# National Passenger Survey

# User guidance

Passenger**focus**
putting rail passengers first

# National Passenger Survey

## User Guidance

## Table of contents

# Background

Passenger Focus (and before it OPRAF and the Strategic Rail Authority) set up the National Passenger Survey (NPS) in 1999. The aim of the NPS was to provide customer views on rail company performance on a consistent basis, so that comparisons could be made between the various companies. Over time, data from the NPS has been built into the franchising contracts with train companies, making the results an important commercial dimension of running a Train Operating Company (TOC). Given this, the sample design, fieldwork standards and accuracy of assigning journeys to specific TOCs are of the greatest importance. In addition, large enough sample sizes are required for each TOC to ensure that performance changes can be seen in the marketplace.

The first NPS was run in Autumn 1999 and it has been run twice a year since then. The first seven waves were undertaken by The Oxford Research Agency, until the contract was offered at competitive tender in Autumn 2002. In December 2002, Continental Research was appointed to run the survey from Spring 2003 until Spring 2007.

This document provides guidance on how to use NPS data and outlines the types of analysis that can be undertaken. Information is provided on the likely accuracy of results.

# Summary of advice

The NPS is designed to generate random samples of passengers for each Train Operating Company.  Used at the Train Operating Company level, the normal rules for calculating sampling error for a weighted sample apply approximately.

The NPS can be used to derive data at a station or route level, which may cover more than one TOC.  The sampling error for this type of data is considerably higher, as different TOCs have very different weights.

The NPS can also generate data at regional level and this is used extensively in the Consultees Report.  For some regions, this involves amalgamating data from several TOCs with different weighting levels and as such this can increase sampling error.

The NPS is available as a single dataset covering the last eleven waves – a full five year period.  For the latest Spring 2006 Wave (Wave 14) this covers Spring 2001 to Spring 2006.  Ad hoc analysis from this dataset is easy to produce and can typically be turned round in a few hours.  It is possible to go all the way back to Autumn 1999, when the NPS started, but this involves slightly longer time scales.

It is also possible to acquire NPS data at respondent level, in SPSS format.  The dataset is very large (all waves now comprise well over 300,000 records with 100+ variables for each) and can just fit onto a CD.  Users need a good understanding of analysing large datasets with weighting to be able to use this facility.

# Sample design

The NPS uses a two stage cluster sample design for each train company (with Silverlink split into Metro and County). The first stage sampling unit is a train station and questionnaires are then distributed to passengers using that station and that train company on a particular day at a specified time. The main purpose of the NPS is to generate robust data for each TOC, although the database does allow for other analyses to be undertaken.

Different sample sizes are set for each Train Operating Company that reflect the complexity of routes and the number of passengers the company carries. The sample sizes for the Spring 2006 Wave (wave 14) range from 250 respondents for Island Line up to 1500 for four London and South East franchises. To arrive at a national dataset that represents all passengers' satisfaction with rail, each TOC is weighted to reflect the number of journeys that it contributes to the national rail network. Therefore TOCs that account for a relatively small number of passenger journeys are down weighted and those that account for a high number of journeys are weighted up (see table overleaf).

| Train operating company | A - Journeys (000's per annum) | B -Sample size | Weighting (A / B) |
|---|---|---|---|
| Central Trains | 22700 | 1491 | 15.22 |
| Chiltern Railways | 11700 | 1050 | 11.14 |
| Southern | 115000 | 1609 | 71.47 |
| South Eastern | 132600 | 1605 | 82.62 |
| Gner | 13700 | 1047 | 13.09 |
| First Great Western | 20200 | 1377 | 14.67 |
| Island Line | 800 | 303 | 2.64 |
| First Scotrail | 56400 | 1027 | 54.92 |
| c2c | 28200 | 1036 | 27.22 |
| Merseyrail | 23800 | 571 | 41.68 |
| Midland Mainline | 9400 | 1058 | 8.88 |
| Silverlink | 35600 | 1038 | 34.30 |
| South West Trains | 141200 | 1905 | 74.12 |
| Thameslink | 41400 | 1043 | 39.69 |
| FGWL (First Great Western Link) | 37300 | 1043 | 35.76 |
| Virgin Crosscountry | 17400 | 1350 | 12.89 |
| Virgin West Coast | 15200 | 1061 | 14.33 |
| Wagn | 33470 | 762 | 43.92 |
| Gatwick Express | 4100 | 651 | 6.30 |
| Arriva Trains Wales | 19000 | 722 | 26.32 |
| Wessex | 9800 | 576 | 17.01 |
| Transpennine Express | 11975 | 1053 | 11.37 |
| One | 100770 | 1677 | 60.09 |
| Northern Rail | 57800 | 1333 | 43.36 |

