Transport Focus

NRPS Spring 2020

Quality Assurance Statement

Client contact: David Greeno
Senior Insight Advisor
david.greeno@transportfocus.org.uk
0300 123 0837

Agency contact: David Chilvers
Managing Director
david.chilvers@dcaweb.co.uk
07973 671 990

Contents

Summary

Fieldwork on the Spring 2020 NRPS wave was suddenly curtailed in early March due to the onset of Coronavirus. This meant that target sample sizes were not achieved and there was no real opportunity (as is normally the case) to use the last 2-3 weeks of fieldwork to target specific sample cells that were most adrift of target.

The lower sample size naturally leads to wider margins of error. The lack of targeting has led to lower weighting efficiencies for most TOCs reducing effective sample sizes and further widening margins of error. A small number of sample cells had to be merged due to low sample sizes.

Apart from widening margins of error, the loss of fieldwork does not appear to have led to any unusual results for the key metrics at national level or for most TOCs. The widening of margins of error has been taken into account in identifying any significant shifts in reported metrics.

However, Merseyrail clearly stands out from other TOCs, with six metrics that show spikes in Spring 2020 (which are not continuations of previous trends) and the TOC featuring on four of the six criteria used to judge whether there are any specific issues with the Spring 2020 NRPS wave. There may be good reasons for the changed perceptions if the TOC has undertaken specific related actions but if this is not the case, we would suggest considerable caution is used when interpreting the results for Merseyrail.

A number of TOCs have low rankings on more than one of the criteria used to judge whether there are any specific issues with the Spring 2020 NRPS Wave. These TOCs are Great Northern, East Midlands Railway, London North Eastern Railway, c2c, Chiltern Railways, Transport for Wales and TransPennine Express. With these TOCs, we would suggest greater than usual care is taken when interpreting the results of the Spring 2020 wave.

For the other TOCs, we see no great evidence that the curtailment of fieldwork has had any major impact on the results and we suggest that the metrics for these TOCs meet the normal NRPS standards, with the caveat that reduced sample sizes and weighting efficiencies will widen the margins of error. These widened margins of error will be in any event used to calculate whether there have been significant changes and we recommend that the NRPS User Guide is updated to reflect these wider margins for this wave.

Further details are available from a full Quality Assurance Report, which are detailed below. The updated User Guidance and Technical Reports will also incorporate the impact of both the reduced sample sizes and the lower weighting efficiencies seen.

Background

Fieldwork on the NRPS Spring Wave 2020 was suspended in early March, following the spread of Coronavirus. As a result, only around 19,500 interviews were completed, approximately 75% of the overall target. To maximise the overall sample size, data from the Network Rail boosts were used to generate data in the main reports. The completion rate ranged from 62% to over 100% for individual TOCs.

The last weeks of fieldwork for NRPS are used to ensure targets are met and to focus fieldwork onto areas where TOC profiles are adrift from target in terms of day of week, journey purpose and station size band. Due to the sudden suspension of fieldwork, no such focus has been possible for the Spring 2020 wave and as a result some TOCs have profiles that are adrift of target. In addition, returns for some individual sample cells are too low to allow weighting for that particular cell and a number of station size band cells have been merged with adjacent cells to facilitate weighting.

As a result, weighting efficiencies are generally lower and subsequent margins of error are generally higher than is normal for NRPS and output needs to be interpreted with more caution than is normally the case. Transport Focus has therefore produced this Quality Assurance statement which can be placed on the website and inserted, as necessary, into NRPS Spring 2020 Wave reports.

The Quality Assurance statement provides a permanent record of:

- What shortfalls in sample sizes occurred?
- Which sample cells were grouped together for weighting purposes?
- What the impact has been on the accuracy and robustness of NRPS statistics?

The document:

- Chronicles what has been done, with relevant justification
- Identifies any areas where there are specific issues with the data
- Confirms that the processes used have been the best that could have been applied given the circumstances

The report provides a permanent record of processes used and outcomes derived which can set the NRPS Spring 2020 Wave in its proper context.

