WORKING PAPER

ITLS-WP-12-01

Linking discrete choice to continuous demand within the framework of a computable general equilibrium model for the analysis of wider economic impacts of transport investment projects.

By

Truong P. Truong[1] **& David A. Hensher**

[1] Honorary Professor of Sustainable Transport Systems

January 2012

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-12-01

TITLE:

Linking discrete choice to continuous demand within the framework of a computable general equilibrium model for the analysis of wider economic impacts of transport investment projects.

ABSTRACT:

Discrete choice (DC) models are commonly used as basic building blocks in 'bottom-up' models which seek to describe consumer and producer behaviour at a disaggregate level, in contrast to continuous demand (CD) models which are used to describe behaviour at a more aggregate level. At a disaggregate level, choice behaviour is defined in terms of commodities differentiated by qualities or attributes. In contrast, aggregate demand behaviour is defined in terms of broadly defined and generically different commodities. In a DC model, the main focus of analysis is not the total quantity of demand, but rather the relative shares or substitution between the choice alternatives, in contrast to a continuous demand model where the focus is on the aggregate substitution between groups of commodities as well as on the income effects. Seen in this way, there is scope for complementary usage of DC and CD models within the framework of a CGE model where DC models are used to describe the preferences for a narrowly defined set of commodities belonging to a particular sector of an economy whereas CD models are used to describe the interactions between these sectors. In this paper, we describe how DC and CD models can be used in such an integrated fashion in a spatial computable general equilibrium model to inquire into the wider economic impacts of a transport investment project in the Sydney Metropolitan Area.

KEY WORDS: ***Bottom-up; top-down; discrete choice; differentiated products; continuous demand; computable general equilibrium (CGE) models; transport infrastructure investment; urban planning; wider economic impacts.***

AUTHORS: Truong P. Truong & David A. Hensher

CONTACT:

INSTITUTE of TRANSPORT and LOGISTICS STUDIES (C37)
The Australian Key Centre in Transport and Logistics Management

The University of Sydney NSW 2006 Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: business.itlsinfo@sydney.edu.au
Internet: http://sydney.edu.au/business/itls

DATE: January 2012

1. Introduction

Discrete choice (DC) models are often used as the basic building blocks in a bottom-up model which seeks to describe consumer and producer behaviour at a disaggregate level. Such models are often rich in details of the choice alternatives, as well as characteristics of the individual decision makers, and therefore can capture the behavioural responses of individuals to economic policies more accurately than can aggregate models of supply or demand. For example, the decision on commuter mode choice in a DC model can be described not only in terms of the attributes of the travel modes (travel time, travel cost, comfort level, convenience, etc.), but also the socio-economic characteristics of each decision maker (income level, age-group, position in family structure, occupation, flexibility of travel time, etc). Similarly, the decision on workplace or residential location in a DC model can be described in terms of the varied economic characteristics of the location (e.g., average wage paid for a particular type of work in that location, average rental cost for different housing types, transport costs between different origin and destinations, etc.). This means that policies which seek to influence decisions at the disaggregate level on mode choice or locational choice can be more accurately analysed if set within the framework of a disaggregate DC model.

At a disaggregate level of choice decision, however, the typical choice set within a DC model often consists of mainly a narrowly-defined sets of varieties, or alternatives, of a particular commodity e.g., different modes of travel, different types of cars. These varieties or alternatives are differentiated mainly by quality attributes rather than simply by market prices (which can be considered as the summary indices for these attributes)[1]. In contrast, continuous demand (CD) models which look at behaviour at a more aggregate level are concerned only with the demand for groups of commodities which are generically different (transport, food, housing, education, health, etc.). These groups are to be 'differentiated' only or mainly via their market prices.[2] The choice decision within a DC model is therefore concerned primarily with the 'fine' substitution between different alternatives or varieties of a particular commodity (often produced within a particular sector of an economy), rather than with the gross substitution between groups of commodities (belonging to different sectors of an economy) and the effect of income or budget level on their demand. Seen in this way, there is scope for complementary[3] usage of DC and CD models within the framework of a CGE model where DC models can be used to describe the preferences for a narrowly defined set of commodities, while CD models are used to describe the interactions between the demands for different groups of commodities.

In this paper, we describe how DC and CD models can be used in such an integrated fashion in a spatial computable general equilibrium model to inquire into the wider economic impacts of a transport investment project. These wider impacts are to be considered *in addition to* the usual impacts on the users of the transport network as considered in traditional (partial) benefit-cost analysis (Graham (2007a,b). In the past, there have been studies which also looked at the issue of using a DC model within the framework of a CGE framework (see for example, Horridge (1994)). However, the

---

[1] Using the Lancaster approach to consumer demand based on commodity characteristics or attributes (Lancaster, 1966). Discrete choice models therefore are often used for the study of the demand for quality-differentiated products. See for example Berry (1994), Berry *et al.* (1995, 2004).

[2] Market price is also often used in a DC model but this plays the role of only one particular attribute among many while in a continuous demand model it is the main (and often only) 'attribute'.

[3] An alternative approach is to use DC model in a 'conditional demand' mode and then trying to relate the choice elasticities to demand elasticities (see, for example, Smith *et al.* (2011)). Although this approach is not inconsistent with our approach (see, for example, section 3.2 below), it tends to restrict the usefulness of a DC model because it implies that DC and CD models are merely *substitutes* which can be used to analyse the same problem, while in fact, DC and CD models are quite fundamentally different and designed to deal with different issues. For example, CD model is not well designed to handle the issue of consumer heterogeneity or product variety, but strong in dealing with the issue of income and relative price effects. The reverse is true for DC models. Therefore DC and CD are more complements rather than substitutes.

approach so far has been limited to the use of a theoretical functional form (such as the linear logit[4]) in a CGE model to replace the use of other alternative functional forms (such as CES) to describe demand for different varieties of a particular activity or commodity (such as transport mode choice or location choice). It has not been extended to the use of an *actual* or *true* DC model in a CGE framework. Here we can define a 'true' DC model as one specified and estimated using *individual-specific discrete choice data* rather than estimated (or calibrated) using only aggregate or average market share data (as is the case of the linear logit model). The use of such a model in a CGE framework presents both challenges as well as potential advantages which will be explained in this paper.

The plan of the paper is as follows. Section 2 introduces the Multinomial Logit (MNL) discrete choice model as the basic structure used in most disaggregate behavioural models of choice behaviour. Sections 3 and 4 explain how a MNL basic structure can be considered as part of a (conditional) demand system for simple and more complex decision structures. Section 5 then illustrates the connection between DC and CD models within a CGE framework, with an empirical example taken from a study of the wider economic impacts of an urban transport investment project, and Section 6 gives some conclusions.

2. Multinomial logit (MNL) discrete choice model

A typical MNL model of discrete choice is specified as follows:

$$Prob_i = \frac{\exp(V_i)}{\sum_{j \in I} \exp(V_j)}; \quad i \in I. \tag{1}$$

$Prob_i$ is the probability of alternative i being chosen from a choice set I, and V_i is the (indirect) utility function of the choice alternative i. The indirect utility function V_i is usually specified as a linear function of all the *attributes* of the choice alternative as well as the *characteristics* of the individual who chooses this alternative[5]:

$$V_i = \sum_{m \in M} \alpha_m A_{im} + \sum_{n \in N} \beta_n B_{in}, \quad i \in I. \tag{2}$$

A_{im} stands for the attribute m of alternative i, B_{in} is the characteristic n of the individual who chooses alternative i; and α_m and β_n are parameters.[6] For example, if i is a mode choice alternative (say "bus") then A_{im} can be variables describing the travel time and cost associated with alternative "bus", and B_{in} are the variables describing the socio-economic characteristics of the individual who chooses this alternative (such as income level, whether owning a car, professional status, residential location, etc.). If we now assume that, because of the existence of other 'unobserved' characteristics of the choice alternative, the indirect utility function is a random rather than deterministic variable, consisting of the deterministic part V_i as specified in (2) and a random error term ε_i; we can define (3).

---

[4] That is, a logit model of choice behaviour which is specified and estimated using market shares data rather than discrete individual choice data (see Oum (1979)).

[5] Although product variety and consumer heterogeneity can be considered as equivalent from a theoretical viewpoint, if we look only at the aggregate (or average) behaviour of a 'representative' consumer, and if the distribution of the heterogeneous consumer preferences can described in terms of symmetrical positions in the attribute space with respect to the various choice alternatives (product varieties) (see for example, Anderson *et al.* (1989), it is still convenient to distinguish between these two concepts from an empirical point of view because in practice, a DC model describes 'consumer heterogeneity' in terms of the characteristics of the decision maker, while product variety is specified in terms of the attributes of the choice alternatives.

[6] In general, the parameters α_m and β_n are assumed to be 'generic', i.e., independent of the choice alternatives, except for the parameter of the constant term, which can be assumed to be 'alternative-specific', and this parameter can be represented by the symbol α_{i0} where $\alpha_{i0} \neq \alpha_{j0}$ for $i \neq j$.

$$U_i = V_i + \varepsilon_i \tag{3}$$

The individual is then said to choose alternative i over all other alternatives $j \neq i$ if and only if $U_i > U_j$ for all $j \neq i$. This means $(\varepsilon_i - \varepsilon_j) > (V_j - V_i)$ for all $j \neq i$. Depending on the distribution of the random error term ε_i, different choice models can be derived. For example, if ε_i's are assumed to be independently and identically distributed (i.i.d.) as a Weibull distribution[7], then the probability of condition $(\varepsilon_i - \varepsilon_j) > (V_j - V_i)$ being satisfied is given by the choice probability function (1)[8].

Often, a (complex) choice decision can involve many different layers or levels of decision; therefore, the basic MNL model can be extended into a nested structure.[9] A decision on residential location, for example, can depend not only on the attributes of the residential location choices (such as distances from the Central Business District (CBD), environmental characteristics of the locations) but also on choices regarding the types of dwelling available within each location, and also the decision on work location. The latter in turn can depend on the types of travel mode choices available for journeys to work from each residential location and to each particular work location. These interrelated choice decisions can be represented by a 'nested' MNL choice structure as shown in Figure 1.

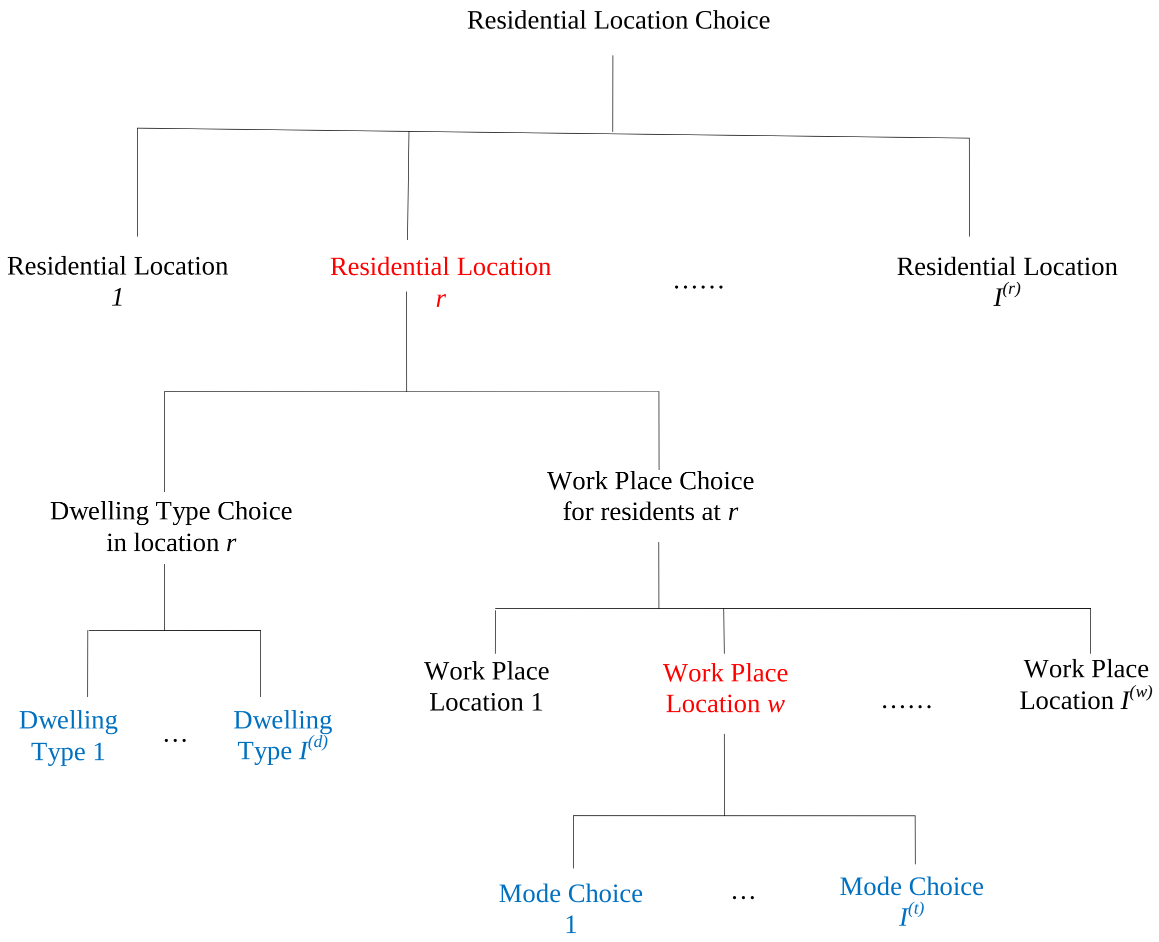
---

[7] If the distribution is normal rather than Weibull (also called extreme value type I distribution) then the choice probability function will take on a different form which is referred to as the 'probit' model.

[8] We limit the discussion to closed form choice models. The literature has advanced significantly with open-form models such as mixed (or random parameter) logit, error components logit, and scaled MNL – see Hensher and Greene (2003), Train (2003) and Greene and Hensher (2010). The focus of the current paper is on integration with a computable general equilibrium framework in order to capture economy wide impacts. In principle, more advanced choice models could be included.

