Monograph:
Fowkes, A.S. and Watson, S.M. (1989) Sample Size Determination to Evaluate the Impact of Highway Improvements. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 282

Published paper

Fowkes, A.S., Watson, S.M. (1989) *Sample Size Determination to Evaluate the Impact of Highway Improvements.* Institute of Transport Studies, University of Leeds. Working Paper 282

Working Paper 282

November 1989

(Version 2, May 1997)

SAMPLE SIZE DETERMINATION TO EVALUATE THE IMPACT OF HIGHWAY IMPROVEMENTS

AS Fowkes and SM Watson

This work was commissioned by the Transport and Road Research Laboratory.

UNIVERSITY OF LEEDS
Institute for Transport Studies

ITS Working Paper 282

November 1989

Sample size determination to evaluate the impact of highway improvements:

AS Fowkes and SM Watson

This work was commissioned by the Transport and Road Research Laboratory.

This report is one of a series produced during a project: "Feasibility of measuring response to new highway capacity", carried out by a consortium of ITS, TPA and John Bates Services on behalf of TRRL.

The views expressed in these reports do not necessarily reflect those of TRRL or the Department of Transport.

Reports in the series are:

The Feasibility of Measuring Responses to New Highway Capacity (by Bonsall PW). TRRL Contractors Report No CR200, due early 1990.

Travellers Response to Road Improvements: Implications for User Benefits (by Mackie PJ and Bonsall PW). Traffic Engineering and Control 30 (9), 1989. Note that this paper was prepared in advance of the TRRL contract.

User Response to New Road Capacity: A Review of Published Evidence (by Pells SR). ITS Working Paper 283, 1989.

Sample Size Determination to Evaluate the Impact of Highway Improvements (by Fowkes AS and Watson SM). ITS Working Paper 282, 1989.

Measuring Impacts of New Highway Capacity: A Discussion of Potential Survey Methods (by Bonsall PW). ITS Working Paper 289, 1989.

Evidence on Ambient Variability and Rates of Change (by Watson SM). ITS Technical Note 257, 1989.

Application of Stated Preference Methods to User Response to New Road Schemes (by Bates J and Wardman MR). ITS Technical Note 258, 1989.

Some Observations on Appropriate Time Scales, Siting and Location of Surveys and on the Use of Panels to Measure Responses to Highway Improvements (by Goodwin PB and Jones PM of TSU, Oxford). ITS Technical Note 259, 1989.

User Response to New Road Capacity: What The Experts Said (by Bonsall PW, Mackie PJ and Pells SR). ITS Technical Note 260, 1989.

Calculation of Costs of Surveys to Measure User Response to New Road Capacity - Worked Examples for Typical Schemes under Alternative Survey Strategies (by Bonsall PW). ITS Technical Note 261, 1989.

Abstract

FOWKES, A.S. AND WATSON, S.M. (1989) Sample size determination to evaluate the impact of highway improvements. Working Paper 282, Institute of Transport Studies, University of Leeds, Leeds.

This paper was prepared for the Department of Transport, as a support document to a main report on the feasibility of measuring responses to highway improvements. The paper discusses the statistical issues involved, particularly as regards the determination of suitable sample sizes. Worked examples are provided, using such data on ambient variability and adjustment factors as were available to us. Some of the data is included as an appendix where it was felt to be otherwise not easily available.

The note asks two sort of questions. Firstly, what is the minimum sample size to take to be a certain percent confident that a given quantity lies in a range of a given width. Secondly, what sample sizes should be taken in Before and After studies so as to be a certain percent confident that a change in a quantity by a given amount will be detected as a statistically significant difference at some chosen significance level.

Three sorts of quantities are discussed:

- total flows past a point, which may be counted by loops, tubes or manually;
- partial flows, such as a particular O-D flow, which require roadside interviews;
- journey times over particular links.

SAMPLE SIZE DETERMINATION TO EVALUATE THE IMPACT OF HIGHWAY IMPROVEMENTS

1. INTRODUCTION

This note will discuss issues relevant to the determination of sample size requirements when the objective is to measure the impact of a highway improvement. Statistical theory will be presented and indications provided as to where the appropriate numbers required for the statistical formulae can be obtained. Worked examples will be provided, on an illustrative basis, using such data as are to hand. We have included in appendices a tabular rendering of such data as may not be easily available.

The note will essentially be asking two sorts of questions. Firstly, what is the minimum sample size to take to be a certain percent confident that a given quantity lies in a range of given width. Secondly, what sample sizes should be taken in Before and After studies so as to be a certain percent confident that a change in a quantity by a given amount will be detected as a statistically significant difference at some chosen significance level.

As well as the two sorts of questions, discussed above, we will apply these to three sorts of variable of interest, namely:

(i) total flows past a point, which may be counted by loops, tubes or manual methods; (NB. AADT = Annual Average Daily Traffic)

(ii) partial flows, such as a particular O-D flow, which require roadside interviews;

(iii) journey times over particular links.

The note does not deal with the case of comparing modelled and forecast values.