Approach

To construct this Report, it has been necessary to analyse the NRPS Spring 2020 data and examine the process used to create weights. A number of issues have been covered including:

- What impact has the smaller sample size and revised weighting made upon confidence intervals?
- What impact have any failures to match TOC profiles had on key metrics?
- What has been the impact of combining cells to create weights?

Reductions in sample size

Table 1 below shows the sample sizes achieved for each TOC in the Spring 2020 wave.

Table 1 – TOC sample sizes achieved compared to target

TOC	Achieved	Target	% target
Avanti West Coast	641	1000	64%
c2c	619	1000	62%
Chiltern Railways	818	1000	82%
CrossCountry	932	1200	78%
East Midlands Railway	637	1000	64%
Gatwick Express	416	500	83%
Grand Central	395	500	79%
Great Northern	314	500	63%
Great Western Railway	1032	1500	69%
Greater Anglia	1016	1300	78%
Heathrow Express	533	500	107%
Hull Trains	427	500	85%
London North Eastern Railway	753	1000	75%
London Overground	1244	1600	78%
Merseyrail	525	700	75%
Northern	1004	1400	72%
ScotRail	1025	1300	79%
South Western Railway	1471	2000	74%
Southeastern	1345	1500	90%
Southern	899	1300	69%
TfL Rail	825	1000	83%
Thameslink	852	1000	85%
Transport for Wales	781	1000	78%
TransPennine Express	668	1000	67%
West Midlands Trains	675	1000	68%
Total	**19847**	**26300**	**75%**

Most TOC sample sizes cluster around the 75% overall average but C2c, Great Northern, East Midlands Railway and Avanti West Coast are all below 65% and Southeastern and Heathrow Express are above 85%. We might expect issues to be more prevalent among those TOCs with the lowest sample sizes.

Merging station size band cells for TOC Building Blocks

NRPS data is grossed up by station size band and passenger numbers at TOC Building Block level. Four station size bands (small, medium, large, very large) are set for each Building Block in a way that tries to allocate around 25% of annual passenger volumes to each of the four building blocks. Sometimes, particularly for TOCs with few stations, the number of size bands is reduced. The NRPS sample attempts to replicate the station size band distribution, to minimise weighting effects.

TOC Building Blocks have target sample sizes that vary from 100 to 800 and so the targets for individual station size bands vary from around 25 to around 200. Even in a normal NRPS Wave, it may be necessary to merge some size bands for a particular TOC Building Block if the sample size is too small to warrant weighting that group.

For the Spring 2020 Wave, the number of such merges has been increased due to the lower overall sample size and failure to be able to use the last 2-3 weeks of fieldwork to target the lowest cells. Table 2 below shows which station size bands have been merged in this wave (the figure in parentheses is the sample size for the size band):

Table 2 – list of TOC Building Block size bands that have been merged

TOC Building Block	**Size bands being merged**
Merseyrail – Wirral	Small (8), Medium (39)
ScotRail – Strathclyde	Medium (16), Large(66)
South West Trains - Metro	Small (5), Medium (107)
South West Trains – outer suburban	Small (25), Medium (148)
TransPennine Express – North West	Medium (23), Large (0), Large (1)

All other things being equal, we might expect key metrics for the TOCs listed above to have greater than average impact from failures to achieve cell targets and the consequent need to merge cells.

There are a large number of Building Blocks where the target sample size is outside the range 50% - 100% but the key issue will be where one block is over sampled and another under.

Table 3 below shows those TOCs where at least one Building Block is below 50% of target (coloured red) and at least one above 100% of target (coloured green).

Table 3 – TOCs with both low and high sample as a % of target

TOC	Building block
Chiltern Railways	Commuter
Chiltern Railways	Oxford
Chiltern Railways	West Midlands
London North Eastern Railway	London - Newcastle/Sunderland and East Yorkshire
London North Eastern Railway	London – Scotland
ScotRail	Rural
ScotRail	Urban
Thameslink	North/South
Thameslink	Kent
Thameslink	Loop
TransPennine Express	North West
TransPennine Express	South
TransPennine Express	North
Transport for Wales	Interurban
Transport for Wales	Mid Wales and Borders
Transport for Wales	Cardiff and Valleys
Transport for Wales	North Wales and Borders
Transport for Wales	South Wales and Borders/West Wales

On the basis of this analysis, we might expect the data to be more stretched for Chiltern Railways, London North Eastern Railway, ScotRail, Thameslink, TransPennine Express and Transport for Wales.

