From CBA to CGE: A Review of Computable General Equilibrium Modelling for Transport Appraisal

Edward Robson¹, Kasun Wijayaratna¹, Vinayak V. Dixit¹

¹Research Centre for Integrated Transport Innovation (rCITI)
School of Civil and Environmental Engineering
UNSW Sydney NSW 2052 Australia

Email for correspondence: e.robson@student.unsw.edu.au

Abstract

In the transport planning process, decision makers require reliable and informative appraisals to facilitate comparisons and determine if a proposal is worthwhile to society. The cost benefit analysis is the most common form of appraisal, but the consumer surplus metric used in cost benefit analyses will only reflect total social welfare if markets operate perfectly. There may be significant uncaptured impacts, known as wider economic impacts, which agencies are beginning to incorporate in appraisals using ad-hoc methods. Computable general equilibrium (CGE) models are an increasingly popular method for assessing the economic impacts of transport, including both direct and wider economic impacts, as they can determine the distribution of impacts among every market and agent in the economy. By simulating the behaviour of households, firms and others from microeconomic first principles, they can provide a measure of welfare that guarantees no double counting and accounts for nth order effects. This paper reviews CGE models that have been applied to transport issues, and discusses the general role of CGE modelling in transport appraisal as well as theoretical and practical concerns regarding CGE modelling practice.

1. Introduction

As planners and engineers in Australia and beyond strive to improve transport networks, the appraisal of proposed transport projects is vital to provide a rational basis for decision making. Cost benefit analysis (CBA), in which the social, economic and environmental impacts of projects are monetised, remains one of the most popular methods to assess and rank projects. In a CBA, each impact is assessed separately, taking care to reduce the risk of double-counting.

A key issue is that there are many interactions in the economy that are not captured in this process. For example, new infrastructure can stimulate economic growth, which in turn generates additional transport demand that alters the benefits of the project and complicates its evaluation. These concerns have been recognised since the genesis of the CBA approach:

*If investment decisions are so large relatively to a given economy... that they are likely to alter the constellation of relative outputs and prices over the whole economy, the standard technique [of CBA] is likely to fail us, for nothing less than some sort of general equilibrium approach would suffice in such cases.* —Prest and Turvey (1965), in a review of CBA state-of-the-art in its formative years.

In other words, significant transport projects can impact demand and supply in other markets, and therefore the transport market should not be treated as independent from the rest of the economy when analysing these projects. The effect of treating transport in this
manner has been assumed to be inconsequential in CBAs until recent years, possibly resulting in incomplete and misleading analyses.

Various agencies around the world have begun to incorporate these uncaptured impacts, known as ‘wider economic impacts’ (WEIs), in CBAs over the past two decades to strengthen the justification for transport projects. In some cases, WEIs can rival traditional (direct) impacts in scale. Most WEIs are estimated with a number of ad-hoc models, which has led to differing assessment practices between jurisdictions and the risk of double-counting impacts.

One particular type of model that has the potential to unify the estimation of WEIs is the computable general equilibrium (CGE) model. A CGE model simulates an entire economy by representing the supply and demand of every market. The central mechanism is that both supply and demand in each market are functions of all prices across other markets in the economy, not just their own price, meaning changes in one market affect all others. Solving a CGE model involves searching for a set of prices that results in equilibrium in all markets simultaneously, i.e. ‘general equilibrium’.

Furthermore, CGE models applied to transport can provide a framework to assess both direct impacts and WEIs within a single model. GDP, prices and other economic measures can be extracted as the models are built from fundamental microeconomic behaviour. This enables agencies to prioritise across transport projects and facilitates comparisons with proposals for government expenditure in other sectors. Planners can also identify the distribution of impacts when agents and markets are spatially disaggregated, and can measure welfare directly from utility functions, rather than use the transport market as a proxy. However, there are questions about what role CGE models should play in appraisal. Data and computational requirements can be prohibitive, especially when spatial detail is necessary. The operation of CGE models also tends to be a ‘black box’ where model mechanics are hidden or difficult to understand.

This paper aims to synthesise the case for applying CGE models in transport appraisal. Section 2 summarises existing transport appraisal methods and issues, including conventional CBA practices, their limitations and the valuation of WEIs within CBAs. Section 3 introduces the concepts underlying CGE modelling and reviews existing CGE models applied to transport. Section 4 explores how CGE models can be applied in appraisal practice. Section 5 identifies theoretical and practical concerns regarding their application and Section 6 outlines directions for future research. Section 7 concludes the paper.

2. Current practice in transport appraisal

2.1 Concepts of cost benefit analysis

CBAs have been the dominant appraisal methodology for transport over several decades, in part due to their intuitive foundations. All of the significant social, environmental and economic impacts of a transport project are first identified and monetised. Future impacts are then discounted to a present value so that decision makers can use metrics such as net present value, benefit–cost ratio (BCR) and internal rate of return to prioritise projects and determine if an individual project is worthy of funding. This allows CBAs to be flexible as the range of impacts evaluated can be scaled to suit each project. A CBA assesses social welfare improvement, rather than financial viability, as many of the benefits and costs of transport projects are not incurred by users.

