General equilibrium models

for transportation economics (*)

By

Johannes Bröcker(**) and Jean Mercenier(***)

(Final version: May 2010)

(*) Prepared for
Handbook in Transport Economics,
de Palma A., R. Lindsey, E. Quinet and R. Vickerman (eds)
Edward Elgar Pub.

(**) Kiel University, Institute for Regional Research
Wilhelm-Seelig-Platz 1, 24118 Kiel, Germany
http://www.uni-kiel.de/ifr/

(***) ERMES, Université Panthéon-Assas (Paris 2)
12 place du Panthéon, 75230 Paris Cedex 05
http://mercenier.u-paris2.fr/

1. Introduction

Applied --or computable-- general equilibrium models (AGE or CGE) build on rigorous modelling of microeconomic agents' behaviours (households, firms,…). These agents are exposed to signals (in prices, quantities…) provided by markets (for goods, assets, production factors…). Agents make decisions by explicit maximization of their own criterion (utility, profits, portfolio returns…). These choices determine their positions on each market. From the interaction between these supply and demand decisions, and conditional on the form of organization that prevails on each market (perfect competition, monopolistic or oligopolistic competition…), new signals emerge that feed back on the optimal decisions of all agents. The general equilibrium (GE) typically describes a stable state of consistency between these individual decisions: when the signals that condition individual choices coincide with those emitted by markets so that there is no incentive for anyone to change position. The computation of a GE therefore consists in determining a system of signals and an allocation between individuals, sectors of activities, regions, possibly time periods… such that all agents are at their optimum yet satisfying their respective constraints (budget, technological…) and that the set of transactions conducted on each market corresponds to the desired set of transactions by all agents simultaneously.

Governments of course have the ability to influence both directly (by taxes, transfers…) and indirectly (by their own demand and supply decisions on individual markets…) the environment that agents face and therefore their behaviours and the resulting resource allocation. It should be clear from what precedes that, in principle at least, any kind of microeconomic behaviour and any degree of disaggregation of agents can be built in an applied GE, and it will always be possible to evaluate and compare equilibria in terms of individual welfare. For this reason applied GE models are today indispensable tools of policy analysis. See Shoven and Whalley (1984) for an introduction and Ginsburgh and Keyzer (1997) for an advanced textbook presentation; see Srinivasan and Whalley (1986), Mercenier and Srinivasan (1994) and Fossati and Wiegard (2002) for illustrative applications. Bröcker (2004) provides an alternative introduction to CGE applications to transport problems. For solution software, computer codes and illustrative applications, see http://www.gams.com/.

There is no free lunch however: computations can be extremely costly. For this reason transportation economics has, until recently, mainly relied on the restrictive cost-benefit approach. The traditional cost-benefit evaluation of a new road, say, measures the benefit by the consumer surplus of users generated by reducing generalized costs, and subtracts building costs in market values and the net increase of technological external costs caused by existing and induced traffic. For this approach to be valid requires the following three conditions to hold: (a) markets are perfectly competitive and cleared by fully flexible prices; (b) welfare distribution is not an issue, that is, each euro counts equally, irrespective of who gets it; (c) technological externalities outside the transport sector are negligible. None of these conditions are particularly appealing to modern economists and policy makers so that with the spectacular development of computing possibilities, the CGE approach is becoming increasingly popular in transportation economics. A typical transport economics application is to study quantitative impacts of transport initiatives like infrastructure investments or pricing policies on economic variables.

It is the aim of this chapter to provide an introduction to the use of the CGE approach in transportation policy evaluation. For this, we start —in Section 2— by a short tutorial on the CGE methodology, and introduce what constitutes the core elements of most –if not all— CGE models. Having set the stage, we then discuss how transport is introduced in applied GE models (Section 3). The chapter closes with a brief conclusion.

2. A short introduction to CGE modelling

Any AGE model builds on a data matrix that accounts for all the transactions operated in the economy during a base period: we therefore begin this section with a short description of how these transaction data are organized. We then describe how preferences and technologies are specified and calibrated so that, in absence of shocks, the model replicates the base-year data set. For this, we first assume perfect competition prevails in a closed economy setting. The basic model is then extended (i) to acknowledge the possible existence of increasing returns to scale technologies and imperfect competition between firms; (ii) to multicountry/region models with trade.

2.1. The base year data set

Consider a closed economy comprising producers, households and a government. Producers are grouped into industries or sectors indexed s,t according to the type of goods they produce; households are grouped according to some characteristic — such as income class — indexed *h*.[1] During a specified period of time, all these agents simultaneously operate on different markets where they make transactions. Table 1 provides a symbolic representation of all these transactions organized in a meaningful way. Incomes (appearing with a negative sign) and expenditures of all agents are displayed so as to make explicit the consistency constraints imposed by the general equilibrium of the economy.[2] It is useful to explore this table in some detail.