Effective sample size and weighting efficiency

We would expect the overall weighting efficiency to reduce in the Spring 2020 Wave of NRPS as the over and under sampling of particular Building Blocks will be stretching the data to meet the rim weighting totals that have been set. And we would expect this effect to be particularly pronounced for TOCs where either cells have been merged (as shown in Table 2) or where particular building blocks have been over or under represented.

We would also expect weighting efficiency to be impaired for TOCs where the overall targets by journey purpose or weekday/weekend have not been met. Table 4 shows the weighting efficiency each TOC has achieved in the Spring 2020 NRPS Wave. Overall, there is a reduction of 4% in the overall weighting efficiency as anticipated above, but there are certain TOCs where the weighting efficiency has dropped more than this (and others where it has actually increased)

Table 4 – weighting efficiency for each TOC

TOC	weighting efficiency Spring 2020	weighting efficiency Autumn 2019	Change
Avanti West Coast	69%	70%	-1%
c2c	71%	76%	-5%
Chiltern Railways	57%	63%	-6%
CrossCountry	77%	74%	2%
East Midlands Railway	58%	66%	-8%
Gatwick Express	49%	48%	1%
Grand Central	84%	55%	29%
Great Northern	80%	84%	-4%
Great Western Railway	50%	50%	1%
Greater Anglia	49%	52%	-3%
Heathrow Express	77%	71%	6%
Hull Trains	68%	59%	9%
London North Eastern Railway	44%	48%	-4%
London Overground	72%	84%	-12%
Merseyrail	35%	57%	-22%
Northern	61%	66%	-5%
ScotRail	52%	59%	-8%
South Western Railway	71%	74%	-2%
Southeastern	53%	55%	-2%
Southern	86%	85%	2%
TfL Rail	65%	75%	-11%
Thameslink	78%	91%	-14%
Transport for Wales	31%	32%	-1%
TransPennine Express	61%	57%	4%
West Midlands Trains	73%	72%	1%
Total	62%	66%	-4%

The TOCs with the biggest falls in weighting efficiency are Merseyrail, Thameslink, London Overground, TfL Rail and East Midlands Railway. The TOCs with the lowest absolute weighting efficiencies are Transport for Wales, Merseyrail and London North East Railway.

Multivariate analysis

If there are instabilities in the data caused by the curtailment of the sampling plan, we would expect the results of the multivariate analysis to change significantly compared to previous waves. We have compared the importance of each metric as a driver for satisfaction with the results from the Autumn 2019 analysis and computed the difference for each factor, the maximum absolute difference and the average difference (there is not a problem comparing Spring with Autumn here as the multivariate analysis uses the last two waves to generate results). We have done the same for the analysis of dissatisfaction. Table 5 below shows the results:

Table 5 – results of comparisons of importance of each metric from multivariate analysis

TOC	sat-biggest	sat-ave	dissat-biggest	dissat-ave
National	3%	0.31%	1%	0.14%
London and South East	3%	0.36%	2%	0.33%
Long Distance	5%	0.79%	4%	0.51%
Regional	4%	0.59%	3%	0.50%
South Western Railway	4%	0.67%	7%	0.70%
CrossCountry	7%	1.01%	7%	0.91%
East Midlands Railway	6%	0.81%	9%	1.24%
West Midlands Trains	9%	1.17%	7%	1.10%
Thameslink	4%	0.80%	16%	1.59%
Great Western Railway	8%	0.85%	8%	1.76%
Transport for Wales	5%	0.93%	9%	1.68%
c2c	13%	1.04%	9%	1.18%
Southeastern	7%	0.88%	22%	1.60%
TransPennine Express	16%	1.61%	4%	0.80%
Northern	14%	1.56%	5%	0.86%
Greater Anglia	8%	1.04%	12%	1.64%
ScotRail	8%	1.70%	10%	0.97%
London Overground	8%	0.95%	18%	1.69%
Avanti West Coast	10%	1.71%	8%	1.20%
Chiltern Railways	12%	1.40%	11%	1.37%
Hull Trains	10%	1.42%	9%	1.84%
London North Eastern Railway	12%	1.33%	19%	1.62%
Merseyrail	9%	1.42%	11%	1.77%
Southern	13%	1.17%	12%	1.87%
Heathrow Express	9%	1.42%	26%	2.40%
Great Northern	16%	2.95%	8%	2.06%
TfL Rail	13%	1.70%	18%	2.30%
Gatwick Express	13%	2.20%	21%	1.93%
Grand Central	19%	2.54%	28%	4.48%