The impacts considered in a transport CBA span users, service providers and the community. While capital and operating expenses may be forecasted using engineering judgement and historical costings, user impacts are far less tangible. The welfare benefit to users is indirectly measured as the change in ‘consumer surplus’; consumer surplus being the difference between what users are willing to pay for a transport service and what they
actually pay. The concept of ‘price’ here refers to the generalised cost of travel—an index of all of the measurable attributes of a transport service, where each attribute is converted into a monetary equivalent using multipliers. The change in consumer surplus is then calculated as the area between the demand curve (a function relating the expected quantity of trips to the generalised cost) and the generalised costs before and after the transport improvement. Knowing the expected demand and generalised costs is enough for this calculation since the demand curve is assumed to be linear between the two points (the ‘rule of a half’). Changes in demand and generalised costs are estimated with external transport models accounting for mode choice and congestion.

2.2 Issues with cost benefit analysis

There are two well-known weaknesses of the conventional CBA methodology: (1) it assumes that all benefits can be measured from the transport system, and (2) it cannot describe the eventual distribution of benefits among economic agents (Bröcker et al., 2010). Until the 1990s, transport networks were treated separately from the broader economy in transport planning—a partial equilibrium perspective. Land-use distributions and economic parameters were assumed to remain constant in four-step demand models, producing fixed trip matrices for route choice (Mackie, 2010).

From a general equilibrium perspective, a change in the transport market will not only affect the demand for transport, but also demand and supply in external markets. This, in turn, will induce second-order effects on the demand for transport and the equilibrium generalised price. Land markets have a particularly strong connection with transport networks (Wegener and Fuerst, 2004). If a transport improvement provides a travel time saving to an individual, they may utilise it by increasing their travel demand or by moving to a preferred residential location that has become more accessible, whilst maintaining a constant daily travel time (Ahmed and Stopher, 2014). Transport demand curves would therefore shift, land values would change and travel time savings would evaporate as people adjust their lives to maintain their travel budget. In other words, as people adjust their travel behaviour from a transport improvement, travel time savings eventually diminish to zero as the benefits from the transport market are transferred to other sectors (Metz, 2008). Despite this, benefits in a CBA are most often reported as travel time savings—up to 80% in road projects according to Mackie et al. (2001).

Proponents of the CBA approach claim that regardless of where they eventually accrue, all impacts to users will be captured by the consumer surplus metric. If markets operate perfectly, where prices equal marginal costs, then transport users will correctly consider the changes in external markets in their valuation of transport (Dodgson, 1973; Jara-Díaz, 1986; Vickerman, 2007a). The willingness to pay for a transport improvement would therefore account for both changes in property prices as well as travel time savings (Sue Wing et al., 2007). However, this assumption is unlikely to be realistic as any source of market imperfection, such as monopolistic firms, taxes and excess demand/supply, would violate it. Even if it were realistic, other economic externalities would still not be reflected. Appraisals nowadays consider these externalities (WEIs) to be separate to user impacts.

The second issue with the conventional CBA methodology is that it can be difficult to determine the eventual distribution of impacts. When general equilibrium effects are ignored, transport models can estimate how much a region will benefit from travel time savings. This distribution is likely to change once markets adjust to the transport improvement as the benefits transfer away from travel time savings and into other markets. The consumer surplus metric, which captures welfare in a perfect system irrespective of its distribution, is an aggregate measure that does not identify how much a particular household will benefit. This appears to be a serious issue regarding the effectiveness of CBAs (Odeck, 1996; Eliasson and Lundberg, 2012; Eliasson et al., 2015) as equity can be as important a consideration as the magnitude of benefits in project prioritisation.
2.3 Wider economic impacts

There are a range of economic impacts that are not captured in the consumer surplus metric, collectively known as WEIs. These occur when the price of transport does not equal its social marginal cost due to market imperfections and technological externalities. WEIs came to the forefront of transport economics during the 1990s as concerns grew about biases in existing appraisal practices and governments sought to justify infrastructure projects on the basis of economic growth. The seminal SACTRA (1999) report in the UK identified sources of WEIs and recommended amendments to appraisal practice, including the requirement of an ‘Economic Impact Report’ to supplement conventional CBAs. Over time, these recommendations have been implemented somewhat haphazardly. The field is still under active development and the nature, relevance, and even existence of WEIs are contentious. Some BCRs for large projects, such as Crossrail in the UK, have shown significant improvement after including WEIs (Banister and Thurstain-Goodwin, 2011).

The most frequently recognised WEIs are agglomeration externalities, labour market effects and impacts in markets with imperfect competition (Wangsness et al., 2016). Agglomeration externalities refer to the relationship between the concentration of economic activity and productivity (Venables, 2007). When businesses locate near each other, there are increasing returns to scale through knowledge spillovers, better access to markets and sharing of facilities. Transport projects reduce the effective distance between businesses, thereby generating productivity benefits that can spread throughout the urban area. Labour markets are a source of WEIs from the presence of taxes, imperfect information and imperfect competition. People entering the workforce, changing jobs or their hours of employment as a result of a transport project, do so on the basis of their net wages; the additional taxes they pay are an otherwise uncaptured benefit. Firms that set prices above marginal costs, as in markets with imperfect competition, are another potential source of WEIs as they have additional scope to improve efficiencies to generate welfare gains. WEIs can also include detrimental impacts to social welfare (Kanemoto, 2013), though this is not often discussed.

