Hidden);

jjtt1: 0..90 (Hidden);
jjtt3: 0..90 (Hidden);

QChoice: Array[1..3] of String[22] (Hidden);

jjxt1: 0..90 (Hidden);
jjxt3: 0..90 (Hidden);
.....
```

## Surveys to evaluate British Rail's rolling stock

Route

```
( data statement for choice fill )
  Compute RanCh[1]:= 'FIRST';
  Compute RanCh[2]:= 'SECOND';
  Compute RanCh[3]:= 'THIRD';

  SplIntro; BRack; DesShow; SameOthers; ChoosePlease;

  Compute jf2:= Fare;
  If ((Typeticket = stas) or (Typeticket = frss)) then
    Compute jf2:= 2* Fare;
  endif;

  If (Class = fi) then
    Compute fact1:= 0.1276;
    Compute fact3:= 0.064;
  else
    Compute fact1:= 0.111;
    Compute fact3:= 0.074;
    If (Purpose = emp) then
      Compute fact1:= 0.125;
      Compute fact3:= 0.09375;
    endif;
  endif;

  Compute Far:= fact1 * jf2;
  Compute frc:= Far * 2;
  Compute rou:= Round(frc);
  Compute jf1:= rou / 2;
  If (jf1 < 0.50) then Compute jf1:= 0.50; endif;

  Compute Far:= fact3 * jf2;
  Compute frc:= Far * 2;
  Compute rou:= Round(frc);
  Compute jf3:= rou / 2;
  If (jf3 = 0) then Compute jf3:= 0.25; endif;

  Compute jtime:= 0.136 * Diff;
  Compute frc:= jtime/5;
  Compute rou:= Round(frc);
  Compute jti1:= rou * 5;
  If (jti1 > 50) then Compute jti1:= 50; endif;

  Compute jtime:= 0.227 * Diff;
  Compute frc:= jtime/5;
  Compute rou:= Round(frc);
  Compute jti3:= rou * 5;
  If (jti3 > 60) then Compute jti3:= 60; endif;
```

## Surveys to evaluate British Rail's rolling stock

```
{ ask hidden questions for level values of design }
  jjtt1; jjtt3;
  jjff1; jjff3;

{ start rankings }
  Compute Flag:= 0;
  For Iter:= 1 to 4 do
    If (Iter = 4) then
      Compute Flag:= 1;
      Compute Last1Rank:=10-(ord(Des1[1])+ord(Des1[2])+ord(Des1[3]));
      Compute CardOrder[Last1Rank]:= 4;
    endif;
    If (Flag = 0) then
      Compute RanFil:= RanCh[Iter];
      Des1[Iter];
      Compute Ch:= ord(Des1[Iter]);
      Compute CardOrder[Ch]:= Iter;
    endif;
  enddo;

  If (CardOrder[4] < CardOrder[3]) then
    Alarm;
  endif;

  Des1Sure;
{ ask hidden for the one ranked last }
  llr1;

{ ready for second pack of cards }
  For Iter:= 1 to 4 do
    Compute CardOrder[Iter]:= 0;
  enddo;

  Compute Flag:= 0;
  For Iter:= 1 to 4 do
    If (Iter = 4) then
      Compute Flag:= 1;
      Compute Last2Rank:=10-(ord(Des2[1])+ord(Des2[2])+ord(Des2[3]));
      Compute CardOrder[Last2Rank]:= 4;
    endif;
    If (Flag = 0) then
      Compute RanFil:= RanCh[Iter];
      Des2[Iter];
      Compute Ch:= ord(Des2[Iter]);
      Compute CardOrder[Ch]:= Iter;
    endif;
  enddo;
  .....
```

EndBlock;