

TRAFFIC GENERATION ESTIMATION - SUGGESTIONS ON  
NEW RESEARCH DIRECTIONS

P.O. Barnard  
Australian Road Research Board

R.E. Brindle  
Australian Road Research Board

*ABSTRACT:*

*An important aspect of planning for new developments is to predict the amount of generated traffic. This is a necessary input both for traffic design and traffic planning activities. In this paper conceptual problems in measuring traffic generation are outlined, then examples given of estimates currently used in Australia. Two land use categories are highlighted, residential and retailing, since these probably represent the best and worst of current practice. Deficiencies with the techniques used to estimate generated traffic for both these land use categories are noted. Suggestions are then made on improvements to current methodology. A major contribution is to show how O-D travel survey data could theoretically be used to obtain traffic generation estimates for residential and retailing developments. It is recognized, however, that current traffic generation methodology is essentially flawed and an altogether new approach may be needed. One new approach based on economic individual choice theory is outlined.*

## I. INTRODUCTION

The ultimate purpose of traffic generation studies is to improve estimates of traffic activity associated with land uses of particular types and magnitudes.

It is useful before embarking on an exploration of traffic generation studies to distinguish between the terms "traffic generation" and "trip generation". Both terms are in common use in transport and traffic engineering, and this may sometimes cause confusion. Consistent with more common practice, this paper uses the term "**trip generation**" for person movements in the context of travel analyses for transport planning. Usually trip generation is concerned with a wider set of characteristics than traffic generation and tends to emphasize both origin and destination aspects of each trip.

"**Traffic generation**" is reserved for measures of vehicle (and sometimes pedestrian) movements associated with various sites or land uses. It can be defined as the number of one-directional vehicle (or pedestrian) movements arriving at or leaving the study area per unit time. Traffic generation is concerned with only one end of each trip.

The distinction is simple and obvious, but is not universally applied. For example, the major U.S. data source on traffic generation as defined here (ITE 1982) is called "Trip Generation".

The two concepts are clearly related. Traffic generation measured at a particular site is the end result of trip decisions made by individuals (purpose, destination, mode etc). In principle, at least, one can be used as an estimator or check on the other. Both involve understanding the relationship between land uses and human activity. "Trip generation provides the linkage between land use and travel" (DoT 1975). The relationship between these two mirrors of land use/activity is used extensively in this paper.

This paper looks at some aspects of current traffic generation practice. In the following section issues related to the measurement of traffic generation are discussed. Section 3 briefly outlines (primarily by example) the methods by which traffic generation rates are currently derived. The final substantive section, which forms the major part of the paper, contains some comments both on immediate possible improvements to current practice and on future longer term research that could be considered.

## 2. TRAFFIC GENERATION MEASUREMENT

### 2.1 USES

Traffic generation is an important tool for both traffic engineers and planners. It is fundamental to:

- (a) traffic design - the design of elements of the traffic system to accommodate traffic generated by activities which the network services, and

- (b) traffic planning - making the necessary arrangements for traffic-generating activities, including site and/or local circulation networks, traffic impact minimisation, access points (the interface between land use activities and the traffic network), provisions for parking, and so on.

Brindle (1984) discusses some simplified case studies which illustrate the use of traffic generation data in traffic impact studies.

## 2.2 UNITS

Ideally traffic generation rates are expressed in terms of vehicle movements to or from the activity per unit of activity per unit time, e.g. vehicle movements per employee per day. It should be stressed that 'vehicle movements' (vm) are one-way, i.e. a vehicle arriving at the site and later departing from it contributes two vehicle movements to that site's daily total.

Some practical difficulties arise in the selection of the appropriate activity unit. Movements occur in most cases as a result of some sort of human activity occurring at a particular location (buying, drinking, meeting, sleeping etc.). The problem is then to measure the amount of such activity at a particular site. Activity, however, is an elusive concept and even when obvious measures exist - for instance, the number of workers occupying an office building - these tend neither to be predictable nor stable. Further, in a traffic planning sense, there are often minimal or no regulations that can be applied to an activity (e.g. there is no control over how many workers occupy a building). Therefore even when a unit of traffic generation expressed in terms of activity is relatively easy to derive it may not be particularly useful in application. As a consequence, while it is recognized that traffic movements arise through human activity, generation rates have been more commonly expressed in terms of land use. There is an obvious close, but by no means exact, correspondence between activity and land use. To provide an example for residential uses, car ownership and number of residents are both known to be better correlated with vehicle trip ends than is the number of dwelling units, but since the latter is most readily obtainable or estimated for development proposals it is usually preferred. In other words in expressing traffic generation rates, activity is proxied by land use measures which tend to be more stable and easier to derived and apply. In doing so, it must be remembered that floorspace, seats, flats and so on do not in themselves generate traffic. The different levels of utilisation of these dimensions of the physical confines of human activity partly explains the variability in the data noted under the next heading.

Typical measures of land use proxying 'activity' are:

- (a) gross floor area (e.g. shopping centre, office)
- (b) nett selling area, or nett customer space (e.g. hotel bars, restaurant)
- (c) number of seats (e.g. theatre)
- (d) number of beds (e.g. hospital)
- (e) number of enrolments (e.g. college)
- (f) number of dwelling units (e.g. block of flats).

## TRAFFIC GENERATION ESTIMATION

### 2.3 VARIABILITY OF DATA

One of the requirements for accurate estimation is that the rate applicable to a chosen measure of land use/activity should itself not be greatly variable. Variability in the data has two aspects:

- (a) Variations within land use/activity types, which suggests that stratification into more precisely defined categories may be fruitful, and
- (b) Variations due to time, which is partly a forecasting problem and subject to some speculation.

