National Centre for Social Research

TRAFNIDIAETH CYMRU
TRANSPORT FOR WALES

Wales National Travel Survey: Sampling approaches

Authors: Peter Cornick, Eva Aizpurua & Shane Howe
Prepared for: Transport for Wales

SEPTEMBER 2023

We believe that social research has the power to make life better.

By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

National Centre
for Social Research
35 Northampton Square
London EC1V 0AX
020 7250 1866

www.natcen.ac.uk

A Company Limited by Guarantee

Registered in England No.4392418

A Charity registered in England and Wales (1091768) and Scotland (SC038454)

This project was carried out in compliance with ISO20252

designbysoapbox.com

Contents

List of Tables

Background

The Wales National Travel Survey

Transport for Wales (TfW) has commissioned the National Centre for Social Research to work as its delivery partner to design a suitable methodology for the new Wales National Travel Survey (WNTS). Scheduled for a 2024 launch, the WNTS will collect data on travel attitudes and behaviour among the population in Wales.

The primary goals of the WNTS are threefold:

1) To monitor changes in travel behaviour and evaluate the degree to which specified targets are being achieved.
2) To provide data that will improve the transport evidence base and regional transport models in Wales.
3) To develop a robust, repeatable data collection strategy that provides more agency and control over the data generation process.

To fulfill these objectives, the WNTS will incorporate two components:

1) A survey that will gather data on household composition, socio-demographic characteristics, attitudes towards travel, and some travel behaviours.
2) A travel diary to collect detailed information on travel behaviour over a specific period.

The Current Report

This report is one of several outputs from the third stage of a process aimed at identifying and designing the optimal survey mode(s) for delivering the WNTS. This work is guided by the National Centre for Social Research's REMoDEL approach, a systematic process for designing or transforming social surveys and for gathering robust evidence around the trade-offs linked with different design options (Cornick, 2021).

Key Stages of NatCen's REMoDEL Approach:

Review the research and information needs, including any design parameters

Evaluate the feasibility of various methodological designs considering different quality dimensions

Model a prototype design for further development

Design and develop the new methodological approach and questionnaire

Experiment by testing the design in a quantitative context

Launch the new survey

The Evaluate stage of the WNTS was completed in July of 2023. This stage evaluated competing survey modes and diary collection tools against information and design parameters identified in the Review stage (Cornick, Aizpurua & Howe, 2023). The current document focuses on a review of the key overriding options for the sample design.

Sampling design decisions

Background and summary

During the Evaluate stage it was agreed that the WNTS will be designed to operate as a 'web-first' survey. This means that the initial contact with selected households will be to invite and encourage respondents to complete the survey online. For those who do not complete the survey online further modal options will be offered, including telephone and face-to-face.[1]

It has also been agreed that the study will consist of two key components: i) a questionnaire which provides demographic information and attitudes towards travel; and ii) a travel diary to collect detailed information on travel behaviour over a specified time period. These will need to be completed in full by respondents for the case to qualify as fully-productive.

This mix of modes, and the inclusion of both questionnaire and non-questionnaire research tools, provides some challenges to the design and operation of the survey which it will be important to consider when devising the optimal sampling approach. The survey will need to work in both self-administered and interviewer-administered environments. While in interviewer-administered surveys it is possible to retain some control over who takes part and when, which means that it is possible to implement complex rules around respondent selection and the data collection task, in self-administered surveys this control is significantly diminished. Indeed, in self-administered surveys complex rules can lead to greater bias if they are not followed or if they deter people from completing the task. As such, it is important that the survey is simple and easy to complete.

However, this requirement for simplicity needs to be balanced against the need to produce a high-quality, robust sample. A key design parameter around the WNTS is that it is possible to make statistical inferences to the general population, and that it meets the standards required to be badged as a National Statistic. As such, it is vital that it conforms to the principles of random probability sampling. (This rules out all non-probability sampling designs, such as quota sampling.)