In a study of appraisal guidelines across 23 countries, Wangsness et al. (2016) found that 15 so far acknowledge the existence of WEIs. The authors found remarkable disparity between the guidelines under review, but many appeared to take inspiration from the UK Department for Transport (2014) guidelines. Many sets of guidelines considered the evidence for WEIs to be less robust than for other impacts. As a result, only 7 used WEIs as components of CBAs, and most recommended that a BCR without WEIs be presented first. Each WEI is generally calculated with a set of equations based on transport and economic changes (e.g. employment densities and generalised costs of travel), using parameters from other studies (e.g. the elasticity of productivity with respect to employment) and rules of thumb as other inputs. The lack of consistency has also led to concerns about double-counting impacts. There appears to be a need to find alternative models and methodologies that agencies find more reliable to evaluate the impact of transport on the economy.

2.4 Existing economic models for transport

Interest in these broad economic effects predates the recognition of WEIs, and a number of models have been developed to simulate interactions between transport and the economy. In terms of appraisal, these models produce outputs that can be used in multi-criteria analyses, economic impact reports and, since the advent of WEIs, equations to estimate economic uplift. Some of the most widely used economic modelling packages (e.g. REMI and IMPLAN) are based on a multiregional input–output (IO) structure that simulates trade flows between markets. These models use coefficients to describe the relationship between outputs and inputs of industries by region, allowing the effects of changes in demand and transport costs to flow through to changes in supply. IO frameworks have been criticised for a lack of flexibility in these coefficients, as well as a lack of supply-side feedbacks in prices and resource constraints (Bachmann et al., 2014). Many of these drawbacks have been
addressed in recent IO developments, such as the RUBMRIO class of models. In addition, some of these modelling frameworks have trended towards simulating a general equilibrium (Wegener, 2011).

Urban modelling requires detailed consideration of both transport and land-use markets as these constitute two of the most important factors of urban spatial development. Land-use transport interaction (LUTI) models link separate transport and land-use models to simulate feedbacks between the two systems over both short- and long-term time scales (Wegener, 2004). These models are detailed, mature, and are empirically-based. However, LUTI models have been described as lacking a strong microeconomic foundation and, like IO models, rely on parameters that are fixed with regard to prices. It can also be difficult to extract welfare for the purposes of a CBA or incorporate details of imperfect markets (Oosterhaven and Knaap, 2003).

In general, economic models for transport tend to be empirically detailed but are restricted in terms of market representation. Most do not produce a metric for welfare and have a limited ability to address the requirements of decision makers who are increasingly interested in capturing the full range of impacts to households, businesses and the economy at large, as well as their distribution. CGE models have been employed for transport projects and policies to bridge this gap, and are the focus for the remainder of this paper.

3. CGE modelling of transport

3.1 Overview of CGE modelling

Modern CGE modelling has evolved from two distinct branches of economics: ‘computable general equilibrium’ (CGE) modelling and ‘applied general equilibrium’ (AGE) modelling (Thissen, 1998; Mitra-Kahn, 2008). These two names are used interchangeably in the literature nowadays. The CGE approach arose from the IO models of Leontief (1941), from which Johansen (1960) introduced a price mechanism that allowed firms, households and investors to substitute between sources of inputs and outputs according to utility and profit maximisation. The AGE approach was developed as an empirical application of Walrasian general equilibrium theory. Scarf (1973) developed algorithms to compute Arrow–Debreu equilibria, which enabled a number of researchers to apply AGE models to practical problems. However, by the mid-1980s, the CGE approach led by the highly influential ORANI model (Dixon et al., 1982) became dominant due to their flexibility and faster solution methods, but not before adopting some of the terms and concepts of the AGE approach.

A CGE model comprises a set of equilibrium equations representing commodity markets in an economy—the model is solved by finding a set of prices and outputs that results in equilibrium in every market simultaneously. All CGE models simulate the behaviour of consumers and the production process through representative households and firms respectively. Households are endowed with primary factors (e.g. labour and capital) which are sold to firms for an income. Firms then transform these factor inputs, possibly with intermediate inputs from other firms, into commodities. Households and firms purchase these commodities to provide utility and to produce further outputs. Households are assumed to be utility-maximising in their behaviour, and firms are profit-maximising. Solving the utility and profit maximisation problems analytically yields supply and demand functions that are then used to compile the equilibrium equations. CGE-style models (as opposed to AGE-style) will typically include representative governments funded by ad valorem taxes on sales and production, investors and external markets. Shoven and Whalley (1992) provide further detail on the construction of CGE models.

Once models are specified, they are implemented in software to access solvers, including GEMPACK (Harrison and Pearson, 1996) and GAMS using the MPSGE syntax (Rutherford, 1999), or are specially coded. Static models are operated by altering exogenous parameters,
known as ‘shocking’ the model, to produce a counterfactual equilibrium for comparative static analysis. Dynamic models trace the transition of an economy over time, most commonly by simulating a sequence of short-term static equilibria where long-term parameters (e.g. capital stock) are adjusted with external models between time periods.

