Department for Transport

Rail Strikes: Understanding the impact on passengers - technical report

April 2023

Department for Transport
Great Minster House
33 Horseferry Road
London SW1P 4DR

This publication is also available on our website at
www.gov.uk/government/organisations/department-for-transport

Any enquiries regarding this publication should be sent to us at
www.gov.uk/government/organisations/department-for-transport

Contents

1. Background and Objectives

The Department for Transport (DfT) commissioned the independent research agency Savanta to undertake a survey of the rail network in England to investigate the impact of strike action on rail passengers. The primary aim of the project was to provide robust evidence on changes to rail passengers' travel behaviours in response to rail strikes, and to measure any related impacts on rail passengers' social, economic, and personal lives.

The research covered four periods of strike action over the summer and early autumn of 2022, with fieldwork being conducted in the weeks following strikes. Each phase focused on understanding the impact of the most recent week of action, and any associated changes to travel behaviours.

Table 1 Fieldwork dates

Phase	Fieldwork	Strike week referenced	Extent of disruption
1	16/07/2022 - 26/07/2022	w/c 20/06/2022	3 strike days (Tues, Thurs, Sat)
2	01/08/2022 - 12/08/2022	w/c 25/07/2022	2 strike days (Weds, Sat)
3	22/08/2022 - 08/09/2022	w/c 15/08/2022	2 strike days (Thurs, Sat)
4	10/10/2022 - 23/10/2022	w/c 03/10/2022	2 strike days (Weds, Sat)

The extent of impacts on the rail network varied across the different strike periods, depending on which staff at different Train Operating Companies (TOCs) were striking, and whether Network Rail employees were also striking.[1] There was wider uncertainty for passengers around strike dates too, and so the impact was likely to extend beyond just the specific days of strike action (including, for example, early closure of lines or late starting on days following strike action). Focusing questions in the survey on the whole week in which strike action took place meant that we were able to capture that broader impact.

This technical document provides detailed information on the overall planning and delivery of this project.

[1] Train Operating Companies are companies that run passenger train services, and lease and manage stations. Each TOC manages a different set of services on routes on the rail network. Network Rail is the organisation that owns, operates, and develops the railway infrastructure in England, Wales and Scotland including track, bridges, and tunnels, and manages some of the largest stations.

2. Methodological Approach

Sampling for this research was designed to provide a representative sample of rail journeys and representative coverage of journeys on all TOCs in England with the exception of London Underground, London Overground, and the Elizabeth Line. It was also designed to provide robust data by sub-groups according to various demographic characteristics. In total, 17,383 questionnaires were completed across four phases of fieldwork between July and October 2022.

There are numerous approaches to surveying within the rail environment. All methods involve finding rail users from the general population of adults, or intercepting rail users making a journey, because no comprehensive sample frame of rail users exists for England or Great Britain. In this instance an on-train methodology with fieldworkers distributing (and collecting) paper self-completion questionnaires or, based on passenger preference, QR codes linked to an online version of the survey was seen to be the most efficient, cost-effective, robust, and timely way of reaching rail users who may have been intending to travel on strike days. This approach is an industry standard for surveys of this kind.

On-train self-completion surveys provide:

- greater coverage of passengers boarding at a wide variety of stations along a route/served by a specific TOC: passengers boarding along the entire route of a train service (including passengers transferring from other routes) prior to and including the journey covered by the fieldworker can participate
- accurate allocation of the TOC serving the route: knowing the service on which self-completion questionnaires were completed (or online links provided) removed risks related to TOC misallocation, which can be an issue when relying on respondents' self-classification

Alternative methodological approaches can be used to target rail users but were not suitable for this project. Using online panels is a common approach taken to market research projects but there are certain groups who tend to be underrepresented on these panels: particularly male respondents and those aged over 65. Furthermore, relying on online methodologies can make it difficult to produce large enough sample sizes at a regional and TOC level.

Sampling plan

The aims of the sampling approach were to:

- produce a sample of rail users who had planned to travel during phases of strike action. As the profile and number of journeys that were planned during strike action is unknown, a sampling plan built around contacting individuals who could possibly have travelled during strike phases was required.
- create a sample of rail journeys being made directly after phases of strike activity, to target users of the railway who could have been impacted during strike activity.
- create a sample of the whole English rail network (with exclusions mentioned below), including all TOCs, and a sample of different routes within each TOC which is as representative as possible of that TOC's routes. Where more routes were served by a TOC than could be covered by the research, a random stratified sample of routes would be selected.
- collect a large enough sample for each TOC to allow for robust reporting at a TOC level, of sufficient size for robust analysis of key sub-groups including TOC, journey purpose by TOC (business, leisure, commuting of different types, etc.), weekend/weekday travel by TOC, degree of flexibility to travel, employment sector including essential workers vs non-essential workers, household income, age, gender, ethnicity, disability, and whether or not the passengers have caring responsibilities
- collect a sample of typical journeys made by adult (aged 16+) users across those routes in weeks following (not during) strike action
- collect a sample across different days of the week and times of day.

How the sampling plan was constructed

A sampling plan was devised to ensure that the goals of this approach were achieved. To robustly examine data at the TOC level, all TOCs were targeted with a planned minimum sample size of 500 responses with that target being reached on all TOCs except one (see Table 2 and footnote 5 below). Given that TOCs vary quite widely in the number of passenger journeys that take place on their routes, collecting a large enough sample for each TOC ultimately meant that the sample over-represents those TOCs with fewer passenger journeys, and under-represents those with more passenger journeys. The effect of over- and under-sampling was adjusted for using scaling weights: see the Weighting chapter below for details.

The rejected option of randomly sampling routes across the network based on proportion of operating journeys would have resulted in some TOCs being missed out entirely, or a sample being collected from TOCs that was too small to produce robust statistics at a TOC level.

In order to cover initial phases of strike activity there was a requirement to mobilise the research as quickly as possible. Coverage of the variance in routes within TOC was achieved for initial fieldwork shifts by selection based on published route maps and in consultation with TOCs. A profile of shifts was identified that covered a range of the typical services and routes operated by each TOC.

For phases three and four of fieldwork, data from the LENNON (Latest Earnings Networked Nationally Over Night) ticketing and revenue system was used to ensure that the final sample of routes was representative, using numbers of passenger journeys made across routes within TOCs. LENNON holds information on the vast majority of rail tickets purchased in Great Britain, and allocates journeys from those ticket sales to TOCs using the mathematical model ORCATS (Operational Research Computerised Allocation of Tickets to Services). ORCATS utilises similar logic to journey planning systems, and identifies passenger 'opportunities to travel' from an origin station to a destination station using timetable information.