[1] The telephone option will be experimented with during the 'Experiment' phase

It will be important, firstly, for the final sample to be representative of the population resident in households in Wales. Alongside this, a key design parameter is that it will be possible to conduct analysis at the region level in Wales. Therefore, the sampling design will need to allow for this regional-level analysis while optimising the representativeness of the sample at the national-level. Because the population of Wales is not evenly spread across regions, this implies that a disproportionate sampling approach is required. A key decision will be to what extent regions with lower populations (such as Mid-Wales) are oversampled. This decision will need to be led by the specific analytical requirements; that is, what level of precision is required for key variables of interest.

A further key driver of the sampling design is whether the unit of analysis, and therefore the sampling unit, is individuals, households or both. This impacts on the sampling design in a number of ways: it impacts on the required sample size, the clustering approach and, perhaps most notably, the approach to respondent selection. Before the sampling design can be finalised it is important to decide whether the sampling unit is individuals or households.

Finally, the travel diary component adds specific challenges to the sampling design. In particular, the number of days covered by the travel diary will directly influence the volume of data collected and therefore the statistical power of any analyses. However, it is widely accepted that it is not possible to recall detailed travel information over a period of longer than 48 hours, and in self-administered surveys it is not possible to (fully) control when people complete the diary. As such, if more data needs to be collected it will need to be done in a second stage of the survey. This would in itself create further challenges in terms of additional non-response to the second stage. Therefore, the question of how many diary days to cover needs to be carefully considered.

The complexity of the task for respondents and the robustness of the sampling design need to be further balanced by a third factor: cost. This adds practical limitations to the sampling design. The aim of this stage is to outline an approach that allocates the resources available to meet the survey objectives identified in the Review stage in such a way as to maximise the accuracy of the data collected using the mode(s) selected in the Evaluate stage.

In the rest of this report we outline the key decisions which will allow us to design the optimal approach. These decisions centre around disproportionate sampling; the sampling unit; and the volume of diary data collected. This is supported by statistical analysis and indicative costings. However, it needs to be emphasised that these statistical analyses and indicative costings are necessarily underpinned by assumptions. These assumptions are around expected response rates, design effects, the volume of data generated by the diary, and the operational tasks. While these are based on research literature, previous experience and expert judgement, they remain estimates. Each of these assumptions will be tested during the Experiment stage so that final decisions on the optimal survey design are based on empirical evidence.

Disproportionate sampling

A key design parameter for the WNTS is to facilitate region-level analysis. However, as demonstrated in Table 1, the population in Wales is unevenly distributed by region. South East Wales accounts for almost half of the adult population, while Mid Wales accounts for less than seven per cent.

Table 1. Wales regional population distribution

Region	Overall population	Percentage	Population 17 or over	Percentage
North Wales	687,098	22.1%	560,893	22.2%
Mid Wales	204,215	6.6%	171,393	6.8%
South West Wales	691,625	22.3%	565,149	22.4%
South East Wales	1,522,472	49.0%	1,225,869	48.6%

When sampling, the most efficient approach is to adopt simple random sampling. This would produce a proportionate sample. That is, the distribution of addresses selected would represent the actual distribution of addresses across Wales. In other words, Mid Wales would account for c.7% of the sample. While this approach is the purest form of sampling, in that all addresses have an equal, non-zero chance of selection, and therefore no sampling error, it is impractical for two reasons. Firstly, it would mean that a very large sample is required to generate a suitable sample size to allow analysis in Mid Wales. Secondly, it is inefficient. A key requirement of the WNTS is that it is cost-effective, and under a purely simple random sampling approach Transport for Wales would be paying for more interviews than is necessary to meet its information needs. It is highly likely that generating a sufficiently large sample in mid Wales this way would exceed the available budget.

Instead, the sample will be stratified by geography (local authority and region) and addresses will be selected from within each strata. In other words, different sampling fractions will be used in different regions to facilitate analysis in the most cost-effective way. This means that addresses will have an unequal probability of selection, but that probability will be known and non-zero. Under this approach, we will oversample addresses in Mid Wales to facilitate analysis in that region, while ensuring the overall sample size remains in budget.

Oversampling is a common technique used in survey research to facilitate analysis of particular subgroups which would otherwise not be represented in sufficient volume.

