

National Rail Passenger Survey:

User Guidance Report

Spring 2014 (wave 30)

Contents

	Page No.
1. Background	1
2. Summary of advice.....	2
3. Sample design	3
3.1 Weighting.....	4
3.2 Accuracy at TOC level	4
3.3 Accuracy at TOC building block level.....	5
3.4 Minimum sample sizes	7
3.5 Route analysis.....	7
3.6 PTE area analysis.....	8
3.7 Other geographies.....	8

1. Background

Passenger Focus (and before it OPRAF and the Strategic Rail Authority) set up the National Passenger Survey (now the National Rail Passenger Survey, NRPS) in 1999. The aim of the NRPS 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 NRPS 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 NRPS 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 (now merged to become BDRC Continental) was appointed to run the survey from Spring 2003 until Spring 2007, 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 NRPS data and outlines the types of analysis that can be undertaken. Information is also provided on the likely accuracy of results.

2. Summary of advice

NRPS 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 NRPS sample is selected for TOC building blocks; typically, these are 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.

NRPS 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.

NRPS can also generate data at regional level and this is used extensively in the Stakeholder Report (formerly known as 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.

NRPS is available as a single dataset covering the last ten waves – a full five year period. For example for the Spring 2014 Wave (Wave 30) this covers Autumn 2009 to Spring 2014. Ad hoc analysis from this dataset is easy to produce and can typically be turned round in a few hours. It is also possible to go all the way back to Autumn 1999, when NRPS started, although analysis of this data takes slightly longer.

It is also possible to acquire NRPS data at respondent level, in SPSS format. The dataset is very large (all waves together now comprise around 810,000 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, NRPS data is available on the NRPS Reportal, an online system which allows access to the basic NRPS 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 NRPS data.

NRPS results for the main station and train factors only for the last 10 waves for all TOCs and building blocks are also available through the NRPS online data tool at:

<http://data.passengerfocus.org.uk/train/nps/question/service-overall/>.

3. Sample design

NRPS 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 distributed to passengers using that station and that train company on a particular day at a specified time. The main purpose of NRPS 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 2014 Wave (wave 30) 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
Abellio Greater Anglia	103929	2313	44.93
Arriva Trains Wales	28528	1097	26.01
c2c	36028	1089	33.08
Chiltern Railways	19402	1146	16.93
CrossCountry	36683	1129	32.49
East Coast	18785	1126	16.68
East Midlands Trains	23167	1123	20.63
First Capital Connect	107253	1805	59.42
First Great Western	92873	3050	30.45
First Hull Trains	721	605	1.19
First TransPennine Express	24893	1092	22.80
Grand Central	770	653	1.18
Heathrow Connect	3349	578	5.79
Heathrow Express	5750	573	10.03
London Midland	60051	1121	53.57
London Overground	123887	1169	105.98
Merseyrail	44909	598	75.10
Northern Rail	106517	1150	92.62
ScotRail	81506	1094	74.50
South West Trains	209611	1944	107.82
Southeastern	162334	1652	98.27
Southern	166197	2179	76.27
Virgin Trains	30195	1238	24.39

(Note that while the above table includes the non-franchised TOCs which take part in NRPS, only franchised TOCs are included within national, regional or sector aggregates for normal reporting.)

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.

3.2 Accuracy at TOC level

At TOC level, the normal rules for assessing 95% confidence intervals with a weighted sample can be applied. Typically these would be as in the table overleaf, based on the worst case scenario of a 50% satisfaction level; satisfaction levels that are considerably away from 50% will be more accurate. This table shows the accuracy of data at TOC level, for analysis run on Spring 2014 results only; combining waves together for analysis will increase robustness and therefore accuracy:

TOC	accuracy +-%
Abellio Greater Anglia	3.1
Arriva Trains Wales	3.8
c2c	3.5
Chiltern Railways	3.3
CrossCountry	3.5
East Coast	3.3
East Midlands Trains	3.3
First Capital Connect	2.9
First Great Western	2.2
First Hull Trains	5.8
First TransPennine Express	4.0
Grand Central	4.6
Heathrow Connect	4.5
Heathrow Express	4.5
London Midland	3.5
London Overground	4.2
Merseyrail	6.0
Northern Rail	3.6
ScotRail	5.9
South West Trains	2.8
Southeastern	2.7
Southern	2.5
Virgin Trains	3.8

All analyses from NRPS are undertaken on weighted data. Weighting increases sampling error and the figures above take account of the “weighting efficiency” that the weighting regime produces.