The four columns are defined as:

- Sat-biggest: the largest absolute change in the importance coefficient of any metric
- Sat-ave – the average absolute change across all metrics
- Dissat-biggest: the largest absolute change in the importance coefficient of any metric
- Dissat-ave – the average absolute change across all metrics

The TOCs have been shown in ascending order of the average rank of these four quantities.

At national level, there is little change in the factors driving overall satisfaction and dissatisfaction, with the maximum change between this wave and the previous wave being 3% and 1% respectively. South Western Railway has the smallest overall change when the rank each of the four measures is averaged and Grand Central the largest (but Grand Central has a small sample size and as the fieldwork is undertaken on train, the data for this TOC normally only covers a small number of distinct services).

Different TOCs feature in different tables, so to try and pull together all the factors that might impact on whether the weighting regime has unduly affected results for this wave, we have combined the following:

- Achieved sample size (lower will be most affected)
- Weighting efficiency (lower will be most stretched by weighting)
- Change in weighting efficiency (lower will be more affected compared to previous waves)
- % of sample target achieved (lower will be more affected compared to previous waves)
- Whether sample cells have been merged (more will be more affected); those with no merged cell have a rank of 9.5, those with 1 have a rank of 21, those with two have a rank of 24. This has been done to allow this factor to be evaluated on the same 1-25 scale as all the other factors
- Average multivariate analysis change. Some TOCs have equal rankings which lead to the figures shown

To enable these dimensions to be aggregated, we have generated the rank for each TOC for each dimension, where 25 is the lowest and 1 the highest and then sorted by the average rank. Table 6 below shows the results:

Table 6 – ranking of each TOC by different factors

TOC	Sample size	Weff	weff change	% target sample	Merged cells	multivariate	average
Merseyrail	21	24	25	15	21	15.75	20.29
Great Northern	25	3	15	24	9.5	19.5	16.00
East Midlands Railway	18	16	21	23	9.5	7	15.75
London North Eastern Railway	14	23	16	14	9.5	15.75	15.38
c2c	19	10	17	25	9.5	11.25	15.29
Chiltern Railways	12	17	19	7	21	13.75	14.96
Transport for Wales	13	25	10	11	21	8.75	14.79
TransPennine Express	16	15	4	21	21	11.25	14.71
Gatwick Express	23	21	8	5	9.5	21.75	14.71
ScotRail	5	19	20	9	21	12	14.33
Avanti West Coast	17	11	11	22	9.5	13.25	13.96
TfL Rail	11	13	22	6	9.5	20.25	13.63
Northern	7	14	18	17	9.5	11.25	12.79
Greater Anglia	6	22	14	10	9.5	11.75	12.21
Great Western Railway	4	20	9	19	9.5	8.75	11.71
Grand Central	24	2	1	8	9.5	24.75	11.54
London Overground	3	8	23	12	9.5	13	11.42
West Midlands Trains	15	7	7	20	9.5	7.75	11.04
South Western Railway	1	9	13	16	24	2	10.83
Hull Trains	22	12	2	3	9.5	14.5	10.50
Southern	9	1	6	18	9.5	17	10.08
Thameslink	10	4	24	4	9.5	8.25	9.96
Heathrow Express	20	5	3	1	9.5	19	9.58
Southeastern	2	18	12	2	9.5	11.25	9.13
CrossCountry	8	6	5	13	9.5	5.5	7.83

On this basis, Merseyrail is the TOC most likely to have been affected by the reduced sample size as it appears in the bottom five TOCs for four of the dimensions and is 15th in terms of what % of the target sample has been achieved. If the data for Merseyrail appears unaffected by the reduced sample, it is likely that the same conclusion can be drawn for other TOCs.