Economic agents and commodities can be identified by their location to give models a spatial dimension. These have been referred to as spatial, multiregional or inter-regional CGE models, depending on the field. The role of transport in the economy is to enable spatially separated entities to physically interact for a cost. Transport is therefore an integral component of spatial models and it is generally necessary for CGE models applied to transport issues to have a spatial dimension, unless the model is used to assess economy-wide policies. A number of interrelated disciplines, such as urban economics and regional science, have adopted the theory and tools of CGE modelling to analyse transport issues, and each have their own set of standard practices.

3.2 Urban CGE modelling

Urban CGE models have been developed within the field of urban economics for the study of economic issues in urban areas. Many of the early urban CGE models were created to understand the relative effects of policies on urban economies and were theoretical only. A series of papers by Anas linked theoretical urban CGE modelling with transport modelling, including Anas and Kim’s (1996) study of the formation and stability of multiple city centres, Anas and Xu’s (1999) study of the effects of job dispersion and Anas and Rhee’s (2006) study of congestion tolls versus urban boundaries to control sprawl. More recently, these models were adapted to analyse the effects of carbon charges (Tscharaktschiew and Hirte, 2010), public transport subsidies (Tscharaktschiew and Hirte, 2012) and the economic impact of speed limits (Nitzsche and Tscharaktschiew, 2013).

Applied urban CGE models tend to focus on the markets and agents relevant to urban microeconomic simulation, such as land markets, and less often on representative governments or external markets as in the standard macroeconomic-style CGE models. In the typical setup, households choose residential and job locations according to discrete choice models, and conditional on those locations, decide how much and where to consume in terms of shopping, housing and leisure. Producers can also be competitors in land markets and regions tend to be geographically small. Locations are sorted by assuming uniform utility or profit across the urban area. Equilibrium in this sense means that not only are supply and demand equal, but agents also have no incentive to change location.

One of the most well-known applied urban CGE models is RELU–TRAN, developed by Anas and Liu (2007). RELU–TRAN comprises two modules—a CGE module and a transport module—which feed into each other and iterate until convergence is achieved in both modules. In the basic structure of the model, employment and consumption of goods in the CGE module generate shopping and commuting trips which provide inputs for the transport module. The transport module then determines mode split and assigns trips to the transport network, returning expected travel costs to the CGE module. The CGE module comprises four economic agents: households (consumers), firms, landlords and developers.

- Households are modelled in a two stage utility maximisation process. In the upper level, households jointly choose their optimal residential location, job location and preferred housing type according to a logit model. In the lower level, given their upper level choices, households maximise utility over the consumption of goods, leisure and housing size, subject to monetary and time constraints.
- Producers minimise costs according to a production function of labour, capital, buildings and intermediate inputs.
- Landlords control the supply of floor space depending on profitability.
- Developers construct and demolish buildings according to demand.
RELU–TRAN has been applied to the urban areas of Chicago, Paris and Los Angeles to model issues such as fuel price increases, cordon tolling and job growth from rail investment (Anas, 2013), and has been extended to model fuel consumption (Anas and Hiramatsu, 2012).

Rutherford and van Nieuwkoop (2011) similarly simulated a transport network, but instead formulated the mode split and traffic assignment problems together with the CGE model as a single mixed complementarity problem. Robson and Dixit (2017) also used a mixed complementarity format to develop an urban CGE model suitable for transport infrastructure appraisal. In their model, discretionary trips were generated in addition to shopping and commuting trips, and freight costs were modelled as proportional to travel times.

Other urban CGE models have introduced additional disaggregate choice structures to simulate behaviour outside the pure CGE framework. The field of computable urban equilibrium (CUE) modelling in Japan, which branched from Anas (1982) in the late 1980s, utilises these extensively. CUE models lie in between CGE and LUTI models in terms of economic consistency (higher for CGE models) and empirical detail (higher for LUTI models)—see Ueda et al. (2012) for a review. TRESIS–SGEM (Hensher et al., 2012; Truong and Hensher, 2012) is the combination of a discrete choice model for travel behaviour (TRESIS) within a CGE framework (SGEM). In TRESIS, location and transport choices are modelled with nested discrete choice structures, given economy-wide variables such as housing prices. The economy-wide variables are then equilibrated using the continuous demand structure of SGEM. The full model was used to estimate agglomeration impacts for a rail project in Sydney.

### 3.3 Regional CGE Modelling

Regional CGE modelling refers to the modelling of discrete regions at a scale larger than urban regions, often spanning an entire country. At the regional scale, urban issues such as land markets, household transport demand and congestion become less significant and are often not modelled. There may also be more emphasis on macroeconomic results. Transport is a margin commodity in many regional CGE models, and the impact of freight on the economy is a common emphasis of regional CGE studies. A major differentiator is in how regional varieties of a commodity are treated:

- ‘Spatial CGE models’ (SCGE models) have emerged from the disaggregation of single-region national CGE models into the spatial dimension and adopt the Armington (1969) assumption of regional varieties being imperfect substitutes.
- Spatial price equilibrium (SPE) CGE models have emerged from the conversion of spatial price equilibrium (SPE) models from a partial equilibrium to a general equilibrium basis and assume that regional varieties are perfectly substitutable.
- New economic geography (NEG) CGE models have emerged from the field of new economic geography, which explains the formation of cities as a balance between increasing returns to scale and increasing transport costs from agglomeration.

