ITLS

WORKING PAPER

ITLS-WP-18-05

Commodity interaction in Freight movement models for Greater Sydney

By
David A Hensher, Collins Teye and Richard B. Ellison

Institute of Transport and Logistics Studies (ITLS), The University of Sydney Business School

March 2018

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-18-05

TITLE: **Commodity interaction in Freight movement models for Greater Sydney**

ABSTRACT: Central to freight movement models is an understanding of where the freight commodities are produced and consumed. An important driver in the production and/or consumption of each commodity is the production and consumption of other commodities. In this paper, these important interactions between commodities are captured in a path based freight model which incorporates models of commodity production and consumption. We identify the key factors driving the consumption and production of each commodity together with their elasticities. To be suitable for forecasting and policy testing, the estimated models are transformed into linked logit models that allow for important policy measures such as accessibility and commodity generation powers to be estimated. The proposed model has been implemented to generate the amount of commodity of each type produced and consumed in each state of Australia with illustrations of how the production and/or consumption of one commodity triggers the production and/or consumption of others commodities. When built into an integrated transport and land use model system, this capability adds a richness to the way in which freight movements influence and hence impact on the performance of the entire transport network, for both passenger and freight.

KEY WORDS: ***freight models, commodity flows, commodity generation model, logit share model, Integrated transport and location model systems***

AUTHORS: **Hensher, Teye and Ellison**

Acknowledgements: This paper contributes to the research program of the Volvo Research and Education Foundation Bus Rapid Transit Centre of Excellence (through the MetroScan project). 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: March 2018

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 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 (and integrated into a system that allows for both freight and passenger movements). However, with the growing acknowledgment of the importance of freight to both the local and national economies, and also the disproportionate impacts of freight related trucks on congestion, pollution, accidents and other road hazards, there is a stronger call for a better understanding of the freight system (Hensher and Figliozzi 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 Figliozzi 2007) capable of capturing all the 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; 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 Figliozzi 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) to answer all 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., Wisetjindawat et al., 2006; Holguin-Veras and Patil (2008) and models that incorporate interaction between agents in the supply chain (e.g., Hensher and Puckett, 2005; Puckett et al, 2007), this paper presents a novel and intuitive approach for estimating the quantity of each commodity type produced and/or consumed in each freight sector. The proposed model captures a very important element in the freight system where the production and/consumption of one commodity has the potential of triggering more production and/or consumption of other commodities. The key parameters governing the behaviour of the model are estimated together with their elasticities, and the resulting model is converted into a linked logit model suitable for forecasting and testing various freight related policies.

The paper is organised follows; Section 2 presents an overview of commodity based models in the literature and in practice. The proposed model, called the Commodity Generation Chain Model (CGCM), is presented in Section 3. The estimation of model parameters and discussion on the key drivers of each commodity production and consumption are then presented in Section 4. In Section 5, the model is converted into a linked behavioural model suitable for forecasting and policy testing. An application of the model for sensitivity and policy testing is presented in Section 6 and how it fits into the overall model architecture of the Sydney freight system, with the way it can be built into an integrated transport and land use strategic model system called MetroScan-TI discussed in Section 7, followed by conclusions in Section 8.

2 Commodity generation models for freight

Among the most common forms of freight models are those that incorporate the flow of commodities between specific origins and destinations in order to determine the movement of freight vehicles. The rationale for the use of commodities to drive models of freight movements is, logically, that demand for freight can be considered to be a derived demand that emanates 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 used 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., 2011).

In the United States in particular, as well as several other countries, national-level data on commodity flows is available from the 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., 2011). However, in Australia no such survey is conducted; hence the 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. A variety of different methods have been applied to estimate these flows including regression (Novak et al. 2011; Wisetjindawat et al. 2006), spatial computable general equilibrium (SCGE) models (Friesz et al.1994,1998; Goldsman and Harker 1990), simulation (Liedtke 2009), and logistics models (Tavasszy et al. 1998; Liedtke and Schepperle 2004; and de Jong and Ben-Akiva 2007), among others. 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 (distance to infrastructure, length of motorways, etc.). These models predicted freight generation and showed that there was a strong correlation with freight attraction. In other words, zones that generate large volumes of commodities also tend to attract large volumes of commodities. These estimates could then be used within commodity-based freight models to predict vehicle flows.

The use of regression models are 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. Of particular interest in the approach adopted by Wisetjindawat et al. is that the models incorporate the flows through the supply chain, meaning that the outputs of one industry become the inputs of another industry. In doing this, it allows for more realistic assumptions regarding how commodities are produced or attracted to a specific zone, given the mix of firms. The use of simulation techniques have also been applied to generate the commodity flows between origins and destinations (Liedtke, 2009). Just as with models developed by Wisetjindawat et al., Liedtke (2009) simulates the decisions of individual firms within the supply chain to estimate the flows of commodities, and then integrates this with a module that identifies how those commodity flows are ultimately transported to generate vehicle flows. Using various different methods, the interactions between different agents in the supply chain in estimating freight movements have also been used by Holguin-Veras and Patil (2008) and Hensher and Puckett (2005).

As demonstrated by the models discussed 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 this previous research that it is not sufficient to treat the flows of commodities as simple flows between zones, but that they must also account for the existence of interactions between commodities, where the rate of consumption of one commodity can trigger the production and/consumption of other commodities in addition to being policy responsive. The primary focus of the current paper is on the commodity generation chain model as well as how it is incorporated into the freight model system and the broader model system that also incorporates passenger and service vehicle models. Following a brief overview of commodity generation models for freight and a review of the related literature, this paper describes the commodity generation chain model (CGCM), including a discussion of estimation approach and a summary of the results. This is followed by the methods to transform the CGCM to linked logit models, and then how it can be integrated into a broader land-use and transport modelling system, MetroScan-TI, is discussed.

3 Commodity generation chain model

The CGCM is modelled at the national level and uses commodity flows between states together with land-use and other commodity attributes to capture the chain reactions triggered by the production and/or consumption of one commodity on other commodities. The main output of this model is the total commodity by type produced/consumed in each state and the evaluated factors governing the generation of these commodities. The proposed CGCM is a path based model (Wright 1918, 1921, 1934), which is a special case of a structural equation model (SEM) (see Wiley 1973). The flexibility of this model type means that the productions and/or consumptions of commodities can be allowed to interact both within and between commodities (commodity productions/consumptions) and with multiple independent (exogenous) observed variables such as land use and a variety of socio-demographic and firm-related variables. Crucially, these are then transformed into linked logit choice models and used for policy testing. The key notation used in the paper is presented below, followed by the model specification.

