A commodity-based production and distribution freight model with application to Sydney, New South Wales

By
Collins Teye and David A. Hensher

Institute of Transport and Logistics Studies, University of Sydney Business School, NSW 2006, Australia

December 2019

ISSN 1832-570X
A commodity-based production and distribution freight model with application to Sydney, New South Wales

The complexity of freight demand forecasting coupled with non-availability of data with the required scale and features often limits its inclusion in demand forecasting. To model freight demand, data are required on various aspects of the freight system including and not limited to commodities, shipments, demand and production cycles, the actors, and how they interact on the supply and logistics corridors and the broad economic influences on freight movements. Available data on many of these aspects of freight, at varying degrees of aggregation spatially, are publicly available for modelling in Australia. However, these data have not been fully utilised to build a freight model suitable for explaining freight movements and their impacts on the local economy. This is largely due to the diverse nature of these datasets, especially relating to the degree of aggregation and the inherent difficulty of combining these datasets in a consistent and unbiased way. This paper provides a novel approach based on the principle of entropy maximisation to combine these diverse datasets to develop a freight behavioural logit model for the state of New South Wales (NSW), Australia. The resulting model is a linked logit model system comprising a Commodity Production Model (CPM), and Commodity Distribution Model (CDM), segmented by commodity type and vehicle class. The focus is on commodity flows and not conversion to vehicle flows. An important outcome is the link between maximising entropy and maximising access to each commodity group by each vehicle class over production and consumption zones. The implementation of the model in practice and the illustration of its key features are presented using NSW as case study.

KEY WORDS: freight models, commodity flows, commodity generation model, logit share model, MetroScan, New South Wales

AUTHORS: Teye and Hensher
Acknowledgements: This paper contributes to the research program of the Volvo Research and Education Foundation Bus Rapid Transit Centre of Excellence. We acknowledge the Foundation for partial funding support.

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (H73)
The Australian Key Centre in Transport and Logistics Management
The University of Sydney  NSW 2006  Australia
Telephone:  +612 9114 1824
E-mail:  business.itlsinfo@sydney.edu.au
Internet:  http://sydney.edu.au/business/itls

DATE: December 2019
1 Introduction
Almost without exception, the design, planning and management of the road network is determined by travel demand largely derived from passenger travel models. The relative neglect of freight modelling or development of analysis tools in the past was typically justified on the assumption that freight constitutes a very small fraction of the daily road traffic. The difficulty and the cost of collecting freight data has also contributed in large measure to the general absence of freight model systems that are sophisticated and behaviourally appropriate. However, with the growing acknowledgment of the importance of freight to both the local and national economies, and also the disproportionate impacts of trucks on congestion, pollution, accidents and other road hazards, there is a stronger call for a better understanding of freight activities (Hensher and Figlioizzi 2007). Network planners or managers are also keen to understand freight movements and their impacts on road capacity so as to better manage congestion and plan for the future. To achieve this, we need innovative freight models (Hensher and Figlioizzi 2007) capable of capturing key behavioural responses and the interaction of actors within the freight system. Freight models are critical to assessing national, regional and local road capacities, economic development initiatives, and for informing the transport planning process.

Freight is however difficult to model due to several factors, among them the non-availability of data on commodities, shipments, demand and production cycles, especially at an origin-destination level; the lack of understanding about the actors and how they interact on the supply and logistics corridors, and the broad economic influences on local freight movements (Hensher and Figlioizzi 2007). These limitations mean that in the short to medium term, modellers may not have the resources needed to develop a freight model system with the level of detail and richness similar to the current state of the art in passenger modelling (e.g., activity-based models or even integrated transport and location model systems using tour based nested logit models) to answer policy questions of interest. The current practice in freight modelling is therefore based on the efficacy of building models using existing data sources to answer as many important policy questions as possible.

Drawing on existing commodity-based freight models that have incorporated the generation and attraction of commodities into freight models (e.g., Anas 1983, Wisetjindawat et al. 2006; Holguin-Veras and Patil 2008 and review in the next section) and models that incorporate interaction between agents in the supply chain (Hensher and Puckett 2005; Puckett et al. 2007), this paper presents a novel approach, implemented for the NSW, of combining diverse pieces of data from various sources on commodity flows, vehicle usage, and the distribution of freight, trip length distribution by commodity type and other relevant information, to construct a model framework of linked logit models suitable for explaining the main interactions between the key stakeholders (shipper-carrier-consumer) in the freight system, and how changes in land-use and transport network conditions affect commodity productions and distributions. To achieve this, the NSW study area is divided into commodity production and consumption areas where commodities can be seen as being produced and transported to their consumption areas for consumption. The key models in the framework include a commodity production model (CPM) and a commodity distribution model (CDM) at the zonal level of the study area and segmented by commodity group and vehicle class. The framework also involves other sub-routines (that are not the focus of this paper) including an empty truck trips model (ETM), conversion of commodities to vehicle trips, and preparing matrices of trips for assignment. The demand sub-models form the cargo flow model (CFM) and its interactions with the assignment models is illustrated in Figure 1.
Figure 1: Supply-demand model framework within which CFM is positioned

2 Brief Literature Review

Among the most common forms of freight models are those that incorporate the flow of commodities between specific origins and destinations to determine the movement of freight vehicles. The rationale for the use of commodities to derive models of freight movements is, logically, that demand for freight can be considered to be a derived demand that comes from the requirement to move specific commodities between different locations. This includes both commodities used as inputs for businesses (raw materials, etc.) but also manufactured goods and food that are ultimately purchased by consumers. However, this being the case, it is critical that the commodity flows that are used as inputs to the freight models are as accurate as possible since this translates directly to the modelled vehicle flows (Novak et al., 2008).

In the United States in particular, as well as several other countries, national-level data on commodity flows is available from a Commodity Flow Survey (CFS). This provides a basis from which commodity flows can be estimated, even if these data are not always directly applicable (Novak et al., 2008). However, in Australia no such survey is conducted; hence data on commodity flows are limited, and the flows of commodities must be estimated from other sources to provide the necessary input for commodity-based freight models. Models on freight can broadly be classified into four main model types; the trend models, zonal freight rates models, Input-Output (I/O) and Spatial Computable General Equilibrium (SCGE) models.