[9] This is also to avoid one of the weaknesses of the basic MNL choice model, the global assumption of 'independence from irrelevant alternative' (IIA) which says that the choice of alternative i over alternative j is independent of the existence of other alternatives. This means the ratio of choice probabilities (Probj/Probj) is independent of the existence of (and therefore also the levels of the attributes of) other alternatives. To overcome this weakness, the choice set must be defined carefully so as not to contain subsets of alternatives which are more 'similar' to each other than are to others in the choice set. This means in general a 'nested' or hierarchical structure is unavoidable to group alternatives together according to degrees of 'similarity', for example, 'blue bus' and 'red bus' are grouped together before being compared to 'train' and 'car', or mode choices are grouped together before looking at location choices.

Figure 1: Nested choice structure



Let a superscript '*t*' denote travel mode choice, '*w*' to denote work location choice, '*r*' to denote residential location choice, and '*d*' to denote housing or dwelling type choice. Consider first the travel mode choice decision. Each travel model choice decision (*t*) can be assumed to be conditional (/) on a particular work-residence locations pair (*wr*). Therefore we can write:

$$Prob_i^{(t/wr)} = \frac{\exp(V_i^{(t/wr)})}{\sum_{j\in I^{(t/wr)}} \exp(V_j^{(t/wr)})}; \quad i \in I^{(t/wr)}. \tag{4}$$

$$V_i^{(t/wr)} = \sum_{m\in M^{(t)}} \alpha_m^{(t)} A_{im}^{(t/wr)} + \sum_{n\in N^{(t)}} \beta_n^{(t)} B_{in}^{(t/wr)}$$

$$\bar{V}^{(t/wr)} = \ln \sum_{j\in I^{(t/wr)}} \exp(V_j^{(t/wr)}) = \sum_{i\in I^{(t/wr)}} \frac{\exp(V_i^{(t/wr)})}{\sum_{j\in I} \exp(V_j^{(t/wr)})} \ln(\exp(V_j^{(t/wr)})) \tag{5}$$

$$= \sum_{j\in I^{(t/wr)}} Prob_j^{(t/wr)} (V_j^{(t/wr)})$$

$\overline{V}^{(t/wr)}$ is referred to as the 'logsum' or 'inclusive value' (Ben Akiva and Lerman 1977, McFadden 2001, Hensher et al., 2005) of all the travel mode choice decisions (conditional on a particular work-residence locations pair) and which stands for the *expected value* of the utility coming from all mode choice decisions. The logsum for mode choice is then used as an explanatory variable in the work place location choice to indicate the interrelationship between these decisions:

$$Prob_i^{(w/r)} = \frac{\exp(V_i^{(w/r)})}{\sum_{j\in I^{(w/r)}} \exp(V_j^{(w/r)})}; \quad i \in I^{(w/r)}. \tag{6}$$

$$V_i^{(w/r)} = \sum_{m\in M^{(w)}} \alpha_m^{(w)} A_{im}^{(w/r)} + \sum_{n\in N^{(w)}} \beta_n^{(w)} B_{in}^{(w/r)} + \gamma^{(w)(t)} \overline{V}^{(t/wr)}$$

$$\begin{aligned} \overline{V}^{w/r)} &= \ln \sum_{i\in I^{(w/r)}} \exp(V_i^{(w/r)}) = \sum_{i\in I^{(w/r)}} \frac{\exp(V_i^{(w/r)})}{\sum_{j\in I} \exp(V_j^{(w/r)})} \ln(\exp(V_i^{(w/r)})) \\ &= \sum_{i\in I^{(w/r)}} [(Prob_i^{(w/r)})(V_i^{(w/r)})] \end{aligned} \tag{7}$$

Note that the decision on work location is conditional on a given residential location choice. Finally, the residential location choice is assumed to depend not only on work place location (*w*) but also on dwelling type choice ('*d*'). The decision on dwelling type choice is given by:

$$Prob_i^{(d)} = \frac{\exp(V_i^{(d)})}{\sum_{j\in I^{(d)}} \exp(V_j^{(d)})}; \quad i \in I^{(d)}. \tag{8}$$

$$V_i^{(d)} = \sum_{m\in M^{(d)}} \alpha_{im}^{(d)} A_{im}^{(d)} + \sum_{n\in N^{(d)}} \beta_{in}^{(d)} B_{in}^{(d)}$$

$$\begin{aligned} \overline{V}^{(d)} &= \ln \sum_{i\in I^{(d)}} \exp(V_i^{(d)}) = \sum_{i\in I^{(d)}} [\frac{\exp(V_i^{(d)})}{\sum_{j\in I} \exp(V_j^{(d)})} \ln(\exp(V_i^{(d)}))] \\ &= \sum_{n\in N^{(d)}} [Prob_i^{(d)} V_i^{(d)}] \end{aligned} \tag{9}$$

and the decision on residential location choice is defined to be dependent on both the logsum of dwelling type choice $\overline{V}^{(d)}$ and the logsum of work location choice $\overline{V}^{(w/r)}$:

$$Prob_i^{(r)} = \frac{\exp(V_i^{(r)})}{\sum_{j\in I^{(r)}} \exp(V_j^{(r)})}; \quad i \in I^{(r)}. \tag{10}$$

$$V_i^{(r)} = \sum_{m\in M^{(r)}} \alpha_{im}^{(r)} A_{im}^{(r)} + \sum_{n\in N^{(r)}} \beta_{in}^{(r)} B_{in}^{(r)} + \gamma_i^{(r/d)} \overline{V}^{(d/r)} + \gamma_i^{(r/w)} \overline{V}^{(w/r)}$$

$$\begin{aligned} \overline{V}^{(d/r)} &= \ln \sum_{j\in I^{(d/r)}} \exp(V_j^{(d/r)}) = \sum_{j\in I^{(d/r)}} [Prob_j^{(d/r)} \exp(V_j^{(d/r)})] \\ \overline{V}^{(w/r)} &= \ln \sum_{j\in I^{(w/r)}} \exp(V_j^{(w/r)}) = \sum_{j\in I^{(d/r)}} [Prob_j^{(w/r)} \exp(V_j^{(w/r)})] \end{aligned} \tag{11}$$

3. Multinomial logit for use as part of a complete demand system

Although the parameters of the choice probability function in a multinomial logit discrete choice model are estimated using *individual* choice data, once estimated, the model can be used to predict *aggregate* or 'representative consumer' demand.[10] This can take two forms: either individual-specific data are fed into the discrete choice model to predict the choice *probability* for each individual and then aggregated up to a level of continuous demand; or alternatively, 'representative' individual data can be fed into the DC model to predict the choice probability for this hypothetical individual, and then using the choice probability to predict market share. Either way, a discrete *choice* model is now being used to predict continuous ('representative consumer') *demand* and a question arises: how and when is this approach valid?

Demand refers to multiple choices either by many individuals with deterministic but heterogeneous preferences on any single occasion, or by a single (hypothetical 'representative') individual with probabilistic preference on many different occasions. Therefore, demand refers to a continuous number (absolute quantity of demand or relative market share) whereas discrete choice refers only to a single discrete (0,1) decision at any one time.[11] Discrete choice, therefore, tends to indicate a *preference* level rather than demand as such. To arrive at the optimal demand decision requires some trade-off between commodities at the extensive margin with a binding budget constraint, rather than a trade-off between *attributes* of a particular choice alternative at the 'intensive' margin where the budget constraint is often *not* binding. Therefore it is more appropriate to interpret the use of a DC model as being applied to situations where only the preferences or shares between different alternatives of a specific decision are to be determined while assuming that the total aggregate level of demand for all the choice alternatives is to be determined outside of the choice model.[12]

Although DC model is fundamentally different from a CD model[13], they can be regarded as parts of a more *complete* and *accurately specified* demand system where 'demand' refers to 'optimal choices', but choices can only be optimal if (1) the constraint is specified accurately (as in a CD model), and (2) preferences are known with accuracy taking into account, not only market factors (such as price and income level), but also other individual socio-demographic characteristics and commodity quality attributes (as in a DC model). Therefore, in what follows, we consider the issue of how to combine the uses of both DC and DC models to describe consumer behaviour in a more accurate manner in the framework of a CGE model.

---

[10] See for example, McFadden and Reid (1975).

[11] Even if the decision may involve more than one decision, e.g. mode choice nested into work location choice, this is still *not* multiple decisions but rather only a single combined decision in a 'nested' structure.

[12] This suggests that concepts such as 'demand elasticities' when applied to a DC model must be appropriately qualified. Schmidheiny *et al.* (2011) for example, referred to the elasticities estimated from a DC model as 'semi-elasticities' to distinguish these from the conventional demand elasticities. Alternatively, elasticities from a DC model can be referred to as 'conditional' demand elasticities because the 'condition' on these choices is that either total *quantity* of the demand for all the choice alternatives is fixed, or the budget level for all choices is given. See, for example, Smith *et al.* (2010).

[13] See, for example, Anderson *et al.* (1989a, p. 163) where it is noted that while the (linear logit) discrete choice model is 'equivalent' to a (CES) continuous demand model when referring to aggregate representative behaviour, the two models are fundamentally different because in the former case, the individual consumer is assumed to buy only *one unit* of the differentiated products while in the latter case, he/she is assumed to spend a *fixed amount of income* (perhaps over time or through repeated exercise) on different varieties of the product.

3.1 Using a DC model to infer the 'effective' or 'generalised' price for each choice alternative and to derive an aggregate price index for all choice alternatives

The choice probability function in a DC model is often specified in terms of the indirect utility function $V_i(.)$ which charactertises the (maximum) utility level associated with each particular choice alternative *i*. This indirect utility function contains not only the 'characteristics' of the choice alternative being considered (such as travel time and travel costs, acess time and egress time, comfort and convenience level, etc. in a mode choice model) but also the (socio-economic) characterstics of the decision makers (income level, professional status, age, sex, household characteristic, etc). Leaving aside the characteristics of the decision makers which are used to account for the heterogeneity of the consumers, the characteristics of the choice alternatives can be considered as different attributes which make up the commodity in question (choice alternative).[14] Each of the attributes can be assumed to have a 'shadow price', and this shadow price is derived from the parameters of the indiract utility function estimated for each DC model. The indirect utility function in a DC model therefore can be written as:

$$V_i = V_i(\pi_{i1},...,\pi_{ik}) = V_i(P_i(\pi_{i1},...,\pi_{ik})); \quad i = 1,...,I \qquad (12)$$

where π_{ik}'s are the implicit or shadow prices of the *k* attributes of alternative *i*, and P_i is an aggregator which combines these attributes and shadow prices into an 'effective' or 'quality-adjusted' price index for the choice alternative. For example, in a mode choice model, different categories of travel time variables can be combined with travel cost (given a value of travel time savings) to obtain a 'generalised price' index for the choice alternative. The concept of generalised price can be extended to include, not only travel time and travel costs, but also comfort and convenience levels, etc. as important attributes of a travel mode. Each of these attributes has a shadow price attached to it which is inferred from the parameters of the indirect utility function (e.g., the shadow price of travel times is inferred from the ratio of the coefficients of travel time and travel costs). Therefore, in estimating the price index for each choice alternative, the travel cost ('market price') must be combined with the shadow costs of other non-market factors to arrive at an 'effective' or 'generalised' price for the choice alternative. This generalised price is in fact inferred from the indirect utility function. Seen in this way, the indirect utility function of a DC model (as shown in equation (2) or (12)) can be rewitten in a 'reduced-form' as follows:

$$V_i = \alpha_{i0} + \alpha_{i1}P_i; \quad i = 1,...,I \qquad (13)$$

where α_{i0} is the original 'alternative-specific' constant, α_{i1} is the co-efficient of the market price or money cost variable associated with the choice alternative[15], and P_i is the aggregator which combines, not only the market price but also all other non-market attributes of the choice alternatives into a 'generalised price' for the choice alternative *i* [16].

Given the generalised prices for all the choice alternatives, an *aggregate* or *average* price index for all choice alternatives can then be computed, using the so-called logsum function introduced earlier (see equations (7), (9), (11)):

---

[14] This interpretation is similar to that in the Lancaster characteristics approach to the theory of consumer demand (Lancaster 1971).

[15] Normally, this coefficient is generic, i.e. the same for all choice alternatives, therefore we can set $\alpha_{i1} = \alpha_1$ for all choice alternatives.

[16] Using the shadow prices of the attributes as estimated from the DC model. In fact, the estimation of the shadow prices of these attributes is the strength of a DC model which helps to add to the behavioural accuracy of the (complete) demand system.

$$\begin{aligned}\overline{V} &= \ln\sum_{i\in I}\exp(V_i)\\ &= \sum_{i\in I}\frac{\exp(V_i)}{\sum_{j\in I}\exp(V_j)}\ln(\exp(V_i))\\ &= \sum_{i\in I}[(Prob_i)(V_i)]\end{aligned} \tag{14}$$

Substituting (13) into (14), we have: [17]

$$\begin{aligned}\overline{V} &= \sum_{i\in I}[(Prob_i)(\alpha_{i0}+\alpha_1 P_i)]\\ &= \sum_{i\in I}[(Prob_i)(\alpha_{i0})]+\alpha_1\sum_{i\in I}[(Prob_i)(P_i)]\\ &= \overline{\alpha}_0+\overline{P}\end{aligned} \tag{15}$$

where:

$$\begin{aligned}\overline{\alpha}_0 &= \sum_{i\in I}(Prob_i)(\alpha_{i0})\\ \overline{P} &= \sum_{i\in I}(Prob_i)(P_i)\end{aligned} \tag{16}$$

$\overline{P}$ therefore can emerge as the aggregate or average price index for all choice alternatives and this index is inferred from the logsum function.