The first CGE models of national economies lacked a spatial dimension and treated transport as a tradeable commodity supplied by a production sector, similar to other sectors. Buckley (1992) proposed one of the first SCGE models to incorporate the costs of transport and wholesaling services explicitly through price margins. Regionally produced goods in Buckley’s model were transported to clearinghouses in other regions where they became available for local production, consumption or export. Each movement between regions incurred a transport cost specific to each origin–destination pair, which was added onto the regional price.

As theory and computing power improved through the mid-1990s, it became feasible to disaggregate models to the extent required for representing transport infrastructure. Bröcker (1998) designed an SCGE model that included a number of simplifying assumptions to make
the model easy to implement. In each region and for each sector, transport agents imported commodities from other regions, which were then combined to form a 'pooled' commodity available for local use. The transport costs themselves were modelled as a price mark-up, known as the 'iceberg assumption', which in effect assumes that an amount of the transported good is used up ('melts') during transit in proportion to freight costs. Prototype models such as these enabled governments to develop SCGE models for planning purposes. PINGO (Ivanova et al., 2002; Vold and Jean-Hansen, 2007), a model for freight movements in Norway, was based on Bröcker’s (1998) model but included an explicit sector to provide transport services rather than assume iceberg costs.

SCGE modelling for transport began to spread globally throughout the 2000s, with many models differing in their methods of simulating the transport sector. For example, Kim et al. (2004) developed a dynamic model of South Korea to evaluate the regional economic impact of four highway proposals. In Kim et al.'s model, the production function for each sector in each region included an accessibility index to approximate the difficulty of transport from other regions. There has also been a strong research output in SCGE modelling from Japan in addition to their urban models, particularly for studying the impacts of earthquakes and other natural disasters. For example, Koike et al. (2000) and Ueda et al. (2001) applied an SCGE model to calculate the damage of the 1995 Kobe great earthquake.

The models described thus far were formulated with special consideration of the transport sector, but there exists a wide range of more general SCGE models that have also been used for transport analysis. Most are direct descendants of the Johansen lineage, derived from CGE models created by the Centre of Policy Studies in Melbourne (e.g. ORANI, MONASH, TERM and MMRF) and GTAP (Hertel, 1997) and its databases. Karplus et al. (2010) used the EPPA model, an international CGE model of emissions built from the GTAP database, to analyse the environmental impact of introducing plug-in hybrid electric vehicles in the United States and Japan. Verikios and Zhang (2015) analysed the effects of urban transport reform in Australia on household income groups, including changes in governance, pricing and market structures, by using the multiregional MMRF model to simulate region-specific changes.

Spatial price equilibrium (SPE) models predict the production and flow of goods between regions in an economy to meet consumer demand, accounting for transport costs. While they are mathematically robust, they do not model market interdependencies and cannot account for cross-hauling between regions since commodities from different regions are perfect substitutes; thus, consumers choose the source with the lowest delivered price. Some CGE models have been developed to incorporate principles from SPE modelling—Elbers (1996) argued that commodities (primary commodities for example) that are not distinguished by their origin should be treated as perfectly substitutable. Roson (1996) developed the MITER model for freight flows in Italy which integrated a CGE model, dispersed SPE model and freight network equilibrium model. In MITER, trade flows and regional demand were assigned to the transport network, from which transport costs were used to update prices and then quantities in the model. Lofgren and Robinson (2002) developed a model that assumed all commodities follow the SPE paradigm of perfect substitutability and one-way trade flows, but within a CGE framework.

The field of new economic geography (NEG) (Krugman, 1991; Fujita et al., 1999) attempts to reconcile theories from regional science and urban economics by explaining that cities form due to a balance between increasing returns to scale, drawing industries together, and increasing transport costs. The concepts of NEG were formalised in a general equilibrium context by Venables (1996) to examine how agglomeration is affected by economic integration and how important an industrial base is for a region. Venables and Gasiopek (1999) later developed a CGE model using NEG theories to analyse the supply-side effects of road projects funded by the European Commission.
Bröcker (1995) developed a similarly structured CGE model to Venables (1996), but in an applied context with a greater focus on how the model could be calibrated and solved. The paper also demonstrated that trade flows in the model would follow a gravity law form, as is commonly assumed in other spatial economic models. The models from Bröcker (1995) and Bröcker (1998) formed the basis of CGEurope, a family of CGE models used within a unified assessment framework for proposed transport projects and policies in the European Union (Bröcker et al., 2001, 2004). Bröcker et al. (2010) later investigated the spatial distribution of impacts generated by the TEN-T transport network projects to determine whether European Union involvement would be justified (when benefits were predicted to spill over into jurisdictions not financing the project) and what role they should play.