The trend models involve extrapolating historical trends into the future, using time series or cross section data to develop models of various degrees of complexities, ranging from simple growth factor methods (Cambridge Systematics et al. 1997) and regression models (Novak et al. 2011; Wisetjindawat et al. 2006) to complex multivariate autoregressive moving average models (e.g. Garrido 2000). These models are usually easy to implement, given available data, and are mainly developed for short and medium term forecasts of freight volumes and activities.
on transportation networks. The main advantage of this class of models is that it is easy to implement, requiring limited data and less time and resources. However, it provides limited insight into causality and limited scope for policy effects, especially demand and price related policies (de Jong et al. 2004; Bröcker 1998). Studies employing these type of models can be found in Novak et al. (2011), Wisetjindawat et al. (2006) and Ellison et al. (2017). Novak et al. (2011) applied spatial regression models to estimate the commodity flows using a combination of employment, population and various spatially-adjusted transport variables (e.g., distance to infrastructure, length of motorways). The use of regression models was also applied by Wisetjindawat et al. (2006) to generate initial commodity productions and attractions. However, in generating the commodity flows, Wisetjindawat et al. embed these initial regressions in a set of spatial discrete choice models where the choices of firms, either as producers or consumers (i.e., firms that attract commodities) are modelled such that individual firms choose between available suppliers. This results in estimates of commodity flows between each firm, and by extension, each pair of origins and destinations.

The zonal freight rate method is similar to the well-known category analysis technique for generating passenger trips at zonal levels. This method uses several zonal characteristics including employment and population data to estimate the quantity of each commodity type produced and/or consumed or the number of freight vehicle trips originating or terminating in each freight analysis zone. These freight rates can either be borrowed from existing sources or developed locally by using an existing commodity flow table (e.g., Quick Response Freight Forecasting Manual; ITE 2003) or by estimating them from vehicle surveys, and are usually segmented by commodity type and/or by vehicle class (de Jong et al. 2004). The outcome of this type of model is the amount of commodity and/or the number of vehicles originating or terminating at each freight analysis zone. Examples of existing freight rates in practice can be found in the Quick Response Freight Forecasting Manual in the US (Cambridge Systematics et al., 1997) and the work by the Institute of Transportation Engineer (ITE 2003) containing trip rates for various types of road vehicles classified by commodity type and by production and consumption zones.

Also, special freight rates are often estimated for special generators like the ports and intermodal terminals using survey methods (de Jong et al. 2004). The main strength of this method is that one can readily build freight models using available rates from existing sources or tables and even where such rates are not available, the data requirement to estimate them are low. The limitations of this model type are similar to the trend type model as it also provides very little insight into causality and limited scope for policy effects (Bröcker 1998). Additionally, these models do not capture the interactions among different commodity sectors of the freight system and thus ignore the fact that commodities do act as inputs in the production of other commodities or as substitutes to other commodities in terms of consumption. This leads to the third class of models, the input-output (I/O) models.

Input-Output related models (Leontief 1936; Oosterhaven, J. 1988; de Jong 2007; Holguin-Veras and Patil 2008; Ivanova et al. 2002) are mainly specified to allow interactions between various commodity sectors, an important feature missing in the class of models discussed so far. These models use as input available input-output tables (Leontief 1936) of the study area, which describe the interactions among various sectors of the economy in monetary units, including final demand by consumers and import and export values (de Jong et al. 2007). A variant of the I/O models are Multiregional input-output (MRIO) models. It is an I/O table with an additional dimension: regions or zones. MRIO models are generally more flexible and allow
the modelling of both interregional and intersectoral interdependencies of interest (de Jong et al. 2004); however the data requirements for MRIO are heavier than for national I/O.

Examples of these models in practice include the Italian national model system for passengers and freight (Cascetta 1997), the Expert-system based predictions of demand for internal transport in Europe (EXPEDITE 2000) model for Norway, and the Strategic Transport Research for European Member States (STREAMS) (Leitham et al., 1999). The main strength of these class of models, as noted in Bröcker (1998), is their ability to fully account for interregional and interindustry interdependencies. However, they are heavily constrained by the need for reliable factors to convert the trade flows in monetary units into commodities in tonnes and splitting the resulting flows into various dimensions including freight analysis zones. Bröcker (1998) also noted that these models are mainly demand driven and unable to account for the effects of supply or level of service variables like cost and capacity variations.

An extended version of an I/O model is the spatial computable general equilibrium (SCGE) model (de Jong et al. 2004; Tavasszy et al., 2002; Bröcker 1998). This class of models, in addition to accounting for interregional and intersectoral interactions, incorporate important economic concepts like competition and economies of scale (de Jong et al. 2004). This enhanced model type is, however, difficult to implement in practice (see Broker 1998; Ivanova et al. 2002; de Jong et al. 2004; Tavasszy et al., 2002). Additionally, they are also limited by the need for reliable conversion factors to convert resulting trade flows in values into commodity flows in tonnes. Broker (1998) proposed various assumptions under which this model type can be made operational.

Other important innovations include the mathematical formulations of intercity freight movements by Harker and Freisz (1986a, b), the origin-destination model of truck trips using observed link counts by Holguin-Veras and Patil (2008) and the fractional split-distribution model developed by Sivakumar and Bhat (2002). These models all have elements of commodity generation built into the system. Harker and Friesz (1986a) proposed a conceptual framework that allows the simultaneous prediction of intercity freight traffic generation, distribution, modal split and assignment of freight movements. The supply-side of the transportation market is based on a neoclassical profit maximisation model whilst the demand-side is represented by a spatial price equilibrium model, and the economic mechanism which integrates the supply and demand sub-models. The theoretical limitations imposed on the model by the requirement that it be capable of solving large-scale problems are also addressed. In Harker and Friesz (1986b), they presented alternative mathematical representations of the intercity freight model. The model is a nonlinear complementarity formulation and was used to develop criteria for the existence and uniqueness of a solution to this model. The fractional split-distribution model developed by Sivakumar and Bhat (2002) is structured to resemble the choice patterns of tours in a logistics distribution channel, where fractions of the commodity from an origin are estimated to be consumed at each destination using a multinomial logit form.

As demonstrated by the models reviewed here, various approaches have been applied to estimate the commodity flows that are used as inputs to generate freight vehicle flows. Furthermore, it is apparent from previous research that the various methods employed depend heavily on available data in the study area. More importantly, none fully captures the diversity of sources of data available to explain the production and distribution of commodities in the freight system. This study introduces a framework for combining several available data sets (aggregate and disaggregate) to model and explain freight activity. The framework is flexible and can be extended to model and explain other aspects of freight activities by converting and
adding available relevant information constraints. For the NSW study area, the framework allowed for the modelling of a linked commodity production model (CPM) and commodity distribution model (CDM) at the zonal level of the study area and segmented by commodity group and vehicle class. The CPM estimates the quantity of each commodity type in each production zone and carried by each vehicle class and is influenced by the demand of the commodity in all consumption zones and also the production capacity of the production zone. The CDM then takes as input the output from the CPM in the form of the total quantity of each commodity produced in each production zone by each vehicle class and distributes them to the various consumption zones based on accessibility measures and the expected demand of each consumption zone. This means that the CPM conditions the CDM whilst the CDM influences the outcomes in the CPM. The outputs from these models can then be converted to vehicle trips by time of day to the transport network (which is not the focus of this paper). The overall model framework, as shown in Figure 1, also involves other sub-routines including an empty truck trips model (ETM) and preparing matrices of trips for assignment.