3.2 DC model as a conditional probabilistic demand model

The logsum can be interpreted as the *expected* indirect utility of demand for *all* choice alternatives, and therefore, we can apply Roy’s theorem to this expected indirect utility function to get a system of *probabilistic* demand for each choice alternative as follows:

$$\begin{aligned}X_i &= -\frac{\partial\overline{V}/\partial P_i}{\partial\overline{V}/\partial M} = -\frac{\sum_{j\in I}(\partial\overline{V}/\partial V_j)(\partial V_j/\partial P_i)}{\sum_{j\in I}(\partial\overline{V}/\partial V_j)(\partial V_j/\partial M)}\\ &= -\frac{\sum_{j\in I}(Prob_j)V_{ji}^{'}}{\sum_{j\in I}(Prob_j)V_{jM}^{'}};\quad i=1,...,I.\end{aligned} \tag{17}$$

Here X_i is the *probabilistic demand* for choice alternative *i*, $V_{ji}^{'}=\partial V_j/\partial P_i$ and $V_{jM}^{'}=\partial V_j/\partial M$; i.e., these are the derivatives of the choice indirect utility function with respect to the generalised price index and with respect to the aggregate expenditure level of demand, respectively. From equation (13) we can see that $V_{ji}^{'}=0$ if $j\neq i$, and $V_{ji}^{'}=\alpha_{i1}$ if $j=i$. The value of $V_{jM}^{'}$, however, cannot be inferred directly from a DC model but must be estimated exogenously of the DC model. Despite the lack of information on the value of $V_{jM}^{'}$, equation (17) can still be used to estimate the *share* of demand for

---

[17] Henceforth, we will assume that α_{i1} is generic rather than alternative-specific and therefore will write the coefficient for price variable as α_1 rather than $\alpha_{i1}$for simplicity of notation.

each choice alternative if in fact the *total* level of demand for all choice alternatives is assumed to be given[18]. Rewriting equation (17) in an alternative form:

$$\frac{X_i}{X_j} = \frac{(Prob_i)\alpha_{i1}}{(Prob_j)\alpha_{j1}}; \quad i, j = 1,...,I. \tag{18}$$

or, when $\alpha_{i1}$is assumed to be constant across all alternatives, i.e. $\alpha_{i1}=\alpha_1$ for all i's, this is simplified to:

$$\frac{X_i}{X_j} = \frac{(Prob_i)}{(Prob_j)}; \quad i, j = 1,...,I. \tag{19}$$

Thus, if total level of demand for all choice alternatives ($\overline{X}$) is assumed to be given, then the demand for each choice alternative can be derived from equation (19):

$$X_i = (Prob_i)\overline{X}; \quad i = 1,...,I. \tag{20}$$

Equation (20) shows that the choice probabilities in a DC model can be used to indicate the demand (quantity) shares[19] if the total level of demand for all choice alternatives is assumed to be given. When a DC model is linked to a CD model, this total quantity of demand will no longer be assumed to be 'given' but is to be estimated within the CD model *given* the aggregate price index $\overline{P}$ as derived from the DC-component. Therefore the combined DC-CD model is now a complete demand system with demand quantity and aggregate price index interconnected, as they should be in the framework of a general equilibrium model.

3.3 Example of a simple linkage between DC and CD models

Consider the simple example of Figure 2 where different mode choice decisions are assumed to be represented by a DC model *conditional* on a total level of travel demand (say, the total number of work trips originating from zone *s* and with destination in zone *z*). Similarly, different dwelling type choices can be regarded as different demands for various housing types *conditional* on a total level of demand for all housing types. Let the mode choice model be defined by equations (4)-(5), and with the indirect utility function for each choice alternative be rewritten in a form similar to equation (13), i.e.:[20]

$$V_i^{(t)} = \alpha_{i0}^{(t)} + \alpha_1^{(t)} P_i^{(t)}; \quad i = 1,...,I^{(t)}. \tag{21}$$

---

[18] And to be estimated by the CD model in the complete (general equilibrium) demand system..

[19] If a_{i1}'s are alternative-specific rather than generic (i.e. constant across all alternatives), then the quantity shares will have to be weighted by these coefficients. The a_{i1}'s being alternative-specific implies that each choice alternative is to be regarded as 'generically different' from each other. This is the extreme opposite to the case when they are considered as similar (when $a_{i1}$is assumed to be constant across all alternatives). In the latter case, the choice alternatives are distinguished only via their 'qualities' and once their price indices P_i's have been 'adjusted' for the quality differences (as implied by equations (12) and (13)) the choice alternatives can be considered as 'perfectly substitutable' and therefore their demand *quantities* can be aggregated. If on the other hand, they are regarded as generically different then the weights a_{i1}'s will stand for their relative (shadow) prices and therefore, in aggregating their demand, instead of adding the *quantities* (which will not be meaningful) one should add their total *values* or expenditures (given by $a_{i1}P_i$ where P_i is the quality-unadjusted market price and a_{i1} is the shadow value of the different qualities).

[20] For simplicity, we have assumed that the coefficient for the price variable is generic rather than alternative-specific (see also the previous footnote). The analysis can easily be extended to the case where some of these coefficients are assumed to be alternative specific. We have also dropped the 'conditional' notation (*/wr*) for simplicity in this section.

The logsum for this choice model is given by:

$$\begin{aligned}\overline{V}^{(t)} &= \ln \sum_{j\in I^{(t)}} \exp(V_j^{(t)}) \\ &= \sum_{j\in I^{(t)}} (Prob_j^{(t)})(V_j^{(t)}) \\ &= \sum_{j\in I^{(t)}} (Prob_j^{(t)})(\alpha_{j0}^{(t)}) + \sum_{j\in I^{(t)}} (Prob_j^{(t)})(P_j^{(t)}) \\ &= \overline{\alpha}_0^{(t)} + \alpha_1^{(t)}\overline{P}^{(t)}\end{aligned} \tag{22}$$

where:

$$\begin{aligned}\overline{P}^{(t)} &= \sum_{j\in I^{(t)}} (Prob_j^{(t)})(P_j^{(t)}) \\ \overline{\alpha}_0^{(t)} &= \sum_{j\in I^{(t)}} (Prob_j^{(t)})(\alpha_{j0}^{(t)})\end{aligned} \tag{23}$$

Both $\overline{\alpha}_0^{(t)}$ and $\overline{P}^{(t)}$ can be estimated from the DC model for mode choice, and $\overline{P}^{(t)}$ can be regarded as an aggregate price index for travel by all mode choices.

In a similar manner, we can also use equations (8)-(9) and the assumption underlying equation (13), to define a DC model for housing (dwelling type choices) and estimate an aggregate price index for all dwellings types:

$$V_i^{(d)} = \alpha_{i0}^{(d)} + \alpha_1^{(d)} P_i^{(d)}; \quad i = 1,..., I^{(d)}. \tag{24}$$

$$\begin{aligned}\overline{V}^{(d)} &= \ln \sum_{j\in I^{(d)}} \exp(V_j^{(d)}) \\ &= \overline{\alpha}_0^{(d)} + \alpha_1^{(d)}\overline{P}^{(d)}\end{aligned} \tag{25}$$

$$\begin{aligned}\overline{P}^{(d)} &= \sum_{j\in I^{(d)}})Prob_j^{(d)})(P_j^{(d)}) \\ \overline{\alpha}_0^{(d)} &= \sum_{j\in I^{(d)}})Prob_j^{(d)})(\alpha_{j0}^{(d)})\end{aligned} \tag{26}$$

$\overline{P}^{(d)}$ is interpreted as the aggregate price index for housing derived from the DC model for dwelling type choices.

$\overline{P}^{(t)}$ and $\overline{P}^{(d)}$ can be used in an aggregate CD model to estimate the *total* levels of demand for travel and housing activities (which have been assumed to be given in the DC models). Let the aggregate demand for these activities ($\overline{X}^{(t)}$ and $\overline{X}^{(d)}$) be described by a constant elasticity of substitution (CES) utility function:

$$U = U(\overline{X}^{(t)}, \overline{X}^{(d)}) = \left[\delta_t(\overline{X}^{(t)})^{\rho} + \delta_d(\overline{X}^{(d)})^{\rho}\right]^{1/\rho} \tag{27}$$

Here the parameters δ_t, δ_d are referred to as distribution parameters, and $\sigma = 1/(1-\rho)$ is the elasticity of substitution between these aggregate activities. Maximising the utility function (24) subject to a budget constraint:

$$\overline{X}^{(t)}\overline{P}^{(t)} + \overline{X}^{(d)}\overline{P}^{(d)} \le M \tag{28}$$

where M is the total expenditure level on travel and housing activities, will result in a system of demand equations of the following form:

$$\overline{X}^{(I)} = M(\delta_I / \overline{P}^{(I)})^{\sigma}\left[\delta_t(\overline{X}^{(t)})^{\rho} + \delta_d(\overline{X}^{(d)})^{\rho}\right]^{-1}; \quad I = \{t, d\}. \tag{29}$$

Equation (29) can also be written in the simpler 'percentage change' form (using a lower case letter to denote percentage change):

$$\begin{aligned}\bar{x}^{(I)} &= m - \sigma\bar{p}^{(I)} - \sum_{J=\{t,d\}}[S_J(1-\sigma)\bar{p}^{(J)}] \\ &= (m - \sum_{J=\{t,d\}} S_J\bar{p}^{(J)}) - \sigma[\bar{p}^{(I)} - \sum_{J=\{t,d\}} S_j\bar{p}^{(J)}] \\ &= (\sum_{J=\{t,d\}} S_j\bar{x}^{(J)}) - \sigma[\bar{p}^{(I)} - \sum_{J=\{t,d\}} S_J\bar{p}^{(J)}] \\ &= \bar{x} - \sigma[\bar{p}^{(I)} - \bar{p}]; \quad I = \{t,d\}\end{aligned} \tag{30}$$

where:

$$S_I = \frac{\bar{P}^{(I)}\bar{X}^{(I)}}{M} = \frac{(\delta_I)^{\sigma}(\bar{P}^{(I)})^{1-\sigma}}{\sum_{J=\{t,d\}}(\delta_J)^{\sigma}(\bar{P}^{(J)})^{1-\sigma}}; \quad I = \{t,d\}. \tag{31}$$

$$\bar{p} = \sum_{J=\{t,d\}} S_J\bar{p}^{(J)} \tag{32}$$

$$\bar{x} = \sum_{J=\{t,d\}} S_J\bar{x}^{(J)} \tag{33}$$

$$m = \sum_{J=\{t,d\}} S_J(\bar{x}^{(J)} + \bar{p}^{(J)}) \tag{34}$$

$\bar{x}$ measures the total income effect on aggregate demands $\bar{x}^{(t)}$ and $\bar{x}^{(d)}$, whereas the term $(-\sigma[\bar{p}^{(I)} - \bar{p}])$ measures the (aggregate) [21] substitution effect between these aggregate demands. From equation (20) we can write:

$$X_i^{(t)} = (Prob_i^{(t)})\bar{X}^{(t)} \tag{35}$$

where $X_i^{(t)}$ denotes the level of demand for choice alternative *i* (i.e. mode *i*) within the (aggregate) travel branch (*t*); $\bar{X}^{(t)}$ is the *aggregate* level of demand for travel (all modes), and $(Prob_i^{(t)})$ is the choice probability for a particular alternative *i* within this aggregate 'branch' of demand. Taking the differential of the logarithmic function (*d* ln) – which is similar to taking the perentage change – of both sides of equation (35) and using a lower case letter to denote percentage change, we have:

$$x_i^{(t)} = \bar{x}^{(t)} + d\ln(Prob_i^{(t)}) \tag{36}$$

The first term on the right hand side of equation (36) measures the income effect for choice alternative *i*, and this is derived from the aggregate demand model (30). The second term on the right hand side measures the *disaggregate* substitution (i.e., choice) effect. This second term can be estimated directly from a DC model as follows. First, assuming that the DC model for travel activities is given by equations such as (4) and (5), taking the differential log of equation (4) we have:

$$d\ln(Prob_i^{(t)}) = dV_i^{(t)} - d\bar{V}^{(t)} \tag{37}$$

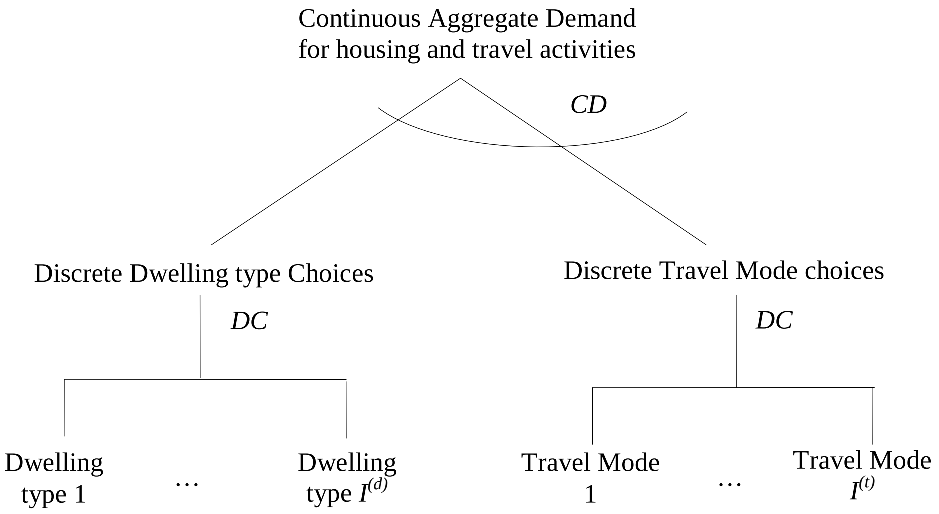
---

[21] To be distinguished from the *disaggregate* or fine substitution effects which describes only the substitution between choice alternatives within a DC model.

The term $(dV_i^{(t)})$ indicates the (absolute) *change* in indirect utility of the choice alternative *i* within the travel branch (*t*), and the term $(d\overline{V}^{(t)})$ measures the absolute change in the value of the logsum of this branch. Both of these terms are functions of the *changes* in the levels of *attributes* of the choice altenratives (as well as those of the consumer) as indicated by equation (5). Similar analysis can apply to the case of disaggregate and aggregate demand for housing activities using equations such as (8)-(9) and linked to equations such as (35)-(37) but applied to the housing activities rather than travel activities.

In summary, a DC model can be used to estimate the substitution effects between choice alternatives of a narrowly defined class of commodities (travel mode choices, dwelling type choices, etc). From the indirect utilities of the DC models, price indices of the narrowly defined choice altneratives as well as the aggregate price indices of broadly defined classes of commodities can be estimated. Using these price indices in an aggregate CD model allows one to estimate not only the aggregate income effect due to changes in income level but also the substitution effects. The linkage of DC models to a CD model therefore allows for a wide range of commodities to be considered in demand analysis with a richer analysis.