However, there are trade-offs associated with disproportionate sampling. It turns the sample from a simple one to a complex one. Any disproportionate sampling would need to be ‘corrected’ in the weighting stage. That is, cases which are overrepresented in the final sample would need to be weighted down to ensure that the final sample is representative of Wales at the overall level. This increases the design effect (DEFF) which quantifies the impact of complex sampling decisions on the precision of estimates when compared to a simple random sample (SRS) design of the same sample size. As the DEFF increases from 1 the precision decreases, which

means the effective sample (NEFF) size (the hypothetical simple randoms sample that would provide the same level of precision as the complex sample design) is reduced.

In other words, oversampling reduces the efficiency of the sample selected.

Table 2 demonstrates how this oversampling may work in practice. Based on a productive sample of 5,000 (selected only as an indicative number ahead of a final determination of the optimal sample size), and assuming an even response rate across Wales, if selected in proportion only 340 productive interviews would be completed in Wales. This has a design effect from sampling of 1. That is, there is no sampling error, so the total effective sample size is 5,000.

However, taking an ‘optimal’ approach to sample allocation by region would increase the scope of conducting regional level analysis with only a relatively small impact on the overall effective sample size. In the examples given, setting a minimum target of 500 productive cases in Mid Wales increases the DEFF due to sampling to 1.01 (overall effective sample size = c.4,940), while setting a minimum target of 700 productive cases in Mid Wales increases the DEFF to 1.04 (overall effective sample size = c.4,790).

While this is for illustrative purposes only, is demonstrates that oversampling is more cost effective than simply increasing the sample size. That is, to achieve a productive sample size of 700 in Mid Wales using a proportional sampling approach would require a total sample of more than 10,000.

Table 2. Wales possible regional targets

Regional target	Responses if no oversampling	Responses with a regional target of 500	Responses with a regional target of 700
North Wales	1,111	1,073	1,025
Mid Wales	340	500	700
South West Wales	1,120	1,081	1,033
South East Wales	2,429	2,346	2,241
Total achieved sample	5,000	5,000	5,000
Design effect from sampling	1.00	1.01	1.04

If we accept that oversampling is required, the next question is what degree of oversampling is optimal. This will depend on the analysis requirements and in particular what degree of precision is required and what sample size is feasible within the given budget.

This, in turn, will be driven by the decision around the most suitable sampling unit.

Sampling and analytical unit

Whether the survey is conducted at the individual- or household-level has been discussed before in earlier reports (Cornick, Aizpurua & Howe, 2023). This takes a central focus as it impacts on the required achieved sample size and, therefore, the optimal approach to oversampling.

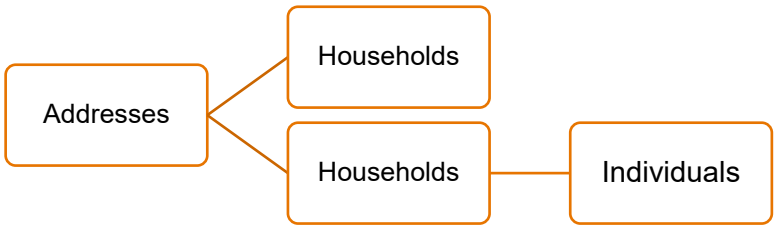
Household-level surveys collect information about entire households. They are typically used when collecting information on topics which impact on households, such as housing conditions (the English Housing Survey) or income distribution (the Family Resources Survey). Typically, in household surveys the entire household, or family group, is treated as the unit of analysis.

Individual-level surveys collect information from individuals, with each individual treated as a separate unit of analysis. They are typically used when collecting information which will vary by person, such as social attitudes (the British Social Attitudes survey), levels of activity (the Active Lives survey), or health behaviours (Health Survey for England).

Some surveys incorporate both the household and individual level, through having both elements. One example of this is the National Travel Survey, which has a household-level questionnaire and an individual-level questionnaire. This facilitates analysis at both the individual and household level, but requires complex administration.

Figure 1 illustrates the different levels of the sampling/analytical units within household-level and individual-level surveys.