#### Variations within Land Use/Activity Types

Just as trip-making characteristics will vary from one person to another, so will the traffic generating characteristics of different offices, shops, dwellings etc.

The basic U.S. traffic generation data source (ITE 1982) briefly discusses this variability for a wide range of land uses. Considering, for example, retail uses (which were found to be more variable in their traffic generation characteristics than office and residential uses), correlations of vehicle movements with numbers of shop employees, gross floor area, and numbers of parking spaces were all poor. In commenting on this, the manual noted that:

'There are many probable reasons for this lack of correlation and range in trip generation rates:

- (a) types of tenants
- (b) method of marketing the centre and tenants' merchandise
- (c) density in the market area
- (d) newness of a centre in a relatively undeveloped market area
- (e) size of centre
- (f) categorisation of centres by type and size'.

The manual goes on to say that, unreliable as it is, no alternative to gross leasable floor area has been found which better describes a centre and calculates trip generation rates. To illustrate the variability in retail traffic generation, the traffic generation of 21 observed centres in the range 4645-9290 m<sup>2</sup> (50,000-100,000 ft<sup>2</sup>) gross leasable area varied from 27 to 174 vehicle movements per 100 m<sup>2</sup> per day around a mean of 85. Traffic generation rates for centres covering the full spectrum of sizes are even more widely dispersed (Figure 1). Any application of selectively-obtained Australian data must recognise that similar variability probably exists here.

These variations within land use categories suggest the possibility of stratifying the category to produce more predictable sub-categories. Travel survey data, for instance, suggests that predictably different generation rates will apply to households having different car ownerships, numbers of resident workers, incomes and so on. This suggests, for instance, that residential traffic generation rates above the metropolitan mean should be applied in areas of obviously higher-

## TRAFFIC GENERATION ESTIMATION

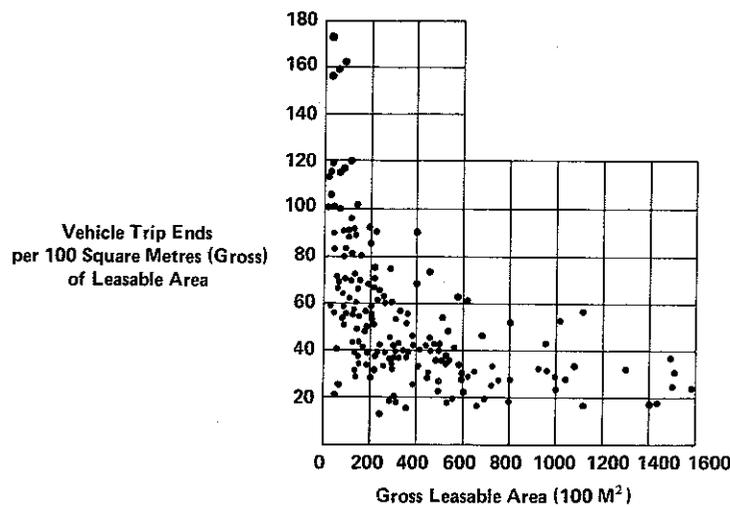


FIGURE 1. : VARIATION IN TRAFFIC GENERATION WITH SHOPPING CENTRE SIZE  
SOURCE : ITE (1982)

than-average income, car ownership etc. Perhaps an immediate use for 'transport study' data in this context would be to provide a basis for factoring mean data up or down to provide more pertinent values in specific areas of known household characteristics.

### Time Variation

There are two time considerations: normal temporal cycles, and changes over time into the future (trends).

Even on the one site, there will of course be variations in traffic generation between hours of the day, days of the week, and months of the year. Traffic design and impact analysis are usually based on normal high conditions, i.e. they anticipate the traffic volumes that would be experienced at the busiest times each week. To allow for special peaks, such as shopping before Christmas, a '30th highest hour' type of criterion would seem to have merit, but available data are not usually good enough to allow this to be done.

The other aspect of data variability is temporal trends in traffic generation rates. An implicit assumption in the use of these rates is that they will remain unchanged over the planning period. This may have important planning implications. For instance, would a suburb planned today for 10 vehicle

movements per dwelling per day look any different if it could be confidently predicted that a rate of 5 or even 15 will apply in the next few years? We shall return to the question of variability in Section 4 of the paper.

### 3. CURRENT ESTIMATES OF GENERATED TRAFFIC.

#### 3.1 DATA SETS

The minimum requirements of a useful data set are:

- (a) Expected daily traffic generation rates for each land use/activity category.
- (b) Expected 'normal high' hourly traffic generation rates.
- (c) Expected traffic generation rates at the peak of the street traffic.
- (d) Indicators of the ranges of uncertainty around these expected values.

There is an abundance of literature on traffic generation characteristics of land use types and specific cases. Most of the reported data are not accompanied by explorations of causal factors; the reader is generally left to extrapolate on the basis of whatever "independent variables" can most easily be specified. The two comprehensive data sets intended for planning applications which are most familiar and accessible to Australian users are the U.S. Institute of Transportation Engineers compilation (ITE 1982), and the reports issued by the Traffic Authority of NSW.

The NSW data was compiled and analysed for a range of land uses using consistent field procedures. Of its type, the NSW data is regarded as being of exceptional quality. While extrapolation of the data to other locations is clouded by the usual uncertainties, the NSW sources are considered to be superior to the U.S. data for use in Australia, both in terms of greater relevance to Australia and the quality of its empirical basis. For most users, the Summary Report provides sufficient detail on the NSW data (Traffic Authority of NSW 1980). There are thirteen detailed reports available for those seeking more information. This data covers a range of employment and retailing activities.