Figure 2: Discrete choice (DC) models of transport modes and dwelling types linked to an aggregate continuous demand (CD) model of transport and housing activities



4. More complex linkages between DC models and aggregate demand models within a CGE framework

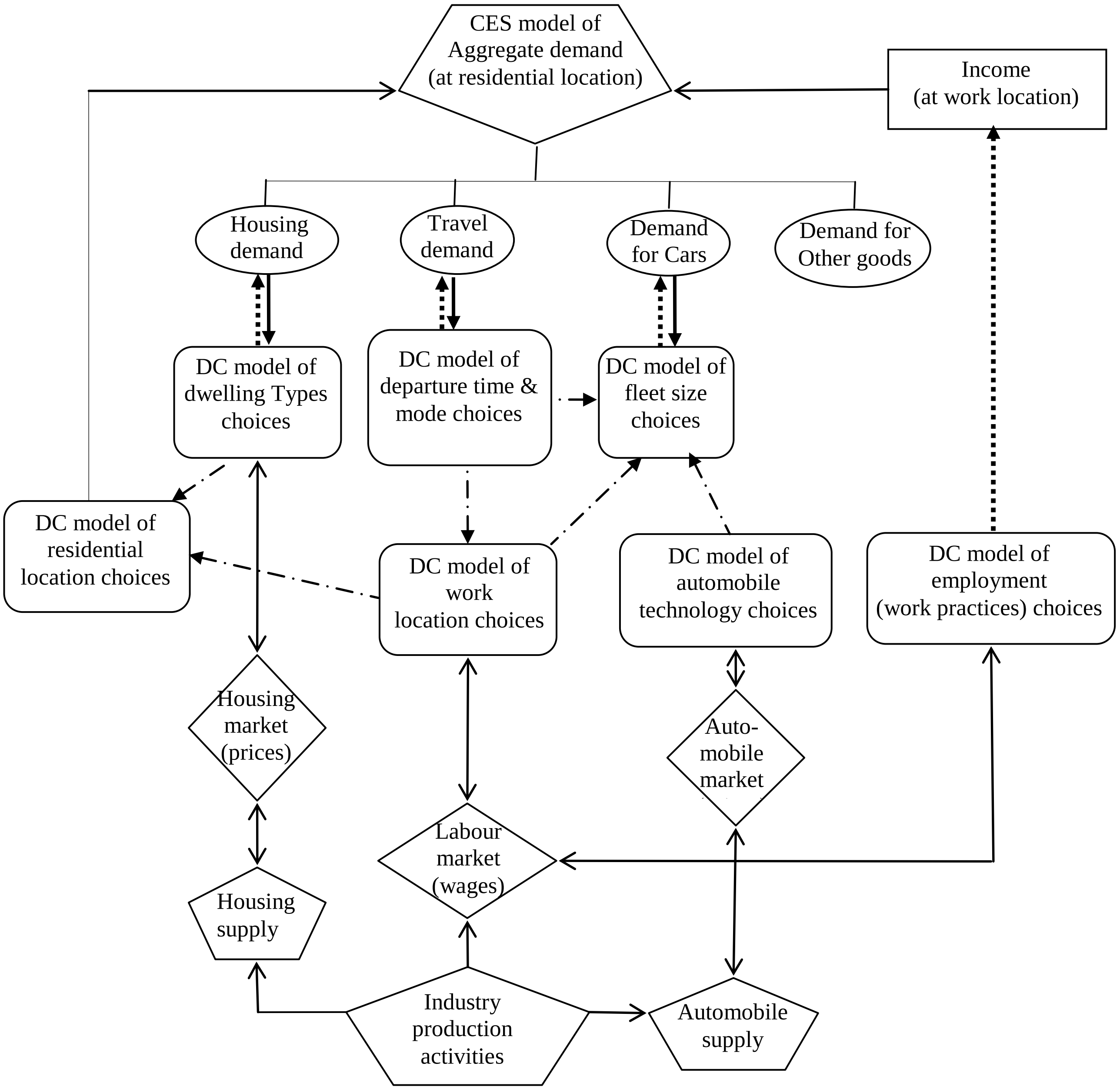
To utilise the strength of each type of model, more complex linkages between DC and aggregate demand models can be assumed, depending on the nature of a particular choice-demand situation. In the previous section, we have considered the simple linkage between housing and travel mode choice activities, with DC models describing the choices within each group of activities and a CD model linking the aggregate demand of these two groups. Consider now a more complex relationship between many activities in a local economy such as described in Figure 1. In this case, activities such as dwelling type choices and travel mode choices can be linked to a continuous demand model of housing, travel, and other goods. Other activities such as residential and work place location choices can be linked, not only to a demand system but also to production activities (supply of employment opportunities in various locations). Figure 3 is a schematic representation of all the possible linkages between DC models described in TRESIS (a Transport Environmental Strategy Impact Simulator (see Hensher (2002); Hensher and Ton (2002)) and a CGE model (SGEM or Sydney General Equilibrium Model) describing the local economy of the Sydney Metropolitan Area (SMA). [22] TRESIS and SGEM have been developed at the Institute of Transport and Logistics Studies (ITLS) in Sydney. Some of the DC models contained in TRESIS have been described in details section 1 (see Figure 1) but in addition, TRESIS also contains DC models of household automobile technology choice, fleet size choice, and employment (or work practices) choice. To link the choice behaviours of these DC models to the SMA economy, we need to extend the aggregate CD demand module within the SGEM model to include not only housing and travel activities (as done in the previous section), but also demand for automobiles and other goods. Secondly, we also need to specify a housing supply function which can be linked to the level of aggregate housing demand as specified by the DC model of housing type choice. The DC model of automobile technology choices can be linked to an automobile supply function.[23] Finally, labour markets in various zones also need to be linked to the work location choice of workers as described in the DC model of work location choice, and with the actual employment opportunities within each zone of the local economy.

---

[22] For a description of the integrated TRESIS-SGEM model, see the Appendix.

[23] An alternative is to assume that supply is given exogeneouly.

Figure 3: Schematic representation of links from discrete choice (DC) modules to continuous demand (CD) modules within a general equilibrium (GE) framework.



Notes: ▭ DC model; ⬭ aggregate quantities; ⬠ aggregate CD model;
⬠ supply activities; ◇ market equilibrium.
→ links (e.g., quantities) from continuous demand (CD) modules,
⇢ links (e.g., prices) from discrete choice (DC) modules.
–·→ links (e.g., via logsums) between discrete choice modules.
↔ general equilibrium links.

5. An illustrative experiment

To illustrate the applicability of the approach described in this paper for linking DC models to CD model within a CGE framework, we consider a simple experiment. In this experiment, we investigate the impacts of a transport investment project in the Sydney Metropolitan Area (SMA) on transport users and on the local economy. The impacts on transport users are traditonally estimated within the TRESIS module using the various DC models which are shown in Figure 1. TRESIS, however, only captures the 'conventional' impacts as considered in standard cost-benefit analysis, i.e. direct impacts on the users of the transport network, but not the indirect impacts on the wider economy which can be measured only if TRESIS is linked to a demand/supply system and imbedded within a CGE framework such as described in Figure 3.

In the experiment, we assume that the New South Wales (NSW) Government will spend money to upgrade a particular rail link in the SMA transport network. This is the so-called North-West Rail Link (NWRL) which is a 23-kilometre rail line between Epping and Rouse Hill. The project involves the construction of six new rail stations along this railway line, with approximately 3,000 park and ride spaces and bus interchange facilities to provide rail access for commuters living in the growing North West region to major employment centres in Norwest Business Park, Macquarie Park, St Leonards, Chatswood, North Sydney and the CBD. It is suggested that by providing rail access through to Rouse Hill, the new line will also support future residential and commercial development in the North West growth centre. The rail link will serve a population of 360,000, which is expected to grow to 485,000 by 2021, and by 2036, the new rail link is expected to service a region with more than 145,000 jobs (see Hensher et al. 2011 for further details[24]). The purpose of our illustrative experiment is to estimate what will be the economic impacts of the investment projects, not only on potential transport users of the network, but also on employment opportunities in the SMA.[25]

5.1 Conventional cost benefit analysis

Using DC models, conventional cost benefit analysis (CBA) looks at the impacts of a transport investment project on users of the transport network. Assume that as a result of the investment projects, users of the transport network experience some welfare improvement in their activities associated with the transport projects (for example, an increase in comfort and convenience level while travelling, a reduction in total travelling time) which can be quantified by the a willingness to pay (WTP) measure. In a DC model of transport mode choice, for example, the WTP measure is captured by the value of the maximum (indirect) utility associated with each choice alternative which will change when the transport project is implemented. The expected value of the maximum utility associated with all choice alternatives is captured by the logsum value. Therefore, the change in the logsum value (from 'before' to 'after' the project) will indicate the extent to which the transport project impacts on all users of the transport neworks travelling in all modes. This change in logsum value can then be converted into dollar terms by dividing it by the marginal utility of money, and this is then referred to as the 'compensating variation' of the transport project (for mode choice activities only): [26]

---

[24] This paper is available on request from the authors.

[25] We confine the analysis of the 'wider economic impacts' of the project to only employment opportunities because we want to illustrate the simple links between DC models (of transport mode choice, work and residential location choices) to economic activities in the local areas using the methodology presented in this paper for linking DC to CD models. The purpose of the exercise is not to conduct an exhaustive cost-benefit analysis of the transport projects, which is beyond the scope of the present paper.

[26] The use of equation (38) as a measure of CV in a DC model is conditional on the assumption that there is no or negligible income effect. In a general equilibrium approach, the income effect is then taken into account via the measure of 'wider economic impacts' as will be considered in the next section. CV is a concept first introduced by Hicks (1939). It refers to the amount of money which must be taken away from the users in the 'new state' (i.e. after a project) to make him/her indifferent between the new and old (i.e. before the project) states. After the project, the 'real income' of the user is indicated by the logsum measured at the new attribute levels. This is normally greater than the 'real income' of the user at the old state (i.e., logsum value at the old attribute levels). Therefore, the difference (as indicated by the right hand side of equation (38)) must be a measure of the CV.

$$CV^{(t)} = \frac{1}{\alpha_1}\left[\ln \sum_{j \in I^{(t)}} \exp(V_j^{(tA)}) - \sum_{j \in I^{(t)}} \exp(V_j^{(tB)}) \right] \tag{38}$$

Here α_1 is the estimated coefficient of the money cost attribute in the joint departure time and mode choice (DTMC) mode in TRESIS which is assumed to represent the marginal utility of money, the superscripts '*A*' and '*B*' in the utility functions are used to indicate the situation '*A*fter' and '*B*efore' the projects are implemented, and the superscript '*t*' refers to the *t*ravel (i.e. DTMC) model.

Since joint departure time and mode choice is only a 'short run' decision which has the capacity to influence medium and long run decisions such as work place location and also residential locations, to measure the welfare change of transport investment, it is important and more appropriate to look at the welfare changes associated with long run decisions rather than with just short run or medium run. In TRESIS, it is assumed that residential location choice is a long run decision whereas work location choice is medium run, and joint departure time and mode choice is a short run decision. Therefore, in measuring the long run impacts of the transport investment project, it is sufficient to look at just the impacts on the residential location choice decision (and this is long run impact will automatically include the medium run work location choice and short run DTMC decisions within it because of the nested structure).[27]

5.2 Wider economic impacts

Conventional CBA of a transport project does not usually include the 'wider economic impacts' (WEI) of a project on the rest of the local and national economies, partly because it often relies on a partial equilibrium approach. With the integration of DC to CD models within a CGE framework, the stage is now set for the analysis of the WEIs of a transport project. In considering the wider economic impacts of the NWRL projects, we focus attention mainly on some local issues, such as the *redistribution* of residential locations and employment opportunities among the various geographical zones of the SMA. [28] Although some of these impacts are captured in the DC models of residential and work place location choices (see previous section), only with their links to a GE model can the WEIs on the local economy be assessed. First, given the improvement in the NWRL rail link resulting from the transport investment projects, the short run departure time and mode choice (DTMC) decisions of workers may change. These changes feed onto the medium and long run decisions of work location and residential location choices (see Figure 3). Work location choice decisions affect the supply of labour in various locations, and it is assumed that the firms' demand for labour in these location will respond to the changes in supply so that there will be new equilibrium in the labour markets in these locations. This will mean a redistribution of employment opportunities among the various locations with some potential impacts on the equilibrium wage level due to agglomeration (or dis-agglomeration) effects (Venables, 2007). However, even without the agglomeration effects, a redistribution of employment opportunities among the various zones of the SMA will still result in a net increase (or decrease) in the total wage bill of the workers arising from the fact that different locations may have already experienced the agglomeration effects in the past, and therefore the wage levels offered in these different locations will be different[29] (for the same occupation and industry). The total wage bill can be used to indicate the total level of labour productivity in the local economy, and therefore, any net changes in the total wage bill can be used to indicate the WEI of the transport project on employment productivity in the local economy.

---

[27] See equations (10)-(11) where the logsums of work location choice and departure time and mode choice decisions are included within the indirect utility functions of the residential location choice decision, therefore, the welfare impacts on work location choice and departure time and mode choice decisions are automatically included within the welfare impacts on residential location choice.

[28] Other impacts such as employment *growth* and economic development generally are not considered due to the limited scope of the paper. The primary focus of the paper is only in presenting a methodology for linking DC models to CD models within the framework of a CGE model and this is achieved in linking a 'bottom-up' model (TRESIS) to a 'top-down' economy-wide model (SGEM) even though the scope of the SGEM model is still limited only to the Sydney Metropolitan Area.

[29] As reflected in the base year data.

Another component of the WEI which can be considered is the effects on residential location choices. With demand for housing (dwellings of different types) being affected by the changes in residential location choices (resulting from the transport investment project), housing prices may change, and this may impact on rent levels in various locations. From the analysis in Venables (2007), however, it can be shown that these impacts are not *in addition to* the impact on wages and employment, but rather they are merely a redistribution of the impacts on workers' wage level to housing rent, on the one hand, and transport costs on the other hand. Therefore, although it would be interesting to record these changes, they are not added to the already measured WEI of the transport project on employment productivity in the local economy.