LENNON contains pre-allocation (sales) and post-allocation (earnings) datasets. The pre-allocation dataset collects total kilometres, journeys and revenue at purchase, and then apportions the data to the appropriate TOCs (in the post-allocation data set) using established allocation factors. For example, in the pre-allocation dataset, a passenger may purchase a ticket between two stations where it would be possible to travel using more than one TOC under the conditions for travel on that ticket type. LENNON then uses allocation factors to apportion journeys (and revenue) to different TOCs which are operating services on that specific flow. Data used for the purposes of this research are drawn from the post-allocation dataset, such that passenger journeys can be assigned to TOCs.

A summary of passenger journeys allocated to routes operated by each TOC over a baseline period of 6 – 12 June 2022 was used to estimate passenger journeys on those routes in a typical week prior to the beginning of strike action. The routes covered in the initial phases of research were checked against these profiles and the final sampling plan was adjusted to ensure representative coverage of services operated. Where there were more routes than could be covered by the number of fieldwork shifts allocated to a TOC, routes were stratified based on passenger journeys made and then randomly sampled to ensure coverage of a representative sample of routes. The resulting sample was spread across the whole of the English rail network, excluding the Elizabeth Line, London Overground, and London Underground services.[2]

Although fieldwork was conducted across four phases, in weeks following strike action, the sample was designed to provide representative coverage of the network as a whole (and of individual TOCs at a total level) across the four phases of fieldwork, rather than within each phase. Samples were not matched across phases of fieldwork, and so data across phases should be treated as a single sample, and interpreted in that way, rather than as separate samples for individual phases.

The sample is representative of journeys being made at the time of recruitment for the survey, during a normal week of rail service: it does not attempt to represent typical rail usage during weeks of strike action, since that was not the goal of the research. We took this approach to understand the extent of impact (or lack thereof) that rail passengers making typical journeys on the network experienced in a prior strike week, including by gathering responses from passengers that would not have been travelling in a strike week.

[2] Additional permissions were required to survey on Merseyrail, which meant that Merseyrail's services were covered in phase four of the research. For sampling purposes Govia Thameslink Railway brands (Southern, Thameslink, Great Northern, and Gatwick Express) were treated separately, and West Midlands Trains was treated as one TOC combining West Midlands Railway and London Northwestern Railway Routes.

Summary of individual TOC/route coverage

All English routes within each TOC were eligible for inclusion within the sampling plan (including mainline routes and branch lines). Across the 22 TOCs that sampling plans were developed for, 14 included all of the routes that that TOC served (within England). The routes covered for all TOCs are detailed in Annex A of this document. This annex also provides details on how the routes covered were selected for the remaining 8 TOCs.

Depending upon the length of the route chosen to be covered, services were identified that, where possible, involved single journeys of between 30 minutes and one hour in length.[3] Schedules were then created to cover between two and eight train services within a six hour 'shift' (see Table 5 in the fieldwork chapter for an example of an interviewing shift). TOCs were allocated a spread of shifts across different times and days: morning/afternoon weekday shifts covered a period of 7am-1pm, afternoon/evening weekday shifts covered 1pm-7pm, and weekend shifts covered either the morning 7am-1pm, afternoon/evening 1pm-7pm or mid-morning to mid-afternoon 10am-4pm.

To the extent that it was possible, all passengers travelling on the services covered by the sampling plan were approached and offered the opportunity to participate in the research, including those who had boarded the train prior to the fieldworker. Fieldworkers recorded the details of refusals to participate to assist in calculating non-response bias (see chapter on Weighting below). Fieldworkers walked the length of the train and back (changing carriages at interim stations where there was no internal connection between carriages) for the entire duration of their shift.

As all passengers travelling on these services had a broadly similar opportunity to participate, the sampling as a whole can be seen as being broadly representative of passenger journeys made by TOC across the interviewing period (once all phases are combined and weighted to account for non-response bias and TOC journey share).

Quality Assurance: Fieldwork Schedules

Fieldwork schedules were created by a team of researchers. These schedules were initially rotated and checked by other individuals within this team before being signed off by the project lead. After these schedules were signed off, further checks were made by area supervisors and individual fieldworkers. Any issues noted at this stage were checked and approved by the project lead and discussed with DfT before any amendments were made.

Prior to phase one fieldwork TOCs were invited to comment on the initial sampling plan to highlight any concerns or issues. No concerns were raised at this stage.

Summary of the sample achieved

309 fieldwork shifts were conducted, with 247 on weekdays and 62 at weekends. In phase one all weekend interviewing was conducted on a Saturday, but from phase two onwards a mix of Saturday and Sunday shifts were conducted. The final achieved sample was

[3] This is the length of the journey for the fieldworker: passengers on journeys of various lengths would still be picked up in this way (filling in the survey during the part of the journey for which the fieldworker was on the train).

17,383 responses (an average of around 56 completed responses per shift). The breakdown of responses and shifts by phase was:

- Phase 1: 3715 interviews from 65 shifts (57 interviews per shift)
- Phase 2: 3560 interviews from 71 shifts (50 interviews per shift)
- Phase 3: 5523 interviews from 92 shifts (60 interviews per shift)
- Phase 4: 4585 interviews from 81 shifts (57 interviews per shift)

Table 2 below provides the breakdown of responses achieved, number of shifts conducted and average response per shift by TOC:[4]

Table 2 Sample size and response rates by Train Operating Company

TOC	Responses	Responses per shift	TOC	Responses	Responses per shift
Avanti	727 (13 shifts)	56	Merseyrail	564 (10 shifts)	56
c2c	639 (16 shifts)	40	Northern	937 (18 shifts)	52
Chiltern	763 (15 shifts)	51	Southeastern	921 (19 shifts)	48
CrossCountry	799 (11 shifts)	73	Southern/Gatwick Express	1400 (29 shifts)	48
East Midlands Railway	924 (14 shifts)	66	Greater Anglia	565 (15 shifts)	38
Grand Central	526 (10 shifts)	53	GWR	1373 (17 shifts)	81
Hull Trains	794 (10 shifts)	79	Heathrow Express	531 (9 shifts)	59
LNER	744 (14 sifts)	53	SWR	1013 (20 shifts)	51
LUMO[5]	420 (8 shifts)	53	TPE	961 (15 shifts)	64
Thameslink/Great Northern	1771 (30 shifts)	59	West Midlands Railway/LNWR	1011 (16 shifts)	63

Strengths and limitations of the sampling approach

The sampling plan provided a representative sample of rail passenger journeys made in the weeks following specific phases of strike action. The data is robust at individual TOC level as well as by different user types, and by journey purpose. It provided robust samples of passenger who had planned to travel by train during the phase of strike activity.

Our approach means that there is a higher chance of sampling frequent passengers, as they will be more likely to be travelling in the weeks after strikes. This is an intentional feature of the approach, since the sample is intended to represent typical rail journeys, and frequent users make more rail journeys. It is important to note, however, that the final sample ultimately captures a wide range of frequent and infrequent rail users, which remains true when looking at the proportion who planned to travel during the previous strike week. Table 3 below summarises this.

