ONS

Quality Review of the National Travel Survey

Some Aspects of Design and Estimation Methods

1. Areas covered by the review

The areas covered by this review are grouped broadly under two headings: those associated with the sample size and sample design of the survey and those associated with maximising response rates and minimising nonresponse bias.

Under sample size and design we have considered in particular:

a) The current sample size and what could be achieved by a 3-fold increase

b) Optimising the design for measuring level and change

c) The impact of the current boost to the sample in London

d) The possibility of boosting the sample size differentially in some areas

e) The choice of regional and other strata

f) The choice and number of primary sampling units

g) The current placing pattern and fieldwork periods

Under response and nonresponse we discuss:

h) Respondent burden and diary length

i) Extent of nonresponse bias

j) Weighting for nonresponse

k) Recovering data from partially responding households

l) Imputation

A summary of recommendations is included at the end.

2. Sample size and design

2.1 Sample size

The National Travel Survey (NTS) has been running at more or less its current size of c. 3300 responding households per year since it became continuous in 1988. The annual data from the survey is little used – for most purposes data from 3 consecutive years are aggregated for analysis. The latest available 3-year dataset relates to 1995-1997. In addition some results are

produced for other 3-year periods - the latest NTS Bulletin covering the years 1997-1999 was published in August 2000.

The current sample size is felt within the Department to be inadequate for some uses, in particular for identifying both current levels and trends in small populations, for example in bicycle use or for regional or sub-regional analysis. The annual sample size is also felt to be inefficiently small by the current fieldwork agency (Social Survey Division of ONS) so one might expect some reduction in unit costs from an increased sample size. A 3-fold increase in the sample size has been mooted in the Department, subject to financial approval.

With a larger sample size some other aspects of the survey design should be reconsidered but here I consider the impact of precision from the survey if all other aspects of the design were to remain unchanged. Results are produced from the survey for a number of different analysis units at household level and below. In particular results are currently produced for households, cars and vans, individual adults and children, journeys and stages.

Assuming the current practice of aggregating 3 years data continues, sampling errors of estimates based on all these different units would reduce by 42% (with a doubling of the sample size leading to a reduction of 29%). Alternatively more up-to-date current estimates could be produced at the same precision as now but from the latest available year.

In considering the effect of this increase in sample size on the ability of the sample to measure change, it is important to consider the estimation methods used. The use of the 3-year fixed aggregates is very limiting and may in some circumstances be less effective than the use of annual estimates to identify change, even with the current sample size. Certainly a move to a more routine use of the rolling 3-year aggregates (e.g. 95-97, 96-98, 97-99 etc.) is recommended even if it is decided not to increase the annual sample size. (This type of rolling estimate is used very successfully by ONS in its regular series of employment and unemployment estimates derived from the Labour Force Survey data.)

To illustrate the advantages of a more flexible use of the NTS data, consider measuring change in the proportion of journeys made by walking. Suppose this variable started to increase (or decrease) by a sustained ½% per year, after a period of stability. Then if this change first occurred in the first year of the NTS' regular 3-year publication cycle, it would take 6 years of sustained change before the survey could safely conclude that a change had occurred with the current annual sample size. With a move to publication of rolling 3-year estimates, such a change could be identified (with 95% confidence) in 4 years whereas if annual estimates were compared, the detection delay would be 5 years. (This analysis uses the sampling error estimates from the 1997 Technical Report[1]).

If the sample size were trebled, then the same sized change would be detected in 3 years, based either on annual data or with the rolling 3-year estimates. Table 1 shows similar calculations for other selected NTS estimates. While the gains from a larger sample for detecting changes would clearly be worth having, the speed at which changes of this order could be safely detected would be improved further by a change to the sample design, discussed in the next section.