Finally, transport network improvement may also affect the decision on automobile ownership (through vehicle type choice and fleet size choice decisions, see Figure 3). These decisions are also closely connected to the decision on consumption of other goods (such as housing and travel activities considered above). The welfare effects of these consumption decisions are only part of the changes in optimal consumption decision flowing on from changes in the wage rate (or income), and therefore are part of the changes in welfare (indirect utility) resulting from employment and productivity changes. Therefore, they are not to be counted as 'in addition' to the welfare effects of changes in employment and wage rate as already considered.

5.3 Results

Tables 1-2 show the effect of the transport improvement in the rail link between zone 1 (Inner Sydney) and zone 10 (Blacktown Baulkham Hills) (See Figure A1 for zone locations) resulting from the NWRL investment project on mode choice decisions of workers living, not only in these two zones, but also in all other zones. Quite clearly, the probability of a worker living in zone 10 choosing rail as the preferred mode of transport (to travel outside of zone 10) will increase following the improvement in this rail link, which allows them to go to and from, not only the inner city (zone 1) but also all other zones (see column and row 10 of Table 1). Network equilibrium analysis implies that workers living in zones other than zone 10 will also be affected (although to a lesser degree). For example columns 5, 7, 8 of Table 1 show that workers travelling *to* zones 5 and 7 (Fairfield_Liverpool and Inner_West_Sydney) will be adversely affected, while those travelling *to* zone 8 (Central_West_Sydney – which is adjacent to zone 10) will be favourably affected. On average the last column of Table 1 shows that train trips from origin zones 1, 4, 5, 7, 9-12 will increase, while those to zones 13-14 will decrease. Train trips to all destination zones will increase (see last row of Table 1), particularly for zones 8-10. Table 2 shows the percentage change in choice probability for other modes of transport[30] between different zones following the improvement in the rail link between zones 1 and 10. Quite clearly, trips by other modes to and from zones 1 and 10 will be adversely affected because of the (mode choice) substitution effect (see rows and columns number 1 and 10 of Table 2). However, trips to and from other zones will also be affected because of a combination of the mode choice substitution and total trip effect (i.e. substitution between destinations and/or origins due to work location and residential location choice effects – discussed below). For example, it is interesting to note that trips to zone 7 (Inner_West_Sydney) are adversely affected, both for train (Table 1) and for other modes (Table 2). In fact, the changes for destination zone 7 in Tables 1 and 2 are similar, indicating that these are dominated by the total trip (work location choice) effect rather than by the mode choice effect (see column 7 of Table 3 below).

Work location choice decisions are affected by the transport network improvement, and this is shown in Table 3. Here it is interesting to observe that as a result of the transport improvement between zone 1 and zone 10, workers may now prefer to work in zones other than zones 1 and 10. In fact, zone 10 is now not preferred as a work place as compared to, say zone 8 (Central West Sydney) or zone 4 (Canterbury_Bankstown). Other zones (5, 7) also lose preference as a work place in addition to zone

---

[30] Because of the symmetry in the MNL choice functional form, all other modes will experience that same substitution effect (percentage change in choice probability) following the shock to train characteristics. However, *on average*, because of the different number of trips by these modes, the effects will not be the same (hence they are not shown in Table 2 as they are shown in Table 1 for the train mode).

10. This shows that the impacts of an investment in a particular link in a transport network may have the potential to flow on to other links to affect other zones.

Residential location decision is also affected by the transport improvement but to a much smaller extent than the work location decision (see Table 4). This is because residential location is considered to be a longer term decision as compared to the work location decision; hence for a one-period simulation (assumed to be one year) the effect on residential location is very small (as can be seen from Table 4). However, even though the magnitude of the change can be small, the direction of the change may still be interesting and significant to show which zones are likely to be affected by the transport improvement. From Table 4, it is seen that even though zone 1 (Inner_Sydney) will now become a preferred place of residence following the transport improvement between zones 1 and 10, zone 10 (and also zone 5), however, have become less preferred places of residence following the improved rail link between zones 1 and 10, perhaps as a result of substitution between zones, but also because of other (more general) equilibrium effects, such as work location decision (which affects residential location decision – see Figure 3). This shows the importance of considering different (discrete choice) decisions within the framework of a more general equilibrium model with detailed interactions rather than separately as partial decisions.

Table 5 shows the changes in fleet size for households living in different zones following the improvement in the rail link between zones 1 and 10 as a result of the NWRL project. It can be seen that households residing in zones 1, 3-6, 8, 10-11, 13-14 will now prefer a single car fleet rather than multi-car fleet for the household, while the opposite is true for households residing in zones 2, 7, 9, and 12. This can partly be explained by the fact that following the improvement in public transport (rail link), people travelling to work and residing in zones 1, 5, 7-11 (see last column of Table 1) will now tend to switch to train and therefore there is less need for the ownership of a second vehicle for travelling to and from work. This encourages a switch to single-car fleet size. Consistent with this switch to single-car fleet size is a switch also to larger-sized car, as seen from Table 6, for people living in zones 1, 3-6, 8, 10-11, 13-14.

To measure the benefits of the NWRL project, Table 7 shows the values of the various types of impacts, which can be expected following the completion of the transport project.[31] To facilitate the identification and distinction between these different types of impacts, we define three different scenarios for the experiment and use the integrated TRESIS-SGEM model to simulate these scenarios. In the first (Base Case) scenario, we assume that there are no (new) agglomeration effects resulting from the improvement in the transport network. This means any changes in work location[32] following from the improvement in the transport network are only to take advantage of the existing wage differentials between different zones (as reflected in the base year data). Considering a land use model such as described in Venables (2007), wage differentials will reflect differences in housing rent plus transport costs. Therefore, following changes in transport costs, and assuming existing wage level are unchanged, the only change which can follow are changes in (equilibrium) housing rent. TRESIS-SGEM does not measure rent directly, however, we can still infer or ‘impute’ the (total)

---

[31] Only for the first year following the completion of the project. The purpose of the experiment is to test run the model using the methodology for integrating DC and CD models within a CGE framework therefore, the model is only run in a ‘comparative static’ mode which can stand for the experiment of a ‘what if’ scenario but not a forecast. Depending on the specific assumptions which are used to define a particular ‘closure’ for the experiment, the results can be considered as ‘short run’ or ‘long run’ results. In this case, we assume no change in ‘long run’ economic variables such as capital and non-labour resource utilisation (in non-transport industries) - except for the (completion) of the infrastructure project, therefore, the only economic variable which is allowed to change is employment location only (Scenario A and B) and total employment for the SMA (Scenario C). Housing (residential locations) activities are allowed to change but this is assumed to involve no ‘wider economic impacts’ in the form of new housing construction which will flow on to other sectors of the economy. Vehicle type choice and fleet size choices are allowed (as are reported in Tables 5 and 6) but they are not linked to any other economic activities in the economy (such as import of cars). These assumptions are made for reasons of simplicity which does not undermine the validity of the methodology used in the model, although future extensions of the model can certainly build upon this initial prototype to allow for more sectors of the economy to be included.

[32] Residential location choice is not influenced by wage level but only by characteristics such as distance from the CBD, and housing type choice. Mode choice influences residential location choice only via work location decisions (see Figure 3) hence in Scenario A, the effect of mode choice on residential location choice is only secondary and tend to be negligible as indicated by a negligible change in the value of the logsum for residential location choices.

change in equilibrium rent from the (total) change in wage bill for the workers who use the network as a whole. This is estimated to be an increase of $3.86 million. This increase therefore can reflect either the value of the improvement in the transport network and/or the total change in housing rent. Either way, these are the benefits which would fall under the heading of 'transport and land use impacts' (TLUI) and therefore the increase in total wage bill in this case is also considered as part of the conventional measure of TLUI. Therefore, there is no 'mark-up' of WEI (increase in wage bill) over the TLUI (housing rent and transport cost impacts) because the two are in fact the same.33

Next, we consider Scenario A where we assume there are agglomeration effects due to the changes in worker locations (as considered in the Base Case).34 However, we continue to assume that agglomeration only affects the redistribution of the total level of employment among the various zones rather than affecting the total level of employment itself. To calculate the agglomeration effects, we first estimate the change in employment density for each zone as a result of the redistribution of employment locations. Next we use the estimated measures of agglomeration elasticities for the Sydney Metropolitan Area (see Appendix for details) to estimate the change in wage rate which results from the changes in labour productivity which follows from the agglomeration effect. The change in the total wage bill in this Scenario is $4.38 million which, as argued in the base case can be considered as either a WEI or in fact is just reflecting the imputed rent plus the benefit arising from reduced transport costs. In addition to the increase in these benefits, however, there is now a new component of TLUI, and this is reflected in the change in the value of the logsum of all residential location choices, estimated by the RLC model at $27.644 million.35 This change in logsum value is traditionally used in TRESIS (in a stand-alone mode) to measure the total benefits of a transport investment project (as explained in the previous section – see equation (33)). In TRESIS-SGEM, however, these benefits must now also include an additional element, represented by the change in total wage bill ($4.38 million) which can stand for either the change in (imputed) rent arising from improvement in transport costs and/or the benefits of these transport costs improvements themselves (as explained for the Base Case). In other words, the total of $4.38 million can be distributed to the benefits of the land and property owners who benefit from the increase in rent value following the transport improvement project (which the government can tax), and/or it can completely or partly remain with the household in the form of increased income (which again, the government can also tax). Since these benefits are not usually included in the logsum value of the residential location choice decision (see Figure 3 where work decision and the labour market is considered to be part of the 'wider economy' which is not included within the (partial) framework of the nested DC models) they are regarded as additional to the conventional RLC benefits, even though they may belong to the same category of welfare improvements (in transport cost and housing sector). If *all* of the extra $4.38 million benefits are 'spent' on housing rent or 'redistributed' to the increase in public transport revenue and private transport expenditure,[36] then total TLUI in this case can be considered as consisting of $27.64 million in RLC benefits plus $4.38 million in transport and housing rental value accruing to both the government and the private sector of the local economy. In this case, the

---

[33] The Base Case, therefore, can be considered as a 'recalibration' of the model and data base to a new equilibrium which takes into account the new transport parameters and with existing wage levels in the SMA economy assumed unchanged. This will then 'recalibrate' the implied values of housing rent.

[34] Changes in work locations are assumed to be matched with changes in employment opportunities within these locations, i.e. the model is '(labour) supply driven' rather than demand driven. Future extension of the model can consider the issue of labour demand by looking at the industry structures in the SMA economy.

[35] This value is derived as follows. The change in the logsum of individual (probabilistic) residential location decision is specific to each household, which is characterised by the occupation of the main worker in the household and the industry in which the worker is employed. As a result, to calculate the *aggregate* change in these logsums, the individual-specific logsum must be multiplied by the total number of household in each (occupation, industry) category and summed up over all categories. The gives an expected value in utility terms of the aggregate change which is then converted to dollar term by dividing this by the coefficient of the money cost term in the mode choice model (which stands for the marginal utility of money in the *system* of interconnected (nested logit) discrete choice models (as described in Figure 3).

[36] For example, land lords can increase the rent after the transport improvement project, and/or public transport providers can increase the fare to recapture some of the (time, convenience) benefit of the transport project to partly cover some of its cost, car drivers can drive more on some links now that these links may become less congested.

‘residual’ WEI will be estimated as zero, and therefore the ‘mark-up’ of WEI over conventional benefits will be 0%. On the other hand, if all of the total wage bill benefit, which go to the workers, are spent on commodities in sectors other than transport and land use, then all of this increase in the wage bill can be regarded as ‘WEIs’ of the project. The mark-up of WEIs over conventional benefits in this case will be \$4.38/\$27.64 = 15.8%, and this mark-up arises partly from a reorganisation of employment opportunities in the economy to take advantage of changing transport network conditions and ‘old’ wage differentials, and partly from the ‘new’ agglomeration effects. Neither of these activities is traditionally considered in conventional TLUI analysis.

Finally, we consider Scenario B where the assumption of constant total employment in the local economy is relaxed. This allows for the productivity improvement - arising from the improvement in the transport network and the ‘new’ agglomeration effects generated from that improvement – to flow on to the wider (although still local) economy, and therefore to generate *additional* output and employment to the economy. The increase in total employment in this case is still very small[37] of the order of .001%. Nevertheless, despite this small increase in total employment, the increase in the total wage bill is now \$0.41 million more than the previous Scenario. In the mean time, however, the TLUI measure of RLC logsum change has now decreased slightly from \$27.644 million down to \$27.373 million. [38] These changes can perhaps be explained by the ‘general equilibrium effects’ in the model where part of the feedback from increased total employment and wage bill to the residential location choice decisions are considered to be negative rather than positive. For example, as a result of the change in total employment, dwelling prices in most zones in Scenario B are now ‘worse’ than in Scenario A, i.e. if there were increases in dwelling prices (percentage increase is positive in Scenario A) the increases are now (slightly) greater in Scenario B, and if there were decreases, then the decreases will now be (slightly) smaller. The direction of dwelling price change therefore is considered to be ‘worse’ in Scenario B as compared to Scenario A (see the last three columns of Table 8).

6. Conclusion

In this paper we have presented a methodology for linking a disaggregate discrete choice (DC) model to an aggregate continuous demand (CD) model, and integrated both types of models into a common framework of a computable general equilibrium (CGE) model. The methodology is important and useful because, from a theoretical viewpoint as well as in practice, both types of models are often designed and estimated to deal with different issues or in different areas of study. For example, a DC model is often designed to look at heterogeneous consumer preferences for a variety of products which are to be differentiated mainly by quality attributes rather than just simple market price, while the CD model is better suited to deal with the question of ‘representative’ consumer demand for aggregate commodities which can be ‘distinguished’ only through aggregate market prices.