3.1 Notation and model specification

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 (Road Freight Movements, Australia 2014). The data used included the quantity of each commodity type (in tonnes) produced and consumed in each state and major cities of Australia. The commodities included in the analysis, and a summary of the quantity of each commodity produced or consumed in 2014, is presented in Figure 1. The other key dataset is the data on population and employment levels in each industry and land use data. These data were sourced from the Australian census, ABS business counts by both employment and revenue, and land-use data derived from ABS meshblock data. Also used was data from the GeoScience Australia NEXUS database that provided details on the existing number of physical structures of various types (such as warehouses, factories, offices, and residential buildings).

Table 1: Variables and symbols

Notation	Meaning
O	Set of commodity production zones indexed by i
D	Set of commodity consumption zones indexed by j
K	Set of commodities with size p
A	Set of cargo consumption variables indexed by i with a_{si} being the quantity of variable i (e.g., jobs in manufacturing) in state s
G	Set of variables expected to influence the generation of commodities indexed by i with g_{si} being the quantity of variable i (e.g., population) in state s
S	Set of states or alternatives indexed by s
Y_{sk}	Estimated production of commodity k in state s
X_{sk}	Estimated consumption of commodity k in state s
Emp_Man	Employment in manufacturing
Pop	Population
Emp_Min	Employment in mining
Emp_Con	Employment in construction
Emp_Elec	Employment in Electricity, Gas, Water and Waste Services
Emp_Tran	Employment in transport
Emp_Agric	Employment in Agriculture, Forestry and Fishing

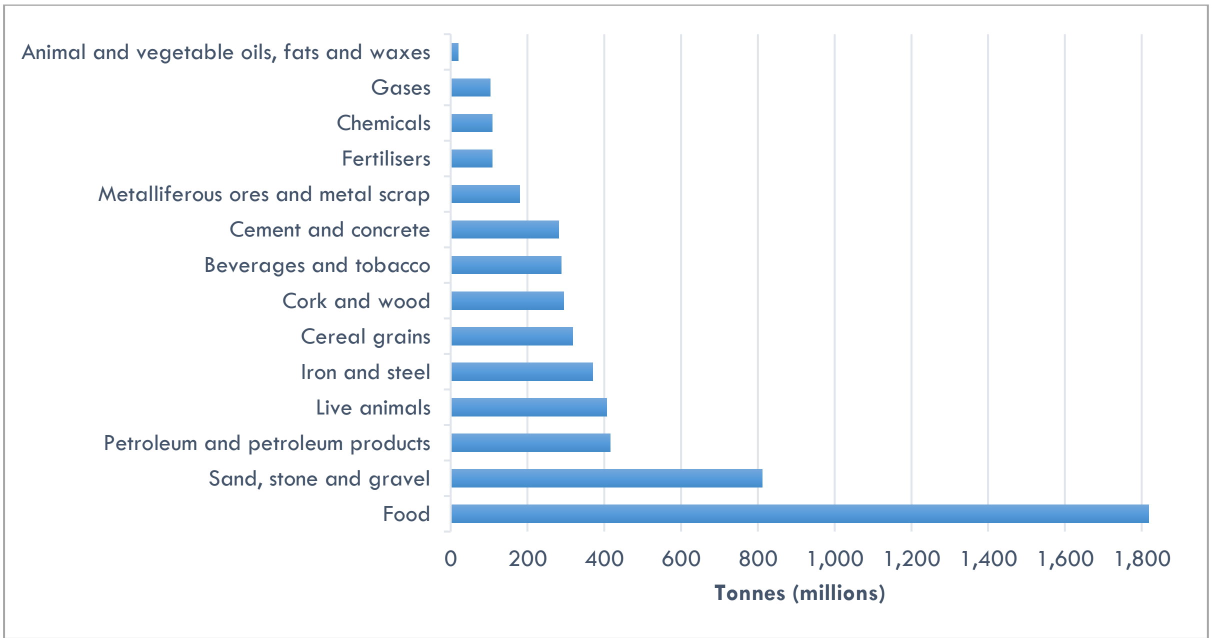

Figure 1: Distribution of each commodity type consumed in Australia 2014

The commodity generation chain model (CGCM) for the production of commodity $k \in K$ in state $s \in S$ is specified in Equation (1):

$$\ln Y_{sk} = \gamma_k + \sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si} \tag{1}$$

Similarly, the consumption of commodity $k \in K$ in state $s \in S$ is specified in Equation (2):

$$\ln X_{sk} = \eta_k + \sum_{l \neq k \in K} \xi_l \ln Y_{sl} + \sum_{l \neq k \in K} \omega_l \ln X_{sl} + \sum_{i \in A} \psi_i \ln a_{si} \tag{2}$$

The parameters $\gamma_k, \eta_k; k \in K$ are respectively commodity specific constants for the production and consumption of each commodity type. Focusing on Equation (1), the first term on the right-hand-side is expected to capture all unobserved factors affecting the production of each commodity type; the second term represents all commodities whose productions can trigger more (or less) production of commodity $k \in K$ in each state. The third term reveals the influence of all commodities whose consumptions also trigger more (or less) production of commodity $k \in K$. The final term captures the production capacity of each state in the production of each commodity type, with higher values expressing higher productive capacity and vice versa. This term is expected to include land use, industry-specific and accessibility variables that drive commodity productions and/or consumptions. Similar interpretations apply for Equation (2) which expresses the amount of each commodity type consumed in each state. The parameters β_l (or ξ_l) capture the rate at which the production of commodity $l \in K$ triggers more (or less) production (or consumption) of commodity $k \in K$. Similarly, the parameters α_k (or ω_k) capture the rate at which the consumption of commodity $k \in K$ impacts the production (or consumption) of $k \in K$. Finally, the parameters θ_i and ψ_i reveal the weight or importance of each non-commodity variable in the production or consumption of each commodity type respectively.

4 Parameter estimation

In estimating the parameters, several model structures were investigated with the most promising models based on available data presented in Table 2, 3 and 4. The tables contain the factors driving the production and/or consumption of each commodity and their interactions. Tables 2, for example, shows the estimated parameters governing the consumption and production of beverages and tobacco (Beverage), cereal grains (Cereal) and food. Table 3, is restricted to the key factors driving the productions and consumptions of chemicals, live animals, animal & vegetable oils, and fertilizer commodities. Finally, the main drivers for the production and consumption of petroleum and related products (Petroleum), gases, metalliferous ores, and iron and steel are presented in Table 4.

As shown in the Tables, all the estimated parameters, except a few specific constants, have the expected signs and are statistically significant at the 95% confidence interval level. The few statistically nonsignificant constants were retained to allow for the transformation of the model into nested logit models, as discussed in Section 5. In Tables 2, the growth in population is the main non-commodity driver of cereal and food consumption, whilst the number of employees in industries such as manufacturing is the main driver for beverage consumption. The amount of food consumed is also driven by the amount of cereal and beverage consumption as shown in Table 2. Also, as expected, the production of beverages is mainly driven by the amount of its consumption, which together with food consumption in turn, governs the quantity of food to produce. The quantity of food produced together with the quantity of cereal consumed then drive the growth in cereal production in Australia, revealing the deep interactions between these commodities.