The ITE data, now in its third edition, covers a wider and more detailed range of land use types but relies on reports of data from practitioners. A summary of the scope of the ITE report is contained in the journal 'Traffic Engineering', October, 1976.

Much detailed Australian data is available but rarely published. Apart from the need to extract and compile this data, there is also the need for more survey work to be done. In cases where the proposed development is 'more of the same' (e.g. expansion of a shopping centre, hospital or college), the present traffic characteristics should be surveyed as a basis for prediction. This would be more reliable than extrapolation from other sources.

To convey the flavour of traffic generation practice estimates currently associated with two land use types, residential and retail, are briefly outlined below. A more comprehensive review is contained in Brindle and Barnard (1985).

## TRAFFIC GENERATION ESTIMATION

### 3.2 RESIDENTIAL TRAFFIC GENERATION

Of all traffic generation estimates, those associated with residential land uses, are possibly, in Australia, the most primitively developed. Residential land uses were not covered in the N.S.W. data. U.S. data suggests 10 vm/dwelling/day for single family residential areas, with rates down to 4.3 vm/dwelling/day for high-rise apartments. Unpublished local data supports the general order of these rates, and suggests that at least under present social conditions a multi-car medium-high income suburb could be expected to generate 10-12 vm/dwelling/day. Other apparent ranges are:

- |   |                     |
|---|---------------------|
| (a) 'Typical' one car households in single family dwellings | 8 vm/dwelling/day   |
| (b) Households in town houses and other larger units        | 6-8 vm/dwelling/day |
| (c) Smaller units and flats                                 | 4-6 vm/dwelling/day |

These rates are not high in comparison with other land uses, but taken over the whole urban area residential traffic generation is significant. It is often not appreciated that a square kilometre of residential development can generate as much as 10,000 vehicle movements each day.

### 3.3 RETAIL TRAFFIC GENERATION

The best published Australian data are found in the reports of the Traffic Authority of New South Wales (1980). Linear regression equations were obtained expressing a number of these measures of traffic generation in terms of gross leasable area. Two of these equations are:  
Peak vehicle movements (in + out). Friday:

$$PVT(Fri) = 184 + 0.0658A \quad (1)$$

Vehicle movements (in + out) at the pm peak, Friday:

$$V(P)(Fri) = 265 + 0.0427A \quad (2)$$

where A = gross leasable area (GLA) in m<sup>2</sup>.

Other equations were derived for Thursday and Saturday, for person movements and for parking accumulation. The following should be noted:

- The NSW equations were based on observations at 34 centres in the range 1,774 m<sup>2</sup> GLA to 77,100 m<sup>2</sup> GLA. The applicability of the equations to other centres, especially outside this range, cannot be affirmed on present knowledge.
- Logically, there is no causal relationship between retail floorspace and traffic generation. Floorspace is an estimator of activity only if the developer has correctly estimated his market and the floorspace it can support. The use of models based on observations of existing viable centres

amounts to assuming that the proposed development being analysed will also be a thriving centre. Since a developer will presumably not deliberately propose an under-utilised centre, the use of the descriptive regression equations as forecasting models is acceptable, as long as the user clearly states: 'Assuming that the centre will be viable .....!'

- (c) The prediction intervals around the traffic generation values calculated from empirical equations are likely to be very wide. For example, the 90 per cent 'prediction' interval around the Friday peak traffic generation for the 20 centres forming the basis of the NSW equation is  $\pm 78$  per cent for the centres of  $10,000 \text{ m}^2$  GLA,  $\pm 43$  per cent for centres of  $20,000 \text{ m}^2$  GLA and  $\pm 31$  per cent for centres of  $30,000 \text{ m}^2$  GLA. Users of such equations would be well advised to always quote confidence intervals as well as mean values.
- (d) Note particularly, by reference to Figure 1 that the generation rate per unit area tends to be higher for smaller centres, so does the range of data observed in different centres of that size. The Traffic Authority of NSW advises that 'great care should be taken in applying the models to small centres. With proposed centres in the range  $0-5000 \text{ m}^2$ , subjective and comparative assessments might be more appropriate'. That is, too many factors influence the traffic generation rate at small local centres, and it would be advisable to conduct specific surveys to determine local values under conditions similar to those being evaluated.

Unpublished Melbourne data gathered at five centres in the range  $1,440-8,800 \text{ m}^2$  GLA (four of them clustered at the lower end of the range) about 10 years ago yields the following equations:

$$\text{PVT} = 35.85 - 0.0015A \quad (3)$$

$$\text{PVT (Fri)} = 36.51 - 0.0019A \quad (4)$$

Equation (3) gives a peak value of  $28 \text{ vm}/100 \text{ m}^2 \text{ GLA}/\text{hour}$  for centres of  $5000 \text{ m}^2$  GLA, which is double the value obtained from the N.S.W. data and nearly 3 times the value obtained from the N.S.W. equations. Despite its age and limitations, this data has been quoted as an authority in recent disputes over shopping centre traffic in Melbourne. Such lack of rigour must cast doubt over many traffic impact assessments.

#### 4. IMPROVEMENTS TO CURRENT PRACTICE AND RESEARCH SUGGESTIONS

Information presented in the previous section well illustrates the uncertainty associated with many traffic generation rates and some flaws in current practice. Some suggestions were made in that section on possible improvements to current practice, such as the use of confidence intervals in addition to mean values. This section more thoroughly examines possible short term and longer term improvements. The short term improvements may be described as extensions to current practice since no basic change in direction is implied. The suggested longer term improvement, however, demands an altogether different approach to estimating traffic generation.