3. Methodology

In formulating the models, the paper assumes that the study area is segmented into freight analysis zones (FAZs - see Figures 2 and 3) where cargo can be seen as originating from one zone and destined to another zone. There are 199 FAZs in the study area as shown in Figure 3. These FAZs are also aggregated into $R$ regions as shown in Figure 4. The aggregation was necessary because data on commodity production and consumption are only available at the regional level. The zones are connected to highway networks so that commodities can be transported from one zone to another using at least one mode of transport. This paper considers three main modes of transport available to carry each commodity; light commercial vehicles (LCV), rigid trucks (RT) and articulated trucks (AT). The data available to the study are summarised in Table 1. The unit of analysis is commodities in tonnes and the modelled period is one year. The main decision variable is the quantity of each commodity group $k \in K$ produced in zone $i \in \mathcal{O}$ and transported to consumption zone $j \in \mathcal{D}$ using vehicle class $v \in V$.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Available information on the freight system</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi$</td>
<td>Observed total commodities (in tonnes) in the system (study area of NSW) per modelled year</td>
</tr>
<tr>
<td>$x_{vk}$</td>
<td>Observed total production (in tonnes) of commodity group $k \in K$ carried by vehicle class $v \in V$ in the study area per year. Three vehicle classes (light commercial vehicles, rigid trucks, and articulated trucks) were identified.</td>
</tr>
<tr>
<td>$x_r$</td>
<td>The total commodities (tonnes) produced within each geographical region $r \in R$. A region by definition contains several freight analysis zones.</td>
</tr>
<tr>
<td>$y_r$</td>
<td>The total commodities (tonnes) consumed within geographical region $r \in R$</td>
</tr>
<tr>
<td>$b_o$</td>
<td>Distance budget (tonne-kms) for each vehicle class. Captures the total weighted distance travelled by each vehicle class $v \in V$ in the study area</td>
</tr>
<tr>
<td>$b_k$</td>
<td>Distance budget (tonne-kms) for each commodity group. This data captures both the weight of the commodity and the distance transported within the study area.</td>
</tr>
<tr>
<td>$\mathcal{O}$</td>
<td>Set of production zones in the study area, indexed by $i = 1,2,3, \ldots 199$</td>
</tr>
<tr>
<td>$\mathcal{D}$</td>
<td>Set of consumption zones in the study area, indexed by $j = 1,2,3, \ldots 199$</td>
</tr>
<tr>
<td>$V$</td>
<td>Set of vehicles available for the transportation of commodities in the study area</td>
</tr>
<tr>
<td>$K$</td>
<td>Set of commodity groups</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>The distance (km) between two freight analysis regions or FAZs</td>
</tr>
<tr>
<td>$r_{iir}$</td>
<td>Equals 1 if production zone $i$ belongs to region $r \in R$ and 0 otherwise</td>
</tr>
</tbody>
</table>
\( \delta_{jr} \) \hspace{1cm} \text{Equals 1 if production zone } j \text{ belongs to region } r \in R \text{ and 0 otherwise}

<table>
<thead>
<tr>
<th>Table 2: Decision variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notation</td>
</tr>
<tr>
<td>( X_{ijk} )</td>
</tr>
<tr>
<td>( X_{ivk} )</td>
</tr>
</tbody>
</table>

The method employed to develop the commodity-based freight model (CFM) is flexible and not restricted by the size and quality of available data. The method also allows data from different sources with different levels of aggregation to be combined in a most consistent way to determine the production and consumption of commodities at the zonal level. The method is based on the principle of entropy maximisation (Jaynes 1957) or information theory (Shannon 1948). The principle asserts that the most likely state of a system at any given time can be found by first defining an entropy function of the system, presenting available information about the system as a set of constraints, and then maximising the entropy function subject to these constraints (Jaynes 1957; Teye 2017). Jaynes (1957) noted that any other state would imply having more knowledge about the system than supplied by the evidence available.

In this study, the state\(^1\) of interest is the quantity of each commodity group \( k \in K \) produced in zone \( i \in O \) and transported to each consumption zone \( j \in D \) using vehicle class \( v \in V \) and is represented by \( X_{ijk} \). Given that the quantity of all commodities in the system is \( x \), the function to maximise is the number of ways that the state of the system can occur (Wilson 1970; Teye et al. 2017):

\[
E = \frac{x^!}{\prod_{i \in O} \prod_{j \in D} \prod_{v \in V} \prod_{k \in K} (X_{ijk}^!)} 
\]

(1)

Thus, we are considering all possible states of the system and only selecting the most likely state consistent with available data on the freight system. Since maximising \( E \) is equivalent to maximizing \( \Lambda = \ln E \), we maximize \( \Lambda \) instead since it is easier to do so (Teye 2017). Thus, the objective function to optimise reduces to:

\[
\Lambda = \ln E 
\]

(2)

Taking the natural logarithm of (1) and applying Stirling's approximation, we have:

\[
\Lambda \approx \ln (x^!) - \sum_{i \in O} \sum_{j \in D} \sum_{k \in K} \sum_{v \in V} X_{ijk} (\ln X_{ijk} - 1) 
\]

(3)

Since \( \ln (x^!) \) is constant, it can be ignored in the optimisation process. Hence, the objective function to optimise becomes an entropy function of the decision variable of interest and is expressed as:

\(^1\) Note that this formulation is still valid even if different states of the system of interest are specified.
\[ \Lambda \approx - \sum_{i \in O} \sum_{j \in D} \sum_{k \in K} \sum_{v \in V} X_{ijkv} (\ln X_{ijkv} - 1) \]  

(4)

### 3.1. Available evidence about the freight system

As noted in the introduction, different study areas may have different sets of available data or information about the freight system. The framework permits these available data to be converted into a set of constraints. In NSW (including Sydney), the available information can be summarised in the following constraints:

1. **Definitional Constraint I:** This constraint links commodity consumption zones to commodity production zones. The constraint therefore allows commodities produced in a zone to be distributed across their consumption zones. This is added as constraint (5).

