23rd Australasian Transport Research Forum
Perth, Western Australia. 29 September – 1 October 1999

Estimating Small Area Commercial Transport Movements

Tim Raimond, John Peachman and Gillian Akers
Department of Transport, New South Wales

Abstract

The Transport Data Centre (TDC) of the NSW Department of Transport has developed a procedure for the estimation of road freight and commercial vehicle movements in the Greater Sydney Metropolitan Region. The procedure brings together a range of existing data sources, in particular regional input-output tables, to produce base year estimates of commercial travel. The key advantages of this procedure are that it offers small area (travel zone) travel estimates and a forecasting capability.

The procedure uses regional input-output tables to generate trip end information via a series of transformations using ABS concordances, vehicle loading information, and detailed industry size and location data. The trip ends are combined with flow estimates from traffic counts and a prior matrix based on the 1991 Commercial Vehicle Survey to produce base year trip tables of commercial traffic at the travel zone level.

Two slightly different procedures are used for estimating heavy and light vehicle movements, a difference necessitated by the difficulty in identifying light commercial vehicles in traffic counts, and in estimating service vehicle trip ends from input-output tables.

Despite some acknowledged limitations of the procedure, it appears likely to offer a pragmatic solution to the requirement to produce base and forecast year estimates of commercial travel at the small urban area level. Subject to the availability of a similar range of data inputs, there is no reason why a procedure such as this could not be implemented in other cities, Australian and elsewhere.

Contact Author

Tim Raimond
Transport Data Centre
NSW Department of Transport
GPO Box 1620
Sydney NSW 2001

Phone: +61 2 9268 2243 Fax: +61 2 9268 2853 e-mail: raimondt@transport.nsw.gov.au

Introduction

Estimation of commercial travel in urban areas has long been a vexed issue for those involved in transport data collection. Collection of commercial travel data is often very expensive, very difficult, and yields results less reliable than those available for personal travel. Nevertheless, effective transport planning and policy development requires a complete picture of travel, one which explicitly includes commercial travel in addition to personal travel.

The Transport Data Centre (TDC) of the NSW Department of Transport has completed the initial phase of a project to produce small area trip tables of commercial travel movements for the Greater Sydney Metropolitan Region (GSMR), and to produce forecasts of these movements. This will enable supplementation of existing forecasts of personal travel available at this same small area level.

The Commercial Transport Study (CTS) recognises the complexity of any attempt to forecast commercial travel at the small area level by adopting a long term, staged approach. In order to achieve the ambitious aims of the CTS, it has been necessary to move beyond traditional sample survey approaches. This paper describes the initial stage of the CTS, which involved the development of a matrix estimation procedure (CTS-EP) combining a range of existing data sources to produce small area travel estimates. One of the keys to the approach is its use of several forecastable inputs, including regional input-output tables and small area employment data, to ensure that the resultant trip table is forecastable.

This first stage of the CTS has involved the development and implementation of the broad CTS-EP methodology. Later stages will concentrate on refining the methodology and extending the range and quality of input data available for the process.

Background

The TDC collects transport data and provides analysis and forecasting services to the NSW Government and private sector clients. Using its Household Travel Survey and Strategic Travel Model, the TDC is able to produce current and forecast year small area trip tables of personal travel for the Sydney Region. However, a more complete picture of travel in Sydney should include estimates of current and forecast commercial travel at the small area level. This is the requirement which is driving the CTS.

Definitions

Commercial travel: TDC defines commercial travel as movements by light or heavy goods vehicles which relate to the transport of freight (either laden or unladen) or the supply of household or business services.

Trip tables: These are tables of the number of vehicle trips between one region and another.

Small area: The "small areas" referred to in the CTS are TDC Travel Zones, of which there are over 1,000 in the GSMR. Travel Zones represent areas of relatively homogenous land use that fit between the Australian Bureau of Statistics (ABS) Census Collector District and Statistical Local Area (SLA) in size.

Choosing an Appropriate Methodology

The aim of TDC's CTS Development project was to develop a procedure to produce base and forecast year commercial vehicle trip tables by vehicle type and time of day for the GSMR. TDC began this process aware of the limitations of its large 1991/92 driver-based Commercial Vehicle Survey (CVS) for producing such outputs. The three main issues precluding use of the CVS were:

- sparse coverage at a zonal level
- defining an appropriate population for expanding the sample survey (the sampling frame was of vehicles registered in the GSMR, not of vehicles operating in the GSMR)
- lack of accurate data to enable forecasting (limited commodity information, and no information on the industry type at the origin or destination).