Figure 1. Household- and individual-level surveys



Household-level

It has been established in earlier stages of this review that the ability to analyse travel behaviour at the household level is not a key requirement of the WNTS. As such, the main advantage of operating the WNTS as a household-level survey is efficiency. That is, household surveys are multiplicative: by collecting information from more than one person per household, a single household can potentially produce three or four times the amount of data than in an individual level survey. When dealing with items of interest that have a low prevalence, such as long-distance journeys, increasing the number of observations is very helpful. Additionally, household-level surveys do not require within-household selection, so can be more efficient in weighting terms.

However, there are a number of limitations to operating the WNTS as a household-level survey. As discussed in the Questionnaire and Travel diary reports (Cornick, Aizpurua & Howe, 2023), and in meetings with TfW, they are less suited to self-administration. Without the input of an interviewer controlling who responds to the survey,

and ‘selling’ the survey to all household members, it would be necessary to accept ‘proxy’ responses (where one household member completes the survey on behalf of another household member). This is naturally inappropriate for attitudinal measures, and may lead to lower data quality on behavioural measures – for instance, one household member may not know in detail the travel behaviour of another. This also significantly increases the burden (both time and cognitive) on the responding household member and we would expect respondents to engage with satisficing or time saving techniques to either avoid questions or omit activities altogether to simply proceed.

Individual-level

In individual-level surveys one eligible respondent is typically selected from all eligible members of a household to complete the survey. This has the advantage of simplicity in sampling and analytical terms, and in terms of the quality of the data collected. That is, by only collecting information about the respondent from the respondent, you reduce the risk of error in the data collected.

The disadvantages of running the WNTS an individual-level survey centre around the challenges of conducting within-household selection in a self-administered survey, and on the increased costs associated with needing to sample more addresses for the same number of interviews as a household-level survey.

Within-household selection

For the survey to conform to the principles of random probability sampling (and therefore it being possible to make statistical inferences to the general population), which is a key requirement of the WNTS, this respondent needs to be selected randomly.

Indeed, a recent meta-analysis concluded that when no within-household selection is carried out the mean absolute bias in sample estimates of demographic representation is large, estimating the sample estimates deviate from population benchmarks by 5 to 10 percentage points (Yan, 2009). Therefore, a within-household selection is necessary.

In true probability surveys within-household selection requires detailed information about household members, including the number of people living in the household and often their sex and age. This information is collected by interviewers who then follow a systematic set of procedures to select a respondent to take part in the survey. This approach is possible and well-established in face-to-face surveys, and we recommend utilising this random selection approach in the face-to-face element of the WNTS if it uses an individual-level approach.

However, within-household selection is more problematic in web surveys. It is generally accepted that true probability methods are too complex for households to implement themselves in self-completion surveys (Battaglia et al. 2008; Reich et al. 1986). Unlike face-to-face surveys, web-first surveys cannot rely on trained interviewers to randomly select an adult in the household to take part in the online element. Instead, the selection needs to be conducted by the respondents themselves.

This can lead to errors in the selection process or nonresponse, resulting in too many or too few of certain types of people in the data (e.g. typically too many female, highly educated, older, and white respondents), and may also lead to biased estimates for other items (Smyth et al, 2019).

Multiple different strategies to conduct within-household selection have been used within self-administered surveys.

Quasi-probability selection

One approach is to use a quasi-probability selection approach. Quasi-probability methods bypass household listing in order to reduce the perceived intrusiveness and sensitivity associated with household listing and to decrease the administration time (Yan, 2009).

The most common of these are the next/last birthday approaches. Under this approach, the invitation letter includes instructions on who within the household should complete the survey – this may be the person with the next/last birthday. However, there is ample evidence from self-administered surveys that these instructions are not followed. Among households where a selection must take place (i.e., two or more eligible adults in the household), as many as one in three participants is not the correct person (Williams, 2014; Park & Humphrey 2014).

This will introduce self-selection bias in the survey estimates with those most interested in the survey topic more likely to take part. It tends to lead to a very skewed sample and a high design effect. This, in turn, means that the survey is not cost-effective: compared to alternative designs, it is necessary to achieve a higher productive sample size to achieve the same effective sample size.

We therefore believe that quasi-probability selection methods are not appropriate for the WNTS.

All adult approach

An alternative approach is to adopt an ‘all adults’ approach. Under this approach, more than one adult is able to take part per household. This avoids the problem of non-random selection within the household, but introduces risks with regard to the accuracy of the household enumeration and the potential for individuals to complete the survey multiple times, motivated by the promise of £10 for each completion.