## TRAFFIC GENERATION ESTIMATION

### 4.1 EXTENSIONS

It should first of all be stated that the NSW data represents an excellent base from which to extend current practice. However, a number of unresolved questions remain and data generalization is required.

#### 4.1.1 Extensions in Methodological Approach

The traffic generation rates from the NSW data and similar studies are usually estimated using only one or two independent variables. There is good reason for this; the basis of traffic planning for development applications must be limited to the set of information available to authorities. A conundrum is immediately apparent; although it is known that traffic generation rates are a function of many factors only a few of these (and perhaps not the best few) can be taken into account.

A by no means comprehensive list of factors affecting traffic generation is (see, in part, Foley 1981):

- a) the types of activity participation facilitated by the development
- b) the likely potential for activity participation facilitated by the development
- c) the intensity with which the development is used for each activity
- d) overall accessibility to the centre
- e) modal competition, especially the supply of public transport services
- f) the supply of parking spaces
- g) proximity and size of similar centres
- h) socio-economic characteristics of people within the catchment area
- i) aesthetics and amenity of the centre
- j) the extent of centre promotion.

Of these only (a) and (b) are generally considered and it is unlikely that (c), (i) and (j) can ever be used. Note, however, that (c) is undeniably crucial in determining traffic generation. For instance, 'factory' employees/100 sq.m (i.e., intensity of use) can range from 0.11 to 3.61 with concomitant impacts on traffic generation (Watters 1981).

#### Inclusion of extra variables

The remaining factors listed (d) - (h) can be determined given data on location and planning parameters which is of course readily available to authorities at the time of a development application. Nevertheless these factors

have generally been excluded from consideration under present traffic generation practice. The value of including these extra factors when estimating traffic generation depends on the degree of accuracy required, how much these factors increase the accuracy of estimates and the cost of collecting this extra information.

Information collection costs will be proportional to the degree of complexity associated with indices developed to represent accessibility, modal competition, centre competition and socio-economic effects. Monetary costs could be minimized by constructing indices which only use data that is now available either through transport or general statistical authorities. Time would be needed, however, for compiling data from these sources, and time is often at a premium when a development application is being considered.

The question of the increased accuracy of estimates awaits research. Nevertheless two preliminary comments can be made. Firstly, trip generation (the mirror of traffic generation) typically is successfully modelled solely as a function of socio-economic variables. It is to be expected that these variables would also be important determinants of traffic generation. Secondly, many authors have noted the dependency of mode split on development location, especially for 'office' and 'factory' land uses, and this therefore would seem to be a fruitful area for research. Clearly, person trips generated by CBD office developments will exhibit very different modal splits to those in outlying suburbs.

#### Separate mode split estimates

The perceived importance of modal competition as a determinant of traffic generation suggests another improvement. It may be better to model generation of person trips, rather than vehicular trips, and then apply mode split estimates to obtain the latter. If modal competition is important the relationship between person trips and land use will be more stable (than for vehicular trips) and therefore easier to predict. Estimates of modal splits might be obtained simply by using current observed (e.g. from transport study data) mode use for the activity in the development area. Alternatively a separate mode split model may be developed.

#### Segmentation by land use type

Tighter estimates of traffic generation may also be obtained by finer division of land use types. This is one method that can be applied to ameliorate the effect of different use intensities alluded to earlier. In particular it is to be expected that the use intensities associated with, for example, motor vehicle factories will be more closely correlated than for factories in general. European results indicate that with current segmentation schemes the variation of traffic generation estimates within a land use type in fact often exceeds the variation between land use types (Tudge 1981).

Finer segmentation schemes, however, pose another problem; that of aggregation. If finer segmentation is adopted many developments will contain more than one land use type. Examples include shopping centre with fast food restaurant, shopping centre with office development and office development with apartments. The total traffic generated ( $G_{\dagger}$ ) by such developments will not in

## TRAFFIC GENERATION ESTIMATION

general be a simple summation of the traffic effects of each separate land use type (i.e.  $G_t = G_{l_1} + G_{l_2}$ ), but some more complex combination (i.e.  $G_t = F(G_{l_1}, G_{l_2})$ ). No research has been done into the form of  $F(\cdot)$ , though aggregation is sometimes a problem even with current coarsely defined land use types. For instance many 'activity' centres built within the last 10-15 years contain shopping areas, offices, leisure facilities, entertainment areas, and restaurants - all land use types defined separately in the NSW data set.

### Distribution of generated traffic

The factors relevant to traffic generation listed early in this section refer only to amount and not to how a given level of generated traffic is distributed. Two dimensions of distribution are important for traffic generation studies - over time and road space close to the generator.

With current practice the variation of generated traffic over time is usually recognized by providing separate estimates for peak periods. It is not clear, however, as to what should constitute the peak period. Many studies record the peak as the period with maximum vm/hour, but is this more important than the peak hour for traffic on the surrounding street system? Seasonal variations, though rarely taken into account, are also important for some land uses and locations. Resort areas, for instance, can experience traffic levels during holiday periods that are several factors higher than for the rest of the year.

Also generally not considered in any sort of precise quantitative manner is the local network distribution of generated traffic, yet this can be crucial in assessing the adequacy of existing roads in coping with increased traffic. Loadings to different entrances can create quite different traffic demands on nearby roads and intersections. Ideally then for major developments a local area traffic model such as LATM or SATURN should be interfaced with traffic generation estimates.