In the next section, we therefore analyse data for:

- Merseyrail – as the most affected TOC
- East Midlands Railway - affected a lot by the first four factors
- Transport for Wales – the TOC with the lowest weighting efficiency
- Northern – a TOC near the average on this table
- Southern – a commuter TOC near the bottom of the table

In analysing the key results for these TOCs, we first of all examine the national picture, to see if there are overall performance changes which might affect all TOCs.

National results

Table 7 below shows the data from the main report nationally, where there are significant changes compared to the Spring wave in 2019, either up or down. The major changes are increase in satisfaction with toilet facilities at stations and power sockets on trains, both from low levels last year, and a decrease in satisfaction with punctuality.

Table 7 – National results – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?
Overall satisfaction with the journey	17098	82	-1	down
Toilet facilities at the station	10909	52	6	up
Attitudes and helpfulness of the staff	12742	79	2	up
Availability of staff at the station	14933	71	2	up
Overall satisfaction with the train	17713	78	1	up
Punctuality/reliability (i.e. the train arriving/departing on time)	17750	74	-3	down
Upkeep and repair of the train	17567	75	2	up
Comfort of the seats	17175	66	1	up
Level of crowding	17662	73	1	up
Availability of power sockets	9206	42	7	up

Merseyrail

Merseyrail is the TOC that appears to be most affected by the curtailed sampling plan, with much lower than average sample size, weighting efficiency, reduction in weighting efficiency and the need to merge cells for one of its two building blocks. If there is going to be an impact on NRPS results, we might expect to see it manifested for Merseyrail.

Table 8 below shows the metrics that show significant change from Spring 2019. We also show the change from Autumn 2019 and it is apparent that these changes are specific to the Spring 2020 wave, as the results are also different from the Autumn 2019 wave.

Table 8 – Merseyrail results – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?	previous	2019 change
Cleanliness	511	89	8	up	7	1
Facilities for bicycle parking	144	63	-16	down	-19	-3
Choice of shops/eating/drinking facilities available	382	40	-11	down	-8	-3

The columns in these and subsequent tables are defined as follows:

- Sample – is the sample size on which the results for this metric are based (this normally involves removing Don't Know and Not Answered responses)
- % sat – is the % giving a very/fairly satisfied/good/well response to the question about that attribute
- Change – is the % change since the Spring 2019 wave – the same wave one year ago
- Sig? – is up or down if the change is statistically significant at the 95% level using the effective sample size on this attribute for both waves
- Previous – is the % change since the Autumn 2019 wave
- 2019 change – is the % difference between the Autumn 2019 wave and the Spring 2019 wave (the change column is the sum of the previous and 2019 change columns)

To investigate this further, we have looked at other metrics. Where the change in Spring 2020 is high (either compared to the Spring or the Autumn wave last year), we might be concerned that the Spring 2020 result may be at least partly driven by the effect of the curtailed sample.

Table 9 below shows all the metrics that meet this criterion. It includes the three metrics with significant changes from the Spring 2019 wave plus five others. Among these five are two metrics with sample sizes of below 100 and the other two are between 300 and 400.

Table 9 – Merseyrail – selected metrics with spikes in Spring 2020

Attribute	sample	% sat	change	sig?	previous	2019 change
Cleanliness	511	89	8	Up	7	2
Connections with other forms of public transport	333	69	-6		-8	2
Facilities for car parking	198	47	-14		-17	2
Facilities for bicycle parking	144	63	-16	down	-19	3
Choice of shops/eating/drinking facilities available	382	40	-11	down	-8	-3
Availability of staff on the train	379	47	-5		-5	0
How well train company deals with delays	93	39	-10		-6	-4
Usefulness of information about the delay	85	67	11		8	3

There may be good reasons as to why these metrics have shown relatively large changes in the Spring 2020 wave but not in 2019, but our initial observation would be that the results for these metrics for Merseyrail should be treated with some caution.

East Midlands Railway

East Midlands Railway features third on the ranking table shown in Table 6. We therefore look at results for this TOC to see if there appear to be any odd results.