[4] In the LENNON system, the Govia Thameslink Railway TOC brands are combined into two sub-brands, Thameslink/Great Northern and Southern/Gatwick Express. As a result, weights were applied using these groupings (see weighting chapter), and the groupings have been used when discussing TOC level data in this report.

[5] LUMO was the TOC where the 500 sample minimum was not reached, but given the very low network journey proportion occupied by this TOC, this does not represent an under-sampling.

Table 3 Proportion of respondents by frequency of train travel

Frequency of travelling by train in the past 6 months	All respondents	Those planning to travel during a strike week
5 or more times a week	17%	25%
3-4 times a week	16%	23%
1-2 times a week	19%	24%
1-3 times a month	26%	17%
Less often	17%	6%
Base	n=17,383	n=8,956

Very irregular users, who may have been intending to make a one-off journey during the week of a strike but who did not subsequently make another journey by rail, will not be captured.

Sampling for this research was designed to be representative at a journey level rather than by individual user. This was achieved by surveying individuals travelling on specific trains, with the aim of capturing representative responses for journeys being made on those trains. There is the possibility that the same person responded to the survey on more than one occasion (across phases), but even in this case a different journey, and the impacts of a different round of strike action, were captured.

It was not possible to cover all routes with the number of shifts conducted. The sampling approach is representative of services operated based on number of passenger journeys made on those services but cannot account for any very localised differences by route. There is no indication that such differences would vary across TOCs in a structured and systematic way which would invalidate the method of selecting routes.

Responses reporting direct impact of strike action on planned journeys do not necessarily indicate that those impacts were experienced on the same TOC that the respondent was travelling on when surveyed. Individuals were not asked for information on the TOCs they had planned to travel on for any particular journey during a strike week. It is not clear that asking respondents to provide this information would have produced a clearer indication of impact across TOCs in any case, as respondents may not have known or remembered who operated the service they planned to travel on, or may have misattributed services to a TOC.

The nature of the sampling approach means that the full sample can only be seen to be truly representative of the network as a whole: individual phases of research cannot be seen as representative of all services within that phase.

3. Questionnaire

The questionnaire was designed to meet the overall objectives of the research.[6] The questionnaire included the following sections/question areas:

- Journey details relating to the journey the passenger was on when contacted to take part in the research. This included details on origin and destination station, journey purpose and ticket usage.
- Intended travel behaviour during most recent strike week (including details of planned/conducted rail journeys by day of week and journey purpose)
- Impact of strikes on respondents' work and home lives, education, and also reported financial impact
- Overall perception of impacts of the strikes on any dependent children/older dependents
- Provision of information related to strike activity
- Longer-term changes in travel behaviours if the strikes continue
- Typical travel behaviour and flexibility to work or study from different locations
- Profiling questions – age, gender, household income, ethnicity, employment sector and disability

Overall, the questionnaire contained a limited amount of routing and was anticipated to take around 10 minutes to complete. Based on those responding online, the median time taken to complete the questionnaire was just under 12 minutes.

Questionnaire Quality Assurance

The questionnaire was reviewed by the project lead and DfT prior to printing and despatch for each phase. The online scripted version of the questionnaire was checked by several executives and DfT analysts across a range of devices before being signed off for use within the survey. All interviewer packs were checked by an executive member of the team prior to despatch to ensure all materials were correct (including the relevant letter of authority/pass for travel to cover permissions to be on trains for that specific shift).

[6] Questionnaires handed out on Merseyrail trains had a very slight amendment to the introductory text. See annex A (Mersyerail section) for details.

4. Fieldwork

Fieldwork was conducted by a national network of fieldworkers, reporting to 14 area supervisors. Only fully trained and experienced fieldworkers conducted this research. Fieldworkers were trained in line with Savanta policies to at least meet Interviewer Quality Control Scheme (IQCS) standards and are all registered on the Market Research Society Interviewer Identity Card (IDD) scheme.

To maximise response, and to be as inclusive as possible, passengers had a number of means of completing the survey:

- Paper and pen self-completion, filled in during their journey and handed back to the fieldworker, or if on a shorter journey via a reply-paid envelope
- A QR code handed to the passenger for them to link to a mobile optimised version of the questionnaire online
- By exception there was also the possibility of asking the fieldworker to assist in completing the questionnaire or to have a telephone interview conducted at a later date. [7]

The table below summarises the number of questionnaires (and percent of the total) that were completed using each method.

Table 4 Completed questionnaires by response method

	All phases	Phase 1	Phase 2	Phase 3	Phase 4
Total	17383	3715	3560	5523	4585
Paper	15498	3213	3066	4958	4261
Paper %	89.2%	86.5%	86.1%	89.8%	92.9%
QR	1885	502	494	565	324
QR %	10.8%	13.5%	13.9%	10.2%	7.1%

All fieldworkers were given a verbal briefing by their supervisor and were provided with written instructions specific to this project. They were also provided with:

[7] Only two respondents provided their telephone details for this purpose and neither of these could be reached.

- A Covid 19 fieldwork guidance document and declaration form: all fieldworkers were required to complete a declaration to state they had no Covid-19 symptoms and had not been in close contact with anyone suffering from Covid before each shift conducted
- Individual train batch sheets to record number of questionnaires/QR codes handed out/completed and to provide any relevant comments relating to the trains they were travelling on
- 100 x Paper self-completion questionnaires per shift
- 100 x QR codes with a link to the online version of the survey per shift

Fieldworkers were provided with a schedule covering a number of trains for them to travel on during shifts lasting approximately six hours. Depending on the routes being covered, questionnaires were handed out on two to eight trains in each shift. The table below shows an example of trains travelled on in a single shift.

Table 5 Example fieldwork shift

Day	TOC	Station Board	Time	Station Alight	Time
Weekday	Avant West Coast	Wolverhampton	06:27	Milton Keynes Central	07:41
		Milton Keynes Central	08:13	Wolverhampton	09:37
		Wolverhampton	09:45	Milton Keynes Central	10:58
		Milton Keynes Central	11.13	Wolverhampton	12:37

Overall, 309 fieldwork shifts were conducted across 22 individual Train Operating Companies.

Once shifts were completed, interviewers were required to inform their supervisor of the number of self-completion questionnaires completed, QR codes handed out and to detail any specific issues that they encountered during their shifts. These details were then passed back to the Senior Associate Director in Fieldwork services who passed this information on to the project team. All self-completion questionnaires were returned to the office via Special Delivery for processing and were ‘booked in’ and checked for completeness on arrival at the office.

Count sheets

As well as handing out and collecting questionnaires/QR codes or offering assistance, fieldworkers also completed count sheets on each train they travelled.

As part of this counting exercise fieldworkers asked all passengers (regardless of whether they participated in the research) whether they were travelling for Commuting, Business or Leisure purposes. They also recorded the individuals’ observable gender and age (with the age groups categorised as 16-34, 35-44, 45-64 and 65+). These counts were then further split by whether the individual took a paper questionnaire, requested a QR code or refused to participate in the research.