DC models are often used to deal with disaggregate or intra-sectoral decisions (such as decisions on mode choices in the transport sector, or land uses in the housing sector), whereas CD models are better suited to handle the issue of aggregate income and substitution effects in demand decisions for different groups of commodities in different sectors of an economy. Both types of models are therefore necessary and useful for the study of policy issues which require a detailed look at the

---

[37] This is partly because the model is run in a ‘short run’ mode where capital investment is not taking place to take advantage of labour productivity improvement to generate additional output and allow the local economy to expand. The issue of economic growth, however, is beyond the scope of the current paper even though from a theoretical viewpoint, it can be incorporated into the model (for example, by considering growth as an exogenous shock rather than endogenous process).

[38] These changes can be explained partly by the ‘general equilibrium effects’ of the model where part of the feedback from increased total employment and wage bill for the workers to the residential location choice decisions can be regarded as negative (perhaps arising out of increased housing prices in some zones to reflect increased demand for various types of dwelling). Average is partly because the model is run in a ‘short run’ mode where capital investment is not taking place to take advantage of labour productivity improvement to generate additional output and allow the local economy to expand. The issue of economic growth, however, is beyond the scope of the current paper even though from a theoretical viewpoint, it can be incorporated into the model (for example, by considering growth as an exogenous shock rather than endogenous process).

individual behavioural responses (such as in the transport land use activities), while at the same time keeping track of the wider economy impacts of these decisions and feedback from the wider economy back to individual decisions.

Although in our experiments we illustrate the usefulness of the methodology mainly in the context of some transport issues (such as how to measure the wider economy impacts of a transport infrastructure investment project), and confine the analysis to this simple exercise, the methodology is potentially applicable to other wider issues (such as how to link transport decisions and policies to the wider environmental objective of greenhouse gas emissions reduction and mitigation of climate change).

Significant policy impacts are often evaluated mostly at the aggregate economy wide level, yet effective policy instruments are often designed only at the disaggregate individual consumer or producer level. Therefore, to be able to combine both a study of effective policy instruments with a study of significant policy impacts, it is important to be able to combine the use of both DC models and CD models (or bottom-up and top-down models) within a common framework, such as that of a computable general equilibrium model.

Origin Zone		Destination Zone No.														
No.	Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Average over all destinations
		Inner_Sydney	Eastern_Subs	StGge_Suther	Canter_Banks	Fairfd_Livrp	Outer_SW_Syd	Inner_W_Syd	Centrl_W_Syd	Outer_W_Syd	Blck_Baulk_H	Lower_N_Syd	Horns_Kuring	Nth_Beaches	Gosfrd_Wyong	
1	Inner_Sydney	0.00	0.00	0.15	0.82	-0.70	0.20	-2.31	1.83	0.25	3.96	0.00	0.29	0.15	0.21	0.38
2	Eastern_Subs	0.00	0.00	0.05	0.71	-0.73	0.14	-2.31	1.72	0.14	12.43	0.00	0.21	0.14	0.03	0.05
3	StGge_Suther	0.00	0.00	-0.03	0.64	-0.90	0.06	-2.60	1.73	0.02	18.82	0.00	0.10	-0.01	-0.03	0.05
4	Canter_Banks	0.00	0.00	-0.22	0.48	-1.06	-0.11	-2.74	1.55	-0.17	1.12	0.00	-0.11	-0.17	-0.21	0.13
5	Fairfd_Livrp	0.00	0.00	0.18	0.83	-0.60	0.31	-2.41	1.90	0.28	17.30	0.00	0.35	0.22	0.19	0.50
6	Outer_SW_Syd	0.00	0.00	-0.08	0.63	-0.87	0.09	-2.62	1.71	0.03	18.84	0.00	0.11	-0.02	-0.08	0.01
7	Inner_W_Syd	0.00	0.00	0.97	1.69	0.08	0.99	-1.30	2.67	1.03	14.55	0.00	1.27	1.15	0.90	0.85
8	Centrl_W_Syd	0.00	0.00	-0.66	0.10	-1.45	-0.47	-3.22	1.11	-0.53	17.75	0.00	-0.52	-0.58	-0.67	0.22
9	Outer_W_Syd	0.00	0.00	-0.08	0.67	-0.85	0.14	-2.64	1.71	0.04	18.53	0.00	0.11	0.00	-0.03	1.56
10	Blck_Baulk_H	4.77	13.30	20.16	3.18	17.13	20.40	11.11	21.06	19.15	-13.54	6.24	9.70	20.28	17.29	5.65
11	Lower_N_Syd	0.00	0.00	0.00	0.73	-0.80	0.20	-2.45	1.78	0.10	4.28	0.00	0.25	0.11	-0.01	0.40
12	Horns_Kuring	0.00	0.00	-0.11	0.61	-0.93	0.05	-2.64	1.69	-0.02	8.20	0.00	0.08	-0.04	-0.11	0.12
13	Nth_Beaches	0.00	0.00	-0.06	0.65	-0.88	0.10	-2.61	1.75	0.04	18.86	0.00	0.11	0.00	-0.05	-0.38
14	Gosfrd_Wyong	0.00	0.00	-0.06	0.66	-0.86	0.11	-2.60	1.76	0.05	19.61	0.00	0.11	0.00	-0.01	-0.17
	Average over all origins	0.25	0.00	0.06	0.07	0.67	0.03	0.62	1.05	2.57	4.00	0.47	0.28	0.13	0.15	0.58

Table 1: Percentage change in the total number of work trips by TRAIN between origin-destination zones following the improvement in the rail link between zone 1 and zone 10.

Origin Zone		Destination Zone No.													
No.	Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14
		Inner_Sydney	Eastern_Subs	StGge_Suther	Canter_Banks	Fairfd_Livrp	Outer_SW_Syd	Inner_W_Syd	Centrl_W_Syd	Outer_W_Syd	Blck_Baulk_H	Lower_N_Syd	Horns_Kuring	Nth_Beaches	Gosfrd_Wyong
1	Inner Sydney	0.00	0.00	0.15	0.82	-0.70	0.20	-2.31	1.83	0.25	-4.57	0.00	0.29	0.15	0.21
2	Eastern _Subs	0.00	0.00	0.05	0.71	-0.73	0.14	-2.31	1.72	0.14	-0.78	0.00	0.21	0.14	0.03
3	StGge_Suther	0.00	0.00	-0.03	0.64	-0.90	0.06	-2.60	1.73	0.02	-0.97	0.00	0.10	-0.01	-0.03
4	Canter_Banks	0.00	0.00	-0.22	0.48	-1.06	-0.11	-2.74	1.55	-0.17	-1.20	0.00	-0.11	-0.17	-0.21
5	Fairfd_Livrp	0.00	0.00	0.18	0.83	-0.60	0.31	-2.41	1.90	0.28	-2.48	0.00	0.35	0.22	0.19
6	Outer_SW_Syd	0.00	0.00	-0.08	0.63	-0.87	0.09	-2.62	1.71	0.03	-0.96	0.00	0.11	-0.02	-0.08
7	Inner_W_Syd	0.00	0.00	0.97	1.69	0.08	0.99	-1.30	2.67	1.03	-5.18	0.00	1.27	1.15	0.90
8	Centrl_W_Syd	0.00	0.00	-0.66	0.10	-1.45	-0.47	-3.22	1.11	-0.53	-1.92	0.00	-0.52	-0.58	-0.67
9	Outer_W_Syd	0.00	0.00	-0.08	0.67	-0.85	0.14	-2.64	1.71	0.04	-1.26	0.00	0.11	0.00	-0.03
10	Blck_Baulk_H	-3.82	0.00	0.14	0.81	-2.65	0.35	-8.14	0.71	-0.86	-0.24	-2.93	-0.27	0.25	-3.44
11	Lower_N_Syd	0.00	0.00	0.00	0.73	-0.80	0.20	-2.45	1.78	0.10	-4.76	0.00	0.25	0.11	-0.01
12	Horns_Kuring	0.00	0.00	-0.11	0.61	-0.93	0.05	-2.64	1.69	-0.02	-1.64	0.00	0.08	-0.04	-0.11
13	Nth_Beaches	0.00	0.00	-0.06	0.65	-0.88	0.10	-2.61	1.75	0.04	-0.93	0.00	0.11	0.00	-0.05
14	Gosfrd_Wyong	0.00	0.00	-0.06	0.66	-0.86	0.11	-2.60	1.76	0.05	-0.92	0.00	0.11	0.00	-0.01

Table 2: Percentage change in the total number of work trips by modes OTHER than train between origin-destination zones following the improvement in the rail link between zone 1 and zone 10.

Origin Zone		Destination Zone No.													
No.	Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14
		Inner_Sydney	Eastern_Subs	StGge_Suther	Canter_Banks	Fairfd_Livrp	Outer_SW_Syd	Inner_W_Syd	Centrl_W_Syd	Outer_W_Syd	Blck_Baulk_H	Lower_N_Syd	Horns_Kuring	Nth_Beaches	Gosfrd_Wyong
1	Inner_Sydney	0.00	0.00	0.08	0.71	-0.59	0.08	-2.13	1.59	0.07	-0.79	0.00	0.27	-0.05	0.30
2	Eastern_Subs	0.00	0.00	-0.03	0.51	-0.51	-0.31	-1.98	1.47	0.57	-0.54	0.00	0.14	-0.05	-0.02
3	StGge_Suther	0.00	0.00	-0.02	0.57	-0.94	0.03	-2.48	1.62	0.14	-0.97	0.00	0.11	0.01	-0.04
4	Canter_Banks	0.00	0.00	-0.21	0.46	-1.04	-0.05	-2.59	1.50	0.06	-0.92	0.00	-0.16	-0.16	-0.05
5	Fairfd_Livrp	0.01	0.00	0.22	0.71	-0.50	0.29	-2.38	1.77	0.37	-1.03	0.00	0.31	0.30	0.19
6	Outer_SW_Syd	0.01	0.00	-0.10	0.57	-0.85	0.10	-2.48	1.62	0.07	-0.84	0.01	0.13	0.00	-0.07
7	Inner_W_Syd	0.00	0.00	0.62	1.45	-0.20	0.42	-1.06	2.18	0.86	-1.49	0.00	1.23	1.13	-0.24
8	Centrl_W_Syd	0.00	0.00	-0.67	0.14	-1.35	-0.45	-3.15	1.07	-0.23	-1.39	0.00	-0.50	-0.18	-0.68
9	Outer_W_Syd	0.01	0.00	-0.08	0.60	-0.83	0.17	-2.67	1.63	0.06	-0.89	0.01	0.18	0.04	-0.03
10	Blck_Baulk_H	0.00	0.00	0.14	0.74	-1.04	0.31	-3.62	1.71	0.12	-0.51	0.00	0.27	0.17	0.32
11	Lower_N_Syd	0.00	0.00	-0.03	0.60	-0.82	0.40	-2.35	1.57	0.05	-1.09	0.00	0.30	0.09	-0.59
12	Horns_Kuring	0.00	0.00	-0.13	0.53	-0.97	-0.05	-2.46	1.61	0.00	-1.10	0.00	0.09	-0.02	-0.04
13	Nth_Beaches	0.00	0.00	-0.07	0.60	-0.87	0.03	-2.43	1.69	0.18	-0.67	0.00	0.13	0.00	0.08
14	Gosfrd_Wyong	0.01	0.00	-0.19	0.58	-0.82	0.08	-2.56	1.69	0.03	-0.88	0.01	0.10	-0.03	-0.01

Table 3: Percentage change in choice probability for a work place location (destination zone) given a residential location (origin zone) following the improvement in the rail link between zone 1 and zone 10.

Residential Zone		Occupation of Worker								
No.	Name	1	2	3	4	5	6	7	8	Average
		Managers	Professnals	TechTrades	CommPersServ	ClericlAdmin	SalesWorkers	MachOperDriv	Labourers	
1	Inner_Sydney	0.001	0.002	0	0	0.001	0.001	0	0	0.001
2	Eastern_Subs	0.001	0.001	0	0	0.001	0	0	0	0.001
3	StGge_Suther	0	0.001	0	0	0.001	0	0.001	0	0.001
4	Canter_Banks	0	0.001	0	0.002	0.001	0.001	0.001	0	0.001
5	Fairfd_Livrp	-0.009	-0.003	0	-0.004	0	-0.002	0	-0.002	-0.002
6	Outer_SW_Syd	0.006	-0.006	0	0.002	0	0.002	0.005	0.002	0.001
7	Inner_W_Syd	0	0.001	-0.001	-0.002	0	-0.001	-0.003	-0.004	0
8	Centrl_W_Syd	0	0.001	0.001	0.003	0.001	0.003	0.004	0.004	0.002
9	Outer_W_Syd	0.004	-0.003	0.001	0	-0.003	0	-0.003	0.001	-0.001
10	Blck_Baulk_H	-0.002	0	-0.001	-0.002	0	-0.002	-0.003	-0.005	-0.002
11	Lower_N_Syd	0.001	0.001	0	0	0.001	0.001	0	0	0.001
12	Horns_Kuring	0	0.001	0	0	0.001	0	0.002	0.001	0.001
13	Nth_Beaches	0	0.001	0	0	0	0	-0.001	0.001	0
14	Gosfrd_Wyong	0.001	-0.006	0	-0.001	-0.002	0	0	0.003	-0.001

Table 4: Percentage change in choice probability for residential locations by occupations and on average (over all occupations), following the improvement in the rail link between zone 1 and zone 10.