Table 10 – East Midlands Railway – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?	previous	2019 change
Punctuality/reliability (i.e. the train arriving/departing on time)	616	73	-9	Down	0	-9
Upkeep and repair of the train	610	64	-7	Down	-3	-4

Unlike Merseyrail, the two metrics with significant changes compared to Spring 2019 do not show such changes when compared to Autumn 2019. For the punctuality measure, the 9% fall this wave compared to Spring 2019 compares to no change when compared to Autumn 2019 so that all the change effectively occurred between Spring and Autumn 2019. For the upkeep and repair of the train, the 7% fall compared to Spring 2019 is made up of a 4% fall between Spring 2020 and Autumn 2019 and a 4% fall between Autumn 2019 and Spring 2019 – so a continuation of a trend.

Analysing East Midlands Railways in more detail, there are some metrics where all the difference compared to Spring 2019 appears to be due to the current wave, as changes compared to Autumn 2019 are similar.

Table 11 – East Midlands Railway – selected metrics with spikes in Spring 2020

Attribute	sample	% sat	change	sig?	previous	2019 change
Overall satisfaction with the station	628	86	3		3	0
Facilities for car parking	202	67	3		3	0
Facilities for bicycle parking	123	72	-4		-4	0
Availability of staff at the station	521	76	4		4	0
Helpfulness and attitude of staff on train	431	82	5		5	1
Toilet facilities	261	49	-4		-5	1
Cleanliness of the outside	543	62	-5		-4	-1

The changes here are more modest and as a result we would not have any major concerns about the impact of the curtailed sampling plan on East Midlands Railways.

Transport for Wales

Transport for Wales has been selected for detailed analysis as it is the TOC with the lowest weighting efficiency. Table 12 shows the metrics with significant changes from Spring 2019.

Table 12 – Transport for Wales – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?	previous	2019 change
Toilet facilities at the station	488	60	20	up	13	6
Attitudes and helpfulness of the staff	528	86	10	up	11	-1
Availability of staff at the station	613	70	9	up	13	-4
Choice of shops/eating/drinking facilities available	623	46	11	up	11	0
Upkeep and repair of the train	760	67	9	up	6	3
Cleanliness of the inside	767	75	9	up	6	3
Cleanliness of the outside	684	71	8	up	6	2
Availability of power sockets	505	46	25	up	16	9

For some of these factors, the trends are continuations of those seen in the previous wave – Toilet facilities up 20 - +7 in 2019 and +13 since Autumn 2019 and similar patterns for the last four metrics in the above table.

Two of the other metrics - Attitudes and helpfulness of the staff and Choice of shops/eating/drinking facilities available – all show the rise just occurring in Spring 2020 with no great change during 2019. The final metric - Availability of staff at the station – shows a 13% rise in Spring 2020 compared to Autumn 2019 after a 4% fall in Autumn 2019. So these three station attributes appear to show a large turnaround between Autumn 2019 and Spring 2020.

There may be a reason for this, for example if Transport for Wales has been investing in stations during the past 6-9 months; but if not, these changes should be treated with caution. As with the previous TOCs, we can identify metrics where all the change appeared to occur in Spring 2020, as Table 13 below shows.

Table 13 – Transport for Wales – selected metrics with spikes in Spring 2020

Attribute	sample	% sat	Change	sig?	previous	2019 change
Ticket buying facilities	470	82	8		6	2
Attitudes and helpfulness of the staff	528	86	10	up	11	-1
Availability of staff at the station	613	70	9	up	13	-4
Availability of seating	729	60	4		5	-1
Choice of shops/eating/drinking facilities available	623	46	11	up	11	0
Availability of Wi-Fi	398	52	11		17	-5
Usefulness of information about the delay	177	31	-14		-14	0
Level of crowding	749	76	5		7	-1

All these metrics show changes for Spring 2020 compared to Spring the previous year, but no real or even negative change during 2019, suggesting that Spring 2020 might be a spike in results due to the curtailed sampling plan.

Northern

We analyse Northern, as it is a TOC in the middle of the ranking table and we would not therefore expect the curtailed sampling plan to have as much, if any, impact compared to those analysed above.