In Tables 3, population growth governs the consumption of live animals, whilst the number of employees in manufacturing (Emp_Man) triggers more consumption of chemicals. The table also shows that the consumption of both animal & vegetable oils and fertilizer are driven by the number of employees in agriculture, forestry and fishing (Emp_Agric). In terms of commodity production, the main drivers of growth are the rate of consumption of the corresponding commodity shown in Table 3. Finally, in Table 4, employment in transport (Emp_Tran), electricity, gas, water and waste services (Emp_Elec), construction (Emp_Con) and mining (Emp_Min), respectively promote more consumption of petroleum and petroleum products, gases, iron and steel and metalliferous ores respectively. The production levels of these commodities are also primarily driven by their rate of consumption as presented in Table 4.

Following the expressions in Equations (1) and (2), the magnitude of the estimated parameters in Tables 2, 3, and 4 (logged before entering into the equations) also represents the mean direct elasticity estimates as all variables (both dependent and independent). The interpretation of each elasticity estimate can be assessed in the usual way (Hensher et al. 2015). Focussing on the estimates in Table 2, the 2.46 estimate for population under cereal consumption model, for example, means that, all things being equal, a 1% increase in population directly leads to about 2.5% increase in the consumption of cereal grains. Similarly, a 1% increase in the number employed in manufacturing triggers a 1.1% direct increase in the consumption of beverages. The direct impact of population growth on food is less severe as food consumption is also driven by the rate of consumption of beverage and cereal grains. Thus, a 1% increase in population growth only results in a 0.4% growth in food consumption. The results also show beverage consumption as a biggest driver of food consumption, with a 1 % increase in beverage consumption expected to result in about 0.6% increase in food consumption. The direct elasticity of beverage consumption is about 1.0, meaning that a 1% increase in beverage consumption directly leads to an equivalent 1% increase in its production, all other things being constant. For food production, a 0.26% or a 0.82% increase is mainly driven by a 1% increase in beverage or food consumptions respectively. Similar interpretations apply for the other commodities in Table 3 and 4. The indirect impacts of an increase (or decrease) in one variable or commodity on others are illustrated empirically in Section 6.

Table 41: Models for the production and consumptions of Beverages, Cereal and Food (t-value in brackets)

	Beverage		Cereal Grain		Food	
	Consumption	Production	Consumption	Production	Consumption	Production
Non-commodity Variables						
Constant	3.502(3.0)	-0.201(1.7)	-20.668(3.9)	3.101(3.2)	0.685(0.7)	-0.884(2.6)
Population			2.462(6.7)		0.418(3.0)	
Emp_Man	1.053(9.9)					
Commodity consumption						
Beverage		1.013(123.9)			0.603(8.0)	
Cereal Grain				0.805(12.1)	0.046(1.3)	
Food						
Commodity Production						
Beverage						0.257(3.4)
Cereal Grain						
Food				0.039(2.0)		0.815(10.0)

Table 3: Models for Chemicals, Live animals, Animals & vegetable oils and Fertilizer (t-value in brackets)

	Chemicals		Live Animals		Animal & vegetable oils		Fertilizer	
	Consumption	Production	Consumption	Production	Consumption	Production	Consumption	Production
Non-commodity Variables								
Constant	-0.207(1.7)			0.038(0.8)	2.471(1.8)	0.469(2.1)	-12.506(5.6)	
Population			1.118(2.7)					
Emp Man	1.542(4.8)							
Emp_Agric					0.755 (5.5)		2.895(2.9)	
Commodity consumption								
Chemicals		1.011(114.1)						
Live animals				0.995(103.5)	0.196(2.3)			
Animal & vegetable oils						0.96(51.2)		
Fertilizer								1.001(236.6)

Table 4: Models for Petroleum, Gases, Metalliferous ores and Iron & Steel (t-value in brackets)

	Petroleum		Gases		Metalliferous ores		Iron and steel	
	Consumption	Production	Consumption	Production	Consumption	Production	Consumption	Production
Non-commodity Variables								
Constant	7.559(10.7)	-0.893(5.4)		-1.735(8.9)	7.237(5.8)		2.443(3.0)	-0.357(1.7)
Emp_Tran	0.551(4.9)							
Emp_Elec			1.294(4.9)					
Emp_Con							1.122(14.4)	
Emp_Min					0.859(6.1)			
Commodity consumption								
Chemicals	0.173(3.5)							
Petroleum		1.056(99.2)						
Gases				1.123(78.9)				
Metalliferous ores						1.027(75.1)		
Iron and steel								0.992(38.3)

5 CGCM transformation into linked logit models for applications

Once the parameters in (1) and (2) are estimated, the resulting models can be used to forecast the amount of commodity of each type produced and consumed in each state. Focusing on the production model in (1), if $\mathrm{y}_k = \sum_{s \in S} Y_{ik}$, the total quantity of commodity $k \in K$ produced in the whole study area (e.g., Australia), then from (1), the total quantity of commodity $k \in K$ produced in state $s \in S$ can be expressed as:

$$Y_{sk} = \mathrm{y}_k \frac{\exp(\sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si})}{\sum_s \exp(\sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si})} \tag{3}$$

The resulting maximum expected utility (Equation 4) can act as a generation power (see proposition 1) of commodity $k \in K$:

$$\ell_k = \text{In} \sum_s \exp\left(\sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si} \right) \tag{4}$$

Proposition 1: The total quantity of commodity produced in the country (Australia) can be expressed in terms of the generation power using equation (1) with some algebraic manipulations as: $y_k = \exp(\gamma_k + \ell_k)$

Proof 1: By definition:

$\mathrm{y}_k = \sum_s Y_{sk}$; then from Equation (1)

$$Y_{sk} = \exp\left(\gamma_k + \sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si} \right)$$

Or

$$Y_{sk} = \exp(\gamma_k) \exp\left(\sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si} \right)$$

Hence

$$\mathrm{y}_k = \sum_s Y_{sk} = \exp(\gamma_k) \sum_s \exp\left(\sum_{l \neq k \in K} \beta_l \ln Y_{sl} + \sum_{k \in K} \alpha_k \ln X_{sk} + \sum_{i \in G} \theta_i \ln g_{si} \right)$$

Or

$$\mathrm{y}_k = \exp(\gamma_k)\exp(\ell_k) = \exp(\gamma_k + \ell_k)$$

Or more generally:

$$y_k = \exp(\gamma_k + \lambda_k \ell_k); 0 < \lambda_k \leq 1 \tag{5}$$

The parameters λ_k are introduced to allow for differential impacts of network and other variables on commodity generation and distribution for each commodity following the theory underlying nested logit models (see Hensher et al. 2015).