Residential Zone		Before project Shares of Fleet sizes		After Project Shares of Fleet sizes		% change in Shares of Fleet sizes	
No.	Name	Single car	Multi-cars	Single car	Multi-car	Single car	Multi-car
1	Inner_Sydney	0.99	0.01	1.00	0.00	0.57	-56.89
2	Eastern_Subs	0.98	0.02	0.98	0.02	-0.10	5.50
3	StGge_Suther	0.81	0.19	0.91	0.09	12.67	-52.83
4	Canter_Banks	0.82	0.18	0.93	0.07	14.18	-63.84
5	Fairfd_Livrp	0.76	0.24	0.82	0.18	8.36	-26.39
6	Outer_SW_Syd	0.61	0.39	0.79	0.21	30.26	-47.00
7	Inner_W_Syd	0.77	0.23	0.74	0.26	-4.81	16.43
8	Centrl_W_Syd	0.73	0.27	0.97	0.03	32.60	-89.42
9	Outer_W_Syd	0.60	0.40	0.24	0.76	-59.57	87.74
10	Blck_Baulk_H	0.71	0.29	0.96	0.04	34.51	-85.01
11	Lower_N_Syd	0.78	0.22	0.92	0.09	16.90	-60.85
12	Horns_Kuring	0.66	0.34	0.61	0.39	-7.72	15.05
13	Nth_Beaches	0.72	0.28	0.80	0.20	11.85	-29.86
14	Gosfrd_Wyong	0.60	0.40	0.88	0.12	47.22	-69.70

Table 5: Fleet size choice for households living in different zones before and after the NWRL project

Residential Zone		Before project Shares of vehicle types			After Project Shares of vehicle types			% change in Shares of vehicle types		
No.	Name	small	Medium	large	small	Medium	large	small	Medium	large
1	Inner_Sydney	0.202	0.455	0.343	0.173	0.468	0.359	-0.029	0.014	0.016
2	Eastern_Subs	0.202	0.455	0.343	0.205	0.453	0.342	0.002	-0.001	-0.001
3	StGge_Suther	0.202	0.455	0.343	0.177	0.466	0.357	-0.025	0.012	0.013
4	Canter_Banks	0.202	0.455	0.343	0.159	0.475	0.366	-0.043	0.020	0.023
5	Fairfd_Livrp	0.202	0.455	0.343	0.188	0.461	0.351	-0.014	0.006	0.008
6	Outer_SW_Syd	0.202	0.455	0.343	0.185	0.462	0.353	-0.017	0.008	0.009
7	Inner_W_Syd	0.202	0.455	0.343	0.214	0.448	0.338	0.012	-0.006	-0.005
8	Centrl_W_Syd	0.202	0.455	0.343	0.121	0.492	0.386	-0.081	0.038	0.043
9	Outer_W_Syd	0.202	0.455	0.343	0.236	0.439	0.325	0.034	-0.016	-0.018
10	Blck_Baulk_H	0.202	0.455	0.343	0.041	0.529	0.430	-0.161	0.074	0.086
11	Lower_N_Syd	0.202	0.455	0.343	0.160	0.474	0.366	-0.042	0.019	0.022
12	Horns_Kuring	0.202	0.455	0.343	0.210	0.451	0.339	0.008	-0.004	-0.004
13	Nth_Beaches	0.202	0.455	0.343	0.185	0.462	0.352	-0.017	0.008	0.009
14	Gosfrd_Wyong	0.202	0.455	0.343	0.179	0.465	0.356	-0.023	0.011	0.012

Table 6: Vehicle type choice before and after the NWRL project

Type of impacts	Description	Scenarios(*)		
		Base Case NO (new) Agglomeration effects and NO change in total employment	A WITH (new) Agglomeration effect but NO change in total employment	B WITH (new) Agglomeration and WITH change in total employment
WEI	Changes in total wage bill ($mill)	3.860	4.379	4.792
TLUI	Changes in total (imputed) rent ($mill)	(3.860)		
	Changes in the logsum (converted to $mill) of the residential location choices.	-	27.644	27.373
	mark-up of WEIs over TLUIs (%)	0%	0-15.8%	0-17.5%
	Explanation	WEI is *not* additional to but overlapping with TLUI	WEI is additional to TLUI	WEI is additional to TLUI

Table 7: Conventional transport -land use impacts (TLUIs) and wider economic impacts (WEIs) of the NWRL transport improvement project. ($2006/per annum)

(*) Notes:

Base Case: no (new) agglomeration effects and no change in total employment for the Sydney Metropolitan Area (SMA) as a whole, only a *redistribution* of existing employment and housing locations between zones assuming the transport project will allow the local economy of the SMA to reach a new equilibrium to take advantage of any *existing* wage differentials between locations following the changes in generalised transport costs between zones as a result of the project ;
Scenario A: (new) agglomeration effects causing (further) wage differentials which can be and redistribution of employment - assuming no change in total employment;
Scenario B: (new) agglomeration effect causing redistribution of employment as well as change in total employment

Residential Zone		Scenario A			Scenario B			Difference (Scenario A – Scenario B)		
No.	Name	Detached house	Semi-detached house	Unit	Detached house	Semi-detached house	Unit	Detached house	Semi-detached house	Unit
1	Inner_Sydney	-0.0109	0.0355	0.0081	-0.0110	0.0361	0.0082	0.0001	-0.0006	-0.0001
2	Eastern_Subs	0.0065	0.0177	0.0040	0.0065	0.0188	0.0043	0.0001	-0.0011	-0.0003
3	StGge_Suther	-0.0141	0.0151	0.0049	-0.0141	0.0160	0.0051	0.0000	-0.0009	-0.0003
4	Canter_Banks	0.0638	-0.0178	-0.0021	0.0638	-0.0138	-0.0017	0.0000	-0.0040	-0.0005
5	Fairfd_Livrp	-0.0768	0.0272	0.0181	-0.0757	0.0311	0.0207	-0.0011	-0.0039	-0.0026
6	Outer_SW_Syd	-0.0072	-0.0017	-0.0011	-0.0052	0.0046	0.0031	-0.0020	-0.0063	-0.0042
7	Inner_W_Syd	-0.2864	0.1456	0.0222	-0.2862	0.1490	0.0227	-0.0001	-0.0034	-0.0005
8	Centrl_W_Syd	0.0927	-0.0334	-0.0107	0.0933	-0.0308	-0.0099	-0.0006	-0.0026	-0.0008
9	Outer_W_Syd	-0.0227	-0.0007	-0.0004	-0.0195	0.0073	0.0048	-0.0032	-0.0079	-0.0053
10	Blck_Baulk_H	-0.0790	0.0345	0.0230	-0.0788	0.0362	0.0241	-0.0002	-0.0017	-0.0011
11	Lower_N_Syd	-0.0036	0.0223	0.0034	-0.0037	0.0230	0.0035	0.0001	-0.0007	-0.0001
12	Horns_Kuring	0.0057	0.0017	0.0012	0.0058	0.0025	0.0016	-0.0001	-0.0007	-0.0005
13	Nth_Beaches	0.0039	0.0027	0.0009	0.0045	0.0046	0.0015	-0.0006	-0.0019	-0.0006
14	Gosfrd_Wyong	-0.0135	-0.0071	-0.0048	-0.0101	0.0012	0.0008	-0.0035	-0.0084	-0.0056

Table 8: Dwelling price changes (%) following the improvement in transport network after the NWRL project

References

Anderson, S., De Palma, A., and Thisse, F. (1988a) The CES and the Logit – Two Related Models of Heterogeneity, Regional Science and Urban Economics, 18, 155-164.

(1988b) A Representative Consumer Theory of the Logit Model, International Economic Review, 29(3), 461-466

(1989) Demand for Differentiated Products, Discrete Choice Models, and the Characteristics Approach, Review of Economics Studies, 56, 21-35

Ben Akiva, M.E. and Lerman, S.R. (1979) Disaggregate travel and mobility choice models and measures of accessibility, in Hensher D.A., and Stopher P.R. (eds.), Behavioural Travel Modelling, Croom Helm, London.Berry, S. (1994) “Estimating Discrete-Choice Models of Product Differentiation, Rand Journal, 25(2), pp. 242-262.

Berry, S. (1994) Estimating Discrete-Choice Models of Product Differentiation, Rand J. Econ. 25 (Summer): 242–62.

Berry, S., Levinsohn, J., and Pakes, A. (1995) Automobile Prices in Market Equilibrium, Econometrica, 63(4), pp. 841-890.

(2004) Estimating Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The Market for New Vehicles, Journal of Political Economy, 112 (1), pp. 68-105.

Dixit, A. K. and Stiglitz, J. E. (1977) Monopolistic Competition and Optimum Product Diversity, American Economic Review, 67, 297-308.

Fiebig, D., Keane, M., Louviere, J., and Wasi, N. (2009) The generalized multinomial logit: accounting for scale and coefficient heterogeneity, Marketing Science, published online before print July 23, DOI:10.1287/mksc.1090.0508

Gitelman, E., and Otto, G. (2010) “Supply elasticity in the Sydney Housing market”, School of Economics, University of New South Wales, December.

Graham, D. J. (2007a). Agglomeration, productivity and transport investment, Journal of Transport Economics and Policy 41, 1–27.

Graham, D. J. (2007b). Variable returns to agglomeration and the effect of road traffic congestion, Journal of Urban Economics 62, 103–120.

Greene, W.H. and Hensher, D.A. (2010) Does scale heterogeneity across individuals matter? A comparative assessment of logit models, Transportation, 37 (3), 413-428.

Hanemann,W.M. (1984) Discrete/Continuous Models of Consumer Demand. Econometrica 52, 541-561.

Hensher, D.A. and Greene, W.H. (2003) Mixed logit models: state of practice, Transportation, 30 (2), 133-176.

Hensher, D.A., Rose, J.M. and Greene, W.H. (2005) Applied Choice Analysis: A Primer Cambridge University Press, Cambridge.

Hensher, D.A., Truong, T.P., Mulley, C., and Ellison, R. (2011) Assessing the Wider Economy Impacts of Transport Infrastructure Investment with an Illustrative Application to the North-West Rail Project, Report to the Director General of Transport, NSW Government, Australia, July.

Horridge, M. (1994) A Computable General Equilibrium Model of Urban Transport Demands, Journal of Policy Modeling, 16(4), 427-457.

Jara-Diaz, S. (2007) Transport Economic Theory, Elsevier Science, Oxford.

Jara-Diaz, S. and Videla, J. (1989) Detection of income effect in mode choice: theory and application, Transportation Research, 23B (6), 393-400.

Mare, D. and Graham, D. (2009) Agglomeration elasticities in New Zealand, Motu Economic and Public Policy Research Working Paper 09-06, Auckland, New Zealand. McFadden, D.

(2001) Disaggregate behavioural travel demand RUM side – a 30 years retrospective, in Hensher, D.A. (ed.) Travel Behaviour Research: The Leading Edge, Pergamon, Oxford, 17-64.

McFadden, D.L. and Reid, F. (1975) Aggregate travel demand forecasting from disaggregate behavioural models, Transportation Research Record: Travel Behaviour and Values, No. 534, 24-37.

Oum, T.H. (1979) A Warning on the Use of Linear Logit Models in Transport Modal Choice Studies, Bell Journal of Economics, 10, 274-387.

Lancaster, K. (1966) A New Approach to Consumer Theory, Journal of Political Economy, 74, 132-57)

Schmidheiny, K. and Brülhart, M. (2011) On the equivalence of location choice models: Conditional logit, nested logit and Poisson, Journal of Urban Economics, 69, 214-222.

Smith, B., Abdoolakhan, Z. and J. Taplin (2010) Demand and choice elasticities for a separable product group, Economics Letters, 108, 134–136

Train, K. (2003) Discrete Choice Methods with Simulation, Cambridge University Press, Cambridge.

Venables, A. J. (2007). Evaluating urban transport improvements: cost-benefit analysis in the presence of agglomeration and income taxation. Journal of Transport Economics and Policy 41(2), 173–188.

Appendix

The structure of TRESIS-SGEM

TRESIS-SGEM model - as the name suggests - consists of two parts: (1) TRESIS module which is a series of nested discrete choice models specifying departure time and mode choice, work location choice, dwelling type choice and residential location choice behaviour for different types of workers-cum-commuters/consumers distinguished by occupations and industries (see Tables A2-3), vehicle type choice and fleet size choice, [39] (2) SGEM module is a spatial general equilibrium model describing some of the basic economic activities within the different geographical zones of the Sydney Metropolitan Areas (Table A1). The Sydney Metropolitan Area is divided into 14 different zones (see Figure A1 and Table A1), each zone is characterised by the total number of dwellings of different types, population and employment of various occupations and in different industries. As well, there are journeys to work information on commuters travelling from each particular zone to a specific work destination with income level, occupation and employment industry specified (the Australian Bureau of Statistics (ABS), 2006).When TRESIS is used on its own, many of the economic variables relating to these activities are assumed to be given ‘exogenously’ in the TRESIS module, for example, housing prices, total level of demand for housing and for travel activities, total employment. The DC models then simply ‘distributes’ these total level of economic activities among the various ‘alternatives’ such as departure time and mode choice (given total number of trips between any origin-destination pair), dwelling types choices (given total level of demand for housing in each residential location), work place location choices (given the total level of employment). When TRESIS is combined with SGEM in the integrated model, these total level variables are now endogenously estimated, with the aggregate price indices to be determined within the various DC models while the equilibrium quantities are to be determined within the SGEM module.

TRESIS Module

Departure time and Mode choice (DTMC) model:

$$\begin{aligned} Prob_i^{(t/wr)} &= \frac{\exp(V_i^{(t/wr)})}{\sum_{j\in I^{(t/wr)}} \exp(V_j^{(t/wr)})}; \quad i \in I^{(t/wr)}. \\ V_i^{(t/wr)} &= \sum_{m\in M^{(t)}} \alpha_m^{(t)} A_{im}^{(t/wr)} + \sum_{n\in N^{(t)}} \beta_n^{(t)} B_{in}^{(t/wr)} \\ \bar{V}^{(t/wr)} &= \ln \sum_{j\in I^{(t/wr)}} \exp(V_j^{(t/wr)}) = \sum_{j\in I^{(t/wr)}} Prob_j^{(t/wr)} (V_j^{(t/wr)}) \end{aligned} \qquad \text{(A1)}$$

where $I^{(t/wr)}$ is the set of departure time and mode choices, $A_{im}^{(t/wr)}$ are attribute variables of these choice alternatives (access, egress, in-vehicle travel times, travel costs, toll charges, frequencies (of public transport modes), etc.), $A_{im}^{(t/wr)}$ are attribute variables of the commuters (socio-economic status, income variable, etc.), $\alpha_m^{(t)}$ and $\beta_n^{(t)}$ are empirical coefficients for these

---

[39] A full blown TRESIS will also include DC modules of work practice choice and vehicle kilometre driven (see Hensher (2002); Hensher and Ton (2002)). However, in this version of TRESIS which is to be linked to a CGE model, these two modules are not necessary and can be replaced by other structures within the CGE.model (such as the labour market (participation rate modelling) and demand for aggregate transport.

variables. 40 The model is 'calibrated' to fit with the initial number of work trips between any origin-destination (O-D) zonal pairs (r-w) using the Australian Bureau of Statistics (ABS) journey to work data (ABS (2006)). Given the initial choice probabilities as described in equation (A1), the percentage change in this probability is given by:

$$d\ln(Prob_i^{(t/wr)}) = dV_i^{(t/wr)} - d\overline{V}^{(t/wr)} \tag{A2}$$

These percentage changes in choice probabilities are to be used in the SGEM module to 'update' the market shares of work trips by different modes for various O-D pairs. From equation (A1), we also get the value for the aggregate departure time and mode choice activity price index:

$$\overline{P}^{(t/wr)} = \sum_{j\in I^{(t)}} (Prob_j^{(t/wr)})(P_j^{(t/wr)}) \tag{A3}$$

which is derived from the DTMC model (see equations (22)-23) in the text). This price index is to be used in the aggregate consumption model of the SGEM module.