Most web-first studies in the UK use a variant of the ‘all adult’ method, with different surveys allowing two, three or four adults per household to take part. This is done to minimise the risk of duplicate or falsified responses. It is simple, low burden for respondents and has proven to be effective across a number of different surveys, such as the Community Life Survey, the Active Lives Survey and the Participation Survey.

As such, we currently recommend this approach as the basis for selection within the push-to-web element of the WNTS. As discussed above, where selection can be administered, such as in the F2F element, it should be.

However, while this all-adult approach has proved to be successful in push-to-web surveys, it is potentially problematic for the WNTS in a few ways: i) it can lead to sample bias if household members who travel more/less than others are more likely to take part; ii) travel behaviour may be clustered within households, so collecting travel behaviour from multiple people in the same household may reduce the representativeness of the data and bias results; and iii) controlling for intra-household clustering can complicate analysis.

Two-stage selection

Given the inappropriateness of quasi-probability selection approaches, and the compromises associated with the all adult approach, it is sensible to explore and experiment with options to randomly select one individual in the web survey.

Two recent selection methods adopt a two-stage approach and are conditional on household size. In part one, any individual within the household can complete the household-element of the questionnaire. In part two, a random selection of the individual to complete the individual-level questions and travel diary element would be made, and that individual would be invited to complete the survey.

These are the Rizzo-Brick-Park method and the Le-Brick-Diop-Alemadi method.

The Rizzo-Brick-Park method (Rizzo, Brick, and Park, 2004)

- Any household member completes the household-element of the questionnaire and is asked the total number of eligible people living in the household.
- If there is only one adult in the household, that adult is chosen for the individual-level interview.
- If there are two or more adults in a household, one adult is randomly selected with a probability equal to the inverse of the total number of adults. If the selected respondent is the same one who has completed the household-element of the questionnaire, the selection process ends and they continue to complete the survey. If the person who completed the household-element is not selected and there are two adults in the household, the informant is told that the other adult is selected for the interview and the selection process ends.
- If the household informant is not selected and there are more than two adults in the household, another selection method (such as the Kish method or the last birthday method) can be used to select an adult after excluding the household contact.

This method takes the advantage of the fact that more than eight in ten households in Wales have two or one adults. As a result, this method has the potential to significantly reduce the effort required to make a selection.

The Le-Brick-Diop-Alemadi method (Le, Brick, Diop, and Alemadi, 2013)

- Any household member completes the household-element of the questionnaire and is asked the total number of eligible people living in the household.
- For households with one or two adults, no additional question is asked and the household contact is automatically selected in one-person households and randomly selected half of the time for two-person households.
- For households with three or four adults, again no additional question is needed and the household contact is selected either 33% or 25% of the time. When the household contact is not selected, the older or the younger of the two adults in a three-person household is selected and the oldest, the youngest, or the second oldest of the other three adults in a four-person household is randomly selected with an equal chance.
- For households with five or more adults, one more question is asked about the number of males in the household and selection is made based on the answer.

The Le-Brick-Diop-Alemadi method was specifically designed for use in countries where households are larger. Of the two, we therefore recommend experimenting with the Rizzo-Brick-Park method.

However, there is little, if any, empirical evidence of the efficacy of these two-stage selection methods, because they have rarely been used. As such, full-scale testing would be required to understand if it offers benefits over the two adult approach.

Indicative efficiency and costs of alternative sample unit approaches

As discussed above, determining the unit of analysis, and therefore the sampling unit, impacts on what selection approaches are viable, what the likely design effect will be and, therefore, the required sample size. All of this will feed into the survey costs.

Table 3 provides illustrative costs for four competing selection approaches. These are household level selection (no within-household selection but requires all household members to take part); selection of one adult using a quasi-probability method such as the next/last birthday approach; the 'all adult' approach capped at two adults; and a two-stage selection approach.