2. SAMPLING THEORY

Suppose X is the name of a particular variable of interest, say a flow or a speed on a particular link over a particular hour on a particular day of the year. We can take measurements x_i from this variable and note that they have average value $\overline{x}$ and a variance which we may take as an estimate of the true variance of X and so denote VAR(X). We shall not worry that our estimate of variance may not be exact, provided we have large (>30) samples. We shall worry that the estimated mean, $\overline{x}$, may not exactly equal the true mean, which we may denote by E(X) or μ.

However, we know that $\overline{x}$ will tend to get closer to E(X) the larger the sample size, n, we take. Hence we can form confidence intervals for E(X) which are of desired width by taking n to be suitably large. A 95% Confidence Interval can be written as

$$95\%\ \text{C.I for } E(X) = \overline{x} \pm z(0.025)\sqrt{\frac{VAR(X)}{n}}$$

where n is the sample size, and 0.025 refers to the proportion of the distribution in each tail outside the Confidence Interval, i.e. 2½% in each tail, leaving 95% in the middle, with z indicating convergence to a Standard Normal Distribution.

In general, a $(1-\alpha)$ Confidence Interval will have $\alpha/2$ in each tail, and so is denoted as

$$(1-\alpha)\ \text{C.I for } E(X) = \overline{x} \pm Z(\alpha/2)\sqrt{\frac{VAR(X)}{n}}$$

When the estimate of VAR(X) is based on a small sample it may be desirable to use the t distribution rather than the standard Normal (z) distribution.

The above method works as long as the error involved in VAR(X) can be regarded at random. This will arise as 'sampling error' which says that each sample of size n from a given population will have a different mean (except by chance) depending on which members of the population are chosen. In addition any measurement error which can be taken as purely random can be included here.

What cannot be included are systematic forms of variation, say relating to the day of the week or month of the year, if we are sampling at a particular time. If we have a Tuesday count in February we will doubtless find a factor to enable us to estimate AADT, but there will be an error variance associated with this factoring which will not be reduced by increasing the sample taken on the February Tuesday. Consequently we should try as far as possible to sample from all periods of interest, and if making comparisons over time (such as in Before and After studies) we should choose sample periods which are as alike as possible.

Where factoring has to take place, statistical theory allows composite variances to be calculated. For example, suppose that:

F_1 takes a particular hour to an average hour

F_2 takes a particular weekday to an average weekday

F_3 takes a particular month to an average month

F_4 has mean 1, but has variance which allows for unexplained variation which we shall call

AMBIENT VARIABILITY

and that coefficients of variation are known (possibly from published sources) for these factors. Let our measured average, $\bar{x}$, be the average speed for a particular hour on a particular weekday in a particular month. Let Y be our estimate of the average speed on any weekday in the year. Clearly we apply all four factors:

$$\hat{Y} = F_1 \; F_2 \; F_3 \; F_4 \; \overline{X}$$

The variance of our new estimate, $\hat{Y}$, is

$$VAR(\hat{Y}) = VAR\ (F_1 \; F_2 \; F_3 \; F_4 \; \overline{X})$$

Variances of products of terms can be a bit cumbersome, but can be handled straightforwardly by repeated use of the formula, for independent A and B:

$$VAR(AB) = VAR(A)\ VAR(B) + [E(A)]^2\ VAR(B) + [E(B)]^2\ VAR(A)$$

Often it will be sensible to make use of an approximate formula which is much simpler to handle, namely (for independent A and B whose CVs are less than about 0.2):

$$CV^2(AB) \simeq CV^2(A) + CV^2(B)$$

where CV denotes 'coefficient of variation', which is defined to be the ratio of the standard deviation to the mean, hence

$$CV^2(A) = VAR(A)/A^2$$

In our example we have

$$CV^2(\hat{Y}) = CV^2(F_1) + CV^2(F_2) + CV^2(F_3) + CV^2(F_4) + CV^2(\overline{X})$$

in which, we know that

$$CV^2(\overline{X}) = VAR(\overline{X})/\overline{x}^2 = VAR(X)/n\overline{x}^2$$

so, solving for n gives

$$n = \frac{VAR(X)}{\overline{x}^2[CV^2(\hat{Y}) - CV^2(F_1) - CV^2(F_2) - CV^2(F_3) - CV^2(F_4)]}$$

Hence to fix $CV^2(\hat{Y})$ at some desired value, with all other terms known, we can find n from this equation. In the simple case where the CV's of all the F's are zero we have

$$n = \frac{VAR(X)}{\bar{x}^2 CV^2(\hat{Y})} = \frac{CV^2(X)}{CV^2(\hat{Y})}$$

So if speeds had a coefficient of variation of 0.2, and we wanted a 95% CI of width 1% of mean speed then

$$95\% \text{ C.I is } \hat{Y} \pm 0.01\hat{Y}$$

$$\text{So } 1.96SD(\hat{Y}) = 0.01\hat{Y}$$

$$\frac{SD(\hat{Y})}{\hat{Y}} = \frac{0.01}{1.96} = CV(\hat{Y})$$

$$n = \frac{(0.2)^2}{\left(\frac{0.01}{1.96}\right)^2} = (0.2)^2(196)^2 = 1537$$

This says that we must take a sample of size 1537 individual speeds, for that hour on that weekday in that month.