# Weighting

Within the sample for each TOC, quotas are set by day of week, journey purpose and size of station. The sampling plan is designed in a way to select larger stations more often and to assign days of week and times of day to selected stations to generate a random sample of passengers. The data is weighted for each TOC by journey purpose, day of week and station size but the weights do not vary greatly and do not unduly affect the effective sample size. So at TOC level, the normal rules for assessing confidence intervals with a weighted sample can be applied. Typically these would be as follows:

| Train operating company | Accuracy |
|---|---|
| Central Trains | +- 2.5% |
| Chiltern Railways | +- 3.1% |
| Southern | +- 2.5% |
| South Eastern | +- 2.5% |
| Gner | +- 3.1% |
| First Great Western | +- 3.1% |
| Island Line | +- 6.2% |
| First Scotrail | +- 3.1% |
| c2c | +- 3.1% |
| Merseyrail | +- 4.4% |
| Midland Mainline | +- 3.1% |
| Silverlink | +- 3.1% |
| South West Trains | +- 2.5% |
| Thameslink | +- 3.1% |
| FGWL (First Great Western Link) | +- 3.1% |
| Virgin Crosscountry | +- 3.1% |
| Virgin West Coast | +- 3.1% |
| Wagn | +- 4.4% |
| Gatwick Express | +- 4.4% |
| Arriva Trains Wales | +- 3.6% |
| Wessex | +- 4.4% |
| Transpennine Express | +- 3.1% |
| One | +- 2.5% |
| Northern Rail | +- 3.1% |

All analyses from the NPS are undertaken on weighted data (except those for PTE areas – please see overleaf).

These figures give 95% confidence limits assuming that the data comes from a random sample. In practice, the NPS is not a random sample, as it employs both clustering (several interviews at one point which widens sampling error) and stratification (by station size which reduces sample error). These phenomena can be assumed to cancel each other out and the program we use to run analyses does actually use the weights that are generated to calculate sampling errors.

## Minimum sample sizes

At TOC level, most analyses are robust enough to stand up to scrutiny. At station level, the combination of smaller sample sizes and greater variation in weights if more than one TOC is involved mean that data is substantially less robust. Ideally, station or route analysis should be based on sample sizes of 100 and certainly at least 50. To reach this level of sample size for some stations or routes, it may be necessary to combine waves.

As an example, the data for Cardiff Queen Street in Wave 14 is based on 115 completed questionnaires. All the questionnaires relate to services offered by Arriva Trains Wales and so all will have similar weights, varying from 12.16 to 27.24. At worst, one questionnaire has a weight of double another and this fairly tight range means that the effective sample size, on which sampling error is based, will be close to 115. For an estimate of 70% from this station, the approximate accuracy limits would be +- 10%.

At another extreme, the data for Newcastle is based on 149 completed questionnaires but covering five different TOCs: First Scotrail, GNER, Northern Rail, TransPennine Express and Virgin CrossCountry. Questionnaires completed at Newcastle have weights varying from 3.86 to 166.55, so at worst one questionnaire has a weight of 43 times that of another. These wide variations in weight will reduce the effective sample size considerably, meaning that an estimate of 70% from this station will have approximate accuracy limits of +- 20%.

Not all stations are sampled in the NPS, typically around 700 out of the 2,500+ stations on the rail network.

## Route analysis

The NPS has always recorded where the passenger boarded and left the train service. For the first time in the Spring 2006 wave, we have recorded the origin and destination of the train service itself; this information is added to the survey record when the passenger journey is checked for validity using RailPlanner. This means that line of route analysis is available for the first time from this data source. The same considerations about sample size apply, although at present we cannot amalgamate waves to generate this type of analysis as the current wave is the first where this information is available. As future waves are undertaken, the facility to produce route analysis for lower volume routes will increase.

## PTE area analysis

Boost shifts are grafted onto the NPS to augment the number of local journeys in the six PTE areas (Greater Manchester, Nexus, South Yorkshire, Strathclyde, West Midlands and West Yorkshire). Some of these are not included in the main NPS database, as this would skew the profile too much towards local journeys.

As there is no data on the profile of journeys within PTE areas, all PTE data is unweighted and is not necessarily a random sample of passenger journeys within the area. As such, all PTE analyses should be treated with caution and viewed as indicative of performance in the area rather than definitive.

## Other geographies

It is not easy to superimpose any other geographies onto NPS data. We do not hold the postcode of the origin or destination of the journey and records can therefore only be aligned with TOCs, stations or routes or combinations of these. The database does contain the Category A-F station segment definition, so analyses can be undertaken by this variable.