Proposition 2: An important aspect of the forecasting process is the forecast for the quantity of commodity of each type produced or consumed in the study area for any given forecast year. If $y_k(0)$; $y_k(\tau)$ is the quantity of commodity $k \in K$ produced in the country in the base year (0) and forecast year τ with corresponding generation powers $\ell_k(0)$, $\ell_k(\tau)$, then the forecast $y_k(\tau)$ from Equation (5) can be expressed as:

$$y_k(\tau) = y_k(0) e^{\lambda(\ell_k(\tau) - \ell_k(0))} \tag{6}$$

Proof 2: Let $\tilde{y}_k = y_k(\tau) + y_k(0)$ be the total volume of commodity $k \in K$ produced in both years. From Equation (5), the volumes produced $y_k(\tau)$, $y_k(0)$ are respectively governed by the maximum expected utility $\ell_k(\tau)$ and $\ell_k(0)$. It should be noted that the estimated parameters γ_k and λ_k remain unchanged (or are assumed fixed) throughout the forecast years. Thus, from Equation (5), the quantity of commodity of type $k \in K$ produced in the country in the base year (0) and forecast year τ can be respectively expressed as:

$$y_k(0) = \exp(\gamma_k + \lambda_k \ell_k(0)) \tag{7}$$

$$y_k(\tau) = \exp(\gamma_k + \lambda_k \ell_k(\tau)) \tag{8}$$

Making γ_k the subject in both equations, and equating them we have:

$$y_k(\tau) = y_k(0) e^{\lambda(\ell_k(\tau) - \ell_k(0))}$$

Alternatively, following the logit framework, the utility of producing in forecast year τ can be expressed as:

$$U_k(\tau) = \gamma_k + \lambda_k \ell_k(\tau)$$

Hence, the quantity of commodity produced in forecast year τ can be estimated using:

$$y_k(\tau) = \tilde{y}_k \frac{\exp(\gamma_k + \lambda_k \ell_k(\tau))}{\exp(\gamma_k + \lambda_k \ell_k(\tau)) + \exp(\gamma_k + \lambda_k \ell_k(0))} \tag{9}$$

and using the definition $\tilde{y}_k = y_k(\tau) + y_k(0)$, we have:

$$y_k(\tau) = (y_k(\tau) + y_k(0)) \frac{\exp(\lambda_k \ell_k(\tau))}{\exp(\lambda_k \ell_k(\tau)) + \exp(\lambda_k \ell_k(0))} \tag{10}$$

Grouping like terms and simplifying we have:

$$y_k(\tau) = y_k(0) e^{\lambda(\ell_k(\tau) - \ell_k(0))} \tag{11}$$

Once the total quantity of each commodity type produced and/or consumed at the state level is known, the next stage is to determine the zones in each state where commodities are actually produced and/or consumed. The volumes of each commodity produced or consumed at the zonal level is determined by the Cargo Flow Model (CFM) which uses the forecast quantity of each commodity produced $Y_{sk}(\tau)$ and consumed $X_{sk}(\tau)$ in the state of NSW as inputs. This model forms part of the larger MetroScan-TI framework (see Section 7). The CGCM is specified and estimated for the whole country, with the states in the country acting as zones.

The model outputs the quantity of each commodity produced and consumed in each state and the factors governing the production/consumption of each commodity.

6 Application of CGCM

This section focusses on applying the technique developed in Section 5 by transforming the model results in Section 4 into linked logit models for forecasting and sensitivity analysis. First, the model was calibrated by adjusting the constants to ensure that the observed commodity of each type produced and consumed in each state is reproduced by the model. For simplicity, we also focussed on the production, consumption and the interactions of cereal, beverages and food in the state of NSW. Figure 3 shows the weights of the main factors governing the production/consumption of each commodity, and Figure 4 shows the calibrated (adjusted) commodity specific constants capturing all unexplained factors in the models.

The calibrated constants ensure that the observed quantities of each commodity produced and consumed are reproduced by the model as illustrated in Figure 2 showing the plot between observed production and consumption of each commodity type. The nearly ‘perfect’ relationship indicates that the model has accurately reproduced the observed data. As noted earlier, this outcome was achieved by adjusting the estimated constant for each commodity type. As these constants are unchanged in the forecast years, values closer to zero are preferred. Larger values indicate greater importance (i.e. weighting) of omitted variables and may negatively affect forecast results given that they remain unchanged even in forecast years. As shown in Figure 4, the consumption of cereal has the highest alternative specific constant, ignoring the signs (with a value of -1.00), with the least being the beverage production (with a value of -0.012), indicating that all things being equal, that the model can more accurately forecast beverage production relatively better than cereal consumption using observed factors. Food production and consumption are also reasonably explained by observed factors as shown in Figure 4.

The key observed factors governing the models are presented in Figure 3. For example, population growth can be seen as influencing cereal consumption, which together with the growth in population and beverage consumption influence the rate of food consumption. The rate of food consumption and beverage production in turn drives the growth in food production. Thus, population has both direct impact and indirect impacts (through cereal consumption) on food consumption and also indirect impacts on cereal and food production, as shown in Figure 3. On the other hand, employment in manufacturing, directly promotes growth in beverage consumption and directly influence food production and consumption, beverage production and cereal production. This variable (employment in manufacturing) influences beverage production through beverage consumption, which in turn affects food production and also in turn affects cereal production as illustrated in Figure 3. The variable also indirectly affects food consumption through beverage consumption, which in turn promotes food production and also in turn drives cereal productions, making population growth a key driver in the consumption and production of many commodities.