Table 14 below shows the metrics that have significantly changed since Spring 2019.

Table 14 – Northern – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?	previous	2019 change
Overall satisfaction with the station	984	73	-6	down	-4	-2
Overall environment	980	66	-7	down	-9	2
Punctuality/reliability (i.e. the train arriving/departing on time)	969	63	-10	down	-3	-7
Connections with other train services	547	62	-9	down	-4	-5
Upkeep and repair of the train	965	66	8	up	9	-1
Comfort of the seats	945	64	7	up	5	2
How well train company deals with delays	299	30	-10	down	-4	-6
Availability of power sockets	479	36	21	up	15	6

Apart from the overall station environment and upkeep and repair of the train, all these significant metrics show continuation of a trend observed in 2019, either up or down. Just two of the eight measures appear to have a spike for Spring 2020.

Only one other factor (i.e. one without a significant change compared to Spring 2019) follows this pattern, as Table 15 shows.

Table 15 – Northern – selected metrics with spikes in Spring 2020

Attribute	sample	% sat	Change	sig?	previous	2019 change
Overall environment	980	66	-7	Down	-9	2
Overall satisfaction with the train	971	69	3		5	-2
Upkeep and repair of the train	965	66	8	Up	9	-1

From this, we conclude that the curtailment of the sampling plan has had little effect on the results for Northern.

Southern

Southern was selected as a TOC low in the ranking shown in Table 12 and where we would be surprised if the curtailed sampling plan had any major impact on key metrics.

Table 16 shows the results for metrics where there has been a significant change in Spring 2020 compared to Spring 2019.

Table 16 – Southern – significant changes from Spring 2019

Attribute	sample	% sat	change	sig?	previous	2019 change
Ticket buying facilities	441	80	6	Up	4	2
Toilet facilities at the station	482	52	10	Up	-2	12
Attitudes and helpfulness of the staff	587	78	5	Up	2	3
Facilities for bicycle parking	173	56	-10	Down	-4	-6
Availability of staff at the station	703	72	8	Up	7	1

Three of the five metrics here show continuation of a trend seen in 2019; the other measures appear to show spikes in the Spring 2020 data. There are no metrics with non-significant changes that show any spikes for Spring 2020.

Summarising the TOCs examined in detail

Looking at the number of metrics where spikes seem to have occurred in Spring 2020 which are not continuations of earlier trend suggests the following outcomes:

Table 17 – summary of TOCs analysed

TOC	# above 5%	# 5% or below	Weighted
Merseyrail	6	0	12
East Midlands Railway	0	7	7
Transport for Wales	5	2	12
Northern	2	1	5
Southern	2	0	4

In this table:

- # above 5% is the number of metrics with changes above 5% that do not appear to be part of a trend
- # 5% or below is the same but where the change shown is 5% of below
- Weighted – applies a weight of 2 to the # above 5% and 1 to those 5% or below

The weighted total tends to decline as you go down the table, which ties in with the top three being TOCs that ranked highly on the various factors, with Northern in the middle and Southern near the bottom. The analysis suggests that the ranking given in Table 6 to identify the TOCs most likely to generate concerns appears to be validated when a fuller analysis is undertaken.

Merseyrail clearly stands out as from other TOCs, with low rankings on a large number of criteria and six metrics that show spikes in Spring 2020 (which are not continuations of previous trends). There may be good reasons for the changed perceptions if the TOC has undertaken specific related actions but if this is not the case we would suggest considerable caution is used when interpreting the results.

A number of TOCs have low rankings on more than one criterion: Great Northern, East Midlands Railway, London North Eastern Railway, C2c, Chiltern Railways, Transport for Wales and TransPennine Express. With these TOCs, we would suggest greater than usual care is taken when interpreting the results of the Spring 2020 wave.

For the other TOCs, we see no great evidence that the curtailment of fieldwork has had any major impact on the results and we suggest that the metrics for these TOCs meet the normal NRPS standards, with the caveat that reduced sample sizes and weighting efficiencies will widen the margins of error. These widened margins of error will be in any event used to calculate whether there have been significant changes and we recommend that the NRPS User Guidance and Technical Reports are updated to reflect these wider margins for this wave.