

NPS User Guidance Report

Updated May 2012



Dave Chilvers
Director
Tel: 020 7490 9111
dDave.chilvers@bdrc-continental.com
providing intelligence

Contents

	Page No.
1. Background.....	1
2. Summary of advice	2
3. Sample design.....	3
3.1 Weighting	4
3.2 Minimum sample sizes	5
3.3 Route analysis	5
3.4 PTE area analysis	6
3.5 Other geographies.....	6



1. 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, and was re-appointed to run the survey for a further four years from Autumn 2007 and again re-appointed in 2011 for a further 3-4 year contract.

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.

2. Summary of advice

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. Increasingly, the NPS sample is selected for TOC building blocks, typically operational subsets of TOC franchise areas which align with internal reporting areas. Used at building block level, the normal rules for calculating sampling error for a weighted sample also apply. All franchised TOCs except c2c use building blocks as part of their sample design.

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 can have very different weights.

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.

NPS is available as a single dataset covering the last ten waves – a full five year period. For example for the Spring 2012 Wave (Wave 26) this covers Autumn 2007 to Spring 2012. 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 NPS started, although analysis of this data takes slightly longer.

It is also possible to acquire NPS data at respondent level, in SPSS format. The dataset is very large (all waves together now comprise around 750,000 records with 1000+ variables for each) and can fit onto a DVD. Users need a good understanding of analysing large datasets with weighting to be able to use this facility.

Finally, NPS data is available on the NPS Reportal, an online system which allows access to the basic NPS data for the past six waves and to the verbatim comments written in by respondents for the latest wave. This system, which is available at <http://www.npsreportal.org.uk/> comes with its own online guidance and help functions. Critically, the analysis system does not display any results based upon sample sizes of less than 50, to minimise inappropriate use of NPS data.

3. Sample design

NPS uses a two stage cluster sample design for each Train Operating Company building block. 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 NPS is to generate robust data for each TOC building block and hence for each TOC.

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 target sample sizes for the Spring 2012 Wave (wave 26) range from 500 respondents for Merseyrail up to 2,750 for First Great Western. 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.

TOC	number of journeys (000's per annum)	sample size	Ratio
Arriva Trains Wales	26420	1189	22.22
c2c	32175	1114	28.88
Chiltern Railways	17768	1192	14.91
CrossCountry	29700	1191	24.94
East Coast	17733	1225	14.48
East Midlands Trains	22317	1219	18.31
First Capital Connect	97672	2000	48.84
First Great Western	83870	3044	27.55
First Hull Trains	679	583	1.16
First TransPennine Express	22372	1175	19.04
Greater Anglia	106689	2454	43.48
Heathrow Connect	2800	609	4.60
Heathrow Express	6299	559	11.27
London Midland	52930	1192	44.40
London Overground	92515	1202	76.97
Merseyrail	40082	635	63.12
Northern Rail	94518	1264	74.78
ScotRail	73238	1230	59.54
South West Trains	190065	2334	81.43
Southeastern	163361	1722	94.87
Southern	162014	2338	69.30
Virgin Trains	23172	1112	20.84

3.1 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 and day of week and for each TOC building block by station size. The weights do not vary greatly except in situations where a building block has been deliberately over sampled to generate a robust sample size for the building block and this means the weighting does not unduly affect the effective sample size.

At TOC level, the normal rules for assessing confidence intervals with a weighted sample can be applied. Typically these would be as follows based on the worst case scenario of a 50% satisfaction level; satisfaction levels that are considerably away from 50% will be more accurate:

TOC	accuracy +-%
Arriva Trains Wales	4.2
c2c	3.0
Chiltern Railways	3.2
CrossCountry	3.4
East Coast	3.6
East Midlands Trains	3.1
First Capital Connect	3.2
First Great Western	2.0
First Hull Trains	6.2
First TransPennine Express	4.5
Greater Anglia	2.7
Heathrow Connect	6.4
Heathrow Express	4.6
London Midland	3.2
London Overground	4.0
Merseyrail	5.4
Northern Rail	3.3
ScotRail	4.6
South West Trains	2.7
Southeastern	2.8
Southern	2.4
Virgin Trains	4.0

All analyses from NPS are undertaken on weighted data. These figures give 95% confidence limits assuming that the data comes from a random sample. In practice, 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.

3.2 Minimum sample sizes

At TOC and TOC building block 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 Southend Central in Wave 26 is based on 33 completed questionnaires. All the questionnaires relate to services offered by c2c and so all will have similar weights, varying from 22.36 to 41.39. At worst, one questionnaire has a weight of less than double another and this fairly tight range means that the effective sample size, on which sampling error is based, will be close to 33. For an estimate of 70% from this station, the approximate accuracy limits would be slightly above $\pm 16\%$.

At another extreme, the data for Newcastle is based on 111 completed questionnaires but covering five different TOCs: First Scotrail, East Coast, Northern Rail, TransPennine Express and CrossCountry. Questionnaires completed at Newcastle have weights varying from 12.97 to 101.39, so at worst one questionnaire has a weight of 8 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 $\pm 13\%$, similar to that for Southend Central even though the unweighted sample size is three times as large.

3.3 Route analysis

NPS has always recorded where the passenger boarded and left the train service. Since the Spring 2006 wave the origin and destination of the train service itself have been recorded in the survey database; 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 NPS data is available.

The same considerations about sample size apply, and waves can be amalgamated to generate analysis if required. Now that we have several waves where we have origin and destination of the train recorded the facility to produce route analysis for lower volume routes is available.

As mentioned, most TOCs are now also divided into building blocks (or routes) at the fieldwork stage. This means properly weighted data is automatically available for certain areas below TOC level.

3.4 PTE area analysis

NPS produces data for the six PTE areas (TfGM, Nexus, South Yorkshire, Strathclyde, West Midlands and West Yorkshire). From Wave 26 onwards, all data for PTE areas has been weighted to the aggregate profile by journey purpose and weekday/weekend from the preceding ten waves. Analysis of this data has confirmed that the profile obtained from NPS journeys, using the derived weights, does not vary significantly from one wave to another and thus the use of these aggregate weights provides stability of results from one wave to another. Comparisons between waves will not be due to differences in sample profile and so conclusions can be drawn about significant changes which are likely to be due to real effects rather than variations in the sample design.

The aggregate profiles will be checked for each PTE each year to ensure that any significant trends in either journey purpose and/or weekday/weekend can be reflected in revised weights going forward.

3.5 Other geographies

Analysis by any other geographies requires each station to be allocated to a unit of that geography and then this new geography can be applied to the NPS data set. We have available the Standard Region of the origin station, so this variable is available for analysis purposes.

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.