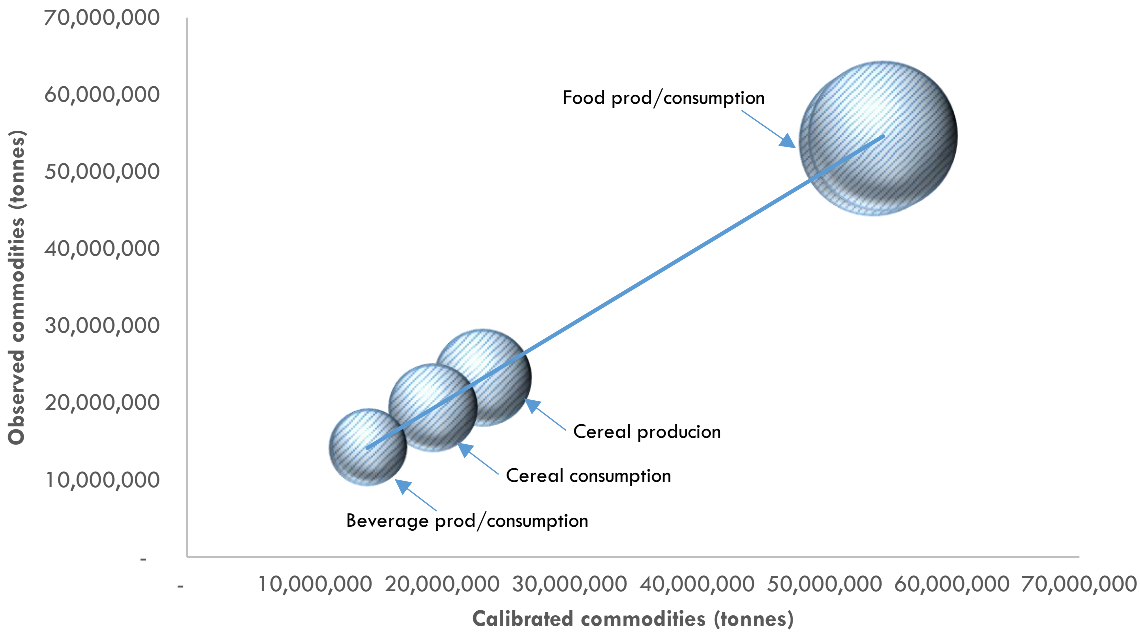

Figure 2: Calibrated vs observed commodities

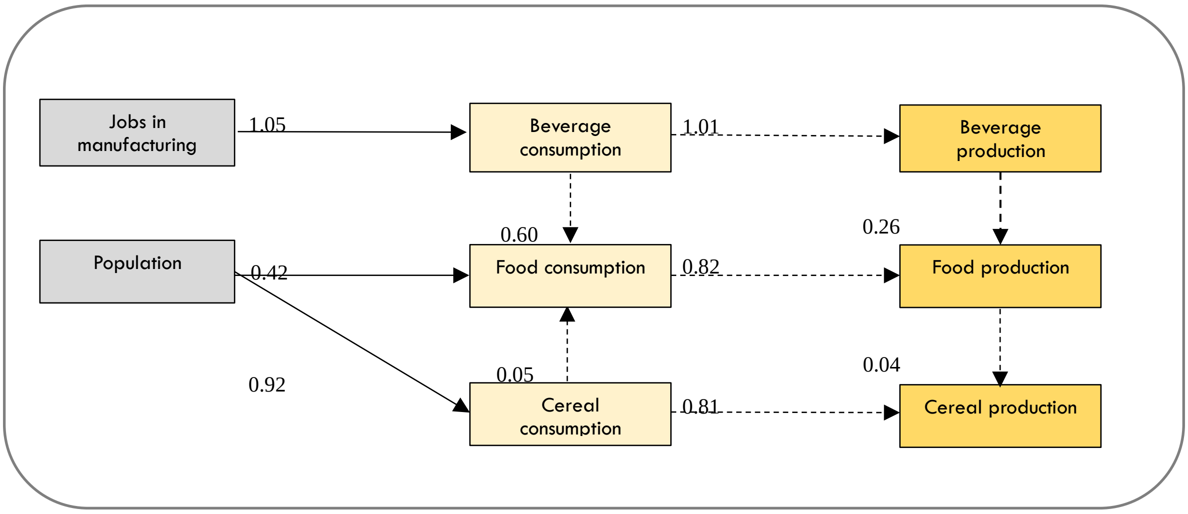

Figure 3: Interdependencies among commodities

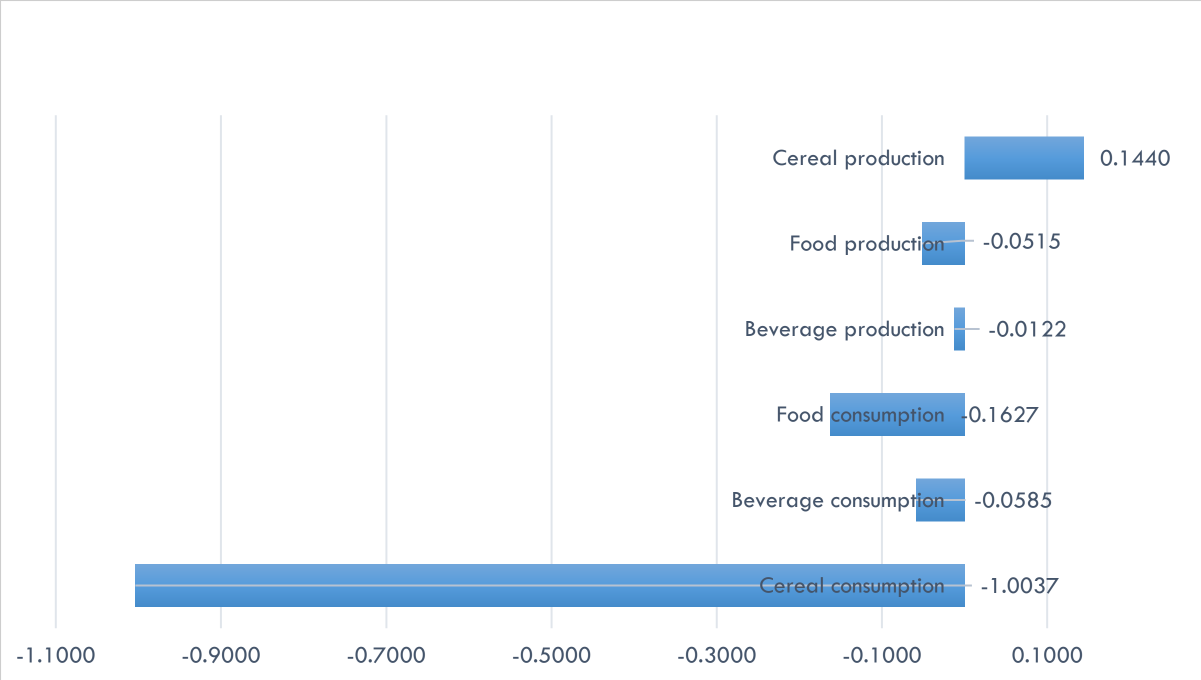

Figure 4: Calibrated commodity-specific constants

The existence of commodity interactions is further demonstrated through sensitivity testing and forecasting outcomes. In forecasting or policy testing, the first set of models to run are the beverage and cereal consumption models. The outputs from these models enter into the food consumption model as inputs, and run as the next model. The third model to run is the beverage production model which takes inputs from the corresponding consumption model. The outputs together with the outputs from the food consumption model become inputs into the food production model, and finally the cereal production model is run using as inputs the outputs from the food production and cereal consumption models as illustrated in Figure 3. The process of running each model is described in Section 5, where the utility for each alternative zone is first computed followed by the logsums (as indicators of expected maximum utility) over all zones. This logsum is then compared with the base (or reference) logsum (from the calibrated model) to compute the total demand for the forecast year using Equation (6). Once the total demand for each commodity is known, they are distributed across the available zones using Equation (3).