	(a)	*(b)*	*(c)*	*(d)*	*(e)*
(1)	$(1+\tau_t^Z)p_t^Z X_{ts}$	$(1+\tau_t^Z)p_t^Z C_{ht}$	$(1+\tau_t^Z)p_t^Z G_t$	$(1+\tau_t^Z)p_t^Z I_t$	*Total demand for good t at market prices:* $(1+\tau_s^Z)p_s^Z\left[\sum_s X_{ts}+\sum_h C_{ht}+G_t+I_t\right]$
(2) = $\sum_t$ (1)	$p_s^X X_{\bullet s}$	$p_h^{Con} Con_h$	$p^{Gov} Gov$	$p^{Inv} Inv$	
(3)	$w_f F_{fs}^{dem}$	$-w_f \overline{F_{fh}^{sup}}$			*Balance of factor market f*
(4) = $\sum_f$ (3)	$p_s^Q Q_s$				
(5) = (2)+(4)	$p_s^Z Z_s$				
(6)	$\tau_s^Z p_s^Z Z_s$		$-\tau_s^Z p_s^Z Z_s$		
(7)		$\tau^{Inc}\sum_f w_f \overline{F_{fh}^{sup}}$	$-\tau^{Inc}\sum_{fh} w_f \overline{F_{fh}^{sup}}$		
(8)		Sav_h	Sav^{Gov}	$-p^{Inv} Inv$	*Investment – saving balance*
(9)	*Total supply of good s at market prices:* $(1+\tau_s^Z)p_s^Z Z_s$	*Balance of household h's budget constraint*	*Balance of Government budget constraint*	*Balance of investors budget constraint*	

Table 1

[1] How finely defined are these industries or household groups is arbitrary, and will depend on the type of analysis. The modeler clearly faces a trade-off here: a finer disaggregation might provide richer answers, but it will require more and possibly less reliable data, it will necessitate additional possibly more questionable assumptions, and it will make the model more difficult to solve and the predictions more difficult to interpret.

[2] Though slightly different in presentation, this table is conceptually identical to what is known in the literature as a Social Accounting Matrix (SAM).

Column (*a*) details the cost structure of sector *s* (there is one such column for each sector) with line (1) reporting payments to industry *t* (there is one such line for each industry) for material inputs bought in quantities X_{ts} at market prices $(1+\tau_t^Z)p_t^Z$, where p_t^Z denotes the producer price and $\tau_t^Z p_t^Z$ a unit *ad valorem* tax levied on the producer's output. Total material input costs are reported in line (2) where we have introduced, as is always possible, an aggregate price index for material inputs, p_s^X, and $X_{\bullet s}$ the number of units of the corresponding aggregate bundle of intermediates, so that $\sum_t (1+\tau_t^Z) p_t^Z X_{ts} = p_s^X X_{\bullet s}$. (The way the latter two aggregate variables are related to the others will be detailed later.) Sector *s* also rents production factors indexed *f* at unit prices w_f in quantities F_{fs}^{dem} as reported in line (3) –with one such line for each *f*. Summing over factors provides the sector's value added at factor costs, reported in line (4), where again we introduce an aggregate price index p_s^Q and the corresponding quantities Q_s so that we have $\sum_f w_f F_{fs}^{dem} = p_s^Q Q_s$. Summing total expenditures on both material and factor inputs defines (reported in line (5)) the sector's output value with Z_s the number of goods supplied at unit price p_s^Z. The good is taxed at *ad valorem* rate τ_s^Z (line (6) reports the amount of taxes levied on the sector's output) so that the market value of output *s* is $(1+\tau_s^Z) p_s^Z Z_s$ reported in line (9).

Column (*b*) reports all ingredients of the household *h* budget constraint (there is one such column for each *h*). Income is earned by supplying factor services to firms (line (3)), and shared between taxes (line (7)), savings (line (8)) and consumption of goods (line (1)). Summing line (1) over all *t* defines *h*'s aggregate consumption expenditure (line (2)), where a consumption basket with unit price p_h^{Con} has been implicitly defined. Line (9) reports the household's balance between accounted incomes and expenditures, which we know should always be null. Column (*c*) similarly reports all ingredients of the government budget constraint. Column (*d*) is associated with a fictitious investor that "spends" the economy's total saving (line (8), the investor's "income") on market goods (line (1)), combining them into investment composites with unit price p^{Inv} such that $\sum_t (1+\tau_t^Z) p_t^Z I_t = p^{Inv} Inv$ (line (2)). The investor's budget balances to zero (line (9)). Now, adding cells (*a*) to (*d*) from line (1) defines total expenditures on each industry's good—reported in cell (1,*e*)— which, by construction, equals the total supplied value of that good—displayed in cell (9,*a*).[3]

Observe that, by construction, factor markets balance (cell (3,*e*) reports the line sum that is null), investment-spending equals the economy's supply of saving (cell (8,*e*) reports the line sum that is null), and all agents satisfy their budget constraints (cells (9,*b*), (9,*c*), and (9,*d*)).

2.2 Specification and calibration

The economy underlying Table 1 is populated by agents that take into account signals provided by markets and make rational choices by optimizing some criterion subject to their technological and/or budget constraints. Our task as applied GE modellers is then to make assumptions on market structures prevailing at base year, to postulate functional forms for preferences and technologies, solve each agent's optimization problem, and set parameter values such that, in absence of shock, each decision maker replicates its base year transaction flows as reported by the data matrix.