\[ \sum_{j \in D} X_{ijkv} = X_{ivk}; \forall i \in O, k \in K, v \in V \]  

(5)

2. **Definitional Constraint II:** This constraint links the total observed commodities in the study area and the mode of transport used to distribute these commodities to the production zones of these commodities. This constraint helps to determine the quantity of each commodity produced in each production zone.

\[ \sum_{i \in O} X_{ivk} = x_{vk}; \forall v \in V; k \in K \]  

(6)

3. **Definitional Constraint III:** This constraint ensures that total commodities produced by all production zones in a region equals the observed quantity of all commodities produced in the region. This constraint therefore links the production capacity of each freight analysis zone to the production capacity of the region it lies in.

\[ \sum_{i \in O} \sum_{j \in D} \sum_{v \in V} \sum_{k \in K} X_{kijv} \delta_{iv} = x_r; \forall r \in R \]  

(7)

4. **Definitional Constraint IV:** This constraint ensures that total commodities consumed by all consumption zones in a region equals the observed quantity of all commodities consumed in the region. This constraint therefore links the consumption capacity of each freight analysis zone to the consumption capacity of the region it belongs.

\[ \sum_{i \in O} \sum_{j \in D} \sum_{v \in V} \sum_{k \in K} X_{kijv} \delta_{jr} = y_r; \forall r \in R \]  

(8)

5. **Budget Constraint I:** The term on the left-hand-side captures the weighted distance of transporting each commodity type between all production-consumption pairs by all vehicle classes and should not exceed the total observed “distance budget” (tonne-km) for that commodity.
\[
\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{V}} X_{ijv} d_{ij} = b_k; \forall k \in K
\]  

The distance budget \(b_k\) of each commodity is the tonne-kms travelled by each commodity in the study area.

6. Budget Constraint II: The term on the left-hand-side captures the weighted distance travelled by each vehicle class between all production-consumption pairs over all commodities and should not exceed the total observed “distance budget” for that vehicle class.

\[
\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{ijvk} d_{ij} = b_v; \forall v \in \mathcal{V}
\]

The distance budget \(b_v\) of each vehicle class is the tonne-kms travelled by each vehicle class in the study area.

In summary, the most likely state of the freight system can be obtained by maximising Equation (4) Subject to constraints (5) to (10) and the following non-negativity constraints:

\[
X_{ijvk} \geq 0; X_{ivk} \geq 0
\]

### 3.2 Solution Summary

Based on Equations (5-11), each decision variable can be estimated by constructing a Lagrangian equation comprising the objective function and the constraints, and enforcing the first order optimality conditions with respect to each decision variable and model parameters as shown below.

\[
L \approx \Lambda + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} \sum_{v \in \mathcal{V}} \eta_{ivk} \left( X_{ivk} - \sum_{j \in \mathcal{J}} X_{ijvk} \right) + \sum_{k \in \mathcal{K}} \sum_{v \in \mathcal{V}} \gamma_{vk} \left( x_{vk} - \sum_{i \in \mathcal{I}} X_{ivk} \right) + \sum_{r \in \mathcal{R}} \phi_r \left( x_r - \sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{I}} \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}} X_{ijvk} k_{ir} \right) + \sum_{r \in \mathcal{R}} \psi_r \left( y_r - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{V}} \sum_{k \in \mathcal{K}} X_{ijvk} \delta_{jr} \right) + \sum_{k \in \mathcal{K}} \beta_k \left( b_k - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{v \in \mathcal{V}} X_{ijvk} d_{ij} \right) + \sum_{v \in \mathcal{V}} \beta_v \left( b_v - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} X_{ijvk} d_{ij} \right)
\]  

9
where the parameters $\eta_{ivk}$, $\psi_r,\varphi_r,\beta_k$, $\beta_v$ are Lagrangian parameters associated with constraints constants (5),(7),(8),(9) and (10) respectively and the parameters $\gamma_{ivk}$ are the Lagrangian parameters associated with constraint (6).

**Proposition 1.** The optimal solution of $\Lambda$ with respect to the commodity distribution variable $X_{ijvk}$ can be expressed in logit form (Hensher et al. 2015; McFadden 1974) as:

$$X_{ijvk} = X_{ivk} \frac{e^{-(\beta_{vik}d_{ij}+\sum_{r=1}^{R}\varphi_r\delta_{ir})}}{\sum_{j\in D} e^{-(\beta_{vik}d_{ij}+\sum_{r=1}^{R}\varphi_r\delta_{ir})}} = X_{ivk}P_{ijvk}$$

(13)

where $\beta_{vik} = \beta_v + \beta_k$ and $P_{ijvk}$ is the probability distribution of commodity group $k \in K$ by vehicle class $\forall v \in V$ from a production zone to all consumption zones. By definition $\sum_{j\in D} P_{ijvk} = 1; \forall k \in K; v \in V; i \in O$

**Proof 1:** The first-order condition for a maximum $L$ with respect to the flow variables $X_{ijvk}$ can be expressed as:

$$-\ln X_{ijvk} - \beta_v d_{ij} - \beta_v d_{ij} - \sum_{r=1}^{R} \varphi_r k_{ir} - \sum_{r=1}^{R} \psi_r \delta_{jr} - \eta_{ivk} = 0$$

(14)

Making the decision variable $X_{ijvk}$ the subject we have:

$$X_{ijvk} = e^{-\beta_v d_{ij} - \beta_v d_{ij} - \sum_{r=1}^{R} \varphi_r k_{ir} - \sum_{r=1}^{R} \psi_r \delta_{jr} - \eta_{ivk}}$$

(15)

Enforcing constraint (5) and making $\eta_{ivk}$ the subject we have:

$$e^{-\eta_{ivk}} = \frac{X_{ivk}}{\sum_{j\in D} e^{-(\beta_v d_{ij} - \beta_v d_{ij} - \sum_{r=1}^{R} \varphi_r k_{ir} - \sum_{r=1}^{R} \psi_r \delta_{jr})}}$$

(16)

Hence, by inserting (16) into (15) and simplifying, we have

$$X_{ijvk} = X_{ivk} \frac{e^{-(\beta_{vik}d_{ij}+\tilde{\psi}_j)}}{\sum_{j\in D} e^{-(\beta_{vik}d_{ij}+\tilde{\psi}_j)}}$$

(17)

where $\beta_{vik} = \beta_v + \beta_k$ and $\tilde{\psi}_j = \sum_{r=1}^{R} \psi_r \delta_{jr}$