The results presented in Figure 5 reveal the impacts of varying population growth or growth in employment in manufacturing on the production and consumption of each commodity type in the whole of Australia, with Figure 6 focussing on the study area of NSW. As expected, the impacts of increasing growth in population and employment in manufacturing on cereal production and consumption is progressively higher than on the other commodities. For example, a 1% increase in population growth resulted in {0.7%, 0.7%}, {1.2%, 1.2%}, and {1.4%, 2.5%} growth in food production and consumption, beverage production and consumption, and cereal production and consumption respectively. However, a 5% growth in population resulted in a {3.2%, 2.3%}, {5.5%, 5.5%} and {9.6%, 12.8%} growth in food production and consumption, beverage production and consumption, cereal production and consumption respectively[1]. Thus, if you use cereal production as a reference we have the ratio of 1:1:1.7:1.7:2:3.6 growth in food production and consumption, beverage production and consumption, cereal production and consumption respectively, revealing that the impacts of the 1% increase in population growth is twice as strong for cereal production than for food production. What is more interesting is the impact of the 5% population growth which yields the ratio 1:073:1.7:1.7:3:4 respectively, showing that the impacts of 5% population growth on cereal production is now 4 times stronger than on food production. This outcome is due to the interactions between the commodities where the growth in one triggers both direct and indirect growth in others, as discussed in Figure 3. In this example, population growth triggers direct growth in food consumption, which in turn drove food production up. The quantity of food produced has direct impacts on cereal production in addition to the direct impact from growth in cereal production, which is also indirectly driven by growth in population, as illustrated in Figure 3. This illustrative example demonstrates the need to properly account for the interactions between commodities when forecasting, as the growth (consumption and production) in one commodity's consumption or production naturally leads to more or less growth in others.

Focussing on the NSW study area, the impacts of population and employment growth on commodity consumptions and productions are shown in Figure 6. Again, the growth in cereal consumption and production is proportionately higher than the other commodities. The 10 growth scenarios, as illustrated in Figure 6, show that cereal consumption is expected to grow from 2% to about 26% for a 1% and a 10% growth in population respectively, which are respectively greater than the 2 to 21% production growth. This further demonstrates that the production of cereal is not only determined by the growth in consumption but also by the growth in other commodities like food production which grew at lower rate. The combined effect means that cereal consumption is expected to be less than cereal production as population grows significantly (more than 2%), and that importation of cereal may be necessary to match the growth in demand.

[1] Some P and C are the same and other are not. This is because, the other factors involved in the P or C are held constant. P is always a function of C and other variables or constant. This analysis looked at impacts of only population, holding everything else constant.