TDC engaged FDF Management to evaluate available methodologies and recommend an approach to suit TDC's particular requirements. The advantages and disadvantages of the many approaches considered are discussed in detail in Raimond (1997). However, the key conclusions were:

- sampling alone cannot produce reliable travel zone level estimates
- the need to forecast commercial travel necessitates reliable collection of commodity and/or industry data.

On the basis of these findings, FDF Management (1996) recommended the development of a "commodity-based" approach to estimating commercial travel, using commodity flow information as a basis for deriving travel. The key advantage of this approach is that it provides an economic basis for producing forecasts of travel.

Before committing resources to the recommended approach, TDC first sought to ensure that the recommendations were in line with the latest developments overseas. Professor Garland Chow, from the University of British Columbia, was commissioned to undertake a review of the North American experience of commercial travel data collection and forecasting in the urban context. This review found that "commodity-based" approaches, similar in many ways to that recommended by FDF, were beginning to be recommended for several new studies in the US, particularly because of the forecastability they offered. Several projects of a pilot nature have been completed (eg New York (List et al 1995) and Puget Sound (Barton-Aschman Associates 1996)) or are underway (eg Portland (Metro and Port of Portland 1996)) using somewhat similar methodologies to that proposed by FDF, though none with the ambitious aims of the CTS to produce travel zone level estimates.

TDC then engaged a consortium led by DJA Maunsell to develop and test a CTS procedure largely based on the recommendations by FDF Management (1996). This project was overseen by an independent expert, Professor Marcus Wigan, to ensure that TDC obtained the best results possible from what was an experimental development project.

The approach which was developed is detailed in the next section. To successfully apply the spirit of FDF's recommendations while being pragmatic about the resources available for such a task, DJA Maunsell (1999) developed a unique methodology which may have some applicability to other jurisdictions within Australia and elsewhere.

To alleviate the significant primary data collection costs associated with obtaining commodity flow information directly from businesses, the CTS uses regional input-output (I/O) tables of flows between industry pairs to derive commodity flows, and then commercial travel. A disadvantage is that this approach does rely on multiple layers of assumptions to convert commodity flows at a regional level to vehicle trips at a zonal level.

To help mitigate this disadvantage, a matrix estimation procedure is used which combines data derived from I/O tables with other known information, such as classified traffic counts, and origin-destination information from TDC's 1991/92 CVS. The matrix estimation procedure can then produce a trip table using all available information, and can take into account the level of confidence the analyst has in each piece of data when reconciling data from different sources.

The Trip Table (Matrix) Estimation Procedure (CTS-EP)

The CTS-EP is actually a combination of two processes, one for heavy goods vehicles (HGVs) (rigid or articulated trucks) and one for light goods vehicles (LGVs). This is necessary because HGVs are almost solely used for the transport of commodities, while a large proportion of LGVs are used for undertaking service travel. In addition, classified traffic count data is not available for LGVs, limiting the data available for matrix estimation.

This paper focuses on the HGV estimation procedure as TDC is currently enhancing its LGV estimation procedure. However, the LGV estimation procedure is discussed briefly at the end of this section.

The CTS-EP for HGVs uses four key data inputs: trips ends, screenline counts, a loaded road network, and a prior matrix. These are combined in a matrix estimation procedure to produce an optimal table of commercial travel movements, by vehicle type and time of day, between travel zones in the Greater Sydney Metropolitan Region. The structure of the CTS-EP procedure for HGVs is outlined in Figure 1.