These count sheets were collected to compare the profile of passengers on the trains with the profile of respondents who completed the survey. This provides an indication of any groups that had lower response rates or were less likely to agree to participate in the research or to complete they questionnaires. Further details on the count sheet exercise

(including details on the level of completion of these sheets) are provided in the Weighting chapter of this document.

Quality Assurance: Fieldwork Checking Procedures

Fieldwork checking was conducted in three distinct phases:

- Ongoing conversation between interviewers/area supervisors and the Associate Director of Fieldwork Services
- Monitoring of every returned/completed questionnaire (including comments made by interviewers on shift detail sheets) to ensure every response fitted within routes and times booked for. This acts as an additional check that fieldworkers actually travelled on the trains they were booked for.
- Telephone back-checking conducted on over 10% of shifts (37 shifts in total) to check respondents were satisfied with the attitude, appearance and manner of the fieldworker and that fieldwork had been conducted as and when expected.[8]

Shifts were identified for inclusion in back-checking based firstly on any concerns over the quality of completion (only one shift was identified on this basis and it passed back-checking), recency of interviewers having been back checked on other projects (all interviewers will be back checked at least once every six months) and finally by random selection. All phases of research were covered by the back checking which was conducted between 26th July and 10th November 2022.

On a number of occasions shifts were either cancelled or moved at short notice, not conducted or curtailed due to punctuality issues (e.g., late trains leading to connections being missed, excessive heat leading to performance issues and guidance for passengers to not travel etc.), or interviewer related issues or illness. Where shifts were cancelled in one phase of research and could not be re-scheduled in that phase they were moved into a subsequent phase of research.

Interviewer notes from all shifts were examined and any shifts where responses were low were checked and replaced if required. Where these notes did not fully resolve any concerns, conversations took place between the Senior Associate Director and the area supervisor/interviewer and shifts were replaced if required.

[8] 18 TOCs were covered with at least one back checking shift with roughly 80% conducted on weekday shifts and 20% on weekend shifts. Given the 100% validation of journeys from the data provided, the back checking focussed on the manner of the interviewer: were they polite and courteous at all times; were they appropriately dressed; did they have any other comments either positive or negative regarding any interaction with our interviewer. There were no issues or concerns from any individuals. All contacts recalled the interviewer. No participant had any negative comments with regards to any interviewer. No negative issues were identified as part of this back checking with largely positive comments made relating to the project in general and the fieldworkers specifically.

5. Data collection, entry, and processing

Project specific editing instructions were written by senior data processing staff. All paper self-completion questionnaires were 100% sight edited in accordance with the editing instructions in order to identify and correct completion errors. These included errors in relation to consistency, journey allocation and following of questionnaire routing etc.

The basic principles of this sight editing were to ensure:

- Every question that did have an answer, should have had an answer.
- Every question that did not have an answer should not have had an answer
- All answers made logical sense e.g., no conflicting responses, or giving multiple responses to a single answer question, stations are stopping points on designated train. Where there are contradictions in responses the questionnaires are looked at in detail and either amended (where it is clear where the contradiction has occurred) or changed to 'missing data' for those questions so as not to provide misleading or false insight
- Questionnaires were completed by the relevant individual – e.g., checks for consistent handwriting across all interviews within a shift etc.

At the sight editing stage, any concerns that fieldworkers had not completed the shift as instructed were referred to the Field Department. If necessary, an urgent back-check was then undertaken to validate the work, or the questionnaires were removed from the sample. There was one instance of this happening during this project.

Any changes made at the sight editing stage were marked in red or green on the completed questionnaire to ensure there is a clear distinction between the original and edited data.

Spoiled self-completion questionnaires, or questionnaires which do not meet quality control procedures were removed from the sample. This includes questionnaires where inappropriate language is written in the open-ended questions or within the body of the questionnaire itself. Fewer than 0.5% of returned paper/completed online questionnaires were rejected.

A member of the research team confirmed coding requirements in advance of analysis, including for the coding of station names as well as full free text questions. Coders were given a briefing on the project, provided with access to the field materials and familiarised

themselves with the subject area in advance. Research on specific comments or abbreviations that were required took place throughout the coding process (e.g., being aware of the 3 letter station codes that some respondents can use rather than writing in the full name).

The first 10% of questionnaires returned were used to create a code frame (other than for station codes where a fill code frame of all UK rail stations had previously been used by Savanta). The initial 10% of open text responses were categorised into key themes and these themes used to create a draft of the code frame text. The code frame text was checked with the project team to ensure it was appropriate and a true reflection of the data set. Additional codes were added where appropriate to keep the level of 'other' answers to a maximum of 10% of the base wherever possible.

Named stations and journeys were individually validated using information available from specific TOCs, dynamic online timetable tools (including historic train information) and in conjunction with interviewer feedback from any shift detail sheets.

All paper questionnaires were entered manually by the in-house data entry team utilising specialist survey analysis software (QPS). Validations and logic, as stated in the editing instructions, were pre-programmed into the individual project set up within the analysis software. This set up was checked by a senior member of the data entry team prior to commencing data entry, using the 'Test paper input' function, and the set-up was amended as required. Before being entered into the system, all paper questionnaires were numbered with a unique serial ID to facilitate identification of any individual questionnaire (so that any issues noted in the data of individual paper questionnaires could be identified and examined in detail).

100% of questionnaires were validated based on their journey profile to ensure:

- That they were travelling on the train and shift required and that this is allocated to the correct TOC
- That the boarding and alighting stations were accurate for the journey being made
- That the specified stations were in the correct direction of travel for that train
- That the time of travel was relevant to the shift being conducted

All paper self-completion data cleaning was completed by returning to the actual questionnaire via its unique ID and amending accordingly rather than based on details contained within the resulting report. This also helps to identify whether data has been mis-keyed or become mis-aligned.

Copies of all data files prior to cleaning and throughout the data cleaning process were maintained electronically and contained in the project Archive folder. Final data files were clearly labelled. No sensitive personally identifiable information was entered or stored as part of this research: name and contact details were collected for fieldwork evaluation purposes only, and these details were not data entered or stored with response data.

In the final data set, there remain some cross-question discrepancies in a small number of responses: e.g., where a respondent reported that they made all of the journeys that they had planned to in one response to one question, but then chose a response that could not be true in that case in a follow-up question. Minor, low-incidence inconsistencies of this sort are a risk of using paper self-completion surveys. These responses were not removed

on the basis that the responses to each question must be taken as the respondents gave them, and the incidence of discrepancy is too low to alter the findings.