To illustrate this, assume that all markets are perfectly competitive and technologies have constant returns to scale. Producer *s* will naturally seek to minimize its production cost conditional on some output target. For example, with a Cobb-Douglas technology, the firm solves, holding Z_s fixed,

$$\begin{aligned}&\underset{\left\{X_{ts},F_{fs}^{dem}\right\}}{\text{Min}} \quad \sum\nolimits_t (1+\tau_t^Z)\,p_t^Z X_{ts} + \sum\nolimits_f w_f F_{fs}^{dem}\\ &\text{s.t.} \quad \ln Z_s = \alpha_{0s}^Z + \sum\nolimits_t \alpha_{ts}^X \ln X_{ts} + \sum\nolimits_f \alpha_{fs}^F \ln F_{fs}^{dem} \qquad \sum\nolimits_t \alpha_{ts}^X + \sum\nolimits_f \alpha_{fs}^F = 1\end{aligned}$$

where all the symbols have been previously introduced (and appear in column (*a*) of Table 1) except the α–coefficients which denote parameters (elasticities and a scale). Most industry observers would however advocate for a less restrictive technology than one that imposes identical substitution elasticities between any pair of inputs. A more realistic alternative consists to group similar inputs into bundles, and to characterize substitutability differently within each bundle. Assume for instance that in industry *s*, it is known that substitution is easy between capital and labour, but that complementarity prevails between material inputs, yet

[3] Our presentation implicitly assumes that all industries are perfectly competitive. Assuming imperfect competition in some industries would only require mild reinterpretation of some variables, as shown later.

that both input bundles account for a constant share of total cost. We could model this by nesting technologies: primary factors would be combined using a constant elasticity of substitution (CES) production function to yield an aggregate factor service called "value added"; material inputs would enter a Leontief-type sub-technology to produce an aggregate "intermediate mix", which would then be combined with value added using a Cobb-Douglas to produce the final good. Endowed with such a technology — illustrated in the left part of Figure 1 — the producer's decision problem looks much more complicated, but we know from Gorman (1959) that it can be decomposed into small and easy to handle sub-optimization problems because all sub-technologies are additively separable. We now show how such a technology can be calibrated to fit the data in Table 1.

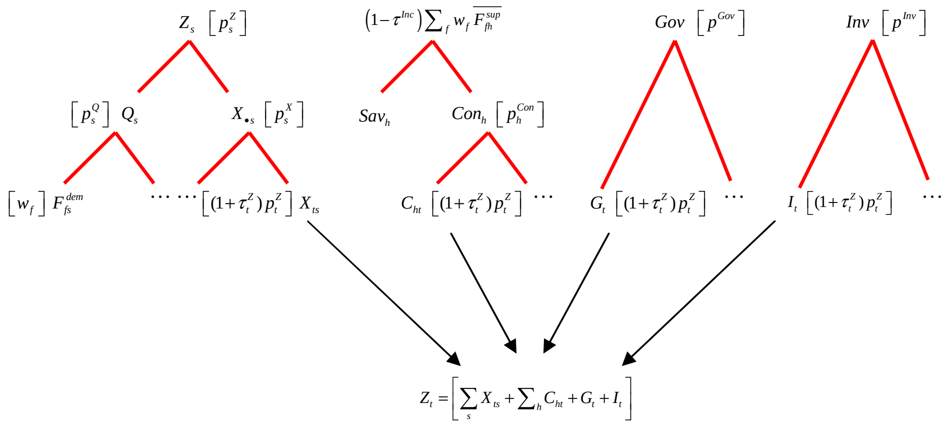

Figure 1

We start with primary factors: the sector s producer's sub-problem consists of choosing the mix of factor services that minimizes costs of producing some specified level Q_s of value added, given market prices for factors and a CES technology. Formally, it will

$$\underset{\{F_{fs}^{dem}\}}{\text{Min}} \quad \sum_f w_f F_{fs}^{dem}$$

$$\text{s.t.} \quad Q_s = \left\{ \sum_f \alpha_{fs}^F \left[F_{fs}^{dem} \right]^{-\rho_s^Q} \right\}^{-\frac{1}{\rho_s^Q}} \qquad -1 < \rho_s^Q < \infty$$

(for a given level of Q_s) where α_{fs}^F and ρ_s^Q are parameters; $\sigma_s^Q = 1/(1+\rho_s^Q)$ is the elasticity of substitution between factors. The first-order conditions of this problem are immediately derived as:

$$\begin{cases} w_f F_{fs}^{dem} = \left[\alpha_{fs}^F\right]^{\sigma_s^Q} \left[\dfrac{p_s^Q}{w_f}\right]^{\sigma_s^Q - 1} p_s^Q Q_s & \forall f \\ \left[p_s^Q\right]^{1-\sigma_s^Q} = \sum_f \left[\alpha_{fs}^F\right]^{\sigma_s^Q} \left[w_f\right]^{1-\sigma_s^Q} & \end{cases}$$

where the second equation, obtained by substitution of optimal factor demands into the constraint, relates the Lagrange multiplier to the factor prices w_f in a way that completely accounts for the technology. It is easily checked that p_s^Q necessarily satisfies

$$p_s^Q Q_s = \sum_f w_f F_{fs}^{dem}$$

(as stated in Table 1, cell (4,*a*)) and can be interpreted as the sector's value added price index. Normalizing this price to unity only affects measurement units of value added and is therefore innocuous; for the same reason, factor prices can in general be set to unity at base year. It is then straightforward to calibrate factor demands to fit the data in Table 1:

$$\begin{aligned} \left[\alpha_{fs}^F\right]^{\sigma_s^Q} &= \frac{w_f F_{fs}^{dem}}{p_s^Q Q_s} \qquad \forall f \\ &= \frac{\text{cell } (3,a)}{\text{cell } (4,a)} \end{aligned}$$

This completes the calibration of factor demands even though we are unable to identify σ_s^Q from the share-parameters α_{fs}^F : values for substitution elasticities have to be provided from outside information.