The RAEM family of NEG-based CGE models of the Netherlands was developed specifically for transport project appraisal. The first version of RAEM (Knaap and Oosterhaven, 2011) was developed in the early 2000s to assess the benefits of a high speed rail link between the west and the north of the Netherlands, in particular to analyse whether the rail link would relieve pressure in the dense west by enabling jobs and residents to relocate to the north. The representation of the transport sector was improved in RAEM-2 (Thissen, 2005) and RAEM-3 (Ivanova et al., 2007) through specific transport sector production functions and trip generation, and the model was also converted from static to recursively dynamic through savings, capital accumulation and technological progress.

### 3.4 Congestion and externality modelling

From the 1990s, a stream of research at the intersection of regional science, urban economics, transport economics and environmental economics has applied CGE models to analyse the externalities of transport, particularly congestion and pollution, and policies to address them. The general equilibrium approach is appropriate for the modelling of externalities as it can allow for linkages between economic agents that do not interact directly. These models represent the transport market in significant detail, often incorporating endogenous congestion (through a congestion index or a household time constraint) and mode choice. As a result of this complexity, most of these models lack a spatial dimension, and hence only aggregate effects on the economy can be determined.

Congestion, accident and pollution externalities were simulated together by Mayeres (2000) to test the efficiency of peak road pricing, fuel taxes and public transport subsidies in the context of lump sum transfers and labour taxes. Households maximised a 10-level utility function comprising consumption, leisure and the time spent consuming transport goods (i.e. travel time), subject to a monetary constraint and time allocation constraints. The utility function also included government spending on public goods and negative contributions from air pollution and accidents. Congestion was reflected in the model as an increase in travel times, which were valued endogenously according to wages and preferences over travel. These travel times were modelled using a single link congestion function that represented the road network.

Parry and Bento (2001) examined how revenue from congestion pricing should be spent. Travel times were modelled as a function of trips, with congestion charges levied per trip. Their CGE model suggested that if the revenue is redistributed as transfer payments, as is commonly assumed in other models, labour supply could be discouraged, leading to a welfare loss outweighing the gain from reduced congestion. On the other hand, an equivalent reduction in labour taxes could lead to a significant welfare gain from the improved efficiency of the labour market. Parry and Bento (2002) introduced suboptimal public transport pricing, congestion on competing routes and fuel taxes into the previous analysis, incorporated accident and pollution externalities into household utility and modelled travel times using a link congestion function.

Vandyck and Rutherford (2014) later integrated aspects from NEG in a CGE model with congestion, agglomeration and unemployment to study the efficiency and equity of road
pricing in Belgium. Similar to other NEG papers, Vandyck and Rutherford argued that congestion pricing should be lowered when there are agglomeration externalities and other inefficient taxes. The study demonstrated that commuters themselves could experience a welfare gain from congestion pricing if the revenue is distributed appropriately.

Other models have gone further in representing the production of transport services and their interaction with industry. Conrad (1997) developed a model to determine optimal levels of transport infrastructure and investment in the context of congestion and taxes. Production in Conrad’s model comprised capital, labour, energy, material and transport, where the transport input itself was a combination of transport services and capital. Transport capital was a function of the stock of trucks and the availability of infrastructure. The productivity of transport was diminished by a congestion index equal to the ratio of transport capital to a baseline value. In a multi-step analysis, Conrad gradually introduced constraints and mechanisms into the model to calculate the optimum level of infrastructure as funded by a fuel tax, which itself incurred a dead-weight loss. Conrad and Heng (2002) used a similar model to examine whether increased fuel taxes to finance the reduction of road bottlenecks could be covered by lowered congestion costs.

4. Applicability of CGE models in transport appraisal

The CGE models in the literature span a variety of scopes and interactions, both for freight and household transport. Many were developed to investigate projects requiring special analysis, for example to determine the impact to GDP from improved freight efficiency. While these models could be applied again to new projects that require it, from the perspective of planners and decision makers, the question is whether CGE models can go beyond the role of standalone assessment and be formalised as part of the appraisal process. Some jurisdictions are beginning to recommend the use of CGE models to derive inputs for appraisal in areas where the standard CBA process is considered lacking.

CGE models simulate the behaviour of all markets in an economy linked together, and thus can trace the effects of a transport improvement flowing through to other markets. The responses of economic agents and markets iterate continuously until equilibrium, yielding the long-term distribution of impacts. This provides a holistic and flexible framework for estimating WEIs as the mechanisms of agglomeration can be incorporated from first principles, including linkages between markets and knowledge spillovers, as well as other market imperfections. Nearly any economic metric can be extracted, including welfare measures such as equivalent variations and consumer surplus, GDP and other economic indicators. As model outputs will reflect both direct and indirect effects working in tandem, it can be difficult to disentangle the change in welfare attributable to WEIs. One method is to separately calculate the direct effects using conventional methods, and then subtract this from the total effects from the CGE model (used in Hof et al. (2011) for CGEurope). In other models with detailed modelling of imperfect markets, individual surpluses from firm profits, taxes and other sources of lost welfare can be summed.