6. Weighting

Weighting is used to account for measurable variance within the achieved data sample and the profile of individuals within the population from which they are drawn. It accounts for variance by assigning a numeric value to each individual respondent such that the resulting sample is representative at a total level. This works on the assumption that responses will be influenced by certain characteristics of respondents, and weighting attempts to filter out that influence. For example, if you were to examine a sample that had 60 female respondents and 40 male respondents, but you knew that there was a 50:50 split in the population as a whole, then the female respondents would need a weighting factor of 0.83 (50/60) and the male respondents a weighting factor of 1.25 (50/40) to bring the sample in line with the population. In effect, if 3 women and 7 men in the original 'unweighted' sample (representing 10% of the total) said that they were very positive about something, the weighted result would be (3x0.83) + (7x1.25) = 11.24 or 11% of the overall sample. The assumption here is that the 11% result would be closer to what we would expect the true proportion to be in the population than the 10%, given the gender distribution of the population.

Due to the complexity of the weights that were required for the data in this project, a Random Iterative Method (RIM) weighting algorithm was applied. A more detailed definition of this approach can be found later in this chapter.

Weighting by size of TOC

The sampling plan (as detailed earlier in this document) was designed to provide robust data at an individual Train Operating Company level. In trying to ensure that all TOCs received large enough sample sizes, some TOCs were allocated more fieldwork shifts than would have been the case if allocating purely based on passenger journey proportions made on those TOCs. In effect, this meant there is an overrepresentation of TOCs carrying smaller numbers of passengers.

To adjust for this, scaling weights were calculated by comparing overall proportions of the samples achieved per TOC with the proportions of operating journeys allocated to each TOC in LENNON data covering all periods of fieldwork. In effect, overrepresented smaller TOCs are weighted down, and underrepresented larger TOCs are weighted up.

Weighting for non-response bias (based on Count sheet data)

As well as accounting for differentials in TOC network size and journey share, a non-response adjustment is also applied to account for differences in the overall profile of rail users observed during fieldwork and the profile achieved in the sample. As noted above in the fieldwork chapter, fieldworkers used count sheets to record profile data for people who took a paper questionnaire, those who took QR codes, and those who refused to participate. Categories recorded were journey purpose (commuter, business, leisure), observable age bracket (under 35, 35-44, 45-64, 65+), and observable gender (male, female). On review, commuter and business journey purpose counts, and counts for the middle two categories for age, were each combined into a single category, giving the following categories used in the final adjustment:

- Age: under 35, 35-64, 65+
- Gender: male, female
- Journey purpose: commuter/business, leisure

We recognise that there is the potential for statistical bias to be introduced through human error when applying this count method, and these counts do not give us a perfect indication of the population profile of rail users. However, there is no other currently available data that would give as accurate a profile of passengers for each TOC, split by age, gender and journey purpose. The justification for the adjustment is that it is likely to have improved the representativeness of the weighted sample overall by correcting for some differences in response for these groups.

Count sheets were examined on return and checked for missing data. On examination, there were some concerns over the completeness of count data on particular shifts, and, where issues were noted, specific count sheets (or sections of count sheets) were excluded from non-response adjustment calculations. Table 6 below lists the completion rates across TOCs.

Table 6 Count sheet completion by TOC

TOC	% of shifts that included useable count sheets	% of useable count sheets with no missing data	% of useable count sheets that included counts on those completing paper questionnaires	% of useable count sheets that included counts on those taking QR codes	% of useable count sheets that included counts on those refusing to participate
Avant West Coast	54%	20%	100%	70%	25%
C2C	57%	47%	96%	65%	76%
Chiltern	90%	28%	100%	56%	54%
CrossCountry	100%	30%	94%	82%	45%
East Midlands Railway	68%	32%	89%	55%	45%
Grand Central	88%	61%	78%	96%	96%
Thameslink/Great Northern	77%	23%	98%	60%	39%
Greater Anglia	45%	54%	100%	85%	78%
GWR	93%	51%	97%	86%	69%
Heathrow Express	37%	50%	73%	100%	80%
Hull Trains	88%	57%	100%	79%	82%
LNER	71%	38%	100%	74%	51%
LUMO	100%	33%	83%	100%	58%
Merseyrail	52%	27%	95%	97%	65%
Northern	82%	45%	75%	87%	80%
Southeastern	77%	34%	93%	59%	74%
Southern/Gatwick Express	68%	50%	97%	81%	63%
SWR	72%	56%	100%	79%	73%
TPE	75%	27%	87%	83%	63%
West Midlands Railway/LNWR	74%	11%	93%	73%	16%

The count sheets provided a mostly consistent picture across TOCs in terms of the types of passengers who were under or overrepresented in the final data due to response bias. However, all TOCs where fewer than 75% of shifts included useable count sheets were further examined to ensure that the data were providing a consistent indicator of response bias:

- Avanti West Coast findings were compared with other Long Distance TOCs (LNER and CrossCountry). Differences in the profiles seen suggested the count sheet data may not have been a reliable profile. Therefore, there was no non-response adjustment of responses for this TOC.
- c2c was compared with Chiltern and found to be consistent.
- Greater Anglia was compared to other London and South East TOCs and found to be consistent.
- Merseyrail was seen to exhibit similar differences between count sheet data and achieved proportions as seen across TOCs as a whole and, therefore, count sheet data was used for weighting.
- Due to a number of partially or incorrectly completed count sheets, Heathrow Express count sheets were not felt to provide robust data (and the differentials in figures were not consistent with what seen on other TOCs). Therefore, there was no non-response adjustment of responses from this TOC.

Random Iterative Method weighting

The final dataset was weighted to reflect the count sheet passenger profiles within TOC, except for in the cases of Avanti West Coast and Heathrow Express, where no adjustment was applied.

A combination of the two weighting adjustments (scaling for TOC journey proportion, and non-response bias adjustment) was achieved using a Random Iterative Method (RIM) weighting algorithm. RIM weighting is a frequently used quantitative market research technique. It is used when sample data is needed to be matched to a known profile amongst a number of characteristics, where there is no known relationship between these characteristics. The technique utilises an algorithm that allows for each characteristic to be weighted to the desired profile at the same time, whilst distorting each variable as little as possible. The RIM weighting algorithm proceeds through a number of iterations in order to match the set target values for all included variables.

RIM weighting works by what is known as an iterative target weighting process. Weights are iteratively adjusted for each case until the sample distribution matches the desired population for the variables that the data are being weighted on. For example, if we want to achieve a 40% female and 60% male weighted sample based on our count-sheet profiles, then weights for each observation are adjusted such that the weighted counts from our observations are 40% female and 60% male. Next, the algorithm adjusts the weights so that the weighed counts of our observations are in the right proportion for our age distribution. This will likely mean that the gender proportions are knocked out of balance with our desired (target) proportions, so the algorithm adjusts the weights again, iteratively. This process continues until all proportions of combinations of the characteristics that are being weighting to match our target population proportions.

Weighting targets

Adjustments for non-response bias were applied within TOC. Therefore, weighting was applied for each TOC, such that they were representative of the overall passenger journeys made on that TOC during our fieldwork period, and accounted for measurable response bias for age, gender and journey purpose within each TOC.

The overall weighting targets are shown in Tables 7 and 8 on the following pages.