Table 1. Years needed to detect a change with different sample sizes and methods

NTS Estimate	Size of Change	Number of Years to Safely Detect change			
		N=3300 Rolling 3 Years	N=3300 Annual	N=10000 Rolling 3 Years	N=10000 Annual
% Households with no car	+/- ½ %	5	6	3	4
% Adults with driving license.	+/- ½ %	3	3	2	2
Miles travelled per person per year	+/- 100	4	5	3	3
Journeys per person per year	+/- 10	3	4	2	3
% of journeys on foot	+/- ½ %	4	5	3	3

2.2 Optimum sample designs for level and change

If the NTS annual sample size were increased substantially, the case for always using data aggregated from 3 years for level estimates would be much reduced. In this situation, it would be appropriate to consider designs that increase the precision of change estimates by overlapping the samples from successive years, i.e. retaining the same "units" in the sample for more than one year. Attempting to re-interview the same households one year later would presumably damage response rates too severely, although perhaps this assumption should itself be tested. However including the same areas in the sample for more than one year serves the same purpose of reducing the variance of estimates of year-on-year change, with larger gains for estimates with larger design effects. If the overlap is limited to the area (primary sampling unit) level, the main disadvantage of these designs are that the variances of estimates made by aggregating data from more than one year are increased.

A number of surveys use a design where half the psus selected in any year are retained and the other half are replaced. Compared with a true panel, this has the advantage that changes to the relative population sizes in the various strata can be accommodated so that no biases accumulate over time. One disadvantage is that if the number of psus used by the design changes over time or if the strata used in the design change, this "rotation" pattern must also change so the overlapping nature of the sample must be suspended for the duration of the change.

It is known that the relative reduction in the variance of a year-on-year change with this rotation design is exactly matched by an increase in the variance of a 2-year aggregate of the same variable[2]. The relative increase in the variance of a 3-year aggregate is somewhat higher, so the viability of this design depends on the relative importance attached to estimates of level (from multi-year aggregates) and change.

As an illustration, consider the effect of retaining half the psus selected in the previous year on two of the variables discussed above. The standard error of the annual change of the proportion of households with no car would be reduced by c.6% if no other aspect of the sample design changed. For the proportion of journeys made on foot, the calculations are more complex as the larger part of the variance of this estimate is due to the "clustering" of journeys within individuals and households, rather than the clustering of households within areas and this will be unaffected by overlapping the areas. The calculations in this case are beyond the scope of what can be attempted here but I would guess that the gains would be comparable.

Much larger gains would be possible for geographically-based variables, such as the proportion of households living more than 27 minutes from a railway station (standard error reduced by 23%) or the proportion of households with a bus service every 15 minutes (20%).

2.3 The London sample size boost

The survey has always achieved a lower response rate in London and particularly in Inner London. Rather than correcting this shortfall by weighting, as is more frequently done, the NTS some time ago opted for a boost to the set sample in London with the aim of achieving responding sample sizes in London and elsewhere that were proportional to their populations.

In 1999, the household response rates (full co-operation) for Inner London, Outer London and the remainder of GB[3] were 49%, 61% and 67%, and for the same year the corresponding sampling rates were in proportion to 1/50%, 1/61% and 1/67%[4]. Thus the responding household sample in that year still very slightly under-represented Inner London, though the representation of outer London and the remainder of GB were very close. For 2000 the relative sampling fractions in both Inner and Outer London were increased and are now in proportion to 1/43%, 1/57% and 1/67%. If the relative household response rates in these three regions are the same as in 1999, it is clear that the responding NTS household sample will over-represent both Inner and Outer London. It is not at all clear why these larger boosts were introduced.

As a device for handling below average response, oversampling runs up against the obvious problem that the sampling fractions have to be fixed in advance of knowing the true response rates and consequently the allocation of the responding sample to the different regions can at best be approximate. Post-stratification is the commonly used alternative approach to correcting for varying response rates, although a problem with household surveys is that the available population data on household populations are partly survey-based, so that individual data may be preferred. Clearly neither method truly addresses the potential nonresponse bias issue and this point is discussed in more detail in Section 3. However by boosting the London sampling fraction in this way it is clear that the achieved sample size in London will be boosted so that London-based estimates will have slightly higher precision at the expense of slightly lower precision for other regions. For national estimates, it is less clear whether, for a fixed sample size, this method or post-stratification would provide greater precision and the variance estimation software needed to assess this is only just becoming available.