Suppose instead that we wish to estimate the average speed in a different month, but must sample in the current month only. We presume there is a factor, F_3, available to convert current month speeds to this other month's speeds, and let us suppose that the source that gives this F_3 value states that $CV(F_3) = 0.002$, which takes into account that (in the data they studied) the relationship between months was not exact, the factor being slightly greater in some instances than others.

We still have $CV(\hat{Y}) = \left(\frac{0.01}{1.96}\right)$ for 95% C.I

of width 1%, but now

$$n = \frac{CV^2(X)}{CV^2(\hat{Y}) - CV^2(F_3)}$$

$$= \frac{(0.2)^2}{(0.0051)^2 - (0.000004)}$$

$$= 1818$$

i.e. a larger sample to overcome this larger variability.

Note that the coefficient of variation for $\hat{Y}$ cannot be controlled to be less than the coefficient of variation of any of the factors which are incorporated in $\hat{Y}$.

3. STUDIES OF TOTAL VEHICLE FLOWS FROM COUNTS

(A) TO DETERMINE AADT TO A GIVEN ACCURACY

Suppose we are counting the flow along a road for 16 hours each day, and wish to know how many days we need to count to obtain 24 hours AADT to within 1%.

Table 4 of Appendix D14 of TAM (March 1985 update) provides 'M-factors' by road type which factor 16 hour counts in stated months to 24 hour AADT, giving associated CV figures which are all 6½%. Clearly the CV for an AADT estimate based on a single count is bound to be greater than 6½%. Our suggested approach here is to derive separate AADT estimates from several days 16-hour counts and then average the result, i.e.

$$\hat{Y} = \frac{1}{n_d} \sum_{1}^{n_d} F_m X$$

where $\hat{Y}$ is our estimate of AADT

n_d is the number of 16-hour counts used

F_m is the M factor, $CV(F_m) = 0.065$

X is our count, subject to measurement error such that

$$CV(X) = \frac{0.05}{1.96} \triangleq 0.025$$

We understand that most of the variation accounted for by $CV(F_m)$ is due to 'site-to-site' variation in the data set from which the F_m factors were derived. However, our view is that 'day-to-day' (or 'ambient') variation is also included in $CV(F_m)$. If this portion of variability could be split off then observations on different days would reduce the variance of $\hat{Y}$.

Let us therefore split F_m into two factors F_3 and F_4, with F_4 having mean unity and merely allowing for ambient variability.

We have $F_m = F_3 F_4$

We know
$$CV^2(F_m) = CV^2(F_3F_4)$$
$$\triangleq CV^2(F_3) + CV^2(F_4)$$

Suppose that we can split the $CV(F_m)$ such that

$$CV(F_3) = 0.048$$
$$CV(F_4) = 0.044$$

(this is for illustration only, but represents our current best guess)

We shall proceed by combining the ambient variability (over days) with the measurement error, and attenuate both by counting over more than one day.

Recapping, we have

$$\hat{Y} = \frac{1}{n_d} \sum_{i=1}^{n_d} F_3 \, F_{4i} \, X_i$$

We will assume that all the counts are taken at a similar time of year, or at least at 'neutral' times of year, such that F_3 can be taken as constant over days:

$$\hat{Y} = \frac{F_3}{n_d} \sum_{i=1}^{n_d} F_{4i} \, X_i$$

$$CV(\hat{Y}) = CV\left[\frac{F_3}{n_d} \sum_{i=1}^{n_d} F_{4i} \, X_i\right] = CV[F_3(\overline{F_4X})]$$

$$\text{Where } \overline{F_4X} = \frac{1}{n_d} \sum_{i} F_{4i} \, X_i$$

and denotes the mean value of the day's count adjusted for ambient variability.

The approximate formula gives

$$CV^2(\hat{Y}) = CV^2(F_3) + CV^2(\overline{F_4X})$$

$$= CV^2(F_3) + \frac{VAR(\overline{F_4X})}{(\overline{F_4X})^2}$$

$$= CV^2(F_3) + \frac{VAR(F_4X)}{n_d(\overline{F_4X})^2}$$

using sampling theory for variances of sample means.

$$CV^2(\hat{Y}) = CV^2(F_3) + \frac{CV^2(F_4X)}{n_d}$$

$$CV^2(\hat{Y}) = CV^2(F_3) + \frac{CV^2(F_4)}{n_d} + \frac{CV^2(X)}{n_d}$$

The values we obtained earlier were

$$CV(F_3) = 0.048$$
$$CV(F_4) = 0.044$$
$$CV(X) = 0.025$$

Clearly, if we know n_d we can determine $CV^2(\hat{Y})$, or equally we can set a value for $CV^2(\hat{Y})$, provided it is greater than $CV^2(F_3)$ and determine sample size n_d.