We next turn to intermediate goods that are combined, assuming complementarity, into an aggregate material input mix. Optimal demands for intermediate goods X_{ts} are derived, for given levels of $X_{\bullet s}$, from cost-minimization taking prices as given using Leontief technologies:

$$\begin{aligned} &\underset{\{X_{ts}\}}{\text{Min}} \quad \sum_t (1+\tau_t^Z) p_t^Z X_{ts} \\ &\text{s.t.} \quad X_{\bullet s} = \frac{X_{ts}}{\alpha_{ts}^X} \qquad \forall t \end{aligned}$$

where α_{ts}^X is now the amount of good from industry t necessary for sector s to produce one unit of aggregate intermediate input. Optimal demands immediately follow as:

$$X_{ts} = \alpha_{ts}^X X_{\bullet s} \quad \forall t$$

and the aggregate intermediate price p_s^X satisfies

$$p_s^X X_{\bullet s} = \sum_t (1+\tau_t^Z) p_t^Z X_{ts}$$

from which we get

$$p_s^X = \sum_t (1+\tau_t^Z) p_t^Z \alpha_{ts}^X$$

where again it is innocuous to choose units of the intermediate bundle so that $p_s^X = 1$. Calibration of the α_{ts}^X is straightforward: from Table 1, we know the amount paid to industry t as a share of total expenses on intermediate goods:

$$\frac{(1+\tau_t^Z) p_t^Z X_{ts}}{p_s^X X_{\bullet s}} = \frac{\text{cell }(1,a)}{\text{cell }(2,a)} \ ;$$

eliminating X_{ts} using optimal demands and p_s^X thanks to normalization, the left-hand side becomes $(1+\tau_t^Z) p_t^Z \alpha_{ts}^X$ .We easily get the base-year tax rate τ_t^Z as:

$$\tau_t^Z = \frac{\text{cell }(6,a)}{\text{cell }(5,a)} \ ;$$

set $p_t^Z = 1 \ \forall t$ at base year (as will be justified soon), it immediately follows that

$$\alpha_{ts}^X = \frac{1}{(1+\tau_t^Z)} \bullet \frac{\text{cell }(1,a)}{\text{cell }(2,a)} \ .$$

We finally turn to the upper-level of the technology, where value added and the aggregate bundle of intermediate goods are combined knowing that expenditure shares are thought to be constant so that the sub-technology is a Cobb-Douglas and the optimization sub-problem writes as:

$$\begin{aligned} &\underset{\{Q_s, X_{\bullet s}\}}{\text{Min}} \ p_s^Q Q_s + p_s^X X_{\bullet s} \\ &\quad \text{s.t.} \quad \ln Z_s = \alpha_{0s}^Z + \alpha_s^Q \ln Q_s + \alpha_s^{X\bullet} \ln X_{\bullet s} \end{aligned}$$

with given prices and output level Z_s. The solution is:

$$\begin{cases} p_s^Q Q_s = \alpha_s^Q p_s^Z Z_s \\ p_s^X X_{\bullet s} = \alpha_s^{X\bullet} p_s^Z Z_s \\ \ln p_s^Z = \alpha_s^Q \ln p_s^Q + \alpha_s^{X\bullet} \ln p_s^X \end{cases}$$

with

$$p_s^Z Z_s = p_s^Q Q_s + p_s^X X_{\bullet s}$$

and the shares are immediately calibrated from the data :

$$\alpha_s^Q = \frac{\text{cell }(4,a)}{\text{cell }(5,a)}$$

$$\alpha_s^{X\bullet} = \frac{\text{cell }(2,a)}{\text{cell }(5,a)}$$

Observe once again that the data provide information on equilibrium values of flows at base-year: we are therefore free to normalize output prices p_s^Z to unity and to define output volumes consistently. Collecting the terms for sector *s*, we get the producer part of Table 2.

We should proceed in a similar way for each of the other agents in this economy (see Figure 1) but, for space-saving reasons, we leave this as an exercise for the reader. The model is then completed by adding equilibrium conditions on each market. Table 2 displays a complete illustrative CGE system. We have there assumed constant factor supplies ($\overline{F_{fh}^{sup}}$) and CES preferences for each household *h*, and parameterized the saving (μ) and income tax rates (τ^{Inc}); *h*'s budget constraint therefore determines its aggregate consumption level Con_h. For the Government, we have assumed Leontief preferences and exogenous real aggregate consumption ($\overline{G}$); given that tax rates have been parameterized, it is the deficit/surplus that will have to adjust to satisfy the budget constraint of the public sector with this specification. (Other specifications are of course possible: clearly, which of the variables are left free to adjust and which are kept fixed will depend on the type of policy explored). Investors are assumed to use CES technologies to combine final goods into a capital aggregate in amount Inv consistent with the economy's supply of savings.