Table 7 Targets weighted to in order to account for TOC size

TOC	% of sample TOC weighted to account for TOC journey share
Avant West Coast	2.5%
C2C	3.4%
Chiltern	2%
CrossCountry	2.9%
East Midlands Railway	2.6%
Grand Central	0.2%
Thameslink/Great Northern	13.6%
Greater Anglia	6.7%
GWR	8.2%
Heathrow Express	0.5%
Hull Trains	0.1%
LNER	2.3%
LUMO	0.1%
Merseyrail	2%
Northern	8.4%
Southeastern	12.1%
Southern/Gatwick Express	10.9%
SWR	14%
TPE	1.9%
West Midlands Railway/LNWR	5.6%

Table 8 Profile targets weighted to within TOC

TOC	Commuter/ Business traveller[9]	Leisure user[10]	Male	Female	Age: Under 30	Age 35-64	Age: 65+
Avant West Coast	41%	56%	51%	49%	35%	53%	12%
C2C	46%	54%	52%	48%	38%	50%	12%
Chiltern	53%	47%	51%	49%	37%	5451%	9%
CrossCountry	47%	53%	47%	53%	39%	48%	14%
East Midlands Railway	50%	50%	46%	54%	42%	47%	11%
Grand Central	17%	83%	47%	53%	27%	61%	12%
Thameslink/Great Northern	57%	43%	53%	47%	36%	54%	10%
Greater Anglia	43%	57%	51%	49%	39%	48%	13%
GWR	33%	67%	52%	48%	44%	43%	13%
Heathrow Express	33%	67%	54%	46%	42%	54%	4%
Hull Trains	39%	61%	50%	50%	31%	58%	11%
LNER	34%	66%	47%	53%	33%	55%	13%
LUMO	22%	78%	48%	52%	32%	47%	11%
Merseyrail	62%	38%	53%	47%	45%	43%	12%
Northern	37%	63%	50%	50%	43%	49%	8%
Southeastern	53%	47%	52%	48%	35%	54%	10%
Southern/Gatwick Express	43%	57%	50%	50%	38%	52%	10%
SWR	39%	61%	52%	48%	38%	49%	13%
TPE	34%	66%	53%	47%	45%	49%	7%
West Midlands Railway/LNWR	43%	57%	50%	50%	45%	49%	7%

In all cases, any missing data (e.g., non/other response to the profiling questions) have been kept at the same proportion as within the raw data, and those factors that were weighted had their relative proportions matched to the profiles outlined above. For example, if on Hull Trains (where there is a 50:50 split for male/female) we had 2% non-response, the weight applied would be 49% male, 49% female and 2% non-response.

Impact of weighting

Weighting has an overall impact on the effective sample size at an aggregate level and also within individual sub-groups. Whilst there is some impact noted here, this is largely caused by the scaling weight applied to account for variances in size of TOC, and overall effective sample sizes remain robust. Table 9 below summarises.

[9] Commuter/Business traveler definition is based on those classifying their journey purpose as: Commuting to/from work, Commuting to education (to/from college/school/university), Escorting a dependant for education or other purposes or On company business (or own if self-employed).

[10] Leisure users definition is based on those classifying their journey purpose as: Travelling to a health appointment (GP, hospital, dentist etc.), Other personal business (job interview, banking etc.), Visiting friends or relatives, Shopping trip, Travelling to/from holiday, Travelling to play sport, Travelling to watch sport or Other leisure trip.

Table 9 Effective sample sizes after weighting

	Unweighted base size	Effective sample size after weighting
All respondents	17383	11400
Commuter/Business Traveller	6638	4832
Leisure user	10535	6451
Age: 16-34	7041	4675
Age: 35-64	7372	4865
Age: 65+	1581	1015
Male	7100	4733
Female	8156	5249
Avant West Coast	727	727
C2C	639	617
Chiltern	763	744
CrossCountry	799	739
East Midlands Railway	924	857
Grand Central	526	497
Thameslink/Great Northern	1771	1674
Greater Anglia	565	519
GWR	1373	1350
Heathrow Express	531	531
Hull Trains	794	725
LNER	744	694
LUMO	420	356
Merseyrail	564	490
Northern	937	864
Southeastern	921	908
Southern/Gatwick Express	1400	1347
SWR	1013	1004
TPE	961	933
West Midlands Railway/LNWR	1011	986

Summary of size of weighting factors applied

A general rule of thumb in survey analysis is to keep weighting factors between 0.5 and 2 (unless there is strong justification for using alternative weights), so that no individual response is treated as too important or reduced to the point of not contributing. In terms of the individual weighing factors applied to the data, 29% of respondents received a weighting factor of less than 0.5 and 12% received a weighting factor of 2 or above. However, this is largely driven by the need to account for differentials in relative size of TOCs (as opposed to the non-response bias weightings), and in this instance the variation in overall size across TOCs and the requirement to produce large enough samples by TOCs for analysis at that individual level means that it is justified. It is important to note

that despite these weightings the effective sample size remains robust and allows for detailed analysis.

Confidence intervals

Our sampling approach means that the result is not a simple random sample, which could only be achieved with a sample frame of every individual who intended to travel by rail during a strike week. To provide a rough indication of how the confidence intervals for results vary according to sample sizes and proportions, table 10 shows what intervals would apply for a random sample. Due to the sample design, the intervals for this sample would be consistently a little larger than those shown here (although the exact intervals for this sample method cannot be calculated). Confidence intervals are provided at a 95% confidence level and based on 10%/90%, 30%/70% and 50% of respondents giving a specific response.

Table 10 Example confidence intervals

Indicative data cell	Sample Size	Confidence Interval (to one decimal place)		
		10%/90%	30%/70%	50%
All respondents	17,383	+/- 0.4	+/- 0.7	+/- 0.7
50% of sample	8,500	+/- 0.6	+/- 1.0	+/- 1.1
25% of sample	4,000	+/- 0.9	+/- 1.4	+/- 1.6
Larger Train Operating Company	1,000	+/- 1.9	+/- 2.8	+/- 3.1
Smaller Train Operating Company	500	+/- 2.6	+/- 4.0	+/- 4.4

Annex A: Detailed Sampling By TOC

The “baseline week” referred to below is the baseline period of 6 – 12 June 2022: passenger journey data from LENNON was used from this week to estimate passenger flows on those routes in a typical week prior to the beginning of strike action (see "How the sampling plan was constructed" in chapter 2). Where journey numbers are stated, these are estimates based on sums of journeys allocated to trains on those routes over the baseline period.

Avanti West Coast

All English routes covered. Sample ensured the number of ‘trains’ covered on the mainline route were in line with journeys on that route in baseline week (around 60%). Glasgow/Edinburgh not covered (outside of England)

Routes covered
London Euston - Blackpool North/Crewe/Scotland
London - Holyhead
London – Manchester via Crewe
London – Manchester via Stoke
London – Liverpool Lime Street
London - Shrewsbury
London – Wrexham General
London – Crewe

CrossCountry Trains

All routes covered other than Scottish routes. Sampling ensured we had coverage both north and south of Birmingham on the Southwest to Scotland mainline route and that the number of trains covered was representative of the proportion of journeys on this core route in the baseline week (around 45%). Newcastle/Aberdeen not covered (route passes into Scotland).