A further refinement is that if there are a fixed number of days, N_d, in the period under review, we can avoid ambient variation, $CV^2(F_4)$, by sampling all days. The term $\frac{CV^2(F_4)}{n_d}$ will then disappear.

More generally, we can multiply this term by the FINITE POPULATION CORRECTION FACTOR (FPCF) defined as:

$$FPCF = \frac{N_d - n_d}{N_d - 1}$$

If we sample only one day this term collapses to unity, while if we sample all possible days ($n_d = N_d$) then the term becomes zero.

Hence all cases are covered by

$$CV^2(\hat{Y}) = CV^2(F_3) + \left[\frac{CV^2(F_4)}{n_d} \quad \frac{(N_d - n_d)}{(N_d - 1)}\right] + \frac{CV^2(X)}{n_d}$$

In order to determine sample sizes we solve for n_d, so

$$n_d = \frac{CV^2(F_4) \quad (N_d/(N_d - 1)) \quad + \quad CV^2(X)}{CV^2(\hat{Y}) - CV^2(F_3) + CV^2(F_4)/(N_d - 1)}$$

While, at first sight, this may appear to overcome the restriction that, for finite n_d

$$CV^2(\hat{Y}) > CV^2(F_3)$$

this is not the case, since if the above is violated we will obtain

$$n_d > N_d$$

which is, by definition, impossible.

Hence the smallest $CV^2(\hat{Y})$ can be is, from our earlier guess, $(0.048)^2$

$$\begin{aligned} \text{if } CV(\hat{Y}) &= 0.048 \\ \text{then } SD(\hat{Y}) &= 0.048\hat{Y} \\ \text{and } 1.64SD(\hat{Y}) &= 0.0787\hat{Y} \end{aligned}$$

Hence the smallest 90% confidence interval for Y, true annual average daily traffic will be

$$\hat{Y} \pm 7.87\%$$

i.e. a 90% confidence interval could not be as accurate as ±5% even if all the available days were sampled.

WORKED EXAMPLE

If the above values for $CV(F_3)$, $CV(F_4)$ and $CV(X)$ hold true (i.e. are 0.048, 0.044, and 0.025 respectively) and if there are (N_d=) 30 days in the period on which the F_3 factor is to be based, how big a sample (in terms of number of days sampled) do we need to take to be 90% sure we know AADT to within ± 10%?

We must start by reviewing our assumptions, the most important being that the M factor CV (as published in TAM) of 0.065 might reasonably be split up into 0.044 from not knowing the average for the survey month exactly from only taking a sample, and 0.048 for the uncertainty in going from this month to an average month (presumed due to site to site variability in the data sets used to determine the M factor). Hence if we sample all 30 days of the month the 0.044 will disappear, and with smaller samples will be reduced as determined by sampling theory.

To get a 90% C.I we need Z=1.64, and to be within ±10% we need

$$1.64\ SD(\hat{Y}) = 0.1\hat{Y}$$

i.e. $$CV(\hat{Y}) = \frac{0.10}{1.64}$$

$$CV^2(\hat{Y}) = 0.003718$$

Using the formula

$$n_d = \frac{CV^2(F_4)(N_d/(N_d-1)) + CV^2(X)}{CV^2(\hat{Y}) - CV^2(F_3) + CV^2(F_4)/(N_d-1)}$$

$$n_d = \frac{(0.044)^2\ (30/29) + (0.025)^2}{0.003718 - (0.048)^2 + (0.044)^2/29}$$

$$= \frac{0.002628}{0.001481}$$

So a two day count would be needed.

If the question had been to find AADT to within $\pm$ 8%, we would have had

$$CV^2(\hat{Y}) = (0.08/1.64)^2 \qquad = 0.002379$$

$$n_d = \frac{0.002628}{0.002379 - 0.002237} = \frac{0.002628}{0.0001425}$$

$$= 18.5$$

In which case we would need to sample for 19 days.

4. BEFORE AND AFTER STUDIES OF FLOW

An additional problem arises with Before and After studies, in that secular growth of traffic may occur, by which we mean traffic growth which would have occurred even in the absence of the scheme which we may presume to have been implemented between the 'Before' and 'After' studies. While it is possible to obtain national average figures for the rate of growth of traffic, the applicability of these figures to particular sites is, at best, dubious.

Our proposal here is to use control sites, unaffected (positively or adversely) by the Scheme. These sites will show what growth would be expected had the scheme not be implemented. In order to best choose suitable control sites we would suggest that two 'Before' surveys are conducted, at a number of sites including the site of interest. Since we are only measuring flow, these 'surveys' may sometimes only entail inspecting the output from permanent counters. The subset of sites whose unexplained growth between the two 'Before' Surveys best matches that of the site of interest should be chosen as our control sites.