Following the contribution of Williams (1977), the level of access of each commodity to all consumption zones from a given production zone and vehicle class can be expressed as:
\[ l_{ivk} = \ln \sum_{j \in D} e^{-\left(\beta_{vk}d_{ij}+\bar{\lambda}_j\right)} \]  

(18)

**Proposition 2:** The Lagrangian parameters \( \eta_{ivk} \) associated with constraint (5) can be expressed in terms of the accessibility measure in Equation (18):

\[ e^{-\eta_{ivk}} = X_{ivk} e^{-l_{ivk} + \sum_{r=1}^{R} \varphi_r \hat{k}_{ir}} \]  

(19)

**Proof 2:** The proof is obvious by inserting Equation (18) into Equation (16) and simplifying.

**Proposition 3.** The optimal solution with respect to the commodity production variable \( X_{ivk} \) can be expressed in logit form and linked with the distribution model through the logsum or accessibility measures (18) as:

\[ X_{ivk} = x_{vk} \frac{e^{l_{ivk} + \bar{\phi}_i}}{\sum_{l \in D} e^{l_{ivk} + \bar{\phi}_i}} = x_{vk} P_{ivk} \]  

(20)

where \( \bar{\phi}_i = -\sum_{r=1}^{R} \varphi_r \hat{k}_{ir} \) and \( P_{ivk} \) is the probability distribution of producing commodity group \( \forall k \in K \) in zone \( i \in O \) and transported over all consumption zones by vehicle class \( \forall v \in V \). By definition \( \sum_{i \in D} P_{ivk} = 1 \); \( \forall k \in K \); \( v \in V \)

**Proof 3.** Similarly, enforcing the first-order condition for a maximum \( L \) with respect to the flow variables \( X_{ivk} \) we have:

\[ -\gamma_{vk} + \eta_{ivk} = 0 \]  

(21)

Making \( \eta_{ivk} \) the subject and making use of Equation (19) we have:

\[ X_{ivk} = e^{-\gamma_{vk}} e^{l_{ivk} - \sum_{r=1}^{R} \varphi_r \hat{k}_{ir}} \]  

The Lagrangian parameters \( e^{-\gamma_{vk}} \) in the above equation can be eliminated by enforcing constraint (6), where \( \sum_{i \in O} X_{ivk} = x_{vk} \). Hence defining \( \bar{\varphi}_i = -\sum_{r=1}^{R} \varphi_r \hat{k}_{ir} \) we have:

\[ X_{ivk} = x_{vk} \frac{e^{l_{ivk} + \bar{\varphi}_i}}{\sum_{l \in D} e^{l_{ivk} + \bar{\varphi}_i}} \]  

(22)

The accessibility to each commodity group by each vehicle class over all production and consumption zones becomes (see Williams 1977):

\[ l_{ivk} = \ln \sum_{l \in D} e^{l_{ivk} + \bar{\varphi}_i} \]  

(23)

**Proposition 4:** The Lagrangian parameters \( \varphi_r \) can be estimated iteratively using:
\[ \varphi_r(t + 1) = \varphi_r(t) + \ln \left( \frac{\tilde{y}_r(t)}{y_r} \right) \]  

(24)

where \( \varphi_r(t + 1) \) is the estimated Lagrangian parameters \( \varphi_r \) at iteration \( (t + 1) \) and \( \tilde{y}_r(t) \) is the estimated total consumption of commodities in region \( r \in R \):

\[ \tilde{y}_r = \sum_{l \in G} \sum_{j \in E} \sum_{v \in V} \sum_{k \in K} \tilde{x}_{kijv} \delta_{jr} \]  

(24)

**Proof 4:** The proof is given in Teye et al. (2017b) who provided a general way of dealing with Lagrangian parameters associated with capacity type constraints.

**Proposition 5:** The Lagrangian parameters \( \psi_r \) can be estimated iteratively using:

\[ \psi_r(t + 1) = \psi_r(t) + \ln \left( \frac{\bar{x}_r(t)}{x_r} \right) \]  

(25)

where \( \psi_r(t + 1) \) is the estimated Lagrangian parameters \( \psi_r \) at iteration \( (t + 1) \) and \( \bar{x}_r(t) \) is the estimated total consumption of commodities in region \( r \in R \):

\[ \bar{x}_r = \sum_{l \in G} \sum_{j \in E} \sum_{v \in V} \sum_{k \in K} \tilde{x}_{kijv} \delta_{jr} \]  

(24)

**Proof 5:** The proof follows directly from Proposition (4) above.

As illustrated through propositions (1)-(5), the estimation of the Lagrangian multipliers are inter-dependent, where estimating one parameter requires the evaluated value of other parameters. A general technique for solving this kind of problem has been developed in Lamond and Stewart (1981) called Bregman's balancing method, which has been shown to converge to an acceptable level of accuracy. Bregman's method is adapted and used to estimate the parameters as shown in algorithm A1 below.

**Modified Bregman’s algorithm (A1)**

1. Initialisation:
   - \( \beta_k = \frac{1}{b_k} \); \( \beta_v = \frac{1}{b_v} \); \( \varphi_r = 0 \); \( \psi_r = 0 \); \( \forall k \in K \); \( \forall v \in V \)

2. Update Logsums
   2.1. Update the logsums over consumption zones in Equation (18)
   2.2. Update the logsums over production and consumption zones in Equation (23)

3. Update flow variables of demands
   3.1. Execute the commodity production model in Equation (22) to update \( \overline{x}_{ivk} \)
   3.2. Using \( \overline{x}_{ivk} \), update \( \overline{x}_{ijv} \) in Equation (17)

4. Update Sensitivity parameters
   4.1. Update the sensitivity parameters \( \beta_k \) by enforcing constraint (9) using Hyman’s (1969) algorithm or a Newton type algorithm.
   4.2. Update the sensitivity parameters \( \beta_v \) by enforcing constraint (10) using Hyman’s (1969) algorithm or a Newton type algorithm
   4.3. Repeat (4.1) and (4.2) until convergence is reached.
5. Update capacity constraints parameters
   4.1. Update the capacity constraint parameters $\varphi_r$ by enforcing constraint (8) using Proposition (4)
   4.2. Update the capacity constraint parameters $\psi_r$ by enforcing constraint (7) using Proposition (5).
6. Repeat steps (2) - (5) until convergence is reached.

The convergence criteria can be based on the convergence of the $\beta$ parameter and small changes in the decision variables between iterations. The overall structure of the solved problem can be presented in Figure 2, where the CPM are influenced by accessibility measures from the CDM, whilst the CDM takes the commodities produced by CPM and distribute them among available consumption zones, making the CPM conditioning the CDM.