In this case, the under-representation of London is so large and has persisted for so long that some boost to the set sample in London, perhaps based on relative response rates over a span of years, seems appropriate. However this should be fine-tuned by applying appropriate regional post-stratification weights.

2.4 Differential sampling by region or other areas.

Some users of the NTS data have expressed a wish for boosted samples in specific areas and from time to time it may be desirable to increase the sample size in some regions or other parts of the country to allow more detailed analysis. This should create no particular statistical problem so long as the data are weighted back to the correct regional proportions, so as to avoid bias in the combined samples.

Differential sampling may affect the precision of the weighted estimates either way. It is well known that the allocation to strata that maximises precision on an individual variable is that the sample size be proportional to the stratum standard deviation. In a multi-purpose survey such as the NTS, this approach is not realistic but it is likely that if some strata or regions are

over-sampled relative to others, that for some variables there will be a gain in precision, while for others there will be a loss. If the decision on which regions to over-sample is essentially arbitrary, then on average one would expect more variables to show a loss in precision, relative to an equal probability design. Kish[5] has derived a method of approximating the average loss of efficiency[6] due to weighting a survey sample. The increase in variance is an increasing function of the coefficient of variation of the weights needed to remove the bias in the sample. So, as the over-sampling of some groups becomes more extreme or as the larger weights are applied to a smaller proportion of cases, the weighting efficiency falls. Elliot[7] shows a graph of the weighting efficiency in the case of over-sampling a single population group. For example if the group to be boosted is over-sampled by no more than twice, the efficiency is no lower than 89%, whereas if it is over-sampled by a factor of ten, then efficiency could be as low as 33%. In an extreme case, applying a large boost to a small population group could result in a weighted sample that is less efficient than omitting the boost entirely when forming national estimates, though that would be very unusual.

As an example consider 4 different sample distributions of the 9,094 households in the 1997-1999 NTS samples across Government Office Regions:

a) a proportional allocation

b) a minimum sample of 700 households in each region

c) an equal sample of 827 households in each region

d) a boost to 2,500 households in London

Table 2. Sampling Efficiency of Four Regional Allocations

GOR	Population %	Proportional	Minimum of 700	Equal Shares	London boost
North East	4.6%	417	700	827	344
North West	12.1%	1096	1016	827	905
Yorks & Humberside	8.8%	797	739	827	658
East Mids	7.3%	663	700	827	548
West Mids	9.1%	830	769	827	685
Eastern	9.4%	853	791	827	704
London	12.2%	1111	1030	827	2500
South East	13.8%	1251	1160	827	1033
South West	8.7%	787	730	827	650
Wales	5.2%	470	700	827	389
Scotland	9.0%	819	759	827	677
Efficiency		100%	97%	92%	90%

2.5 Choice of Strata

The current strata were introduced in 1998 following research[8] by ONS into optimum strata based on 1991 Census data. The sampling frame is the small-users sub-file of the Postcode Address File (PAF) and Postcode Sectors are used as primary sampling units. These units are formed into strata as follows:

i) Government Office Regions (GORs) divided into the old Metropolitan and non-Metropolitan Counties, Inner and Outer London, and the Central Clydeside Conurbation and the remainder of mainland Scotland (18 groups). Scottish Islands are excluded from the sampling frame, as is common on many GB surveys.

ii) Proportions of households in the area with no car (3 groups)

iii) Proportions of heads of household classified as socio-economic group (SEG) 1-5 or 13 (ranking of areas).

Before 1998 the survey used a different regional breakdown into 31 groups that incorporated a population density breakdown; then used car ownership in two groups and finally a ranking on the proportion of employed persons who travel to work by car.

Barton used NTS data from 1992-1994 for 13 variables covering a range of analysis units. Although the proposed and now implemented set of strata provided some improvement for some of these variables, others showed some deterioration, with this being particularly marked for two of the survey variables: mean commuting journey distance and the proportion of journeys that were walking. The first appears to have suffered from the dropping of the population density breakdown in Metropolitan regions and the second from the replacement of the "Car to work" variable with the SEG variable. The analysis was done using a breakdown of Standard Region but the survey has since converted to GOR. Barton did not recommend any change in the NTS stratifiers as a result of his analysis.