For each of these we have made assumptions around the DEFF associated with the design, the response rates to the push-to-web and face-to-face stages, the strike rate (the number of interviews that a face-to-face interviewer will complete per day), and the level of coding and editing required (minimal – assumes a pre-coded diary). All of these assumptions are based on research literature, previous experience and expert judgement. However, they remain estimates and the figures in this table should be treated as illustrative only – they are not formal costings. We will produce formal costings once the initial design has been fully modelled.

Each of these assumptions will then be tested during the Experiment stage so that final decisions on the optimal survey design are based on empirical evidence.

The first step to producing a formal costing is to determine the required sample size. This is based on the level of precision required. For the purpose of this table, we have set the requirement for an estimate of 50% in Mid-Wales to be accurate to +/- 4 percentage points, once the complex sample design has been taken into account. This would require an effective sample size of 600 in Mid Wales. This was to i) ensure that analysis can be conducted in the smallest region with confidence; and ii) to set a baseline for comparing the different design options. The sample sizes and costs presented below are based on each design achieving the same final effective sample size.

This demonstrates that the inefficiencies associated with the quasi-random design mean that it needs to produce the largest productive sample size to generate an effective sample size of 600 in Mid Wales. This, in turn, means that the cost is the highest. Given the high cost and high level of sample bias expected, this option can be ruled-out.

The lowest cost option is the household-level design. This is because of the low expected DEFF and the benefits of generating multiple interviews in one household. However, to achieve this we would need to accept a high level of proxies, which would lead to a high level of bias in the data which it is not possible to control or correct with weighting. Given the concerns around how this would operate in the context of a travel behaviour survey, we would caution against this approach, unless there is a requirement for household-level analysis.

Of the two remaining options, the 'all adult' approach, limited at two adults per household, has a marginally lower cost. As such, this would be our recommended approach (supported by genuine probability sampling in the F2F element).

However, as noted above, the two-stage selection approach offers a number of potential benefits over the all-adult approach. There is little empirical evidence on the efficacy of this approach, so our assumptions in this

area are less robust than for the other options. As such, we recommend experimenting with this approach alongside the all adult approach with the aim of determining whether it offers benefits in efficiency and sample quality over the 2 adult approach.

Table 3. Indicative costs of alternative designs

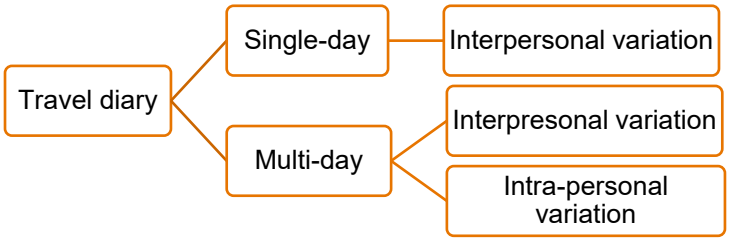
	Household level	Individual level		
Assumptions	**All adults**	**1 adult (Quasi-random selection)**	**1 adult (Two-stage selection)**	**Up to 2 adults**
Assumed DEFF	1.2	1.6	1.25	1.3
Mid-Wales required base size (to achieve a NEFF of 600)	720	960	750	780
Estimated min total productive sample size required	5,500	7,400	5,800	6,000
Cost Per Interview	£165 - £195	£185 - £215	£190 - £220	£170 - £200
Illustrative cost	£990,000	£1,480,000	£1,189,000	£1,110,000
Level of bias	High (proxies)	High (selection effects)	Low	Low

Number of diary days collected

The number of diary days collected is another crucial aspect we need to revisit as this directly impacts on the analytical possibilities with the travel behaviour element of the survey.

A multi-day diary allows for the collection of interpersonal (differences between individuals) and intra-personal (differences within individuals) variation. Intra-personal differences are likely to have become greater since the pandemic due to hybrid working practices and only possible with repeated observations.

Figure 2. Single- and multi-day travel diaries



As discussed in greater detail in the Diary Evaluate report (Cornick, Aizpurua & Howe, 2023) single-day travel surveys are commonly used in national travel surveys (e.g., Denmark, Finland). The choice of a one-day survey is influenced by logistical factors, including the challenges persuading all people to take part. As noted, these are ill-suited to providing information on travel variation as they only collect a single day of travel behaviour. In contrast, multi-day diaries provide more data, and offer more opportunities to detect infrequent journeys or activities alongside different travel behaviours within respondents.