### 4.1.2 Data Generalization

One of the most pressing needs in traffic generation research in Australia is to extend the geographical coverage of current data bases. Constant reference has been made throughout this paper to the relative excellence of the NSW data, but can the estimates obtained in Sydney accurately predict traffic generated by developments in other capital cities and in country areas?

The best way to test the spatial transferability of the NSW models and develop traffic generation estimates in other areas would be to conduct a number of new surveys along similar lines to the NSW surveys. In important matters of detail, guidelines adopted for these new surveys should be the same as those used in NSW (e.g. definition of peak periods, measurement of GLA). At least some of these new surveys should be more comprehensive than the NSW data to allow investigation of omitted factors listed in Section 4.1.1.

Conducting new surveys, however, is an expensive option and therefore unlikely to be considered on a wide scale. The expense is particularly high when information on person trips and not just vehicular trips is required. Many of the NSW models rely on the former. Less expensive options need to be explored.

One option is to use data on traffic generation that has already been collected in the many 'once off' studies conducted for traffic authorities, local government bodies and research institutions. A problem with using this data in a comprehensive comparability research study is that it is not uniform with respect to factors included or methods of collection. Perhaps in the future some sort of standard format should be devised for use in these 'once off' studies.

Alternatively, developers themselves could be surveyed and asked to supply information on characteristics of centres and degrees of use. The reliability of such data must be questioned, however. In particular developers have a vested interest in understating vehicular usage of centres. Difficulty with this approach was the principal stimulant initiating the NSW data collection effort.

A third option, that will be explored extensively in the remainder of this section, is to use data collected for purposes other than studying traffic generation, but which nonetheless may prove useful in this regard. Preferably this data should be available on a relatively comparable basis in a number of spatial areas and not require much modification to enable application in a traffic generation context. Such data is available in the origin-destination (O-D) travel surveys conducted by transport and highways authorities in major centres during the past 20 years. In the remainder of this section the suitability of using O-D travel survey data for studying traffic generation will be explored for two land use types: residential and shopping. Familiarity with O-D travel survey data is assumed.

#### Residential traffic generation

Residential traffic generation rates are used for road and traffic planning in new housing development areas. In the U.S. it has been common practice to use a common residential traffic generation rate for all new developments with adjustments only being made to reflect differences in the mix of housing types. No account is taken of differences in development densities or socio-economic characteristics of prospective residents. In Australia these rates have been estimated by a number of methods ranging from the judgement of local government engineers to erroneous application of area wide household trip rates to site specific counts in comparable residential developments.

O-D travel surveys would seem to represent a particularly fruitful source of data for calculating residential traffic generation rates. At the lowest level of disaggregation personal travel data collected in these surveys pertains to trips ( $T$ ) undertaken by individual  $k$  by mode  $m$  for purpose  $p$  at time  $t$  from area  $i$  to area  $j$  (i.e.  $T_{kmpjtij}$ ). This trip data can easily be summed across individuals in a household, time periods and origin-destination areas. By restricting  $p$  to trips to and from the home (i.e. home-based trips) and  $m$  to vehicular modes, an estimate of the daily total number of home-based vehicular trips made by a household can be obtained ( $T_{HBV}$ ).

Two adjustments need to be made to  $T_{HBV}$  for it to represent an accurate measure of residential traffic generation. Firstly,  $T_{HBV}$  has to be inflated to account for trip under-reporting in home interview travel surveys. Secondly, service visitor trips need to be added.

## TRAFFIC GENERATION ESTIMATION

It is well known that individuals tend to under-report their travel in home interview surveys. The extent of under-reporting can be determined from two major sources. One involves comparing results from a home interview travel survey with a higher quality diary survey where both surveys are similar in terms of sampling frames and temporal and spatial coverage (see Barnard 1984a). The other is to check survey results with cordon counts (see Barnard 1985). This has been normally done in Australian O-D travel surveys. Both these sources indicate typical under-reporting rates of 20-30%.

Estimates of the number of visitor and service trips can also be obtained from O-D survey data. Visitor trips are represented in the home interview travel survey data as trips made to homes in zone z by individuals living in other zones. Similarly service trip information to homes in zone z can be obtained from the commercial vehicle survey data collected at the same time.

Table I summarizes the steps required to obtain residential traffic generation rates from O-D travel survey data. Reid (1982) has already conducted a study along these lines for an 18,000-acre development in Southern California.

Table I  
Calculation of Residential Traffic Generation Rates  
From O-D travel survey data

Item	Data source
home-based vehicular-trips per household per day	home interview travel survey
add visitor trips survey	home interview travel
adjust for trip under-reporting comparisons	screenline count
add service trips	commercial vehicle survey

The data source is particularly rich. Not only can these generation rates be developed separately by dwelling type but can also be related to household characteristics. The later could involve applying non-metric regression-like statistical techniques in a similar manner to the NSW modelling work. The data source is also sufficiently large to allow selection of households only from those areas with similar characteristics to areas of new residential development.

### Shopping traffic generation

Application of O-D travel survey data to shopping traffic generation is more complex than for residential traffic generation primarily because although these surveys allow effective individual identification of homes, apartments, etc. they rarely allow identification of individual shopping centres. It is still possible, however, as will be shown below, to build models similar to those developed in NSW.