A much larger annual sample should make it possible to re-instate both population density and "Car to work" strata into the design which should improve precision on these two and any correlated variables. Whatever decision is taken on sample size, I would recommend that a similar analysis be undertaken once small area data from the 2001 Census is made available.

Alternatively, some users, particularly in local government, have expressed reservations about the volatility of the small area data derived from the survey. With relatively few areas selected independently in each region in each year, such volatility in even the 3-year aggregates is inevitable. A larger sample size would make it possible to switch over completely to either county or some other sub-regional stratification. For example each Local Education Authority in England and each Unitary Authority in Wales and Scotland could form a separate stratum, as in the Local Labour Force Survey. This would almost certainly result in some loss of precision in the national and regional estimates but it would address this specific concern. The size of both the gains at a local level and the losses at national and regional level could be estimated from existing data or this alternative stratification could be evaluated alongside the extension of stratification suggested in the last paragraph.

2.6 Primary Sampling and other Sampling Units

Until 1999, the design sampled 240 PAF postcode sectors as primary sampling units, then 21 delivery points (dps) within each sector. This generated a set sample size of 5040 dps. Some dps contain more than one household, and standard Social Survey Division procedures are used to sub-sample these households. This generates a bias against such households, which is not currently corrected for by any weighting but probably should be as such addresses are concentrated in inner city areas, particularly inner London[9]. Also some delivery points do not contain a private household and as these and multi-household addresses are not identified on the frame, the eligible sample size is subject to random variation.

To compensate for falling response rates, the 2000 sample design selects 252 primary sampling units and then selects 23 dps per sector – this generates a set sample of 5796 dps.

For a two-stage sample such as is used on the NTS (ignoring the effect of sampling households within dps), a simple (approximate) formula relates the design effect and hence the sampling error of a survey estimate to the intra-psu correlation, rho, for that variable and psu. Rho values for each area type are population constants and thus, once these have been estimated, one can predict the impact of varying the number of psus used in the design (and the sample of dps taken from each) on the sampling errors of survey estimates. This method is straightforward for household-level variables but for other analysis units such as cars, individuals and journeys, the models are no longer simple and components of the variances associated with different units must be estimated before such predictions can be made. Bob Butcher undertook such an analysis on the NTS in the 1980s and more recently similar estimates were produced for a small number of NTS estimates by Social Survey Division when the design of a possible Integrated Household Survey was being investigated. Using the results from this later work and the sampling errors reported in the 1997 Technical Report, it is possible to estimate broadly what the impact on the sampling errors of varying the number of psus, while holding the total sample size constant.

Table 3, below provides estimates of the relative reductions in standard errors that would result from a 50% increase in the number of sector-psus per year, if the sample size were held constant. For selected household variables, estimates are reported for each estimate but for other units, individual predictions will be unreliable without extracting the relevant variance components for each variable, so averages are reported.

Table 3. Effect of a 50% increase in psus

NTS Estimate	Relative Reduction in se
Household proportions	
No car	5%
1 car	2%
2 cars	4%
3+ cars	0%
Owns bicycle	5%
Under 6 minutes to bus	7%
27 minutes or more to railway	17%
Bus every 15 minutes	14%
Single person household	3%
Pensioner household	2%
Full-time workers	2%
Household means	
Vehicles	5%
Cars	5%
Bikes	7%
Persons	5%
Average (15 household estimates)	*6%*
Average (10 car estimates)	*1%*
Average (45 person estimates)	*2%*
Average (17 journey estimates)	*3%*

These gains seem modest for all but two estimates so, as fieldwork costs would presumably increase with such a design, it does not seem worth pursuing. By the same token however, a similar 1/3 increase in the number of dps selected per psu (hence a reduction from 252 to 189 psus) would produce the same modest increases in standard errors for most variables so

would be worth considering if the finances for the survey came under pressure. Of course the relationship between sample size, number of psus and field cost is not simple so the cost and effect on precision would have to be considered jointly.