It has been previously decided the diary will be incorporated within the survey software, and we have discussed how two days could be collected from recall. As such, we have assumed there will be no drop-off encountered for the 48-hour recall period. However, it is likely that recall over a period of longer than 48 hours would lead to significant error. Therefore, if it is required to gather travel behaviour information on a period of more than 48 hours, it would be necessary to collect this information in a separate second stage of data collection. This would inevitably lead to drop-off between the stages.

Table 4 provides estimates of the expected volume of data that would be collected via one, two, three and four day diaries. For this, we have assumed a similar level of journey recording per day as achieved in the NTS, and have assumed a 40% drop-off rate between days two and three and days three and four. This is based on an illustrative example of 5,000 achieved interviews. The table includes expected unweighted amounts of data, along with the effective sample (NEFF) size (the hypothetical simple randoms sample that would provide the same level of precision as the complex sample design). There are incremental gains from additional days, with the fourth day providing the most sizeable increases. The DEFF from weighting two days of diary data, rather than four is likely to be higher due to the uneven spread of days per week, although recent analysis suggest we may be able to mitigate this with varying mail out dates.

Table 4. Amount of diary days collected.

Item	1 day diary	2 day diary	3 day diary	4 day diary
Total days of diary data recorded	5,000	10,000	13,000	16,000
Approx days of diary data per day of week	714	1,429	1,857	2,286
Approx days of diary data per day of week NEFF	621	1,242	1,615	1,988
Approx total journeys recorded	9,450	18,650	24,230	29,630
Approx total journeys NEFF	8,217	16,217	21,070	25,765
Approx average journeys per day of week NEFF	1,174	2,317	3,010	3,681

From a purely statistical perspective the greater the number of observations the more statistical power is gained. However, as before there are practical and respondent-led limitations to moving to a three- or four-day diary. Firstly, a second stage would add significant cost and time to the data collection process. Secondly, the drop-off between the first and second stages is currently unknown, and may require additional incentivisation. Thirdly, there is a high risk of satisficing as respondents seek the easiest way to complete the survey.

It remains unclear at this stage whether progressing to a three or four day diary would be cost effective. We would recommend focusing our attention on maximising data quality and user experience of a 48-hour recall diary, whilst simultaneously experimenting with how best to administer the 3rd and 4th day of diary collection and to understand whether data quality will be an issue as expected.

References

Cornick, P. (2021). *The NatCen Remodel Approach. A systematic approach to transforming social surveys*. Briefing Paper. Centre for Social Survey transformation.

Cornick, P., Aizpurua, E., & Howe, S. (2023). *Wales National Travel Survey: Evaluate diary report*. Report prepared by NatCen Social Research for Transport of Wales.

Cornick, P., Aizpurua, E., & Howe, S. (2023). *Wales National Travel Survey: Review report*. Report prepared by NatCen Social Research for Transport of Wales.

Cornick, P., Aizpurua, E., & Howe, S. (2023). *Wales National Travel Survey: Questionnaire Assessment*. Report prepared by NatCen Social Research for Transport of Wales.

Le, K. T., Brick, J. M., Diop, A., and Alemadi, D. 2013. *"Within-household Sampling Conditioning on Household Size."* International Journal of Public Opinion Research, 25, 108-118.

Park, A. & Humphrey, A. (2014). Mixed-mode surveys of the general population: results from the European Social Survey mixed-mode experiment. NatCen Social Research

Rizzo, L., J. Michael Brick, and Inho Park. 2004. *"A Minimally Intrusive Method for Sampling Persons in Random Digit Dial Surveys."* Public Opinion Quarterly, 68: 267-274

Smyth, Jolene; Olson, Kristen M.; and Stange, Mathew, *"Within-Household Selection Methods: A Critical Review and Experimental Examination"* (2019). Sociology Department, Faculty Publications. 753.

Williams, J. (2014). Community Life Survey: Investigating the feasibility of sampling all adults in the household. London: TNS BMRB.

Yan, Ting. (2009). A Meta-analysis of Within-Household Respondent Selection Methods. AAPOR. 17.


National Centre
for Social Research