**Figure 2: Model solution architecture**

### 3.2 Economic interpretation of the entropy objective function

Teye et al. (2017b) have shown the link between maximising entropy and various accessibility measures, reinforcing the contribution of Wilson (1969, 1970). This paper confirms the existence of such link through the propositions below. The proposition shows that maximising entropy is equivalent to maximising access of each by each commodity group and each vehicle class to all production and consumption zones.

**Proposition 6.** The objective $\Lambda$ can be expressed in probabilistic form as follows:

$$\Lambda_1 = - \sum_{k \in K} \sum_{v \in V} x_{vk} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} Q_{ijk} \ln(Q_{ijk})$$

(25)

where

$$Q_{ijk} = P_{ijk} P_{ijk} = \frac{e^{-(\beta v d_{ij} + \psi_r \delta_{jr} + \varphi_r k_{ir})}}{e^{l_{ik}}}$$

(26)

13
where $Q_{ijvk}$ is the probability distribution of producing each commodity group in each production zone and distributing them among consumption zones by each vehicle class. From Equations (13) and (20) $\sum_{i \in D} \sum_{j \in D} Q_{ijvk} = 1; \forall k \in K; v \in V$.

**Proof 6.** Equation (17) can be expressed as in terms of the probabilities $Q_{ijvk}$ as follows:

$$X_{ijvk} = X_{ivk}P_{ijvk} = (x_{vk}P_{ivk})P_{ijvk} = x_{vk}Q_{ijvk}$$

(27)

where

$$Q_{ijvk} = P_{ivk}P_{ijvk} = \frac{e^{-(\beta_{vk}d_{ij}+\psi_{sr}+\varphi_{r})}}{e^{l_{vk}}}$$

(28)

From equation (4), the objective function can be expressed as:

$$\Lambda \approx -\sum_{i \in D} \sum_{j \in D} \sum_{k \in K} \sum_{v \in V} x_{vk}Q_{ijvk}(\ln(x_{vk}Q_{ijvk}) - 1)$$

(29)

Expanding and simplifying the above equation reduces to:

$$\Lambda \approx -(x_{vk}\ln(x_{vk}) - x_{vk}) - \sum_{k \in K} \sum_{v \in V} x_{vk} \sum_{i \in D} \sum_{j \in D} Q_{ijvk}\ln(Q_{ijvk})$$

Note $(x_{vk}\ln(x_{vk}) - x_{vk}) = \ln(x_{vk}!)$ and is a constant, so it can be ignored in the optimisation process. Hence optimising $\Lambda$ is equivalent to optimising $\Lambda_1$.

$$\Lambda_1 = -\sum_{k \in K} \sum_{v \in V} x_{vk} \sum_{i \in D} \sum_{j \in D} Q_{ijvk}\ln(Q_{ijvk})$$

(30)

**Proposition 7:** Optimising the entropy function in equation (4) is equivalent to maximising access to all production and consumption zones of each commodity group by each vehicle class.

**Proof 7:** Inserting the probabilities in Equation (28) into equation (30) we have:

$$\Lambda_1 = -\sum_{k \in K} \sum_{v \in V} \sum_{i \in D} \sum_{j \in D} Q_{ijvk}\ln\left(\frac{e^{-(\beta_{vk}d_{ij}+\psi_{sr}+\varphi_{r})}}{e^{l_{vk}}}\right)$$

(30)

Expanding and simplifying, the above equation can be re-expressed as:

$$\Lambda_1 = \sum_{k \in K} \sum_{v \in V} x_{vk}l_{vk} + \sum_{k \in K} \sum_{v \in V} \sum_{i \in D} \sum_{j \in D} X_{ijvk}(\beta_{v}d_{ij} + \beta_{k}d_{ij} + \psi_{sr}+\varphi_{r}h_{ir})$$

(31)
The above equation is the Lagrangian equation of the following optimisation problem with weighted accessibility as the objective function.

$$\max \Lambda_2 = \sum_{k \in K} \sum_{v \in V} x_{vk} l_{vk}$$

subject to constraints (7)-(10).

Thus, Proposition 7 shows that maximising entropy is equivalent to maximising access to each commodity group by each vehicle class to all production and consumption zones.

4. Data

This section uses a set of common variables to aid in the definition and discussion of the models. The key variables and symbols are defined in Table 1. The commodity flow data used for this analysis is sourced from the Australian Bureau of Statistics (Survey of Motor Vehicle Use, 2014). The data includes total quantity (in tonnes) of each commodity group produced and transported between all production and consumption zones by each the three vehicle classes. This data also includes the distance budgets (total weighted distance travelled in tonne-km) for each of the three vehicle classes and for the each commodity group. Another key data source is the total quantity (tonnes) of all commodities produced and consumed in each of the 27 regions of the study area shown in Figure 4. These data were sourced from Australian Bureau of Statistics (Road Freight Movements, Australia 2014). Network data such as transport service levels and distance travelled by each vehicle were extracted from existing transport models in the study area.
Figure 3: The 199 freight analysis zone system for NSW
5. Estimated Results

5.1 Commodity Distribution Model (CDM)

This model takes as input the quantity of each commodity produced in each production zone (output from the CPM) and distributes them among the consumption zones. The key variables driving the distribution of commodities in the study area are the distance variable (as a proxy for transport cost, noting a high correlation with travel time) and the alternative specific constants associated with each consumption zone. The Lagrangian parameters $\beta_{vk}$ associated with the distance variable captures the sensitivity of transport cost in the distribution of each commodity group by each vehicle class. The estimated values are shown in Figure 5. All the estimated values have the expected signs and as shown in the Figure 5, each commodity group is more likely to be transported during less congested traffic conditions. The results also show variable negative impacts of transport cost on the choice of vehicle class to transport each commodity group. The estimated value of these distance coefficients ensure that modelled average ‘trip length’ (average kilometre travelled) by each vehicle class and each commodity group matched the observed level as shown in Figure 6 and Table 3. Thus, the average distance travelled by each vehicle class and each commodity group is reproduced by the model, satisfying the budget constraints (9) and (10).

The two main commodity carrying vehicles are the rigid and the articulated trucks with a carrying share of about 95% of all commodities in the study area. The results show that for almost all commodity groups rigid trucks are mostly used for relative shorter distance
distributions compared with articulated trucks. For example, distance coefficient associated with rigid trucks is (0.0294) which is significantly higher than that of articulated trucks (0.0042) under the ‘Food and live animals’ commodity group as shown in Figure 5. This means that Food distributions involving short distance trips favour the use of rigid trucks whilst long distance trips favour the use of articulated trucks. The results also suggest that even for the same vehicle class, some commodities are transported over longer distances than others resulting in significantly different magnitudes in the distance coefficients for the same vehicle class.