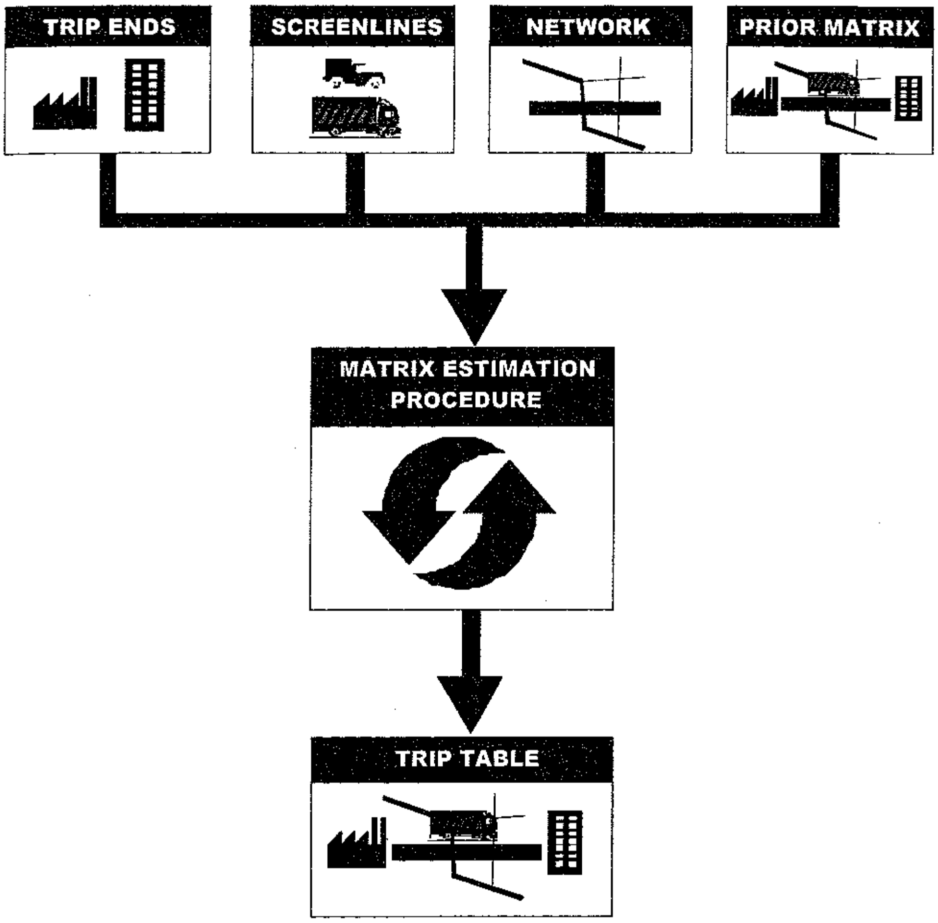

Figure 1 Outline of the CTS Estimation Procedure for HGVs

Input 1: Trip Ends

Trip Ends are the number of trips generated from and attracted to a location, in this case, to and from a travel zone. Traditionally, these would be generated by trip data collected from a sample of truck drivers. The truck driver would record the locations where goods were picked up and delivered. A limitation of this approach is that a sample of truck drivers is never large enough to include trips to or from all travel zone combinations in the study area. Thus, when the data is weighted, those zones with no trips between them remain zero. The sample size required to make reliable travel zone estimates using such survey data is prohibitive.

Another issue is that using a survey to estimate the number of trips to and from travel zones offers no explanatory information about why those trips occurred, and thus offers very limited scope for producing forecasts of travel.

To address these issues, the trip ends used in the CTS have been derived from commodity flow information provided by regional input-output tables. The procedure used is outlined in Figure 2.