Social Survey Division of ONS has done some work on the relationship between rho and area size[10] for selected Census variables. This suggests that for areas significantly smaller than sectors (1/3 the size or less), rhos are substantially larger. Thus cost savings would need to be substantial for designs using smaller areas as psus in order for such designs to be more efficient than the current one.

2.7 Field period and placing pattern

The NTS uses a complex placing pattern in an attempt to ensure that all days in the year are equally represented in diaries completed by the sample. The field periods, unusually, run mid-month to mid-month in an attempt to ensure that any unusual travel patterns associated with bank holidays are not under-represented. The placing pattern rules were relaxed somewhat in September 1999 in an attempt to slow the decline in response rates. Prior to this date, all addresses were pre-assigned to a diary start day and few exceptions were permitted even if that led to a total nonresponse by the address. The relaxed rules mean that there are controls in place for the whole of each interviewer's workload of addresses but the interviewer has some leeway in terms of assigning addresses to start-days. Other diary surveys conducted by Social Survey Division of ONS (the Family Expenditure and National Food Surveys) use a still looser placing pattern. The unusual field periods are known to cause some problems to the Department as data are aggregated to strict calendar periods.

The stricter placing pattern in operation prior to September 1999 would undoubtedly be superior in the presence of full response (or random nonresponse) but practical issues should always be considered in survey research and it seems superficially plausible that pre-assigning the start day of the diary to each sample address may have increased both non-contacts and refusals. However, according to the NTS research unit in Social Survey Division, there is no evidence that either type of nonresponse has declined since the rules were relaxed.

The rationale for the unusual field period is the observed variation in numbers of journeys recorded by week within the field period, which falls off a little in "week" 3 and substantially in "week" 4 of the fieldwork "month". It was though that, as bank holidays tend to occur near the start or end of calendar months, the decline in journey recording at then end of the field period could cause a bias.

The evidence in support of mid-month to mid-month field periods is limited. Certainly, in 1999, most bank holidays fell in "weeks" 2 or 3 of the field period but there was no evidence that reported total travel (numbers of journeys) was unusually high or low in weeks containing these bank holidays. Indeed, in view of the observed variation in weekly totals it is unlikely that any bias of this type would be detectable in a sample of this size (although this could be investigated more thoroughly by aggregating data from more than one year). Nevertheless, unless the unusual field period causes serious disruption or extra cost to the Department, it seems reasonable to maintain it as a way of avoiding this hypothetical bias.

3. Response and Nonresponse

The risk of nonresponse bias in the face of the sustained decline in NTS response rates since the early 1990s is perhaps the most serious issue facing the survey. Such declines have been observed over a wide range of surveys both in this country and abroad so seem likely to

reflect a gradual change of public attitudes to surveys. If that is true, the decline is unlikely to be reversed in the short-term by any general change is survey procedures. Nonetheless, it is important to continue to investigate methods of improving or maintaining response on the NTS and other surveys.

Alongside this, serious consideration should be given to adopting methods to reduce any nonresponse biases present in the responding sample. Adjustment methods have been common practice in National Statistics Offices in most developed countries for many years and in recent years most important UK surveys have introduced some nonresponse adjustment. At the same time, methods of adjustment have become more sophisticated and easier to apply as more research has been published and more software has appeared. There is now little excuse for not adopting some form of adjustment on the NTS.

3.1 Diary Length and Respondent Burden

The survey uses a prospective 7-day diary to collect travel details for all members of responding households. Currently the Department does not use the NTS to investigate patterns of travel within the 7-day period, it simply treats this as a sample of 7 days. The Department is not aware of any external user who makes use of this feature of the data.

Unsurprisingly, response rates are widely held to be lower for diary surveys, as these place a greater demand on respondents. To counteract this, a number of surveys that use diaries offer an incentive or a reward to respondents but the NTS has never done that. In view of the limited use made of the current 7-day diary, it is worth considering whether adequate data could be collected using a shorter diary or even by retrospective recall.