Having made this selection, the variability of growth rates between sites between the 'Before' and 'After' surveys can be expected to be at least as great as that observed between the two 'Before' surveys for the selected sites. Let us denote this variability $VAR(F_g)$, with associated coefficient of variation.

$$CV(F_g) = \sqrt{\frac{VAR(F_g)}{F_g}}$$

Where F_g represents the factor to be applied to the site of interest to take account of secular growth.

In principle, all the other factors discussed in Section 3 as being relevant to the accuracy of a count at a single point in time are also relevant here. However, it is clear that we should avoid the use of such factors wherever possible by sampling at similar times in both the 'Before' and 'After' surveys. The simplest case would probably be to sample after one year in exactly the same month and exactly the same days of the week, and for exactly the same length of time as in the 'Before' survey. This leaves only the secular growth factor to take into account.

In this simple case, consider that as in Section 3 we require to know how many days to sample for, now both before, n_b, and after, n_a. Suppose we denote the scheme effect as S, which might be a figure such as 1.2 to denote a 20% increase. We can set out the situation as follows:-

True 'Before' [redacted] flow : μ_b

True 'After' [redacted] flow : $\mu_a = SF_g\mu_b$

We wish to measure the effect due to S:

$$S = \frac{\mu_a}{F_g\mu_b}$$

In our estimate of S we will have 'Before' and 'After' flow count (average) measures $\overline{X}_a$ and $\overline{X}_b$ and an estimated F_g with coefficient of variation as discussed above

$$\hat{S} = \frac{\overline{X}_a}{F_g \overline{X}_b}$$

$$\text{where } \overline{X}_a = \frac{1}{n_a} \sum_{i=1}^{n_a} X_{ai} \qquad \text{and } \overline{X}_b = \frac{1}{n_b} \sum_{i=1}^{n_b} X_{bi}$$

Such that $VAR(\overline{X}_a) = VAR(X)/n_a$
and $VAR(\overline{X}_b) = VAR(X)/n_b$

assuming underlying variability has not increased.

The exact way in which sample sizes will be determined will depend on the question asked. If we are merely required to have both Before and After samples each to a given accuracy (so as to spot a scheme effect) then the theory of Section 3 is sufficient. However, it is often the case that the question is posed in the form of requiring Before and After sample sizes so as to be a certain percentage confident that a given sized change will be detected as statistically significant at a stated significance level. Note that the practitioner must now supply the statistician with three percentages:

(i) the confidence level - say 90%

(ii) the percentage change to be spotted - say 10%

(iii) the significance level for the test of difference - say 5%

In technical statistical terms (i) gives the Type 2 error as 10%; (ii) gives the mean for the alternative hypothesis as 110% of the original value; and (iii) fixes the Type 1 error at 5%.

Suppose that the Before and After samples are taken at identical times of year and on identical days, then there will be no factors to apply. We will still have F_4 for ambient variability, having mean 1 and some unknown coefficient of variation. In Section 3 we have used 0.044 for our $CV(F_4)$ and we shall use that again here, purely for illustration.

Our Null Hypothesis will be that there has been no change, i.e. H_o: $\mu_a = \mu_b$

Let us take measurements, X, before and after and arrive at Sample means, $\overline{X}_b$ and $\overline{X}_a$, from samples of size n_a and n_b. The variability in these sample means will be determined by the adequacy of each days measurement in reflecting the true value for that day, and day to day variation. Since ambient variability and measurement error are usually taken to be proportional to mean it follows that if the scheme causes an increase, then variability in absolute terms will be higher for the After survey than for the Before survey. The coefficient of variation will be constant, though.

Setting the problem out formally, let us choose (for illustration)

Null Hypothesis, H_0: $\mu_a = \mu_b$

Alternative Hypothesis, H_1: $\mu_a = (1 + k)\ \mu_b$

Type I Error: Prob $(\overline{X}_a - \overline{X}_b > C | H_o) = \alpha_1$

Type II Error: Prob $(\overline{X}_a - \overline{X}_b < C | H_1) = \alpha_2$

From the Central Limit Theorem

$$C = Z_1 SD(\overline{X}_a - \overline{X}_b)$$

$$k\mu_b - C = Z_2 \; SD(\overline{X}_a - \overline{X}_b)$$
$$\therefore \; k\mu_b = (Z_1 + Z_2) \; SD(\overline{X}_a - \overline{X}_b)$$

Skelton (1982) also arrives at this result.
Adding our factors F_4 we have

$$k\mu_b = (Z_1 + Z_2) \; SD \; (\overline{F_4X}_a - \overline{F_4X}_b)$$

where Z_1 and Z_2 are the Z values for the Type 1 and Type 2 errors, on one tailed tests, here giving 1.28 and 1.64, summing to 2.92, and k = 0.1. in reverse order!

$$\text{Hence} \quad VAR(\overline{F_4X}_a - \overline{F_4X}_b) = \frac{0.01 \; \mu_b^2}{(Z_1 + Z_2)^2}$$

$$\text{Now} \quad VAR(\overline{F_4X}_a - \overline{F_4X}_b) = VAR(\overline{F_4X}_a) + VAR(\overline{F_4X}_b)$$