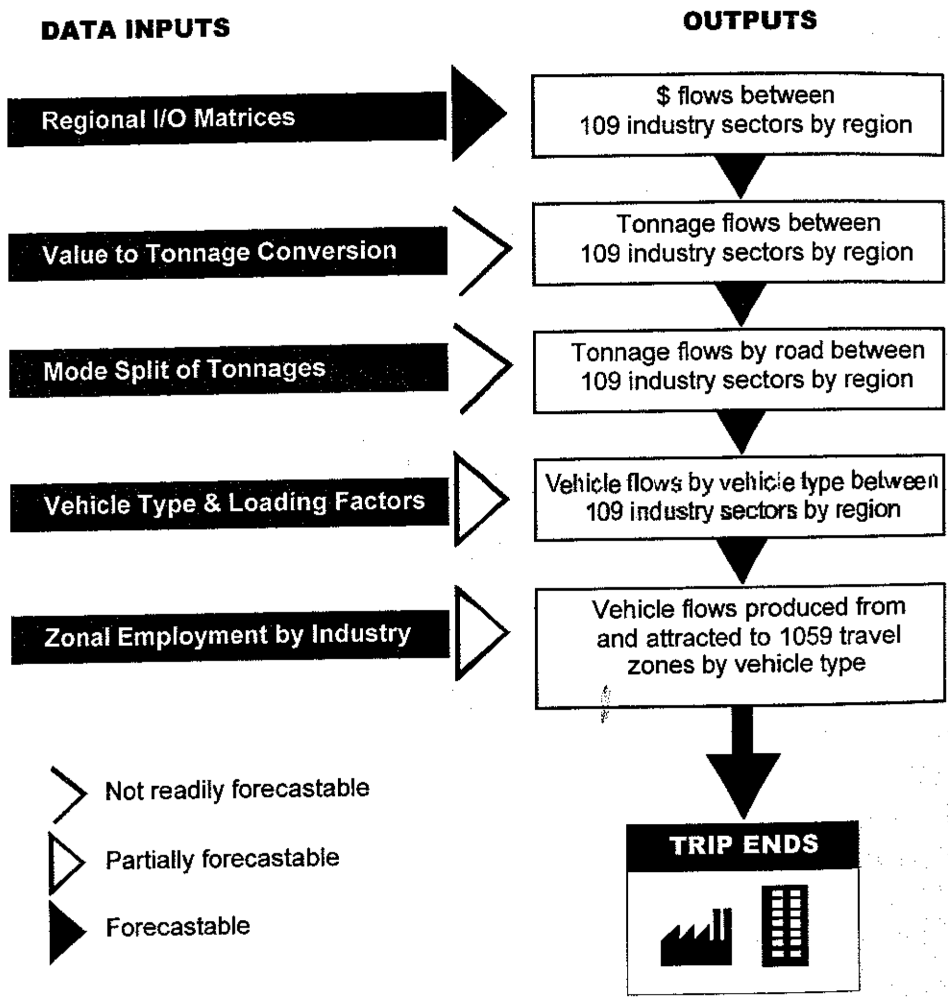

Figure 2 Derivation of Trip Ends for CTS-EP

The starting point is three 109 industry sector regional input-output (I/O) tables, one each for Sydney, Newcastle and Wollongong. These tables show the value, in dollars, of flows between each industry sector within and between each of the regions (including imports, exports, and the household sector).

The *first step* towards converting these value flows between industry sectors to trip ends is to convert them from dollar flows between sectors to tonnage flows between sectors. This involves an intermediate step of working out the value of the different commodities originating from each industry, then converting those commodity values into tonnages.

The breakdown of the value flows from particular industry sectors into values of individual commodities can be estimated from a variety of ABS sources, for example the Make Matrix from the Australian National Accounts Input-Output Tables (ABS, 1997:1).

The commodity flows can then be converted to tonnages using a range of data sources which have both value and tonnage information, such as ABS (1997:2) *Manufacturing Production, Australia: Principal Commodities Produced*, the ABS free board value for exported commodities, FDF Management estimates and others. Commodities are related to the producing sector, and the consuming sector tonnages are then derived using the I/O table relationships.

The output of these conversions is the tonnage flow of goods between each of the 109 industry sector pairs in each of the I/O regions.

The *second step* is to distribute the tonnage flows between modes. The FreightInfo data set (FDF Management, 1997) was used to calculate the mode share by for each industry sector by region (inter and intra-regional trips). Mode shares for each industry were usually assumed constant throughout each region, though for some industries such as mining, dairy and metal manufacturing, sub-regional breakdowns were made. It would also be possible to use CVS data for this purpose.

From this point on, the focus is on the road mode only.

The *third step* is to convert the tonnage flows between each industry sector in the three regions into vehicle flows between each sector. Weigh in motion (WIM) data from a range of Roads and Traffic Authority (RTA) sites in and around the GSMR was used to calculate the average Gross Vehicle Mass for laden rigid and articulated vehicles. The CVS and Survey of Motor Vehicle Use (SMVU) are other possible sources of this data, although the WIM data has the advantage of geographic precision, which can be retained through into the matrix estimation procedure.

The average tare weights for rigid and articulated vehicles were obtained from Austroads and subtracted from the gross weights to produce an average load per vehicle type. The logistics patterns of HGVs are also taken into account using a 'tour' factor and a 'backloading' factor derived from observed travel behaviour in the CVS. The CVS data was also used to identify the mix of vehicles used to carry each commodity, and thus the mix of vehicles used by each industry sector.

Two vectors can then be obtained for each vehicle type – the number of regional productions by industry and the number of regional attractions by industry. Analysis indicates that the conversion from tonnages to trips is the most critical step in the process of producing trip ends. As a result, TDC plans to devote more resources to developing an improved vehicle loading model in the future.