3.3 Accuracy at TOC building block level

The figures in the table on the next two pages show the 95% confidence intervals for each TOC building block, again showing the worst scenario for an estimate percentage of 50% satisfied, based on Spring 2014 only. As with the accuracy figures for TOCs, these estimates take into account the weighting efficiency of the sample for each building block. Estimates closer to 0% or 100% will have tighter confidence intervals than those shown here. Typically, the range for a 70% figure will be about 90% of the figures shown here and the range for a 90% figure will be about 60% of the figures shown here:

Building block	accuracy +/- %	Building block	accuracy +/- %
Abellio Greater Anglia - Mainline	4.8	London Midland - London Commuter	5.3
Abellio Greater Anglia - Intercity	8.3	London Midland - West Coast	7.4
Abellio Greater Anglia - West Anglia Outer	6.5	London Midland - West Midlands	5.0
Abellio Greater Anglia - West Anglia Inner	8.8	London Overground - Dalston - Croydon (formerly Southern - TfL)	7.3
Abellio Greater Anglia - Stansted Express	6.2	London Overground - Gospel Oak - Barking	6.8
Abellio Greater Anglia - Metro	10.7	London Overground - Richmond/Clapham - Stratford	6.6
Abellio Greater Anglia - Rural	9.0	London Overground - Watford - Euston	5.7
Arriva Trains Wales - North Wales	6.0	Merseyrail - Northern	9.4
Arriva Trains Wales - South Wales	5.5	Merseyrail - Wirral	7.0
Arriva Trains Wales - Valley	6.1	Northern Rail - Lancashire & Cumbria	10.0
c2c	3.5	Northern Rail - Manchester & Liverpool	5.8
Chiltern Railways - North	6.6	Northern Rail - South & East Yorkshire	10.6
Chiltern Railways - South	3.8	Northern Rail - Tyne Tees & Wear	12.8
CrossCountry - Birmingham - Manchester	9.7	Northern Rail - West & North Yorkshire	6.2
CrossCountry - Birmingham - North East & Scotland	6.1	ScotRail - Interurban	5.5
CrossCountry - Birmingham - South Coast	7.7	ScotRail - Rural	13.8
CrossCountry - Birmingham - South West	8.9	ScotRail - Strathclyde	9.5
CrossCountry - Birmingham - Stansted	9.2	ScotRail - Urban	7.1
CrossCountry - Nottingham - Cardiff	14.3	South West Trains - Island Line	10.5
East Coast - London - East Midlands/East of England	8.4	South West Trains - London	4.1
East Coast - London - Scotland/North East	6.9	South West Trains - Mainline	8.4
East Coast - London - Yorkshire	6.2	South West Trains - Metro	6.3
East Coast - Non-London Journeys	5.6	South West Trains - Not Managed By SWT	10.8
East Midlands Trains - Liverpool - Norwich	6.8	South West Trains - Portsmouth	14.9
East Midlands Trains - Local	7.2	South West Trains - Reading/Windsor	9.2
East Midlands Trains - London	4.3	South West Trains - Suburban	8.8
First Capital Connect - Great Northern	4.9	South West Trains - West Of England	9.8
First Capital Connect - Thameslink Loop	7.0	Southeastern - High Speed	7.8
First Capital Connect - Thameslink North	4.8	Southeastern - Mainline	5.5
First Capital Connect - Thameslink South	7.8	Southeastern - Metro	3.3
First Great Western - London Thames Valley	3.4	Southern - Gatwick Express	5.3
First Great Western - Long Distance	3.4	Southern - Metro	3.7

First Great Western - West	4.8	Southern - Sussex Coast	3.5
First Hull Trains	5.8	Virgin - Birmingham - Scotland	17.7
First TransPennine Express - North	5.1	Virgin - London - Liverpool	8.7
First TransPennine Express - North West	7.7	Virgin - London - Manchester	6.5
First TransPennine Express - South	8.7	Virgin - London - North Wales	12.8
Grand Central - London - Bradford	7.9	Virgin - London - Scotland	5.6
Grand Central - London - Sunderland	5.7	Virgin - London - Wolverhampton	7.3
Heathrow Connect	4.5		
Heathrow Express	4.5		

3.4 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 30 is based on 19 completed questionnaires. All the questionnaires relate to services offered by c2c and so all will have similar weights, varying from 44.26 to 45.39. A very tight range like this means that the effective sample size, on which sampling error is based, will be close to the unweighted sample size: in fact the effective sample size for Southend Central in wave 30 is 19 – the same as the actual number of questionnaires that were completed. For an estimate of 50% from this station, the accuracy limits would be +/- 22.5%.

At another extreme, the data for Lancaster is based on 72 completed questionnaires but covering three different TOCs: First TransPennine Express, Virgin Trains and Northern Rail. Questionnaires completed at Lancaster have weights varying from 2.27 to 136.2, so at worst one questionnaire has a weight of around 70 times that of another. These wide variations in weight will reduce the effective sample size considerably (in this case, to 29), meaning that an estimate of 50% from this station will have an accuracy limit of +/- 18.2%, quite similar to that for Southend Central even though the unweighted sample size is much larger.

3.5 Route analysis

NRPS 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 NRPS 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.

3.6 PTE area analysis

NRPS 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 NRPS 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.

3.7 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 NRPS 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 NRPS 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.