The reader should observe that:

(a) All coefficients in this economy of Table 2 have been calibrated on the base year data set, except substitution elasticities for which we rely on outside information: econometric estimates should in principle be used, but this could be extremely tedious if, as is usually the case, there are many sectors, households and factors. Also, results tend often to be quite robust to small changes of these substitution elasticity values. For this reason, CGE models often rely (arguably excessively so) on “guestimates” (meaning: an educated guess) and favour *ex post* sensitivity analyses: see below.

(b) All agents are, by construction, on their budget constraints in this economy; by Walras’ law, one market equilibrium condition is redundant and could be dropped from the system. Therefore, only relative prices are determined, not absolute price levels: a numéraire good has to be arbitrarily chosen, and all values are expressed in units of that good.

(c) A general equilibrium of this economy is an allocation (quantities produced, consumed…) supported by a vector of prices that solves a square non-linear system of equations. By construction, with unchanged levels of exogenous variables ($\overline{F_{fh}^{sup}}, \overline{Gov}$) and policy parameters (τ_t^Z, τ^{Inc}), the computed equilibrium will replicate the base year data. (It should be clear that numerous different model specifications can be made consistent with the same base year data by calibration. Calibration is therefore only a convenient way to force consistency on a specific model choice, it does not validate nor provide a selection mechanism.) To analyze the impact of a policy change, the model can be simulated by altering the relevant policy parameter/variable and computing the new equilibrium. Results are then reported as per-cent deviations from initial equilibrium values.

(d) *Ex post* sensitivity analysis consists to recalibrate the model and perform the same policy experiment for alternative values of some (in particular guestimated) parameters within a reasonable range, and check whether the policy conclusions remain qualitatively unchanged. If this is not the case, then additional statistical work is presumably called for to identify a most accurate value for that parameter.

(e) Rarely mentioned by CGE modellers, a problem arises from the possibility that equilibria may not be unique (see Kehoe, 1991). Obviously, the whole benchmarking-calibration exercise is on a different logical level in a world with multiple equilibria, and it is not clear what the comparative statics policy exercises really mean in such circumstances: which is the “relevant” equilibrium to pick among the set of possible solutions? It is remarkable that no case of multiple solutions has been reported to be encountered in calibrated applied GE models of *competitive* economies, so that, to date, whether or not non-uniqueness of equilibria is more than a theoretically possible occurrence remains an open question.

Table 2: A simple CGE model of a closed perfectly competitive economy

- Producer *s*

$$\begin{cases} X_{ts} = \alpha_{ts}^{X} X_{\bullet s} & \forall t \\ p_s^X = \sum_t (1+\tau_t^Z) p_t^Z \alpha_{ts}^X \end{cases}$$

$$\begin{cases} w_f F_{fs}^{dem} = \left[\alpha_{fs}^F\right]^{\sigma_s^Q} \left[\dfrac{p_s^Q}{w_f}\right]^{\sigma_s^Q - 1} p_s^Q Q_s & \forall f \\ \left[p_s^Q\right]^{1-\sigma_s^Q} = \sum_f \left[\alpha_{fs}^F\right]^{\sigma_s^Q} \left[w_f\right]^{1-\sigma_s^Q} \end{cases}$$

$$\begin{cases} p_s^Q Q_s = \alpha_s^Q p_s^Z Z_s \\ p_s^X X_{\bullet s} = \alpha_s^{X\bullet} p_s^Z Z_s \\ \ln p_s^Z = \alpha_s^Q \ln p_s^Q + \alpha_s^{X\bullet} \ln p_s^X \end{cases}$$

- Household *h*

$$Sav_h = \mu\left(1-\tau^{Inc}\right)\sum_f w_f \overline{F_{fh}^{sup}}$$

$$p_h^{Con} Con_h = (1-\mu)\left(1-\tau^{Inc}\right)\sum_f w_f \overline{F_{fh}^{sup}}$$

$$\begin{cases} (1+\tau_t^Z) p_t^Z C_{ht} = \left[\alpha_{ht}^C\right]^{\sigma_h^C} \left[\dfrac{p_h^{Con}}{(1+\tau_t^Z) p_t^Z}\right]^{\sigma_h^C - 1} p_h^{Con} Con_h & \forall t \\ \left[p_h^{Con}\right]^{1-\sigma_h^C} = \sum_t \left[\alpha_{ht}^C\right]^{\sigma_h^C} \left[(1+\tau_t^Z) p_t^Z\right]^{1-\sigma_h^C} \end{cases}$$

- Government

$$p^{Gov}\overline{Gov} + Sav^{Gov} = \sum_s \tau_s^Z p_s^Z Z_s + \tau^{Inc} \sum_{fh} w_f \overline{F_{fh}^{sup}}$$

$$\begin{cases} G_t = \alpha_t^{Gov\bullet} \overline{Gov} & \forall t \\ p^{Gov} = \sum_t (1+\tau_t^Z) p_t^Z \alpha_t^{Gov} \end{cases}$$