As before, we can decompose $VAR(\overline{F_4X}_a)$ etc.

$$VAR(\overline{F_4X}_a) = \frac{VAR(F_4X_a)}{n_a}$$

$$= \frac{(\overline{F_4X}_a)^2}{n_a} CV^2(F_4X_a)$$

$$= \frac{(\overline{F_4X}_a)^2}{n_a} \; [CV^2(F_4) + CV^2(X_a)]$$

We can substitute μ_a for $\overline{F_4X}_a$ giving

$$\left[\frac{k\mu_b}{Z_1+Z_2}\right]^2 = \left[\frac{\mu_a^2}{n_a} + \frac{\mu_b^2}{n_b}\right] CV^2(F_4) + \frac{\mu_a^2}{n_a} CV^2(X_a) + \frac{\mu_b^2}{n_b} CV^2(X_b)$$

If we take the accuracy, in terms of CV, to be equal in both the Before and After surveys, then

$$\left[\frac{k\mu_b}{Z_1+Z_2}\right]^2 = \left[\frac{\mu_a^2}{n_a} + \frac{\mu_b^2}{n_b}\right] [CV^2(F_4) + CV^2(X)]$$

We will usually wish to take equal sample sizes, for efficiency, so we can solve for $n_d = n_a = n_b$

$$\left[\frac{k\mu_b}{Z_1+Z_2}\right]^2 = \left[\frac{\mu_a^2 + \mu_b^2}{n_d}\right] [CV^2(F_4) + CV^2(X)]$$

$$n_d = \frac{(\mu_a^2 + \mu_b^2)[CV^2(F_4) + CV^2(X)]}{\left[\frac{k\mu_b}{Z_1+Z_2}\right]^2}$$

The test assumes H_o: $\mu_a = \mu_b$, so we can further simplify

$$n_d = \frac{2[CV^2(F_4) + CV^2(X)](Z_1+Z_2)^2}{(k)^2}$$

eg. $k = 0.1$, $CV(F_4) = 0.044$, $CV(X) = 0.025$, $Z_2 = 1.28$, $Z_1 = 1.64$

$$n_d = \frac{2(0.00194 + 0.00063)(2.92)^2}{0.1}$$

$$= \frac{0.043826}{0.01} = 4.38$$

i.e. one should sample for 5 days before and 5 days after.

In this way variance expressions for S could be developed, set to desired levels and solved for the sample sizes. This will be greatly facilitated if we can assume sample sizes to be equal in both the 'Before' and 'After' surveys, i.e. $n_a = n_b$. If the before survey has already been done, the size of the after survey can be determined by inserting the known value of n_b.

If other factors, e.g. adjusting for the effects of month (if both samples cannot be taken in the same month) have to be included then sample sizes will increase, as in Section 3, and become infinite if too high an accuracy is required.

5. COUNTING PARTIAL FLOWS

In this section we shall look at the situation where we are only interested in part of the flow passing a particular site, e.g. we may be interested only in vehicles with particular origin/destination characteristics. In this case there will be some unknown proportion of the flow, P, which we will need to estimate so as to multiply it by the estimated total flow passing the site. We may still wish to count for more than one day, in order to obtain a sufficiently accurate estimate of total flow. However, the exigencies of carrying out roadside interview surveys strongly suggests that we should not interview at a given site on more than one day. This could be very unfortunate for our sampling, since the OD mix may vary considerably from day to day.

The best we can probably do is to confine our interest to specific sorts of days (e.g. 'weekdays' or 'Sundays') and to interview accordingly. We will then be in a position to make the vital assumption that our estimate of P will apply to the total flow counted. Let μ be the true total flow of traffic per day. Let P be the true proportion of this which has the attribute of interest (e.g. if for a specific OD pair).

$$\text{The daily flow of interest} = P\mu$$

We can try to get a good estimate of μ by surveying over n_d days, and we can try to get a good estimate of P by taking a sample of size n. The optimal mix of n_d and n will depend on the relative costs of the two sorts of sampling. This will depend on such things as how long the interviews each take, and so we cannot sensibly even give rules of thumb here. We shall merely indicate the statistical framework involved.

The variance of a sample proportion is given, for the Hypergeometric distribution as is relevant here with sampling without replacement, is

$$\frac{P(1-P)(N-n)}{n(N-1)}$$

Here we do not know P, so use our estimate $\hat{P}$, and assume that N is much greater than n.