The last time this was examined systematically was in 1985.[11] An experiment was mounted to test the effect of using a one-day versus a 7-day diary on the survey. This found a small but not statistically significant improvement in response rate (from 75% to 78%) from the shorter diary. An analysis of the variation between days for a range of travel variables had earlier suggested that the number of households would need to rise by 40% to retain the same level of precision for most variables, if a one-day diary were adopted. This surprising conclusion was the main reason for mounting the experiment at a time when low response rates were less of an issue.

The experiment concluded that, although a one-day diary would be cheaper to administer, the overall fieldwork cost would rise by 22% to cover the extra households. There were some indications that data quality might be somewhat better with the one-day diary but the evidence for this was not overwhelming.

Although many of these figures might well be different if the experiment were repeated today, it seems unlikely that the critical response differential would be so great as to justify this level of additional cost. Nevertheless if response rates continue to decline, it is worth considering mounting further experiments along these lines into shorter diaries or the use of one day retrospective recall.

3.2 Extent of Nonresponse Bias

Bushnell[12,13] reported the results of the comparison of 1991 Census data for NTS respondents and nonrespondents and attempts to build weighting adjustment schemes based on the results. Refusal was and remains the dominant source of nonresponse and total nonresponse rates were higher among the following types of household:

- 3 or more adults with no children

- couples with non-dependent children
- female heads
- widowed or divorced heads
- heads born outside the UK
- unemployed or inactive heads
- heads with low qualifications
- London

However differential response rates by Census characteristics were less marked in this survey than in several other household surveys that were included in the study.

3.3 Weighting for Nonresponse

Bushnell's second report attempts to construct weighting schemes based on this Census comparison. These attempts were limited by a lack of agreement between Census and survey reports on the economic status of the household head and the fact that qualification data are not collected on the NTS. Nonetheless the report proposed two alternative schemes and found, as stated in the summary that "bias was reduced across a range of household characteristics". Despite this, the final paragraph from the discussion concluded:

> "Taking all factors into account, when considering the cost of applying weights (both in terms of a decrease in precision and possible extra financial considerations) it would seem that these direct methods of estimating weights to compensate for nonresponse are unlikely to be cost effective".

As the report argued that that any reduction in precision would be minimal, this appears to hinge on a fear that extra financial costs will be incurred by introducing weights.

The Report from the Government Statistical Service (GSS) Task Force on Weighting and Estimation[14], endorsed by the old GSS(M) Committee, discusses a number of different approaches to nonresponse weighting on GSS surveys and concludes that "Calibration" weighting should be the preferred method. This essentially uses available population data to control both random and non-random variation in the responding sample.

This method has now been used on a number of ONS surveys, as well as on the DSS's Family Resources Survey and it is the proposed method of weighting the General Household Survey in future. The variant used or proposed on both the Family Expenditure Survey and the General Household Survey is to apply initial nonresponse adjustment factors, derived from the 1991 Census comparison, then to calibrate these weights to control the sample totals on age, sex and region. Extensive work on variances has indicated that any loss of precision from the introduction of nonresponse adjustment weights is more than compensated for by gains in precision from the final calibration.

We recommend that the Department continues to investigate the effects of nonresponse on the survey, by repeating the Census-based comparison using 2001 data; and that it introduces appropriate nonresponse weighting on the survey in future.

3.4 Partial Response

The latest (1999) nonresponse rate of 34% from the survey includes 8% who are classed as partial respondents. This figure has remained remarkably stable over the years despite the increase in other types of nonresponse. Although interviewers are encouraged to collect as much data as possible for these cases, currently no data from these partial respondents is used in deriving NTS results. The Department has recently undertaken an analysis[15] of the extent

and nature of partial response to the survey. Partial respondents are of two main types – nonresponse by at least one person in the household and failure to complete the travel diary by at least one person. In 1996/1998 the latter group accounted for 77% of the partial responders.

The report concludes that little useful data can be recovered from the individual incomplete diaries, though it is not clear whether the diary data from other household members would be of value in individual-level analysis. The report also identifies some systematic differences between the non-diary data of the partial respondents and the full respondents, suggesting that nonresponse biases could be reduced by including this data in NTS results.