- Investor

$$p^{Inv} Inv = \sum_h Sav_h + Sav^{Gov}$$

$$\begin{cases} (1+\tau_t^Z) p_t^Z I_t = \left[\alpha_t^{Inv}\right]^{\sigma^{Inv}} \left[\dfrac{p^{Inv}}{(1+\tau_t^Z) p_t^Z}\right]^{\sigma^{Inv} - 1} p^{Inv} Inv & \forall t \\ \left[p^{Inv}\right]^{1-\sigma^{Inv}} = \sum_t \left[\alpha_t^{Inv}\right]^{\sigma^{Inv}} \left[(1+\tau_t^Z) p_t^Z\right]^{1-\sigma^{Inv}} \end{cases}$$

- Equilibrium conditions

$$Z_t = \left[\sum_s X_{ts} + \sum_h C_{ht} + G_t + I_t\right] \qquad \forall t$$

$$\sum_h \overline{F_{fh}^{sup}} = \sum_s F_{fs}^{dem} \qquad \forall f$$

2.3 Introducing increasing returns to scale and imperfect competition.

In many sectors, increasing returns to scale technologies and imperfect competition cannot be assumed away. We show how this complication can be dealt with in an applied GE model.

2.3.1 The individual firm's increasing returns technology

With increasing returns to scale technologies, obviously, output scale matters and we need distinguish between individual firms and sector aggregates: we identify firm related variables by lower-case letters, while upper-case letters refer as before to industry aggregates (though for notation ease we here drop the sector index).

The most convenient is to introduce a distinction between variable inputs and fixed inputs. Variable inputs will typically include all of the intermediate inputs and some of the factor inputs, even though, to simplify the exposition, we shall neglect material inputs in what follows. Fixed quantities of some primary inputs are required to operate the firm at any positive level of output. Therefore, the total demand for a factor f by an individual firm can be expressed as: $f_f^{dem} = f_f^v + \overline{f_f^F}$, where superscripts v and F refer respectively to "variable" and "fixed" factors. The individual technology is then written as

$$z = F(\ ...\ , f_f^{dem} - \overline{f_f^F},\ ...) \quad , \quad f_f^{dem} \geq \overline{f_f^F} \quad \forall f$$

where z is the firm's real output, and $F(...)$ is linearly homogenous. The individual firm's problem is then to minimize costs of producing a specified target output level z:

$$\begin{aligned} &\underset{k^v, l^v}{\text{Min}} \quad \sum_f w_f\, f_f^v + fx \quad &&\text{s.t.} \quad z = F(\cdots, f_f^v, \cdots) \\ & &&\text{s.t.} \quad fx = \sum_f w_f\, \overline{f_f^F} \end{aligned}$$

where fx denotes the total fixed cost; this immediately yields the optimal input mix of variable inputs:

$$\frac{\partial F(\cdots, f_f^v, \cdots)}{\partial f_f^v} = \frac{w_f}{v}$$
$$vz = \sum_f w_f\, f_f^v$$

where v denotes the marginal (or variable-unit) cost which differs from the average (or total unit) cost due to the presence of fixed inputs by firms.

2.3.2 Imperfect competition and prices

Imperfect competition can take many different forms. Within a sector, goods may be assumed homogeneous or differentiated; this will bare consequences on the type of competition that can prevail in that industry. Firms will always be assumed to maximize profits, but the optimal price-cost margins will depend on whether the firm's strategic variable is assumed to be its selling price or its production scale (a firm can't of course choose both). Also important is whether the firm is assumed to expect, and therefore to take into account when making its optimal decisions, a strategic reaction by competitors to changes in its own behaviour. In all cases, industry concentration will matter: the equilibrium outcome of an oligopoly game will in general significantly differ from the one to emerge from a large group assumption. In applications, firms will most generally --if not always-- be assumed symmetric within a sector: that is, they will share the same technology and have the same size, so that they charge the same price albeit for possibly differentiated products. This is quite convenient because Herfindahl industry concentration indices are supplied by most statistical agencies, and can be shown to be the inverse of the number of firms under the symmetry assumption. Hence, using this outside information, it is possible to calibrate variables related to the individual firm from data on industry aggregates.

It is clearly beyond the scope of this chapter to detail all the possible alternative modelling of imperfectly competitive markets. For illustrative purpose, let us assume that products are homogeneous within the sector, and that the competitive game is "Nash in output" (or *Cournot-Nash*), that is: firms choose the level of their production scale to maximize profits, expecting no reaction from their competitors (a reasonable assumption if the number of competitors is large enough). Formally, the individual producer seeks to

$$\operatorname*{Max}_{z} \quad prof(z) = p^{z}(Z)z - (vz + fx).$$

where $prof(z)$ is the firm's profits and $p^{z}(Z)$ is the *equilibrium market price*. Observe that the former depends on the firm's output z and the latter on the aggregate supply Z in that industry. Solving the maximization problem with respect to z yields the famous Lerner pricing rule:

$$\begin{aligned}\frac{p^{z}-v}{p^{z}} &= -\frac{d\log p^{z}(Z)}{d\log z} \\ &= -\varepsilon^{C}(z, p^{z}(Z))\end{aligned}$$