The *final step* is to distribute these productions and attractions spatially across the GSMR. These can be distributed to the travel zone level by assuming that the number of trips produced from or attracted to an industry in a travel zone is directly related to the number of employees in that industry working in that travel zone. Using TDC's Journey to Work data set, which contains industry of employment by travel zone, the number of trips by vehicle type produced from and attracted to each zone can be derived.

This completes the preparation of the first input to the CTS-EP. This input, as with others to the CTS-EP, can be assigned a level of confidence by the analyst which reflects the analyst's relative confidence in the accuracy of the various data inputs.

Input 2: Screenline Counts

The second input to the CTS-EP is classified traffic counts. They are used in the matrix estimation process to inform the assignment of trips from production zones to the appropriate attraction zones, and may even help revise the levels of productions and attractions in particular zones, depending on the relative weights attached to the data

Three hundred and six screenlines were mapped manually such that there was at least one screenline between each possible origin-destination pair. In the context of matrix estimation, a "screenline" can be any number of count locations and the selection of their location can significantly influence the estimation process. Where possible, RTA classified traffic counts were used for each screenline, though in some cases where only annual average daily traffic (AADT) counts were available (counts of all vehicles), this data was factored down to rigid and articulated counts, with factors differing depending on the road type.

The matrix estimation procedure used had the capacity to place a level of confidence in each screenline count, which is a useful feature if the user is confident in the accuracy of some or all of the counts. If there is less certainty about the trip ends, these will be altered to a degree by the estimation process to more closely match the observed traffic counts.

Input 3: Road Network

The third input was a model of the road network available to heavy vehicles, including capacity and speed flow curves for each link. Added to this was TDC's standard two-hour morning peak private vehicle matrix, which was manipulated into a 24-hour matrix, of travel between more than 1,000 travel zones.

The result is that the road network used behaves as a "loaded" network, with link speed and travel times reflecting the amount of private traffic already on each link. The estimation procedure uses the "composite cost skims" from this loaded network, which is a measure of the impedance between each zone pair.

Input 4: Prior Matrix

The final input into the CTS-EP was a "prior matrix" or the "best guess" matrix from which the estimation procedure begins. The prior matrix provides a starting point for linking the productions and attractions to the origins and destinations by providing some estimates of OD trips. The prior matrix was built using data from TDC's 1991/92 CVS.

The CVS was designed to provide reliable travel estimates at the SLA level, thus the first step in the development of the prior matrix was to construct an SLA level trip table from the CVS. Some OD pairs in the trip table had no movements between them because those movements were not sampled. This is a problem for matrix estimation procedures because zero cells in a prior matrix will remain zero throughout the estimation procedure. To overcome this problem, some zero cells in the SLA trip table were manually seeded, that is, values were entered in based on professional judgement.

The SLA level matrix was then expanded to a zone level matrix using the same factors applied when splitting SLA trip ends to zonal trip ends, that is, zonal level employment data was used.

For each zone to zone movement in the prior matrix, a level of confidence can be set in the data. For example, some OD pairs where significant movements were sampled in the CVS can be treated with a high degree of confidence. In the production of experimental CTS estimates, the whole prior matrix was given a very low level confidence in relation to the other inputs. This partially reflected TDC's desire for the production of a trip table which was heavily reliant on the I/O derived data and traffic counts, and also allowed some checking of the matrix produced by the CTS-EP with a trip table produced solely from the CVS.

The Matrix Estimation Procedure

The matrix estimation module, MVESTM, part of the TRIPS transport modelling software, was used to generate an optimal commercial travel trip table from the inputs provided. A hierarchic estimation procedure was used because of the size of the matrix (1059 by 1059 zones). This means that TRIPS aggregates the 1059 zones into a smaller number of districts for the first level of estimation, then these district matrices were used to control the next level of estimation at the zone level.

In simple terms, the estimation procedure uses the loaded network to identify the three best routes between each OD pair, assigning each a probability of use. The prior matrix of OD flows is then adjusted by new information about where trips are produced from and attracted to, and in what quantities. The traffic counts are then used to refine the predicted flows with known traffic counts between particular zones. The relative confidence of the analyst in the various inputs will help determine the degree to which the traffic counts and trip ends are adjusted to reconcile with each other.

A more technical explanation of the general matrix estimation process is available in the TRIPS manual (MVA 1996).