Routes covered
Penzance - Edinburgh Waverley
Aberdeen – Plymouth
Bournemouth - Manchester Piccadilly
Birmingham - Newcastle
Stansted Airport - Birmingham New Street
Reading - Birmingham
Nottingham - Gloucester
Birmingham New Street – Nottingham
Cardiff Central - Birmingham New Street
Birmingham New Street - Manchester Piccadilly

Great Western Railway

All main routes (including North Cotswolds) were covered and four of 19 branch lines were covered. This split reflects the relative numbers of passenger journeys made on these branch lines (in the baseline period) compared to mainline services. Sample structured such that train coverage reflected journeys by route type: Mainline 66%, North Cots 14%, Branch 19% and Bristol/Southampton 1%.

Main Routes covered
Paddington Mainline
North Cotswolds
Bristol - Southampton

Branch lines covered
Reading- Redhill
Basingstoke - Reading
Newquay – Par
Looe – Liskeard

Branch lines not covered
Exmouth - Exeter
Bristol Temple Meads - Severn Beach
Slough - Windsor & Eton Central
Falmouth Docks - Truro
Marlow - Maidenhead
Paignton - Newton Abbot
Barnstaple - Exeter St David’s
Henley-on-Thames - Twyford
Okehampton - Exeter St David’s

Branch lines not covered
Weymouth - Castle Cary
Redhill - Gatwick Airport
Gunnislake - Keyham
Greenford - West Ealing
St Ives - St Erth
Paddington London - Heathrow Rail

Transpennine Express

Routes covered
Manchester Victoria - Newcastle
Manchester Piccadilly - Scarborough
Manchester Piccadilly - Redcar Central
Manchester Piccadilly - Hull
Manchester Airport - Cleethorpes
Manchester Airport - Glasgow Central/Edinburgh
Liverpool Lime Street - Manchester Victoria

Chiltern Railways

Routes covered
Aylesbury - Marylebone London
Aylesbury Vale Parkway - Marylebone London
Oxford - Marylebone London
Birmingham Moor Street - Marylebone London
Marylebone London - Banbury
Leamington Spa - Stratford-upon-Avon

Greater Anglia

All routes carrying at least 27,000 passenger journeys in the baseline week were covered. Six of 18 smaller routes were covered, ranging in size from 15,000 passenger journeys to 2,000.

Analysis after phase two demonstrated that there was under representation of the Liverpool Street to Stansted route, and three shifts were added to account for this. The remaining 12 routes were stratified on size and a random route chosen between one and six and then the sixth shift was chosen from here.

Routes covered
Liverpool Street London - Stansted Airport
Liverpool Street London - Colchester
Liverpool Street London - Southend Victoria
Liverpool Street London - Norwich
Cambridge - Norwich
Cambridge - Stansted Airport
Peterborough - Ipswich
Norwich – Lowestoft
Norwich - Great Yarmouth via Lingwood
Ipswich - Lowestoft
Wickford - Southminster

Routes not covered
Braintree - Shenfield
Cambridge – Ipswich
Norwich - Sheringham
Norwich - Great Yarmouth via Cantley
Hertford East - Broxbourne
Ipswich - Felixstowe
Ipswich - Harwich Town
Colchester - Clacton-on-Sea
Walton-on-the-Naze - Thorpe-le-Soken
Sudbury - Marks Tey

South West Rail

Due to the number of routes, for sampling purposes these were initially categorised into 100,000+ passenger journeys, 50,000-99,999 journeys and less than 50,000.

All routes with 100,000 passenger journeys were included and made up roughly 26% of trains covered.

Routes from 50,000-99,999 and under 50,000 were stratified on size and routes randomly selected. After phase two coverage, four routes were chosen from 50,000-99,000 journeys to ensure coverage of around 62% of trains, and three routes were chosen from routes with fewer than 50,000 journeys, accounting for roughly 12% of trains.

Routes covered
Waterloo loop
Waterloo - Weybridge
Waterloo - Hampton Court

Routes covered
Waterloo London - Woking
Waterloo - Windsor and Eton Riverside
Waterloo London - Weymouth
Guildford - Waterloo
Waterloo London - Reading
Dorking - Waterloo
Waterloo London - Portsmouth Harbour (via Guildford)
Waterloo London - Exeter St Davids
Waterloo London - Portsmouth Harbour (via Basingstoke)
Southampton Central - Portsmouth & Southsea
Salisbury - Southampton Central

Routes not covered
Waterloo - Chessington South
Alton - Waterloo
Shepperton - Teddington
Lymington Pier - Brockenhurst
Yeovil Junction – Westbury

Heathrow Express

Only one route served – covered across times of day and days of week.

West Midland Trains

All of the eight routes carrying 10,000 or more passenger journeys were covered by the sampling plan. Of the eight smaller routes three were covered through random selection: this is in line with their overall passenger numbers compared to the TOC as a whole.

Samples in phases one and two oversampled routes with under 20,000 passenger journeys. There was also a need to boost Euston (London Northwestern services) for phases three and four to ensure coverage.

Routes covered
London Euston - Northampton
London Euston - Birmingham New Street
London Euston - Tring
Bromsgrove - Lichfield Trent Valley
London Euston - Crewe
Redditch - Four Oaks
Birmingham New Street - Liverpool Lime Street
Birmingham New Street - Shrewsbury
Birmingham Snow Hill - Kidderminster

Routes covered
Birmingham Moor Street - Stratford-upon-Avon
Birmingham New Street - Rugeley Trent Valley

Routes not covered
Birmingham New Street - Hereford
Leamington Spa - Nuneaton
Northampton - Stafford
Watford Junction - St Albans Abbey
Bedford - Bletchley

LUMO

Limited routes, all covered across times of day and days of week.

Southeastern

Routes covered
Charing X – Gravesend
Charing X – Dartford (via Bexleyheath) s
Charing X – Sevenoaks
Charing X – Dartford (via Woolwich)
Charing X – Hastings
Victoria – Ramsgate
Cannon Street – Orpington
Charing X – Hayes
Victoria – Dover
Charing X – Ramsgate (via Tonbridge)
St Pancras – Margate
St Pancras – Dover
Strood – Tonbridge
Sheerness – Sittingbourne
London Victoria – Ashford

Thameslink and Great Northern

Sampled as two separate TOCs. All routes covered across both TOCs

Routes covered
Brighton - Bedford
Cambridge - Brighton

Routes covered
Horsham - Peterborough
St Albans City - Sutton (loop)
Rainham - Luton
King's Cross London - Peterborough
King's Cross London - Cambridge
London Kings Cross - Cambridge
Brighton - Victoria London
London Blackfriars - Sevenoaks
Moorgate London - Welwyn Garden City
Moorgate London - Stevenage
King's Cross London - Kings Lynn

Southern and Gatwick Express

Sampled as two separate TOCs.