This time the base observation is taken to be vehicular trips between zones  $i$  and  $j$  for purpose  $p$  (obtained simply by summing over individuals, time periods, etc). Further summing over origin zones  $i$  enables an estimate of the total number of trips to zone  $j$  to be obtained. We assume this estimate has been adjusted to account for trip under-reporting and traffic by commercial vehicles. As  $j$  becomes very small its size will be equivalent to individual centre sites so the following relationship can be written:

$$\sum_i \sum_j T_{ijp} = G_{d\ell} = \beta_1 + \beta_2 A_{d\ell} + \epsilon_{d\ell} \quad (5)$$

where  $G_{d\ell}$  = generated traffic (vms) at development site  $d$  of land use type  
 $A_{d\ell}$  = area of development site  
 $\epsilon_{d\ell}$  = an additive error term  
 $\beta_1, \beta_2$  = parameters to be estimated.

Mostly only large shopping centres are given a separate zone in transport studies. For example in Adelaide the West Lakes, Noarlunga, Tea Tree Gully and Kilkenny shopping centres were all reserved special zones in the 1977 Metropolitan Adelaide Data Base Study (Pak Poy and Assoc. 1978). This is despite the availability of techniques such as geocoding which theoretically allow even small centres to be spatially identified. These techniques, however, have yet to be applied in Australian transport data collection and typically a zone  $j$  will encompass many developments ( $d_{j1\ell_1}, d_{j2\ell_1}, \dots, d_{jq\ell_1}, d_{j1\ell_2}, d_{jg\ell_2}, \dots, d_{js\ell_t}$ )

where  $d_{jq\ell_i}$  refers to the  $q$ th development site of land use type  $\ell_i$  in the  $j$ th zone.

Concentrating on a particular land use category  $\ell_i$ , dropping the subscript for simplicity, and assuming  $p = \ell$  (or an aggregation of  $p = \ell$ ):

$$\begin{aligned} 2(\sum T_{ijp}) &= G_{d_{j1}\ell} + G_{d_{j2}\ell} + \dots + G_{d_{jn}\ell} = (\beta_1 + \beta_2 A_{d_{j1}\ell} + \epsilon_{d_{j1}\ell}) \\ &\quad + (\beta_1 + \beta_2 A_{d_{j2}\ell} + \epsilon_{d_{j2}\ell}) + \dots + (\beta_1 + \beta_2 A_{d_{jn}\ell} + \epsilon_{d_{jn}\ell}) \end{aligned} \quad (6)$$

OR

$$\frac{2}{n} \sum T_{ijp} = \frac{1}{n} (G_{d_{j1}\ell} + \dots + G_{d_{jn}\ell}) = \beta_1 + \beta_2 \bar{A}_{j\ell} + \bar{\epsilon}_{j\ell} \quad (7)$$

where:  $\bar{A}_{j\ell} = \frac{1}{n} \sum_{c=1}^n A_{d_{jc}\ell}$

$$\bar{\epsilon}_{j\ell} = \frac{1}{n} \sum_{c=1}^n \epsilon_{d_{jc}\ell}$$

The problem then reduces to developing a method for deriving unbiased and efficient estimates of the  $\beta$ s. These estimates would be directly comparable with

## TRAFFIC GENERATION ESTIMATION

those derived from conventional traffic generation sources.

The properties of grouped data such as that hypothesized in equation (7) are well known (see, for example, Kmenta 1971). Ordinary least squares will yield unbiased but inefficient estimates of the  $\beta$  s. This is because the error terms  $\epsilon_{jl}$  will be heteroskedastic i.e. not exhibit constant variance over all  $j$ s (as a result of the differing number of centres in each zone). Furthermore, conventionally calculated variances for the  $\beta$  s will be biased, leading perhaps to incorrect inferences about the population coefficients. Fortunately revised expressions for calculating the  $\beta$  s and associated variances can be derived which will yield efficient and unbiased estimates and that are easily evaluated.

Of course the  $\beta$ s estimated from equation (7) will nearly always be less efficient than those derived using data on individual centres, even if the  $\epsilon_{jl}$  are homoskedastic (due to there being an equal number of centres in each zone). There will be only no loss in efficiency if there is no variation in centre size areas ( $A_{jq}$ ) within a zone. In general the loss in efficiency will increase as the variation in centre sizes within a zone increases relative to the variation in average centre sizes between zones.

In practice there would appear to be no insurmountable problems in estimating equation (7). The numerator of the dependent variable, as has been demonstrated, can be obtained from O-D travel survey data. Normally data on shopping floor space in each zone is also obtained in these studies. To the author's knowledge no information is retained on the number of centres in each zone but such data could readily be obtained from local authorities. It would also probably be necessary to assume that vehicular travel to individual stores that are not part of a centre (e.g. corner stores) is empirically insignificant.

In summary, O-D travel surveys would seem to offer an innovative and exciting data source for examining traffic generation associated with at least some land uses. This data source directly meets three of the requirements for useful traffic generation estimates listed in Section 3.1. The value of this data, however, in practical studies of traffic generation remains to be demonstrated.

### 4.2 NEW DIRECTIONS

The preceding suggestions for improving traffic generation estimates, however sophisticated they may at first appear, in reality represent only relatively minor extensions to existing methodology. It is argued in this section that currently used methods are essentially flawed in approach. Some suggestions are also made on a new approach for estimating traffic generation.