where $\varepsilon^C(z, p^z(Z))$ measures the market *equilibrium price* elasticity with respect to the individual firm's output z: except in extremely simplified cases, this elasticity is a complicated object. It will typically depend on preference parameters underlying demand functions (i.e., substitution elasticities) as well as on market shares for which data are available at base year; $\varepsilon^C(z, p^z(Z))$ can therefore be calibrated. Assuming zero profits to prevail at base year between a known (from base-year Herfindahl indices) number of symmetric firms, and normalizing $p^z(Z)$ to unity, the variable unit cost v can be determined using the Lerner equation; the level of fixed costs fx follows then immediately. See Mercenier (1995a, 2002) for elaborations on this. In the simulations, firms within a sector will often be allowed to respond to changes in profitability by (costlessly) entering/exiting the market: the equilibrium number of competitors is determined by imposing zero supra-normal profits (the output price then equals the average production cost).

At this stage, it should be mentioned that non-convexities in production technologies generically imply that the equilibrium will not be unique. Mercenier (1995b) presents a numerical example of multiplicity in a large-scale applied GE model calibrated on real world data. It seems therefore that in this generation of CGE models, non-uniqueness of equilibria is not a theoretical *curiosum*, but a potentially serious problem. Disregarding this could lead to dramatically wrong policy appraisals.

2.4 Multi-country/region model with trade.

Our previous model lacks realism in that it assumes no trade with other countries or regions. Depending on the focus of the analysis, trade can be introduced either by setting a number of single-country models together and letting them interact, or by assuming that the country under consideration is so small that it does not affect equilibrium in the rest of the world: foreign prices and incomes are then treated as exogenous. In both cases, the modeller has to decide whether goods in an industrial category produced in different countries are identical from the customers' viewpoint.

One most popular assumption (known as the Armington (1969) assumption) is that goods from the same sector are differentiated in demand by countries of origin. The main justification for this specification is that, because of data restrictions and/or to simplify computations, the modeller works with highly aggregated sectors of activity; even if products are identical across countries at a very fine level of industry disaggregation, the composition of the aggregate basket of goods is unlikely to be identical across regions. The specification is attractive because it accounts for the large amount of cross-hauling (i.e., two-way trade in identical goods) observed in the data, and for the fact that even at fine levels of activity disaggregation, most countries produce goods in all product categories.

The simplest way to implement an Armington system is by assuming that all domestic agents buy units of a common composite basket composed of goods from all geographic origins. The composition and the price of this Armington good result as usual from cost minimization. To see how this is done, let i, j index countries or regions, and let E_{ijs} be the flow of sector s goods exported from i to j at prices $p_{is}^{E} = (1+\tau_{is}^{Z})\, p_{is}^{Z}$.[4] Assuming a CES aggregator, import demands by region j result from:

$$\underset{\{E_{ijs}\}}{\text{Min}} \ \sum_i p_{is}^{E} E_{ijs} \quad \text{s.t.} \quad E_{\bullet js} = \left\{ \sum_i \alpha_{ijs}^{E} \left[E_{ijs} \right]^{-\rho_{js}^{E}} \right\}^{-\frac{1}{\rho_{js}^{E}}} \qquad -1 < \rho_{js}^{E} < \infty$$

for given export prices and aggregate demand levels $E_{\bullet js}$; this yields:

$$\begin{cases} p_{is}^{E} E_{ijs} = \left[\alpha_{ijs}^{E} \right]^{\sigma_{js}^{E}} \left[\dfrac{p_{js}^{Arm}}{p_{is}^{E}} \right]^{\sigma_{js}^{E}-1} p_{js}^{Arm} E_{\bullet js} \\ \left[p_{js}^{Arm} \right]^{1-\sigma_{js}^{E}} = \sum_i \left[\alpha_{ijs}^{E} \right]^{\sigma_{js}^{E}} \left[p_{is}^{E} \right]^{1-\sigma_{js}^{E}} \end{cases} \tag{1}$$

with $\sigma_{js}^{E} = 1 \big/ (1+\rho_{js}^{E})$ the substitution elasticity, p_{js}^{Arm} the unit-price of the Armington aggregate, and $E_{\bullet js}$ is the amount of the sector s Armington good demanded by country j:

$$E_{\bullet js} = \left[\sum_t X_{jst} + \sum_h C_{jhs} + G_{js} + I_{js} \right].$$

The market equilibrium condition for good t in our model of Table 2 then becomes:

$$Z_{jt} = \sum_i E_{jit} \qquad \forall t\,.$$

---

[4] We could of course assume a specific ad valorem tax/subsidy rate on exports; this would however require amending the Government budget constraint, without bringing any additional insight.

Given the base-year bilateral trade data-matrix, we know the expenditure flows $\left(p_{is}^{E}E_{ijs}\right)$, as well as $\left(p_{js}^{Arm}E_{\bullet js}\right)=\sum_{i}\left(p_{is}^{E}E_{ijs}\right)$; set $p_{js}^{Arm}=1$ at base year, and pick values of the substitution elasticities σ_{js}^{E} from outside trade-econometric evidence. The bilateral trade share-parameters can then immediately be calibrated as:

$$\left[\alpha_{ijs}^{E}\right]^{\sigma_{js}^{E}}=\left[p_{is}^{E}\right]^{\sigma_{js}^{E}-1}\bullet\frac{\left(p_{is}^{E}E_{ijs}\right)}{\left(p_{js}^{Arm}E_{\bullet js}\right)}.$$

Observe that with the above specification, even the smallest country faces endogenous terms of trade and enjoys some market power, though perfect competition can prevail among producers (and indeed implicitly prevails in our exposition as implied by our reference to Table 2) so that firms do not take advantage of this market power. In many sectors where production involves fixed costs, firms tend to choose specific product varieties and to specialize, taking advantage of their market power on the chosen niche. The previous framework can easily be extended to account for this possibility.