Currently, only the Netherlands recommends the use of CGE models to estimate WEIs in transport appraisal guidelines (Wangsness et al., 2016). Instead of applying formulae to estimate welfare uplift as a proportion of direct impacts, this approach allows for both negative and positive WEIs depending on market conditions. It has also enabled the Netherlands to monetise more WEIs than in any other guidelines as each market imperfection can be incorporated into the behaviour of the CGE model. These include the WEIs of agglomeration externalities, labour market effects and impacts in markets with imperfect competition, as well as impacts from inefficient land-use regulation. There are other appraisal guidelines which recognise CGE models, but only for use in a supplemental economic impact analysis (Transport for NSW (2013) for example). This view relates more to mainstream CGE modelling which focuses on macroeconomic results rather than welfare.
CGE models can also offer a different perspective for measuring the total change in welfare from a transport improvement. Household agents are simulated directly in a CGE model and transport demand can be generated from household activities. This adds significant detail to their behaviour, enabling changes in welfare to be measured from utility functions at the household level. For example, welfare can be calculated accounting for the desire of a traveller to maintain a constant travel time budget (Anas, 2015), whereas a CBA would assume all travel time savings are valued at a constant rate with an exogenous value of time. Metz’s (2008) concern with the conventional CBA approach was that the focus on travel time savings is misleading as they tend to evaporate when behaviour adjusts, and that it is accessibility that is actually valued. This perspective is accommodated in a CGE framework as travel demand, as well as demand in all other markets, is elastic. Depending on the formulation, household utility in a CGE model will rise from a transport improvement due to increases in consumption and leisure, even if travel demand also rises as a result to negate the travel time savings that may be initially present.

Despite their advantages, it is impractical to supplant CBAs with CGE models for appraisal. Formulating, calibrating and running a CGE model takes considerable time, data and effort, and the level of spatial detail cannot match that of a CBA. For example, appraising a small road intersection upgrade would be straightforward with a CBA, but would hardly warrant the use of a CGE model as the effects would likely be negligible in the wider economy. The CGE results would also be highly spatially aggregated, and disaggregating the model would lengthen run times to an impractical extent. A CGE analysis is most worthwhile when the transport improvement is expected to impact the economy—determining when this applies requires experience. Another significant problem is the lack of transparency (‘black box’ nature) of CGE models—their operation tends to be difficult to understand, and practitioners may not be happy to take outputs as given without questioning their derivation (Vickerman, 2007b). Unfortunately, this is an inherent problem of CGE modelling as the solution process involves significant computations. Koopmans and Oosterhaven (2011) suggest one method of alleviating concerns is to run the model with and without certain behaviours active to estimate the contribution of each behaviour to the final outcome. Finally, appraisals usually require a time series of costs and benefits, which can only be delivered with a CGE model if it is dynamic.

For now, aside from generating WEIs, the best use of CGE models in appraisal may be to assess welfare from the household perspective, which can then be compared with welfare from conventional CBAs. This form of appraisal would have no double counting of benefits and would provide an analysis of nth order effects. A CGE appraisal would also enable policy makers to determine the distribution of impacts among markets and economic agents, and would facilitate comparisons with investments in other sectors of government spending. It may even be possible to integrate CGE models and conventional CBA methods to an extent, for example by deriving parameters for the CBA such as values of time from the CGE model.

5. Issues with CGE modelling

Calibrating a CGE model requires the specification of a ‘benchmark’ dataset representing transactions between all agents in an economy, typically in the form of an input–output table or social accounting matrix. This dataset can be difficult to obtain or expensive to create, more so if it is spatially disaggregated. As CGE models are calibrated to replicate the benchmark dataset when no shocks are applied, there is an assumption that the benchmark dataset represents an economy at equilibrium. Some models are calibrated to time-series data, but this is the exception rather than the norm due to the substantial data requirements. Statistical estimation of parameters is difficult due to the large number of observations required as well as their partitioning into price and quantity variables (Shoven and Whalley, 1992). Miyagi (1998) and Ando and Meng (2009) claim that the calibration methods of CGE
modelling are less data intensive than comparable econometric methods since large data samples are not required for regression formulae.

Validation of CGE models is another area of concern to modellers. It is rare to see the forecasts of CGE models tested with external time-series data. CGE models will replicate the benchmark dataset perfectly due to the calibration process, but from here, the best that is usually done is to test the consistency of the model to check for errors in coding and data-handling. Sensitivity analysis may also be performed to test the robustness of the model to errors in parameter estimation. For modellers who are only interested in the qualitative effects of economic changes, this level of validation may be adequate, but others may find it unsatisfactory for the precise calculations required in an appraisal. Dixon and Rimmer (2013) suggested that a model can be tested by its ability to replicate historical data. Kehoe et al. (1995) analysed the performance of a CGE model of Spain, 10 years after its estimation, finding that its results were generally accurate. Kehoe (2003) later found that CGE models applied to NAFTA performed poorly, emphasising the need for ex-post evaluations of models to inform future models and improve confidence in the field. Partridge and Rickman (2010) advocated for time-series calibration with historical data validation to become standard practice.