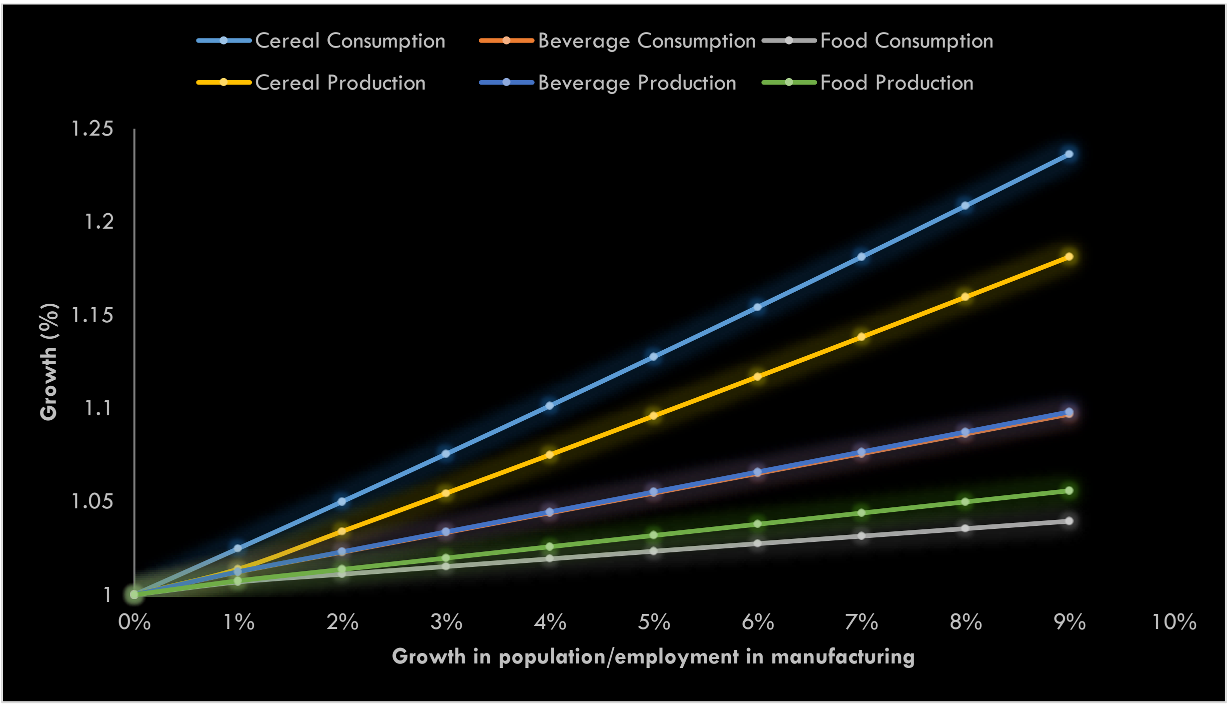

Figure 5: Distribution of commodity production and consumption in Australia

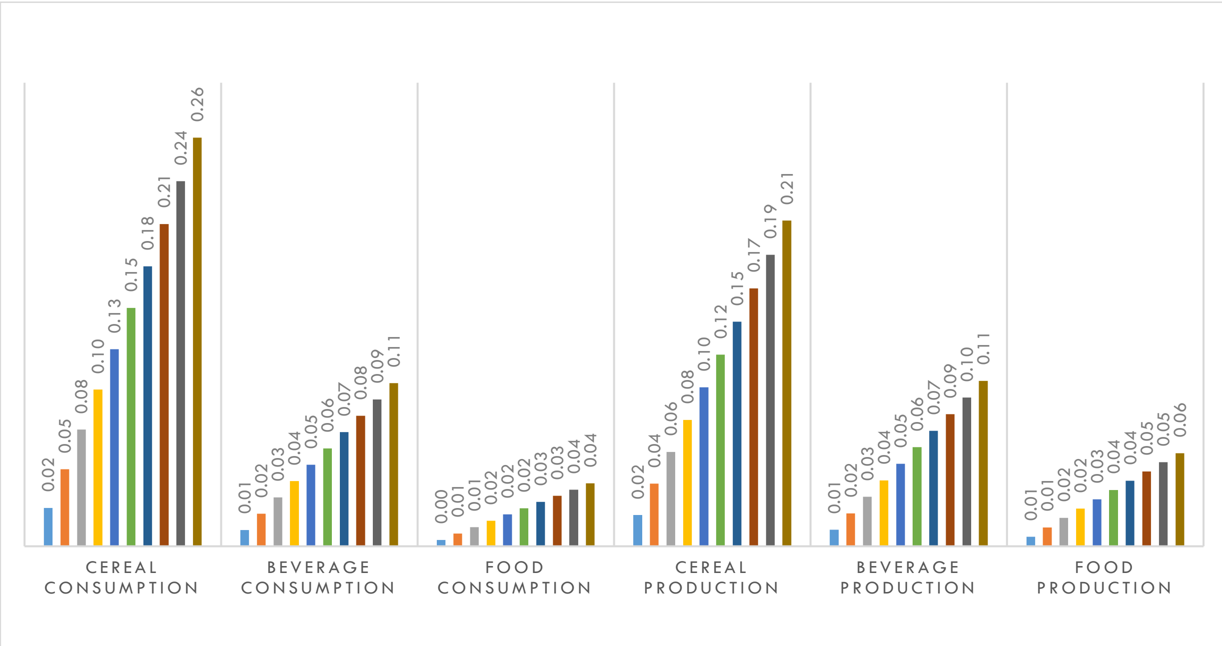

Figure 6: Impacts of growth in population/employment in manufacturing

Finally, we investigated the impacts of unobserved factors by increasing the calibrated constants in turn and observing the impacts of commodity production and consumption. In particular, we increase each constant by 1%, 2% and up to 10% in turn, with the findings summarise in Figure 7. The evidence reveals the potential impacts of unobserved factors, in particular on cereal consumption, where a 1% increase in these factors can results in almost an equivalent 1% decrease in commodity consumption, all other things being constant. The impact on cereal production is less severe at about a 0.6% reduction for every 1% increase in the unobserved factors. From the figure we see that the impact of unobserved factors on both food production and consumption is less severe, and almost negligible on beverage consumption and production. Further research is required to identify other potential cereal production and consumption factors and to include them in the modelling exercise. The models for the other commodities are robust enough to be carried forward in forecasting as shown in Figure 7.

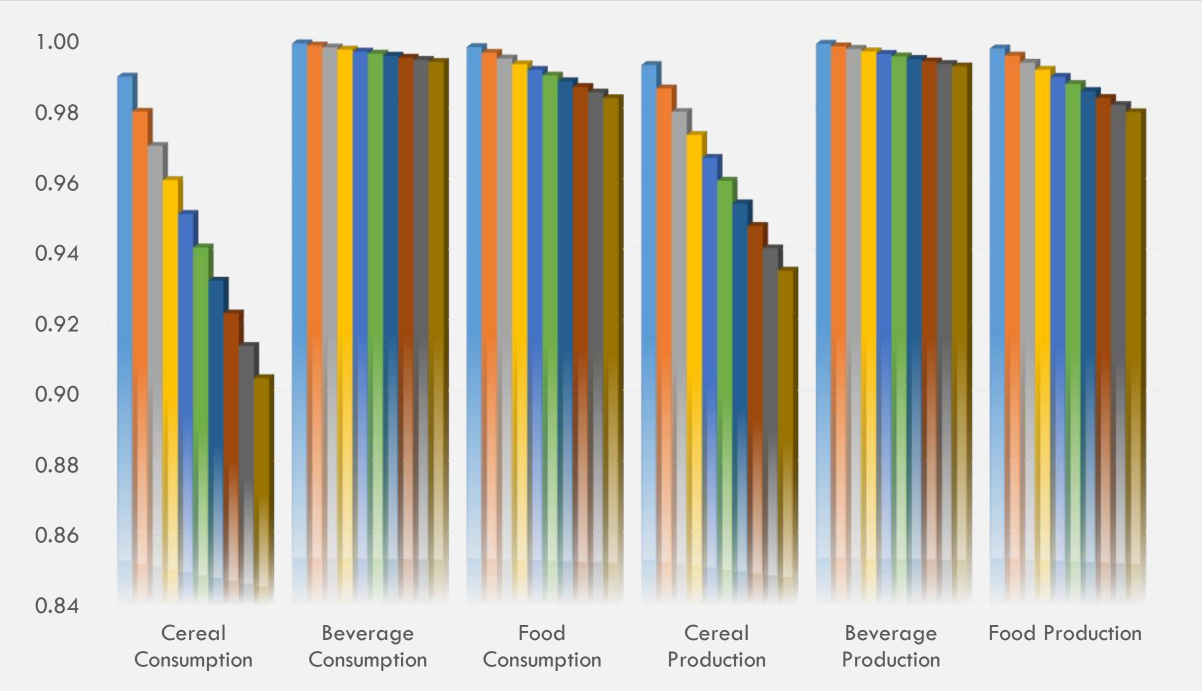

Figure 7: Impacts on unobserved factors in the production and consumption of commodities in NSW

7 Use of commodity models in a freight model system and MetroScan-TI

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 is imbedded within the larger MetroScan-TI framework that is being developed by the Institute of Transport and Logistics Studies at The University of Sydney. In addition to the freight models, this system 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 8.