Outputs

Separate matrices of trips between each OD are produced for rigid and articulated vehicles. These matrices are then 'time-sliced', or split into four times of day: morning and afternoon peaks, business peak (between morning and afternoon peaks) and nighttime. The CVS was used to provide the time profiles of travel for vehicle types by SLA which were used to time-slice the matrix.

Checking

One of the problems with a procedure that uses all available data inputs is that by definition there are no external data sources with which to check the results. There are some basic sense checks that can be made however:

- the traffic volumes created by the estimated matrix can be compared with observed screenline counts, although some of these are inputs to the procedure

- the trip length distribution of the estimated matrix can be checked against distributions such as that available from the CVS

The traffic volume checks reveal virtually all estimated flows to be within one standard deviation of the observed screenline counts. The comparison of trip length distributions with the CVS revealed a good fit for both rigid and articulated vehicles. The CTS-EP did appear to underestimate short trips for rigid vehicles, but it did identify some expected peaks representing travel to and from Wollongong and Newcastle which were not observed in the CVS.

In addition, there are a number of sensitivity tests which can be conducted to determine which procedures, assumptions or data items in the CTS-EP are the most critical. An @Risk analysis was conducted, and found that the process was most sensitive to changes in the tonnage to vehicle trips conversion, suggesting that this was an area for closer investigation.

Another check was to rerun the estimation procedure using different relative weights for the various inputs. This proved that the results are heavily dependent on the relative confidence levels given to different data inputs.

Light Goods Vehicle Procedure

The procedure for estimating LGV movements in Sydney differs from the HGV approach for two main reasons. Firstly, not all LGVs carry commodities – some travel to provide services to business or households. While service industries are included in the I/O tables, they were not able to have commodities associated with them. Thus, trip ends for service vehicles cannot be derived in the same way as for commodity-carrying vehicles. Secondly, existing methods of manual and automated traffic counting cannot reliably distinguish light commercial vehicles from passenger vehicles, thus there are no traffic counts available for LGVs.

Trip Ends

The trip ends for commodity-carrying LGVs are produced in the same way as for HGVs. The trip ends for service vehicles are one of the key data gaps identified during the CTS project. As a result, TDC has undertaken a Service Vehicle Attraction Rates (SVAR) study to measure attraction rates of service vehicles to both households and businesses. These rates will then be applied to household and industry base data and forecasts to produce base year and forecast trip ends for service vehicles.

Gravity Model

The lack of traffic counts means that there is little value in using a matrix estimation procedure. Until traffic counts for LGVs can be reliably obtained, a gravity model, based on a function of the level of employment by industry in each travel zone, will be used to distribute the trip ends.

Further Developments of the CTS

The CTS is now at a stage where the estimation framework for HGV and LGV base year trip tables has been developed and demonstrated using available data, or synthetic data where no data is currently available. The next steps will involve collecting new data that is currently not available, refining the conversion processes to improve the quality of information that is finally input to the CTS-EP, and forecasting key inputs to allow the production of forecast trip tables. Some of the key areas TDC would like to focus on include:

Data Collection

- *Industry research* to augment I/O tables and conversion processes with more detailed data about the commodities produced and attracted to industries in Sydney, the freight systems and modes used, the origins and destinations of freight movements, and the values of commodities.
- *Classified traffic count* program to be reviewed with RTA with a view to providing an ability to distinguish LGVs.
- *Driver/Operator surveys* to update logistics trends for vehicle loading model

Process Refinements

- Improvements to TDC network to include turn bans and toll penalties.
- Collate I/O tables at a sub-regional level to provide more accurate geographic estimates.
- Develop and implement a procedure to forecast I/O tables
- Develop a vehicle loading model using all possible data sources.
- Improve dollar to tonne conversion by searching for superior data sources The heterogeneity of goods within a given commodity class makes this process somewhat problematic, and ways to make it more sophisticated, perhaps breaking it down geographically, will be considered.
- Enhance TDC's employment projections by breaking employment into industry types.

Discussion and Conclusions

The TDC has developed the CTS as a method to fulfil its particular requirements for base year and forecast trip tables at the small area level. It was these admittedly ambitious requirements which forced TDC to consider an approach which departed from the traditional driver survey and scaling up of travel by some economic growth factor.