Throughout this paper minor criticisms have been levelled at existing methodology for calculating traffic generation rates. The principal criticism has not been mentioned. It is that current methods lack any sort of theoretical foundation. Essentially the relationships derived between generated traffic and land use variables are of a type described by statisticians as functional rather than causal. Functional relationships can be distinguished by the characteristic that the only justification for their existence is that they happen to fit well to existing data. This is no cause for concern in situations where the underlying causal variables move together with the functional variables as apparently is assumed to be the case in most short term planning analyses. Problems may arise, however,

when data representative of one point in time is then used to predict change. When major changes are introduced into a previously stable system, such as construction of a major shopping centre in an area previously serviced only by small shops, it is most unlikely that causal and functional relationships will vary similarly. In situations such as these little confidence can be placed on predictions derived from models comprising functional relationships estimated on cross-sectional data. It is also improbable that estimates derived from functional relationships will be amenable to spatial transference.

To develop a more behaviourally based methodology for estimating traffic generation it must first be realized that observed patterns of traffic are an outcome of many individual decisions regarding whether or not to travel, where to go, what mode to choose etc. A causal model, therefore, needs to reflect these individual decisions, distilling the main influencing factors. This implies that the unit of analysis should not be aggregate numbers of trips observed at activity centres (i.e. the destination end of trips made), but rather individuals choosing to make or not to make trips. Important characteristics of trips to be modelled are origin and destination points, mode of travel, timing and perhaps route.

An integrated theory of individual behaviour in these decision circumstances has been recently devised in the economics literature. (e.g. Domencich and McFadden 1975). It assumes each individual acts as if to maximize utility but only a portion of the individually obtained utility concomitant with a particular choice alternative is observable by an outside analyst. This means that although individuals follow a deterministic decision process their choices must be modelled probabilistically. Depending on assumptions made about the unobservable utilities a number of estimatable model forms can be derived. One of these is the well known multi-nominal logit (MNL) model.

In complicated decision processes such as those requiring study in traffic generation measurement any number of variants of sets of estimatable models conforming to this theory can be constructed. It is largely left to the analyst to choose a particular modelling structure having regard to plausibility, data and computational requirements. An example of a structure having convenient properties that may be applied to measuring traffic generation is outlined below. Written in estimation order the set of models are:

$$P_{k \in M_{kat}} = \frac{\exp \lambda^m X_{ktam}}{\sum_{m' \in M} \exp \lambda^m X_{ktam'}} \quad (8)$$

$$P_{ka \in A_{kt}} = \frac{\exp(\lambda^a X_{kta} + \theta^a C_{kta})}{\sum_{a' \in A} \exp(\lambda^{a'} X_{kta'} + \theta^{a'} C_{kta'})} \quad (9)$$

## TRAFFIC GENERATION ESTIMATION

$$P_{kt \in T} = \frac{\exp \lambda^m X_{ktam}}{\sum_{t' \in T} \exp(\lambda^t X_{kta'} + \theta^t C_{kta'})} \quad (10)$$

where  $C$  denotes a composite utility term defined as the natural logarithm of the denominator of the previously specified model. For example:

$$C_{kta} = \ln \sum \exp \lambda^m X_{ktam}$$

the subscripts  $m$ ,  $a$  and  $t$  refer to the chosen mode, activity centre and trip alternatives.\* The subscripts  $m'$ , etc. are similarly defined for non chosen alternatives.

$A_{kt}$  refer to the set of activity centres from which individual  $k$  chooses given that he decides to make a trip.

$M_{kat}$  is similarly defined for methods of travel.

$X_{ktam}$ , etc. are vectors of variables defined for individual  $k$  and alternative  $t$  a  $m$ , etc.

$\lambda^m$ ,  $\lambda^a$  and  $\lambda^t$  are row vectors of coefficients

$\theta^a$  and  $\theta^t$  = coefficients.

Examples of variables that may be included in these models are for  $X_{ktam}$  modal times and costs to activity centre  $a$  for individual  $k$ ; for  $X_{kta}$ , GLA and other centre attractiveness variables; for  $X_{kt}$ , socio economic variables - the travel influences on the decision of whether or not to make a trip being represented in  $C_{kt}$ .

The data demands of this structure would not appear overwhelming. A major report prepared for the U.S. Transportation Research Board (Tye et al 1982) notes that single models of the type shown in equations (8) - (10) may be successfully estimated with less than 100 observations. Furthermore the data needs can primarily be serviced by interviewing individuals within existing centres, in the catchment area of the proposed centre, rather than the more traditional interviewing location of the individual's home. This technique, known generally as 'choice-based' sampling, greatly reduces survey costs. Within centre interviews could be kept short by seeking information only on the residential locality of the individual, method of travel used to access the centre and a list of socio-economic descriptors. Modal travel times and costs are obtainable from network sources and centre attractiveness variables from a physical inventory of the sites. This in centre data would need to be supplemented by home interviews to include those not using a centre during the time period under consideration and data indicating overall usage levels for each centre. If only vehicular traffic generation was being considered the model of equation (8) could be omitted and

---

\* In the simplest terms  $t$  can be thought of as a binary decision - either a trip is made or is not made. More complicated model structures would include  $t$  as a polychotomous variable to account for multi-trip tours (see Barnard 1984b).

analysis restricted to estimating equations similar to (9) and (10).

It is apparent that models (8) - (10) address only one of a number of sources of traffic generation. The models presented are for personal travel stimulated by the activity participation facilitated at the centre (we term this 'customer traffic'). Other sources of generated traffic are trips by staff employed at the centre, freight deliveries and commercial support traffic. The authors know of no comprehensive data enabling analysis of the relative contributions from these sources to total generated traffic. It is probable (if only because of the past neglect of staff, freight and commercial traffic) that customer traffic dominates for many land use categories. If this is so the elaborate model structure of equations (8) - (10) may be wasted for these other sources and simple techniques may be used, perhaps not dissimilar to those currently applied. On the other hand if significant amounts of non customer traffic are associated with a particular land use category it may be worthwhile developing separate model structures. This has the disadvantage of imposing substantial extra data demands. It should equally be recognized, however, that different mixes of traffic generated from customer, staff, freight and commercial sources implies different planning decisions. The parking demands of the four groups, for example, are obviously nonuniform.