There are also concerns regarding the theoretical foundation of CGE models. The normative aspects of general equilibrium theory have faced criticisms with regard to the realism of assumptions, non-uniqueness of equilibria and the use of representative agents. However, if CGE models are taken as an empirical framework to model the economy in terms of prices, quantities and agents, they will have value if they can accurately predict the state of the economy and can be modified to incorporate more realistic behaviour as theory becomes available. The principle behind their use in appraisal is that prices will tend towards equilibrium over time, which provides a consistent basis for measuring impacts and comparing proposals. CGE models seem particularly well suited to assessing infrastructure as the simulated ‘shocks’ have a physical interpretation—they represent the change in infrastructure, which is an exogenous factor that drives the model. In any case, the behaviours underlying CGE models are based on the same economic theory underlying conventional appraisals, including the basic optimisation models used to derive values of time. CGE models are in essence an extension of conventional appraisals through the introduction of a price mechanism to simulate movements in all markets.

6. Future research directions

As highlighted throughout the review, some CGE models are currently more suitable for appraisal than others. For example, RELU–TRAN would be appropriate for an urban infrastructure project as transport is simulated as a network. RAEM, on the other hand, lacks the integrated network representation but introduces detail in modelling agglomeration economies and macroeconomic behaviour. Partridge and Rickman (2010) describe a number of features that future CGE models should incorporate to improve their usefulness for regional policy analysis. The location behaviour of agents should be influenced by the attractiveness of a region, whether by allowing for regionally-differentiated taxes in the model, or by adding some factor to account for the consumption of local amenities. Urban models such as RELU–TRAN do include parameters for regional desirability, but these parameters are generally calculated as residuals rather than being linked to government expenditure. Features from NEG such as mechanisms for agglomeration externalities are necessary in some circumstances. Models should also have well-specified linkages to account for openness between regions, with friction. This would particularly apply to labour markets, where the assumptions of perfect mobility and full employment should be relaxed, and labour forces should be differentiated by skill.

In general, CGE models for transport require further integration with transport models to account for a project’s actual impact on the transport network. Most CGE models incorporate
changes in transport infrastructure as a transport industry ‘technological improvement’ or change in capital stock, both of which neglect network impacts by assuming that transport infrastructure is equally effective regardless of its position in the network. Future models should correspondingly feature smaller regions (to the extent possible with available computing power) to identify the local impacts of changes in infrastructure.

The models examined in this review provide the ingredients to construct a full model for transport appraisal:

- Models from urban economics account for land markets and migration, and demonstrate how discrete choice models can be integrated.
- Spatial CGE models describe the behaviour of macroeconomic agents.
- Models from the SPE literature describe methods of linking transport network models and CGE models.
- NEG models account for the behaviour of regional markets, particularly in terms of imperfect markets, and resulting patterns of spatial development.
- Models from congestion and externality modelling have detailed representations of household demand for transport and provide methods to account for community impacts, such as pollution and government spending.

Constructing a model with all of these aspects is a formidable task as there would be obstacles in finding a solution (there would be multiple equilibria) and interpreting the solution, as well as data and computational requirements. Nevertheless, this review should provide a menu of the wide variety of behaviours and applications that have already been studied in CGE models for transport, such that future modellers can understand how they have been accomplished and incorporate the aspects relevant to their project.

7. Conclusion

Technological advances have made the application of sophisticated simulation models of economic and transport systems viable in recent decades. This has provided the opportunity to improve on past methods of transport appraisal by relaxing assumptions inherent in the static formulae used to estimate impacts. One type of model that has become increasingly popular to analyse transport projects and policies is the CGE model. These models simulate every market in an economy through the actions of consumers, producers and other economic agents. Being built from microeconomic first principles, it is possible to extract a rich array of outputs and represent relationships between transport and the economy that would be difficult to simulate in any other model. However, CGE models are still unknown to many in the transport planning domain, and therefore this paper aimed to discuss their potential application in the appraisal process.

There are two well-known shortcomings of conventional CBAs: firstly, metrics from conventional CBAs do not account for imperfect markets and externalities, and secondly they do not provide the long-term distribution of benefits. Both of these are highly relevant to transport planning nowadays. For the former, transport projects are often justified in part by their potential for economic development, and a range of ad-hoc methods have been used to estimate WEIs in recent appraisal practice. For the latter, knowing the distribution of benefits is valuable as equity has a significant influence over project prioritisation due to its social and political importance.

CGE models can provide a unified framework to estimate WEIs as well as the distribution of both direct and indirect impacts of transport improvements. Models have been developed for a range of spatial scales and behaviours, including urban models (with land markets, discrete choice structures and links with transport network models), regional models (freight-oriented and with backgrounds from conventional CGE modelling, SPE modelling and NEG modelling) and non-spatial CGE models of externalities (e.g. congestion and pollution). Nearly any linkage between transport and the economy can be simulated, and parameters
that would be static in a CBA can be made endogenous. Utility can be measured at the household level after responses of economic agents and markets have iterated until equilibrium. Metrics such as equivalent variations, consumer surplus and GDP can then be extracted. However, CGE models cannot replace CBAs as the sole method of appraisal as they are costly to build and are more spatially aggregated. At this stage, it may be most appropriate to use CGE models to extend conventional CBAs, integrate their outputs with CBAs (e.g. to estimate parameters for CBAs) or to use them as an alternative method of appraisal for comparison. There are also a number of issues regarding calibration and validation that may need to be resolved before CGE models are acceptable to transport planning practice.

References