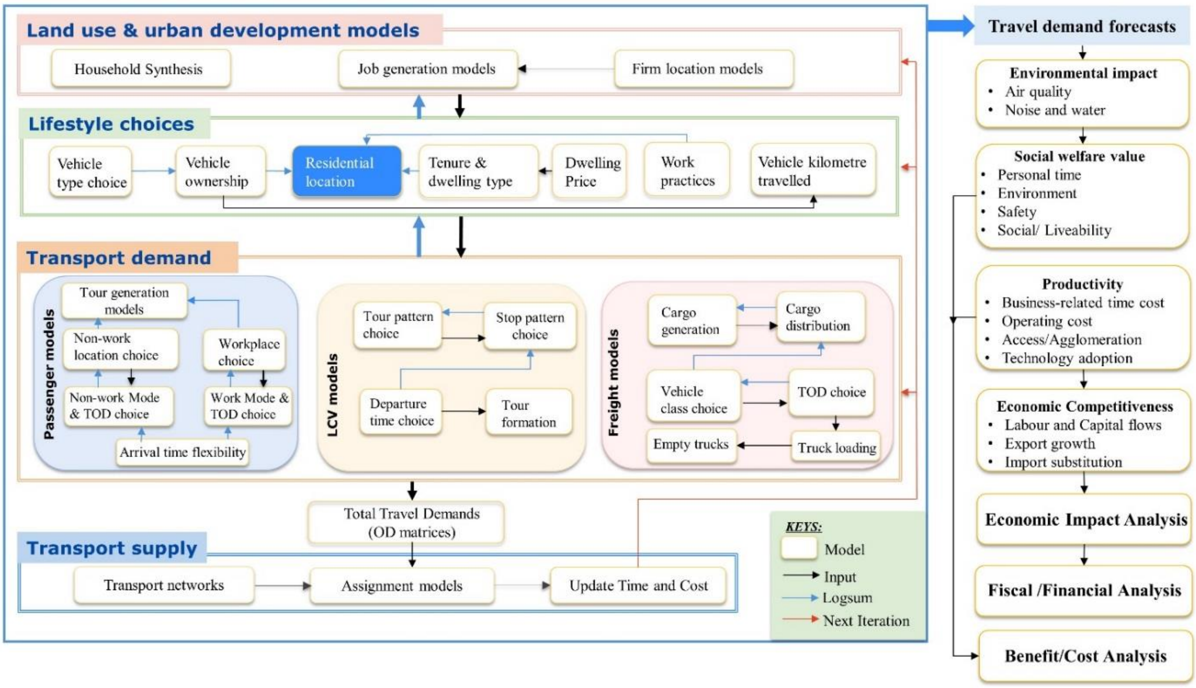

Figure 8: Overall framework of MetroScan demand and supply models

The freight choice models within which the commodity models discussed in this paper are embedded, 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 vehicle 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. Application of MetroScan using the model system in this paper is given in Ellison et al. (2017) and other applications of MetroScan include Ho et al. (2017) and Hensher et al. (2018).

8 Conclusions

This paper develops a commodity-based model capturing the dynamics in commodity production and consumption, and how changes in the production or consumption of one commodity triggers a chain reaction in the production and/or consumption of other commodities. The results are then used in a series of linked logit models to explain freight generation and movements and distribution patterns in Greater Sydney. The models are implemented in a fuller model system (called MetroScan-TI) 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 focus on the production and consumption of commodities (by class), and especially the way such commodities 'feed' off each other in the supply chain, reinforces the important role of what is being moved in the definition of the freight task, and how this might be embedded into a strategic level transport and land use model system as a way of enhancing the richness of outputs associated with multi-modal and multi-sectoral decision making and policy advisory processes.

References

de Jong, G. and Ben-Akiva, M. A micro-simulation model of shipment size and transport chain choice,*Transportation Research Part B: Methodological* **41**(9): 950–965. (2007)
URL: *http://linkinghub.elsevier.com/retrieve/pii/S0191261507000549*

Ellison R.B., Teye, C., Hensher, D.A. Modelling Sydney's light commercial service vehicles. *Transp. Res. A* 96, 79-89 (2017).

Ellison, R., Teye, C., and Hensher D.A. Commodity-based heavy vehicle model for Greater Sydney, paper presented at *The Fifth International Choice Modelling Conference* 3 – 5 April 2017, Cape Town, South Africa (2017).

Friesz, T. L., Suo, Z.-G. and Westin, L. , A Nonlinear Complementarity Formulation of the SCGE Problem, Proc. of the 1993 International Workshop on Transportation and Spatial CGE Models (R. Roson, editor), University of Venice (1994).

Friesz, T. L., Suo, Z.-G. and Westin, L. Integration of Freight Network and Computable General Equilibrium Models, Chapter 12 in *Network Infrastructure and the Urban Environment* (L. Lundquist, L.-G. Mattson and T. J. Kim, editors), SpringerVerlag (1998).

Goldsman, L and Harker, P. T. A Note On Solving General Equilibrium Problems With Variational Inequality Techniques, *Operations Research Letters*, 9, 335- 339, (1990)

Hensher, D.A., Rose, J.M. and Greene, W.H, *Applied Choice Analysis,* Second Edition, Cambridge University Press, Cambridge, (2015).

Hensher, D.A., Puckett, S.M. Refocusing the modeling of freight distribution: development of an economic based framework to evaluate supply chain behavior in response to congestion charging. *Transportation* 32(6), 573–602 (2005)

Hensher, D., Figliozzi, M.A. Behavioural insights into the modeling of freight transportation and distribution systems. *Transp. Res. B* 41(9), 921–923 (2007)

Hensher, D.A., Ellison, R., Ho, C. and Weisbrod, G. How well does BRT perform in contrast to LRT? An Australian case study using MetroScan_TI, *Bus Rapid Transit, edited by* Fiona Ferbrache, Edward Elgar Publisher. Chapter 3 (2018).

Ho, C. Hensher, D.A. and Ellison, R. Endogenous treatment of residential location choices in transport and land use models: introducing the MetroScan framework, (paper presented at *The Fifth International*

Choice Modelling Conference 3 – 5 April 2017, Cape Town, South Africa), *Journal of Transport Geography*, 64, 120-131 (2017).

Holguín-Veras, J. and Patil, G.R. A Multicommodity Integrated Freight Origin-destination Synthesis Model. *Netw Spat Econ* 8: 309 (2008)

Jaynes, E.T. Information Theory and Statistical Mechanics, *Physical Review***,** vol. 106**,** no. 4, pp. 620-630 (1957).

Joreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85-112). New York: Seminar.

Keesling, J. W. (1972). Maximum likelihood approaches to causal flow analysis. Unpublished dissertation, University of Chicago, Department of Education.

Lamond, B., and Stewart, N. F.Bregman's Balancing Method, *Transportation Research Part B*. vol. 15, pp. 239-248 (1981)

Liedtke, G., Schepperle, H.: Segmentation of the transportation market with regard to activity-based freight transport modelling. *Int. J. Logist. Res. Appl*. 7(3), 199–218 (2004)

Novak DC, Hodgdon C, Guo F, and Aultman-Hall L Nationwide Freight Generation Models: A Spatial Regression Approach. *Network and Spatial Economics* 11 (2011).

Puckett, S.M., Hensher, D.A., Rose, J.M., Collins, A. Design and development of a stated choice experiment for interdependent agents: accounting for interactions between buyers and sellers of urban freight services. *Transportation* 34, 429–451 (2007).

Sivakumar, A., Bhat, C. Fractional split-distribution model for statewide commodity-flow analysis. *Transp. Res. Rec*. 1790, 80–88 (2002).

Tavasszy L, Smeenk B et al. A DSS for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research* 5(6): 447-459 (1998).

Teye, C., Bell, MGH, Bliemer, M., Urban intermodal container terminals: The Entropy maximisation facility location problem, Transportation Research Part B, 100, 64-81, (2017)

Wisetjindawat, W., Sano, K. A behavioral modeling in micro-simulation for urban freight transportation. *J. East. Asia Soc Transp. Stud*. 5, 2193–2208 (2003).

Wisetjindawat, W., Sano, K., Matsumoto, S. Commodity distribution model incorporating spatial interactions for urban freight movement. *Transp. Res. Rec*. 1966, 41–50 (2006).

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 69-83). New York: Seminar.

Wright, S. The nature of size factors. *Genetics* Vol 3, 367-374 (1918).

Wright, S. Correlation and Causation. *Journal of Agricultural Research* Vol 20, 557585 (1921).

Wright, S. The method of path coefficient. *The Annals of Mathematical Statistics*, Vol. 5, No., 3. 161–215 (1934).

Yves Rosseel lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/ (2012).