Light commercial vehicles (LCV) carry only 5% of all commodities in the study area. As expected, they are mainly used for carrying commodities where the rigid or articulated trucks are not suitable or economically feasible, and as shown in Table 3, they are the main vehicles for carrying commodities under ‘tools of trade’. This may explain where the distance coefficient associated with this vehicle are relatively more sensitive than those of the articulated but less sensitive than rigid trucks. In summary the estimated distance sensitivity parameters (or Lagrangian parameters associated with the budget constraints) have the expected signs and expected order of magnitudes. These values have also reproduced the observed distances travelled by each vehicle class and also the observed the distances that each of the commodity group is transported within the study area.

Table 5 summarises the commodity attraction factors or the Lagrangian parameters associated with constraint (8). These constants ensure that the total consumption of commodities in each of the 27 regions of NSW are reproduced. These constants are estimated relative to region 1 (Capital region) with higher values (ignoring the sign) reflecting higher weight or importance of region specific factors influencing commodity consumptions. The highest and lowest region specific constant (ignoring the sign) relative to the Capital region (reference region) is associated with Sydney South-West (2.3928) and Central Coast (-0.1415) respectively, implying region specific factors play significantly more role in commodity consumptions in Sydney South-West than in Central Coast as shown in Table 5.
Figure 6: Comparing modelled and observed average trip length by vehicle class

Figure 5 Estimated distance sensitivity parameters
Table 3: Comparing modelled and observed budgets (average distance transported) by commodity groups

<table>
<thead>
<tr>
<th>Commodities</th>
<th>Estimated TL Budget</th>
<th>Observed TL Budget</th>
<th>% diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and live animals</td>
<td>170.2453</td>
<td>170.4132</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Beverages and tobacco</td>
<td>107.2427</td>
<td>107.3332</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Crude materials, inedible, except fuels</td>
<td>48.88222</td>
<td>48.87547</td>
<td>0.0%</td>
</tr>
<tr>
<td>Mineral fuels, lubricants and related materials</td>
<td>49.06279</td>
<td>49.03285</td>
<td>0.1%</td>
</tr>
<tr>
<td>Animal and vegetable oils, fats and waxes</td>
<td>305.3979</td>
<td>305.4383</td>
<td>0.0%</td>
</tr>
<tr>
<td>Chemicals and related products n.e.s.</td>
<td>179.3624</td>
<td>179.521</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Manufactured goods classified chiefly by materials</td>
<td>80.48097</td>
<td>80.52627</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Machinery and transport equipment</td>
<td>78.38943</td>
<td>78.41402</td>
<td>0.0%</td>
</tr>
<tr>
<td>Miscellaneous manufactured articles</td>
<td>187.1557</td>
<td>187.3443</td>
<td>-0.1%</td>
</tr>
<tr>
<td>Tools of trade</td>
<td>25.29552</td>
<td>25.37367</td>
<td>-0.3%</td>
</tr>
<tr>
<td>Commodities and transactions n.e.s.</td>
<td>139.8615</td>
<td>139.8482</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Table 4: Quantity of each commodity group (tones) carried by vehicle class

<table>
<thead>
<tr>
<th>Commodity</th>
<th>LCV</th>
<th>RigidTruct</th>
<th>Artictrucks</th>
<th>allvehs</th>
<th>% share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and live animals</td>
<td>1136</td>
<td>25322</td>
<td>55185</td>
<td>81643</td>
<td>1.4%</td>
</tr>
<tr>
<td>Beverages and tobacco</td>
<td>9</td>
<td>2466</td>
<td>8944</td>
<td>11419</td>
<td>0.1%</td>
</tr>
<tr>
<td>Crude materials inedible, except fuels</td>
<td>2394</td>
<td>128307</td>
<td>54457</td>
<td>185158</td>
<td>1.3%</td>
</tr>
<tr>
<td>Mineral fuels, lubricants and related materials</td>
<td>179</td>
<td>4488</td>
<td>31128</td>
<td>35795</td>
<td>0.5%</td>
</tr>
<tr>
<td>Animal and vegetable oils, fats and waxes</td>
<td>0</td>
<td>353</td>
<td>3849</td>
<td>4203</td>
<td>0.0%</td>
</tr>
<tr>
<td>Chemicals and related products, not elsewhere specified</td>
<td>266</td>
<td>4810</td>
<td>10138</td>
<td>15214</td>
<td>1.7%</td>
</tr>
<tr>
<td>Manufactured goods</td>
<td>2469</td>
<td>53564</td>
<td>37816</td>
<td>93849</td>
<td>2.6%</td>
</tr>
<tr>
<td>Machinery, transport equipment</td>
<td>1944</td>
<td>14534</td>
<td>16754</td>
<td>33232</td>
<td>5.8%</td>
</tr>
<tr>
<td>Miscellaneous manufactured articles</td>
<td>599</td>
<td>1921</td>
<td>6144</td>
<td>8865</td>
<td>6.9%</td>
</tr>
<tr>
<td>Tools of trade</td>
<td>22077</td>
<td>8177</td>
<td>1920</td>
<td>32175</td>
<td>68.6%</td>
</tr>
<tr>
<td>Other commodities, not elsewhere specified</td>
<td>1456</td>
<td>81566</td>
<td>42320</td>
<td>125342</td>
<td>1.2%</td>
</tr>
<tr>
<td>Unspecified</td>
<td>583</td>
<td>6279</td>
<td>122</td>
<td>6985</td>
<td>8.3%</td>
</tr>
<tr>
<td>Total</td>
<td>33,111</td>
<td>331,788</td>
<td>268,778</td>
<td>633677</td>
<td>5.2%</td>
</tr>
</tbody>
</table>

Table 5: Consumption zone constants and observed vs estimated commodity consumption