Some of the advantages of the proposed approach to estimating traffic generation compared to existing methods are listed below:

- i) the models have an inbuilt theoretical base (i.e. maximization of utility).
- ii) since individual behaviour is directly modelled the probability of successful spatial transference of estimated forms and parameters is increased.
- iii) modal competition is explicitly recognized in the modelling structure.
- iv) total traffic generated at new centre is recognized in part to be due to extra trips and in part to redirection of existing trips that otherwise would have been made to other centres.
- v) a mechanism exists for the inclusion of pertinent socio economic influences.

Of these points, probably (iv) is the most interesting and has until recently received least attention (the other points have been extensively discussed elsewhere in this paper). The conventional method of calculating the traffic impacts of a new centre simply adds estimated traffic generated by that centre to existing network flows. This method may significantly overstate the actual impact since a portion of the 'generated' traffic would be using the network in any case to access other centres. Limited U.S. survey results suggest that less than 1/3 of 'generated' traffic by new shopping centres consists of new trips, the remainder being diversions of trips already on the road network (Slade and Gorove 1981). Because the proposed modelling structure separates the destination decision from the decision of whether or not to make a trip, these effects can easily be calculated on an individual centre by centre basis.

This separation would also seem important from an equity viewpoint. It has become reasonably common practice in recent years to charge developers for alterations to the network necessary to accommodate the expected generated

## TRAFFIC GENERATION ESTIMATION

traffic. This does not recognize that the new development may be diverting traffic from older congested activity centres. That is, the social costs of a new development are recognized and charges levied accordingly, but possible social benefits are ignored.

In summary, individual choice modelling would seem to offer a much more comprehensive and attractive framework for estimating traffic generation than conventional approaches. It imposes, however, increased data and computational burdens.\* Further the performance of individual choice models in practice in this context is hitherto an unknown.

### 5. CONCLUSION

This paper has examined current traffic generation practice and identified some areas of deficiency. Possible ways of correcting these deficiencies were outlined. Most of these suggestions represented extensions to the currently used approaches for estimating traffic generation. It was recognized, however, that this approach is basically flawed. A new framework for estimation, based on individual choice theory, was developed. This new approach incorporated some attractive features, but increased computational and data requirements. Because of this and its unproven efficacy in practice, initial testing should be restricted to relatively large developments. Meanwhile continued research effort should be devoted to extending existing methods, particularly for residential developments, along the lines indicated.

---

\* Nevertheless software packages to estimate individual choice models are now available on micro computers and the continuing development of this technology will reduce computational costs over time.

## REFERENCES

- BARNARD, P.O. (1984a) Use of an activity diary survey to examine trip under-reporting in a home interview travel survey, Ninth Australian Transport Research Forum, Adelaide, May.
- (1984b) Towards a model of complex urban food shopping patterns. 12th PTRC Summer Annual Meeting, Transportation Planning Methods Seminar, Brighton, England, July.
- (1985) An examination of trip under-reporting in Australian transport study home interview surveys, Australian Road Research Journal (forthcoming).
- BRINDLE, R.E. (1984). Traffic impact analysis in Traffic Engineering Practice, (edited by K.W. Ogden and D.W. Bennett), Department of Civil Engineering, Monash University, (3rd ed.).
- BRINDLE, R.E. and P.O. BARNARD (1984) A review of Australian traffic generation practice with suggestions on future research. Australian Road Research Board, AIR 408-1, Vermont Victoria.
- DEPARTMENT OF TRANSPORTATION (U.S.)(1975) Trip Generation Analysis. Washington.
- DOMENCICH, T.A. and D. McFADDEN (1975). Urban Travel Demand. North Holland/American Elsevier Publishing Company, New York.
- FOLEY, S.P. (1981). Traffic generated by shopping centres in Adelaide, Australian Road Research Journal, Vol.11, No.2, June, pp. 3-12.
- INSTITUTE OF TRANSPORTATION ENGINEERS (1982). Trip Generation, (3rd Ed.). Washington.
- KMENTA, J. (1976). Elements of Econometrics. Macmillan Publishing Co. New York.
- PAK POY AND ASSOCIATES (1978). Metropolitan Adelaide Data Base Study. Phases 1-5, reports prepared for the South Australian Department of Transport and Highways Department, Adelaide.
- REID, F. (1982). Critique of the ITE trip generation rates and an alternative basis for estimating new area traffic, Transportation Research Record, No. 874, pp. 1-5.
- SLADE, L.J. and F.E. GOROVE (1981). Reductions in estimates of traffic impacts of regional shopping centres. ITE Journal, January, pp 16-18.
- TRAFFIC AUTHORITY OF NEW SOUTH WALES, (1980). Land Use Traffic Generation Data and Analysis, (various reports). Sydney.
- TUDGE, R. (1981). Shopping centre trip generation in Europe, Presented at the ARRB/DMR Workshop on Traffic Generation, Sydney, November 1981.
- TYE, W.B., SHERMAN, L., KINNUGAN, M., NELSON, D. and TARDIFF, T. (1982). Application of Disaggregate Travel Demand Models, NCHRP Report 253, Transportation Research Board, Washington.
- WATTERS, B. (1981). Quality of data and problems of forecasting, presented at the ARRB/DMR Workshop on Traffic Generation, Sydney, November.