- Southern: All six routes that carry more than 40,000 passenger journeys were covered. Of the remaining nine smaller routes three were included through stratification and random selection.
- Gatwick Express: one route covered at different times of the day/days of the week.

Southern routes covered
London Victoria - Ore
London Victoria - Bognor Regis
Brighton - Southampton Central
Brighton - Hastings
Brighton - Littlehampton
London Victoria - Epsom Downs
London Bridge - Caterham
Ashford International - Eastbourne
London Victoria - East Grinstead
London Bridge - Tattenham Corner
London Bridge - Uckfield
Haywards Heath - Lewes

Southern routes not covered
Brighton - Seaford
London Bridge - Beckenham Junction
London Bridge - East Croydon
London Victoria - West Croydon
Victoria London - Beckenham Junction
Portsmouth & Southsea - Bedhampton

Northern Rail

This is a large operator. The routes were stratified into those carrying 60,000 or more passenger journeys, 59,999 to 20,000 journeys, 19,999 to 10,000 journeys and less than 10,000 journeys. Both routes carrying over 60,000 passenger journeys were included in the sample. The services within the other bands were stratified on size for phases 3 and 4 and routes randomly selected to be representative of the number of journeys carried on routes of those sizes. Overall, 16 routes were covered within the sampling plan.

Phases one and two routes covered
Bradford – Leeds
Liverpool – Manchester
Sheffield to Huddersfield
Middlesborough – Whitby
Leeds – Doncaster
Wigan – Liverpool
Blackpool – Preston
Preston – Burnley
York – Bradford
Newcastle – Middlesborough

Phases three and four routes covered
Leeds - York
Carlisle - Leeds
Leeds - Doncaster
Crewe - Stockport
Warrington Central - Manchester Oxford
Manchester Victoria – Halifax

The following routes were not directly chosen as part of the sampling process, but some may have been covered as sub-routes within the route map-based sampling for phases one and two:

Additional routes covered
Carnforth - Leeds
Ilkley - Leeds
Wakefield Kirkgate - Nottingham
Carlisle - Silverdale
Carlisle - Newcastle
Liverpool Lime Street - Manchester Piccadilly
Doncaster - Sheffield
Leeds - Hull
Saltburn - Darlington

Additional routes covered
Ilkley - Bradford Forster Square
Preston - Bradford Interchange
Middlesbrough - Newcastle
Blackpool North - Preston
Scarborough - Hull
Liverpool Lime Street - Wigan North Western
Lincoln - Doncaster
Chester - Stockport
Moorthorpe - Sheffield
Crewe - Mauldeth Road
Stoke-on-Trent - Manchester Piccadilly
Hull - Doncaster
Bolton - Preston
Cononley - Kirkstall Forge
Leeds - Rawcliffe
Windermere - Wigan North Western
Wigan Wallgate - Stalybridge
Leeds - Wakefield Kirkgate
Leeds - Mirfield
Thornaby - Bishop Auckland
Blackpool South – Preston
Bolton - Manchester Piccadilly
Sheffield - New Mills Central
Buxton - Stockport
Hadfield - Manchester Piccadilly
Whitby - Thornaby
Newcastle - Chathill
Blackburn - Bolton
Huddersfield - Barnsley Interchange
Liverpool Lime Street - Warrington Central
Heysham Port - Lancaster
York - Pontefract Baghill
Wigan Wallgate - Southport
York - Micklefield
New Mills Central - Manchester Piccadilly
Chesterfield - Nottingham
Chester - Warrington Bank Quay
Walsden - Rochdale
Newcastle - Durham
Colne - Rose Grove
Scunthorpe - Doncaster
Huddersfield - Halifax
Ormskirk - Preston
Wigan North Western - Salford Crescent
Wressle - Hessle
Rochdale - Moston
Ince - Moorside
Romiley - Guide Bridge
Wigan Wallgate - Kirkby

C2C

Routes covered
Fenchurch Street London - Southend Central
Fenchurch Street London - Shoeburyness
Fenchurch Street London - Grays

East Midlands Railway

All five routes carrying 10,000 or more passenger journeys were covered. Five of the six routes carrying less than 10,000 but more than 4,000 passenger journeys were covered. Three very small routes with 1,103 or fewer passenger journeys were not covered.

Sampling ensured that the three busiest routes (St Pancras to Sheffield, Leicester to Nottingham, and Liverpool Lime Street to Norwich) accounted for roughly 55% of trains covered. Map based sampling for phases one and two provided good coverage of smaller routes.

Routes covered
St Pancras International London - Sheffield
Leicester - Nottingham
Liverpool Lime Street - Norwich
St Pancras International London - Nottingham
Nottingham - Skegness
Nottingham - Lincoln
Worksop - Nottingham
Crewe - Derby
Matlock - Derby
St Pancras International London - Melton Mowbray

Routes not covered
Doncaster - Peterborough
Habrough - Lincoln
Cleethorpes - Habrough
Barton-on-Humber - Habrough

LNER

All routes mostly in England covered except for four routes carrying fewer than 1,200 passenger journeys. Examination of the map based sampling for phases one and two

demonstrated good overall coverage with a need to increase the number of shifts on more popular routes to ensure their coverage was representative of passenger numbers.

Routes covered
King's Cross London - Edinburgh Waverley
King's Cross London - Leeds
King's Cross London - Lincoln
York – Sunderland
Newcastle - Edinburgh Waverley
Aberdeen - Edinburgh Waverley
Leeds - Harrogate

Routes not covered (all except Newcastle to Inverness are limited services with fewer than 1,200 passenger journeys):

Routes not covered
Newcastle - Inverness
Kings Cross London - Hull
Glasgow Central - Edinburgh Waverley
Leeds - Bradford Forster Square
York - Middlesbrough
Leeds - Skipton

Grand Central

Only two routes: both covered.

Routes covered
King's Cross London - Sunderland
King's Cross London - Bradford Interchange

Hull Trains

Only one route served – covered across all times of day and day of week.

Merseyrail

Merseyrail was only covered in phase four of the research.

At request of Merseyrail, questionnaires on this TOC had a small amendment to the introductory text with the inclusion of the following wording:

‘Industrial action can impact rail services and stations in a variety of ways. Operations can be impacted even if staff from particular unions have not opted to carry out industrial action. Keeping this in mind, would you be happy to answer some questions relating to your travel?’

Routes covered – all during phase four
Chester - Liverpool Lime Street
Ellesmere Port - Liverpool Lime Street
Hunts Cross - Southport
New Brighton - Liverpool Lime Street
Sandhills - Ormskirk
West Kirby - Liverpool Lime Street