<table>
<thead>
<tr>
<th>No</th>
<th>regions</th>
<th>constants</th>
<th>observed</th>
<th>estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Capital Region</td>
<td>0</td>
<td>15418</td>
<td>15418</td>
</tr>
<tr>
<td>2</td>
<td>Central Coast</td>
<td>0.381779</td>
<td>17650</td>
<td>17650</td>
</tr>
<tr>
<td>3</td>
<td>Central West</td>
<td>1.34402</td>
<td>64642</td>
<td>64642</td>
</tr>
<tr>
<td>4</td>
<td>Coffs Harbour - Grafton</td>
<td>0.363562</td>
<td>10233</td>
<td>10233</td>
</tr>
<tr>
<td>5</td>
<td>Far West and Orana</td>
<td>1.562873</td>
<td>20187</td>
<td>20187</td>
</tr>
<tr>
<td>6</td>
<td>Hunter Valley exc Newcastle</td>
<td>0.976786</td>
<td>27151</td>
<td>27151</td>
</tr>
<tr>
<td>7</td>
<td>Illawarra</td>
<td>1.261929</td>
<td>54958</td>
<td>54958</td>
</tr>
<tr>
<td>8</td>
<td>Mid North Coast</td>
<td>1.784914</td>
<td>16529</td>
<td>16529</td>
</tr>
<tr>
<td>9</td>
<td>Murray</td>
<td>0.942633</td>
<td>17101</td>
<td>17101</td>
</tr>
</tbody>
</table>
5.2 Commodity Production Model (CPM) estimated results

This model predicts the quantity of each commodity type produced in each of the production zones in NSW, Australia. The key outcome of the commodity production model is the generation power variable (which summarises the variables used in the CDM and includes accessibility measures and zonal specific constants reflecting the overall attractiveness of each consumption zone) and commodity production constants reflecting zonal specific factors influencing the production of each commodity in each zone. These production constants are presented in Table 6. The majority of the values are close to zero, implying that the policy variables (i.e., distances) used in the models provide a reasonable explanation of the levels of commodity production in each zone. In summary, the results are intuitive and follow expectations. The estimated constants ensure that the observed quantity of all commodities produced in region is reproduced and satisfies capacity constraint (7).

Table 6: Results for the NSW Commodity Production Model

<table>
<thead>
<tr>
<th>No</th>
<th>regions</th>
<th>constants</th>
<th>observed</th>
<th>estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Capital Region</td>
<td>0</td>
<td>14817</td>
<td>14817</td>
</tr>
<tr>
<td>2</td>
<td>Central Coast</td>
<td>0.491058</td>
<td>17908</td>
<td>17908</td>
</tr>
<tr>
<td>3</td>
<td>Central West</td>
<td>1.419636</td>
<td>66890</td>
<td>66890</td>
</tr>
<tr>
<td>4</td>
<td>Coffs Harbour - Grafton</td>
<td>0.612089</td>
<td>10887</td>
<td>10887</td>
</tr>
<tr>
<td>5</td>
<td>Far West and Orana</td>
<td>1.417906</td>
<td>16890</td>
<td>16890</td>
</tr>
<tr>
<td>6</td>
<td>Hunter Valley exe Newcastle</td>
<td>0.881385</td>
<td>22393</td>
<td>22393</td>
</tr>
<tr>
<td>7</td>
<td>Illawarra</td>
<td>1.399044</td>
<td>57789</td>
<td>57789</td>
</tr>
<tr>
<td>8</td>
<td>Mid North Coast</td>
<td>1.651522</td>
<td>14251</td>
<td>14251</td>
</tr>
<tr>
<td>9</td>
<td>Murray</td>
<td>0.832826</td>
<td>15929</td>
<td>15929</td>
</tr>
<tr>
<td>10</td>
<td>New England and North West</td>
<td>2.315832</td>
<td>41985</td>
<td>41985</td>
</tr>
<tr>
<td>11</td>
<td>Newcastle and Lake Macquarie</td>
<td>2.091869</td>
<td>63457</td>
<td>63456</td>
</tr>
<tr>
<td>12</td>
<td>Richmond - Tweed</td>
<td>1.08809</td>
<td>14442</td>
<td>14442</td>
</tr>
</tbody>
</table>
6. **Use of commodity models in a freight model system**

Although the generation of commodity flows as discussed in earlier sections of this paper are of interest by themselves, these models have been developed to be incorporated into a set of freight models that convert the commodity flows to vehicle flows by class and departure time, as well as a related model to identify empty vehicles. The full freight model for NSW is embedded within a new regional demand model (Called R-Tresis), and well as a component built into MetroScan for the Greater Sydney Metropolitan Area (GSMA), an integrated multi-modal transport and land use modelling system, developed at The University of Sydney (Ho et al. 2017). For the GSMA application, we aggregate zones outside Sydney GMA as external zones. In addition to the freight models, MetroScan incorporates a set of behaviourally rich models for modelling transport and land-use related decisions including models for light commercial service vehicles, as distinct from freight-carrying light commercial vehicles (Ellison et al. 2017), and passenger models as shown in Figure 7.
The freight choice models within which the commodity models discussed in this paper are used, use a combination of the commodity flows and a variety of firm data to estimate the likely decisions made with which to service the commodity flows. It is important to emphasise that these models are not run in isolation, but instead contain links to many of the other models within the broader system. Of particular importance to the commodity models (and freight models more broadly) are the models that predict the location decisions of households, firms and workers. The inclusion of these models means that the likely consumption patterns across zones are estimated endogenously and so allow for changes to transport, infrastructure and land-use patterns to in turn influence freight transport without the need for external forecasts on which commodity models frequently rely. Furthermore, the generation of freight vehicles flows also has an influence on subsequent decisions by individuals and firms through their effect on travel times on the road network. However, it must be emphasised that the large number of interactions between the models mean that the freight models must be further calibrated within the full model system to ensure subsequent changes to residential and other location decisions are considered. This means there are additional complexities in calibration over and above standard requirements.

6. Conclusions

This paper has proposed and developed a framework for modelling freight production and its distribution to consumption locations by combining both aggregate and disaggregate available information about the freight system in a consistent way. The key features of the framework is illustrated using the state of New South Wales, Australia as a case study. This is the essential first step in the development of models of truck movement, which is highly dependent on the spatial identification of commodity production and consumption flows. In ongoing research, the models are being implemented in a fuller model system for Sydney within MetroScan and RTesis for NSW that incorporates a full range of individual decisions, firm location decisions,
passenger travel decisions, service vehicle travel decisions that together provide fully endogenous inputs for applying the commodity and freight models described in this paper. The research work can be extended by incorporating land use and socio-economic variables into the framework to help explain zonal specific production and consumption capacities.

References


URL: [http://linkinghub.elsevier.com/retrieve/pii/S0191261507000549](http://linkinghub.elsevier.com/retrieve/pii/S0191261507000549)


Teye C, Bell MGH and Bliemer MCJ (2017b) Locating urban and regional container terminals in competitive environment: An entropy maximising approach, *Transportation Research Part B*, http://dx.doi.org/10.1016/j.trb.2017.08.017