If this were to be done, some further weighting of the sub-sample of households (or individuals) with travel diaries would be desirable. Using the household data, it should be possible to produce an effective model of the propensity to complete a travel diary, which could be used to weight the diary data to make it commensurate with the household data, prior to any general nonresponse weighting.

3.5 Imputation

There was no time to consider this aspect of the survey's procedures in this review. It should be considered later.

4. Recommendations

Some of these recommendations are contingent on the decision of whether or not to increase the overall annual sample size. Recommendations that depend on the sample size increasing are marked (*).

1. The Department should consider making more use of rolling 3-year aggregates when assessing change or reporting trends (2.1).
2. *The Department should assess the relative importance of estimates of current level and change then consider introducing a panel element into the design (2.2).
3. The Department should reconsider the size of the current London boost and should post-stratify the sample to the correct regional distribution (2.3).
4. The Department should repeat the assessment of optimal stratifiers once the 2001 Census data become available (2.4).
5. *The Department should consider (a) reinstating the population density and car-to-work stratifiers or (b) deep geographical stratification to improve small area estimates (2.4).
6. The current use of postcode sectors or similar-sized areas as primary sampling units should be retained. Some variation in sample size within sectors could be countenanced (2.6).
7. The Department should apply weights to correct the bias against multi-household addresses (2.6).
8. The Department should retain the current placing pattern but keep it under review by collecting relevant response data (2.7).
9. The Department should retain the current mid-month to mid-month field periods, unless this significantly increases their costs or limits their output options. A more thorough analysis of the impact should be undertaken before any change is made (2.7).

10. If the decline in response rates continues, the Department should consider a range of measures including incentive payments, the use of a shorter diary and a switch to retrospective recall to eliminate the need for diaries completely (3.1).
11. The Department should introduce nonresponse weighting on the survey, taking note of the recommendations of the GSS Task Force on Weighting (3.3).
12. If possible, the Department should commission ONS to repeat the Census-based assessment of nonresponse bias after 2001 (3.3).
13. The Department should continue to investigate the potential use of data from partially-responding households with a view to using as much of this data as possible in NTS estimates (3.4).
14. The Department should assess the quality and impact of its imputations and the use of other statistical approaches (3.5).

Dave Elliot

Office for National Statistics

September 2000

---

[1] National Travel Survey Technical Report 1997 (1999) Freeth, S. et al. ONS, London

[2] Holt & Farver (1992) The use of composite estimators with two-stage repeated sample designs. *Journal of Official Statistics 8, 405-416.*

[3] The survey frame excludes Scottish islands

[4] These are based on PAF address totals from December 1999. Thus they make the simplifying assumption that all households in multi-household addresses are included in the sample. In practice, since multi-household rates are higher in London, the sampling fractions in London will be slightly lower.

[5] Kish, L. (1965) *Survey Sampling.* Wiley

[6] The *weighting efficiency* is the variance (of an estimator) in an unweighted sample divided by the variance in a weighted sample of the same size and design.

[7] Elliot, D. (1991) Weighting for nonresponse – a survey researcher's guide. ONS

[8] Barton J. (199?) Investigating stratification options for the NTS. ONS report to DETR.

[9] Barton, J. (1997) Multi-household procedures for social surveys. Survey Methodology Bulletin 40, ONS.

[10] Bruce, S., Lound, C. & Elliot, D. Designing surveys using variances calculated from Census data. ms submitted to GSS methodology series

[11] Butcher, R. & Eldridge, J. (1990) The use of diaries in data collection. *The Statistician, 39, pp25-41*

[12] Bushnell, D. (1994) The National Travel Survey: Report of the 1991 Census-linked study of survey nonrespondents. OPCS unpublished report to DETR

[13] Bushnell, D. (1995) Weighting the NTS to compensate for nonresponse. OPCS unpublished report to DETR

[14] Elliot, D. (1999) Report of the Task Force on Weighting and Estimation, GSS Methodology Series, 16. ONS

[15] Hird, D. (2000) National Travel Survey: partially responding households database 1995/1996/1998