Our estimate of the flow of interest will be

$$\hat{Y} = \hat{P}\overline{F_4X}$$

$$CV^2(\hat{Y}) \doteq CV^2(\hat{P}) + CV^2(\overline{F_4X})$$

$$= \frac{VAR(\hat{P})}{\hat{P}^2} + \frac{VAR(\overline{F_4X})}{(\overline{F_4X})^2}$$

$$= \frac{(1-\hat{P})(N-n)}{n(N-1)\hat{P}} + \frac{VAR(F_4X)}{n_d(\overline{F_4X})^2}$$

$$\doteq \left[\frac{(1-\hat{P})}{\hat{P}n}\right] + \frac{CV^2(F_4) + CV^2(X)}{n_d}$$

Solving for n_d we have

$$n_d = \frac{CV^2(F_4) + CV^2(X)}{CV^2(\hat{Y}) - \frac{(1-\hat{P})}{\hat{P}n}}$$

If we add the complications of possibly wishing to apply a factor F_3, and possibly having a finite number of days N_d to sample from, we have

$$n_d = \frac{CV^2(F_4)(N_d/N_d-1) + CV^2(X)}{CV^2(\hat{Y}) - CV^2(F_3) + CV^2(F_4)/(N_d-1) - (1-\hat{P})/\hat{P}n}$$

6. MEASURING TRAVEL TIMES

The two main methods of measuring journey times are Moving Observer (MO), and Number Plate Matching (NPM). In the first case, a car or cars are driven along the link(s) of interest and the journey time noted. Common sense dictates that the MO should try to travel at the average speed of other vehicles using the road. This is sometimes difficult and corrections for cars overtaken or overtaking are available. Experience at Leeds suggested they were beneficial, but did not make much difference. The question is how many runs should be made to adequately determine the journey time. Currently, the COBA9 manual offers such advice.

Statistically, the question revolves around how accurate each MO measurement is. Experience in Leeds suggested that MO's, at 15 minute intervals were sufficient to give values which followed NPM results very closely through the morning peak. Since NPM is much more expensive, it is presumed here that it will be used only when there are special reasons, such as needing to know the spread of travel times on a given link. The conduct of NPM surveys is a specialist issue, complicated by considerations of how much of the traffic will actually match, and what will be the extent of spurious matching. The matter will not be further considered here.

If we wish to use MO's, but want to survey on several days so as to reduce measurement error and allow for day to day variation, then the problem is similar to that of Section 3.

If X_i, is the speed measured on day i, and F_4 is a (dummy) factor to allow for ambient (day-to-day) variability, then our estimate of average speed for several days runs is

$$\hat{Y} = \frac{1}{n_d} \sum_{i=1}^{n_d} F_4 X_i = \overline{F_4 X}$$

$$CV^2(\hat{Y}) = CV^2(\overline{F_4 X})$$

$$= \frac{VAR(F_4 X)}{n_d F_4{}^2 X^2}$$

$$= \frac{CV^2(F_4 X)}{n_d}$$

$$CV^2(\hat{Y}) \triangleq \frac{CV^2(F_4)}{n_d} + \frac{CV^2(X)}{n_d}$$

Where $CV(F_4)$ represents day to day variability of speeds, and $CV(X)$ represents measurement error in that the MO estimate will not equal the actual average speed on that day. Solving for n_d

$$n_d = \frac{CV^2(F_4) + CV^2(X)}{CV^2(\hat{Y})}$$

where the denominator will be determined by the desired accuracy required (as demonstrated in Section 3).

7. ESTIMATING HOURLY TRAFFIC FLOW

In the preceding sections it has been presumed that we have been dealing with a day's traffic, e.g. what is AADT?, has AADT changed etc. We may, however, wish to measure flows in individual hours, and any day's data can obviously be broken down

into shorter intervals. If we are interested in just one hour each day (e.g. the 'peak' hour) then we still have a 'days' problem since we need to decide on how many days to count this hour. The basic framework is as set out in Section 3, but clearly the appropriate factors will differ. Our understanding is that suitable factors are not readily available to practitioners.

The best data available to us is from Idris (1981) which may be interpreted as showing that, during heavily trafficked hours, the day to day (non-systematic) coefficient of variation might be about 0.05. If we continue to assume a coefficient of variation of our ATC of 0.025, then we have

$CV^2(X)$	=	0.025^2	:	accuracy of count
$CV^2(F_4)$	=	0.05^2	:	day to day variability
$CV^2(F_3)$	=	0	:	no factors involved

Suppose we wish to measure the true underlying hourly flow (say between 9 and 10 on weekdays in February 1989) Y, to within 5% with 90% confidence, then

$$CV^2(\hat{Y}) = \left[\frac{0.05}{1.64}\right]^2 = 0.00093$$

There were 20 weekdays, so $N_d = 20$

Using the formula from Section 3

$$n_d = \frac{(0.05)^2\ (20/19) + (0.025)^2}{0.00093 + (0.05)^2/19} \triangleq 3$$

So we should count the hour's traffic for 3 days, preferably chosen randomly from the 20 weekdays in February.

References

IDRIS, M K (1981) "The collection and interpretation of traffic flow data with particular reference to West Yorkshire and Pakistan Traffic Counts". MSc Thesis, Department of Civil Engineering, University of Leeds

Skelton, N (1982) "Determining appropriate sample size when two means are to be compared", Transport Engineering and Control, pp29-37.

Data Appendices

Secular Growth

Local evidence from historic traffic counts may be available giving year on year growth and associated CV's. Otherwise use evidence from "similar" sites, or published sources eg. TAM (fairly out of date) or Transportation Statistics.