It has been argued that the CTS method TDC has adopted to produce trip tables is merely substituting deterministic uncertainty (e.g. the use of various assumptions to translate I/O flows into tonnages, loadings and trips) for statistical uncertainty (e.g. sampling error) (Chow 1997). This is certainly true to a degree.

However, the ability of the CTS-EP to combine a range of data from different and unrelated sources minimises the risks associated with deterministic uncertainty. In addition, the argument itself is outweighed by TDC's requirements, which make a sampling approach on its own unworkable.

Traditional sample survey approaches are unable to provide reliable travel zone level estimates of commercial travel because of the great variability associated with commercial travel. The CTS approach of distributing regional trips derived from I/O tables by using detailed employment by industry data for every travel zone helps to ensure reasonable travel zone level estimates.

The I/O driven approach to determining travel also provides an overarching control total, which accounts for all goods moved to, from or within Sydney. With traditional survey approaches, it is extremely difficult to define the population, and thus to expand to a population with any degree of certainty.

Trip ends based on commodity flows between industries also offer a sound behavioural basis for forecasting travel. Traditional approaches simply factor base year trips by an economic growth rate to produce forecasts. The CTS approach allows the use of industry-based economic forecasts to project growth in commodity movement, and thus in trips. Knowledge of the predicted spatial spread of industry also allows for the predicted growth to be spread spatially.

The CTS does have other advantages over traditional approaches. For example, much of the data required is produced during the normal course of business by other organisations, thus the procedure does not incur unnecessarily large primary data collection costs. Traffic counts are collected as a matter of course by the State road agency, employment and I/O data is collected by the ABS, and regional I/O tables are produced by a private sector organisation in NSW.

The CTS is an ambitious project that is only in its early stages. Considerably more work is required to enable greater confidence to be held in base year estimates. The multi-layered assumptions required to generate trips from value flows between industries is the major issue to be addressed. In particular, analysis has identified the vehicle loading, tour and backloading assumptions as critical to the estimation of trip ends. There is a need to identify improved data sources for these conversions. In some cases, this may require the collection of primary data.

In addition to improving the base CTS estimates, another key piece of work will be to develop the forecasting capability of the CTS by producing forecasts of I/O tables and forecasts of employment by industry type.

Despite the acknowledged difficulties with the CTS-EP developed by TDC, it offers a pragmatic solution to TDC's requirements to produce base and forecast year estimates of commercial travel at the small area level. Subject to the availability of a similar range of data inputs, there is no reason why a procedure such as this could not be implemented in other cities, Australian and elsewhere.

Acknowledgments

The Transport Data Centre is indebted to the many consultants who have assisted in the development of the CTS-EP. In particular we acknowledge the significant contribution of DJA Maunsell, Oxford Systematics, FDF Management, and Prof Garland Chow.

References

Australian Bureau of Statistics (1997:1) *Australian National Accounts: Input-Output Tables*, Catalogue No. 5209.0, Australian Government Publishing Service, Canberra.

Australian Bureau of Statistics (1997:2) *Manufacturing Production Australia: Principal Commodities Produced*, Catalogue No. 8365.0, Australian Government Publishing Service, Canberra.

Barton-Aschman Associates (1996) *Analysis of Freight Movements in the Puget Sound Region*, consultancy report for Puget Sound Regional Council, Seattle.

Chow G (1997) *Review of North American Commercial Transport Data Collection and Modelling Experience*, consultancy report to the Transport Data Centre, Sydney.

DJA Maunsell (1999) *Commercial Trip Table Estimation Procedure*, consultancy report to the Transport Data Centre, NSW Department of Transport, Sydney.

FDF Management (1996) *Evaluation of Methods for the Collection of Data on Commercial Transport*, consultancy report to the Transport Data Centre, Sydney.

FDF Management (1997) *FreightInfo: A National Database of Freight Flows*, FDF Management, Melbourne.

List, George F., Mark A. Turnquist, John Mbwana and Julian Wolpert (1995) *Analysis of a Dedicated Commercial Transportation Corridor in the New York Metropolitan Area: Final Report*, University Transportation Research Centers Program, Region 2.

Metro and the Port of Portland (1996) *The Collection and Analysis of Commodity Flow Information for Metro and Port of Portland*, Request for Proposal, Portland, Oregon

MVA (1996) Trips Version Seven: Matrix Estimation, MVA, London.

Raimond T (1997) Urban Freight Data Collection, *Papers to the Australasian Transport Research Forum*, Adelaide.