Similarly with dwelling type (DwT) choice model, work location (WLC) choice and residential location (RLC) choice models, they are given by equations similar to (A1)-(A3) except that with WLC and RLC models, the logsum(s) of lower level decision is included also in the utility function of the choice model according to the nested linkages (see Figure 1 and equations (6)-(7), (10)-(11)).

SGEM Module

The SGEM module consists mainly of an aggregate continuous demand (CD) model describing the aggregate quantities of demand for various commodities referred to by the discrete choice models, such as aggregate demand for housing, for travel activities, and for other goods:

$$\overline{x}^{(c)} = \overline{y} - \sigma[\overline{p}^{(c)} - \overline{p}]; \quad c = \{t, d, o\} \tag{A4}$$

Here $\overline{x}^{(c)}$ stands for (the percentage change in aggregate demand for)[41] commodity(c) where c = t (travel), d (housing, and o (other goods); $\overline{p}^{(c)}$ is the corresponding percentage change in aggregate price, $\overline{p}$ is the share weighted average of all $\overline{p}^{(c)}$, $\overline{y}$ stands for the percentage change in income or after tax wage level, and σ is a CES substitution elasticity. For travel activity, $\overline{x}^{(t)}$ is related to departure time and mode choice (DTMC) decisions via equation such as (36) which says $\overline{x}^{(t)}$ is simply the quantity weighted average of all the individual DTMC quantities $x_i^{(t)}$. Similarly for housing activities, $\overline{x}^{(d)}$ is simply the quantity weighted average of all the dwelling type choice decisions $x_i^{(d)}$ as determined by the dwelling type choice (DwTC) model. The aggregate price indices $\overline{p}^{(c)}$ are inferred from the values of the indirect utilities of the DC models as described by equations such as (21) and (24).

In addition, SGEM also contains equations which specify production and other government activities (such as taxation and spending) if these are referred to by the DC models. For

---

[40] For full details of these variables and coefficients, see Hensher (2002); Hensher and Ton (2002).

[41] See also equation (30) in the main text. Note the convention that a lower case letter is used to denote percentage change and a bar on top of a letter is used to denote *aggregate* level of demand, hence from here on we will drop reference to these terms wherever this can be implied. Equation (A3) is an extension of equation (30) where we include 'other goods'. This goods can also be disaggregated or extended into various types of commodities (such as demand for automobiles) if a DC model (for vehicle type choice for example) specifically requires reference to such an aggregate commodity.

example, with dwelling type choice activities, a total level of supply for dwellings of different types need to be specified which can either be in the form of a special supply function (using exogenously estimated supply elasticities) or in the form of basic input-output (I-O) activities[42]. In this skeletal version of SGEM, we have chosen to use the supply function approach to describe the supply for dwellings of different types rather than resorting to a comprehensive I-O approach:[43]

$$d_i = \delta + \varepsilon_i^d p_i^d; \quad i \in I^{(d)} \tag{A5}$$

Here d_i stands for the percentage change in the supply of dwellings of type *i*, p_i^d stands for the percentage change in the price of dwellings of type *i*, δ is a constant which can be used to stand for any exogenous shocks, and ε_i^d is the elasticity of dwelling supply.[44] Equilibrium condition in the dwelling market requires:

$$d_i = x_i^d; \quad i \in I^{(d)} \tag{A6}$$

For work location choice activities, initial employment levels for different occupations and industries at various locations taken from the ABS (2006) data are used to calibrate the work location choice model. When the decision on work location choice changes (as driven by changes in the characteristics of the transport network), employment levels in various zones are assumed to change also to match with the changes the supply of labour, i.e. firms are influenced by the work location decisions of workers to move production activities to where there are easy access to labour supply. This means transport investment has a capacity to influence the distribution of employment opportunities across different zones with implications for the level of labour productivity and wages in the local economy. This is because on the one hand, workers prefer to work in locations where wage level is high and firms prefer to locate where accessibility to labour supply is high. High wage level may be the result of so-called ‘agglomeration’ effect where economies of scale and scope in production activities are generated by the greater interactions between workers in places where the level of employment density is higher. Employment density (or accessibility to employment) can be defined as follows (Venables, 2007):

$$U_{iz} = \frac{E_{iz}}{\left(\sqrt{A_z / \pi}\right)^{\alpha}} + \sum_{s}^{s \neq z} \left[\frac{E_{is}}{(d_{zs})^{\alpha}} \right]; \quad i = 1,...,I; \quad z, s = 1,...,Z.. \tag{A7}$$

Here, U_{iz} is a measure of employment density for industry *i* in zone z, E_{iz} is the level of employment in industry *i* in zone *z*, d_{zs} is the distance between zone *z* and zone *s*, A_z is the land area of zone *z*, so that $\sqrt{A_i / \pi}$ is an estimate of the average distance between jobs *within* zone *z*, α is a ‘distance decay’ parameter which can be empirically estimated, but often assumed to be = 1 (see Graham, 2007a; Maré and Graham, 2009). Assuming that increases in employment density (agglomeration) can generate economies of scale and scope in production activities which increases output and labour productivity, this means we can then postulate a statistical relationship of the following form:

$$\ln(W_{izo}) = \beta_i + \varepsilon_i \ln(U_{izo}) \quad z = 1,...,Z; \quad o = 1,...,O.. \tag{A8}$$

---

[42] By this, we mean activities which can be described in terms of an input-output matrix which shows how much of different inputs (intermediate goods as well as primary factors) are necessary to produce a unit of output of an activity. These are the types of activities described in computable general equilibrium models.

[43] The I-O approach would require more data which can be considered in a future extension of SGEM.

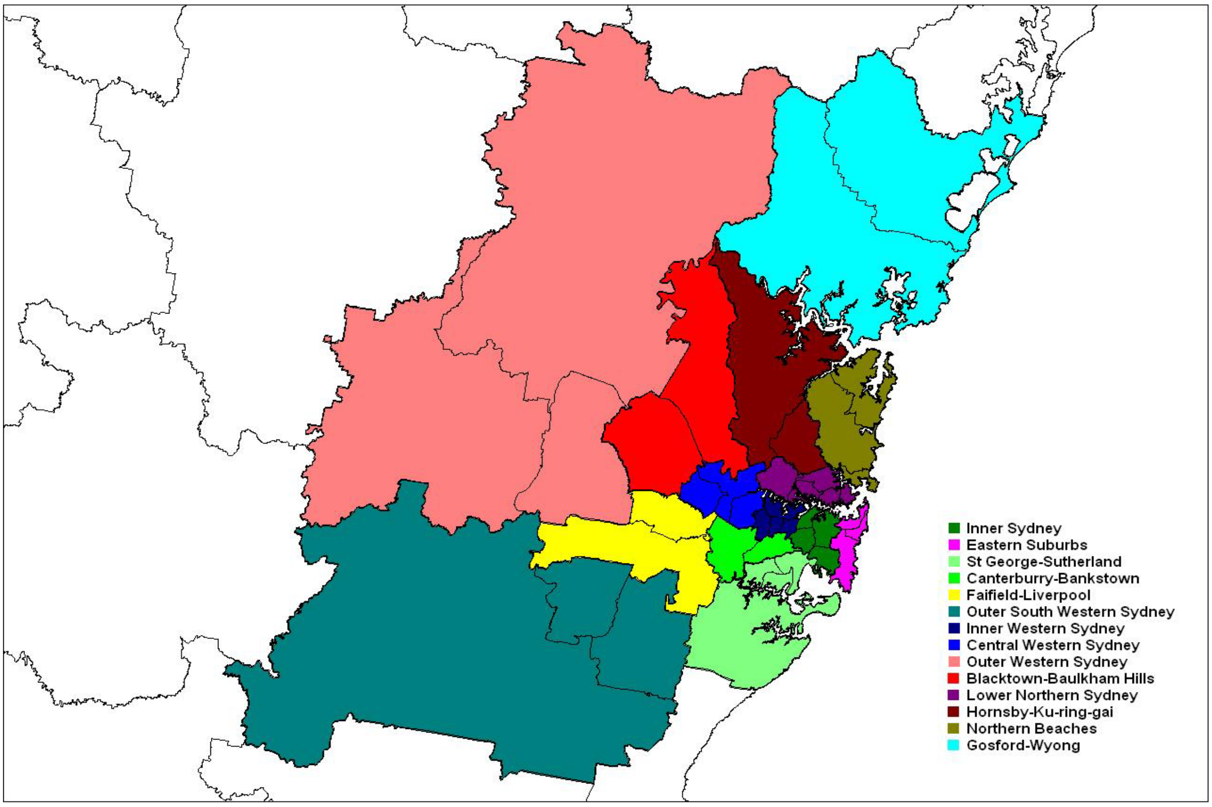
[44] Taken from Gitelman and Otto (2010) for the various local government areas of Sydney.

where (W_{izo}) is the wage level for workers of occupation *o* in industry *i* located in zone *z*, and (U_{izo}) is the employment density for this particular occupation *o* in industry *i* and zone *z*, β_i is an industry specific constant term, and ε_i is a measure of the 'agglomeration elasticity'.[45] The percentage change in the level of wage rate (as indicated by the left hand side of equation (A8)) can then be related to the percentage change in income level of the worker as used in equation (A4) via an equation such as:

$$\ln(W_{izo}[1-\tau]) = \bar{y}_{izo} \quad z = 1,\ldots,Z; \quad o = 1,\ldots,O.. \tag{A9}$$

where (τ) is the income tax rate, and $\bar{y}_{izo}$ is the (percentage change in) the income level of worker of occupation *o* in industry *i* and zone *z*. Any change in after-tax income level as described by equations (A8)-(A9) has a potential to affect the levels of demand for housing and transport activities as described by equation (A4) and these demand levels are then allocated to the various alternatives in housing and transport alternatives as indicated by the discrete dwelling type choice and departure time – mode choice models. This is how discrete models are linked to continuous demand models in TRESIS-SGEM.

Figure A1: TRESIS-SGEM zones for the SMA



---

[45] β_i is and ε_i are to be estimated empirically, see Hensher *et al.* (2011).

Zone Number	Short Name	Long Name	Employment number in 2006	Journeys to work (daily) in 2006
1	Inner Sydney	Inner Sydney	396498	324963
2	Eastern Suburbs	Eastern Suburbs	63497	84867
3	StGrge Sutherlnd	St George Sutherland	88236	161571
4	Canter. Bankstwn	Canterbury Bankstown	73698	111422
5	Fairfld Liverpl	Fairfield Liverpool	85463	133351
6	Outer SW Syd.	Outer South West Sydney	52938	76538
7	Inner W Syd.	Inner West Sydney	58332	75597
8	Central W Syd.	Central West Sydney	147300	153270
9	Outer W Syd.	Outer West Sydney	81518	119032
10	Blcktwn Blk Hills	Blacktown Baulkham Hills	119490	155674
11	Lower N Shore	Lower North Shore	180611	176070
12	Hornsby Kuringai	Hornsby Kuringai	63148	83344
13	Northern Beaches	Northern Beaches	67010	77238
14	Gosford Wyong	Gosford Wyong	74950	90778
Total	SMA	Sydney Metropolitan Area	1552689	1823716

Source: TRESIS (Hensher (2002); Hensher and Ton (2002)) and ABS (2006)

Table A1: Geographical zones in the Sydney metropolitan area with employment levels and journeys to work in 2006

Occupation Number	Short Name	Long Name
1	Managers	Managers
2	Professnals	Professionals
3	TechTrades	Technicians and Trades Workers
4	CommPersServ	Community and Personal Service Workers
5	ClericlAdmin	Clerical and Administrative Workers
6	SalesWorkers	Sales Workers
7	MachOperDriv	Machinery Operators And Drivers
8	Labourers	Labourers
9	Others	Others

Source: ABS (2006)

Table A2: Different labour occupations considered in TRESIS-SGEM

Industry Number	Short Name	Long Name
1	Agr_For_Fish	A,"Agriculture, Forestry and Fishing"
2	Mining	B,"Mining"
3	Manufacturng	C,"Manufacturing"
4	ElyGasWatWst	D,"Electricity, Gas, Water and Waste Services"
5	Construction	E,"Construction"
6	Wholes_Trade	F,"Wholesale Trade"
7	Retail_Trade	G,"Retail Trade"
8	Accom_Food	H,"Accommodation and Food Services"
9	TranPostWare	I,"Transport, Postal and Warehousing"
10	InfoMediaTel	J,"Information Media and Telecommunications"
11	FinanceInsur	K,"Financial and Insurance Services"
12	RentHirRealE	L,"Rental, Hiring and Real Estate Services"
13	ProfSciTech	M,"Professional, Scientific and Technical Services"
14	Admin_Supprt	N,"Administrative and Support Services"
15	PubAd_Safety	O,"Public Administration and Safety"
16	Edu_Training	P,"Education and Training"
17	HlthC_SoAstn	Q,"Health Care and Social Assistance"
18	Arts_Recrtn	R,"Arts and Recreation Services"
19	OthServcs	S,"Other Services"
20	Others	Inadequately described or not stated

Source: ABS (2006)

Table A3: Industries in TRESIS-SGEM