Roadway manual suggests that CV's associated with secular growth between two years, n years apart, are

n	CV %
1	8
2	11.33
3	13.90
4	16.07
5	18.00
6	19.75
7	21.37
8	22.88

Source: Roadway Manual

Historic Evidence of Secular Growth

Motor vehicle flow at average point (1977 = 100)

1982	1983	1984	1985	1986
111	112	113	114	116

(all trunk and classified roads) P69

can be disaggregated as:

Built up roads	1982	1983	1984	1985	1986
Trunk	104	106	109	113	120
Principal	109	110	109	111	112
B roads	116	119	126	130	136
C roads	125	125	125	129	137

Source: Transportation Statistics

Conversion Factors (examples)

HOURS	FACTOR	SD
12-16 HRS	1.160	0.037
12-24 HRS	1.233	0.040
16-24 HRS	1.054	0.012

(conversion weekday to average weekday; Trunk, A, B, C roads)

DAYS	TABC	SD
Monday	1.050	0.026
Tuesday	1.060	0.051
Wednesday	1.030	0.045
Thursday	1.005	0.025
Friday	0.876	0.074

(conversion to a neutral month)

MONTHS	MAIN URBAN	CV (%)
January	1.188	9.0
February	1.190	6.4
March	1.066	4.8
April	1.042	2.9
May	1.013	3.2
June	1.007	4.6
July	0.976	4.3
August	0.982	4.8
September	0.972	2.5
October	0.980	3.9
November	1.030	3.9
December	1.068	7.3

Unless otherwise stated, factors refer to trunk roads.
Other tables of factors are also available.

Source: Roadway Manual

Kent Traffic Flow Information

Monthly average flow levels, (sd), (cv).

	A28			B2077		
J	10827	(409)	(3.78)	4083	(239)	(5.85)
F	11392	(613)	(5.38)	4133	(184)	(4.44)
M	11864	(577)	(4.86)	4484	(137)	(3.05)
A	12415	(341)	(2.74)	4583	(163)	(3.56)
M	12778	(700)	(5.47)	4843	(207)	(4.27)
J	13317	(529)	(3.97)	4813	(196)	(4.07)
J	13856	(625)	(4.51)	4758	(265)	(5.57)
A	15610	(961)	(6.15)	4712	(186)	(3.94)
S	13061	(661)	(5.06)	4650	(189)	(4.06)
O	12781	(600)	(4.70)	4507	(164)	(3.63)
N	12464	(372)	(2.98)	4401	(146)	(3.32)
D	12495	(529)	(4.23)	4401	(145)	(3.29)

	A21			A2		
J	32721	(1188)	(3.63)	23215	(864)	(3.72)
F	33936	(1967)	(5.80)	23666	(691)	(2.92)
M	35203	(2026)	(5.75)	24454	(696)	(2.85)
A	36846	(1501)	(4.07)	24639	(733)	(2.98)
M	37412	(2204)	(5.89)	25476	(853)	(3.35)
J	38068	(1795)	(4.72)	25515	(653)	(2.56)
J	38699	(1975)	(5.10)	25393	(886)	(3.49)
A	39363	(1628)	(4.14)	23850	(877)	(3.68)
S	38344	(1783)	(4.65)	24959	(710)	(2.85)
O	37239	(2089)	(5.61)	24926	(676)	(2.71)
N	35135	(1381)	(3.93)	24921	(772)	(3.10)
D	34340	(2759)	(8.04)	25455	(514)	(2.02)

Source: Kent County Council (Highways and Transportation)

Other information also available from Kent County Council:

i) A road factors with ± 1 standard deviation for 1978-1988

ii) Monthly factors for M/A/B/C in Kent for 1988 with ± 1 standard deviation

All monthly factors relate to the appropriate AADT figure.

IDRIS (using completely infilled data)

Road Code	MON	TUES	WED	THUR	
D402N	6642	6705	6944	6928	(mean daily flow)
	567	540	636	812	(sd)
	(8.5)	(8.1)	(9.2)	(11.7)	(cv)
D402S	7025	7254	7407	7429	
	453	620	555	846	
	(6.5)	(8.5)	(7.5)	(11.4)	
E302E	7480	7595	7614	7711	
	720	946	904	1070	
	(9.6)	(12.5)	(11.9)	(13.8)	
E302W	7344	7509	7402	7409	
	1265	1214	985	1043	
	(17.2)	(16.2)	(13.3)	(14.1)	
J401E	12093	12187	12476	12922	
	1601	1220	1255	1706	
	(13.2)	(10.0)	(10.1)	(13.2)	
J401W	11962	12248	12465	12948	
	1088	1060	733	1349	
	(9.1)	(8.7)	(5.9)	(10.4)	
C302N	6130	6323	6391	6424	
	593	607	565	808	
	(9.7)	(9.6)	(8.8)	(12.5)	
C302S	6121	6274	6369	6285	
	993	834	737	877	
	(16.2)	(13.3)	(11.6)	(13.9)	

Source: Idris (1987)