Let N_{is} be the number of firms producing differentiated varieties of good *s* in country *i*; assume that firms operating within the same country and sector are symmetric (same technology and same market shares, hence, same price) and let e_{ijs}^{f} be an individual *i* firm's sales to market *j*. As in the Armington case, this demand e_{ijs}^{f} can be derived from utility maximization in region *j* provided preferences are amended to acknowledge the existence of product varieties as follows:

$$\begin{aligned}E_{\bullet js}&=\left\{\sum_{i}\sum_{f}^{N_{is}}\left[e_{ijs}^{f}\right]^{-\rho_{js}^{E}}\right\}^{-\frac{1}{\rho_{js}^{E}}}\\&=\left\{\sum_{i}N_{is}\left[e_{ijs}\right]^{-\rho_{js}^{E}}\right\}^{-\frac{1}{\rho_{js}^{E}}}\end{aligned}$$

where the second equality takes account of the symmetry assumption between firms.[5] Cost minimization then yields:

[5] These preferences, associated with the name of Dixit-Stiglitz (1977), are also known as "love-for-variety" preferences because they acknowledge increasing returns in utility with respect to the number of available varieties. Similar technologies can be used by firms to combine intermediate inputs with differentiated varieties

$$\begin{cases} p_{is}^{E} e_{ijs} = \left[\dfrac{p_{js}^{DS}}{p_{is}^{E}} \right]^{\sigma_{js}^{E}-1} p_{js}^{DS} E_{\bullet js} \\ \left[p_{js}^{DS} \right]^{1-\sigma_{js}^{E}} = \sum_i N_{is} \left[p_{is}^{E} \right]^{1-\sigma_{js}^{E}} \end{cases}$$

where p_{js}^{DS} is the (Dixit-Stiglitz) price aggregator. Though this expression may look very similar to system (1) it is actually quite different: with entry/exit of firms into the industry due to zero supra-normal equilibrium profits, N_{is} will be an endogenous variable.

3. Introducing transport in CGE models

3.1 Transport in single-region models

So far, nothing has been said about transport. How does it enter the scene? At a first sight, transport is just one or a subset of commodities, produced by one or a subset of industries, consumed by households and used as an input by firms. These transport related sectors can be differentiated by transport object (passengers versus freight, bulk versus container), by distance class (short versus long), by mode and other characteristics. Typically, demand would be specified by some form of nesting as illustrated in Figure 1. A household could, for example, choose between consumption of travel services and consumption of other goods, and then, conditional on having chosen travel, how much of his private car and of public transportation to use. Thus, apparently, nothing has to be added to what has been explained so far.

At least three aspects of transport need however a special treatment:

(a) Transport demand and benefits generated by transport do not only depend on monetary cost, but also on time needed for travel or freight.

(b) Transport generates negative externalities within the transport sector itself—

---

to yield increasing returns to specialization. Such a technology, first introduced by Ethier (1982), is a key element in the "new economic geography" and in many models of the endogenous growth literature.

congestion—as well as outside the transport sector.

(c) Transport is in most cases not utility generating by itself, but it is instrumental for other activities such as working, shopping, tourism, visiting friends or events etc. Similarly, transport is not directly an input of firms, but instrumental for buying, selling or exchanging information.

Subsection 1.1 introduces transport demand of households depending on monetary cost as well as on travel time. Transport is then just a consumer good like other goods, and we disregard it being possibly instrumental for other purposes. In subsection 1.2 in contrast, we introduce commuting as a means to labour income earning. Transport demand of firms is treated in sub-sections 1.3 and 1.4 deals with transport related externalities.

3.1.1 Travel demand of households

An average UK citizen spends 87 minutes per day travelling and 14% of his or her total expenditure on transport (UK national statistics online, figures for 2005). If an hour travel time is valued at the hourly wage, time costs and monetary costs of travelling are of a similar magnitude, and the former can obviously not be neglected. Also, many transport policy measures mainly affect travel time, not monetary cost, and thus the time component is essential for policy evaluation. The household's allocation of time between work and leisure should now be modelled, and its labour supply therefore endogenized. See Jara-Diaz (2000) and Hensher (this volume, Ch. 19) for a review on allocation and valuation of travel time.

Commentaire [u1] : Editor: I did not put this into the list of references. Check whether reference is correct.

Let production factor $f = L$ be labour so that F_{Lh}^{sup} is the amount of labour the household decides to supply; T_h^l and $\overline{T_h}$ are respectively demand for leisure time and total time endowment. Finally, denote $t_{hs} \geq 0$ as the travel time associated with each unit of the consumption of good *s*. $t_{hs} = 0$, unless good *s* is travel. Sticking to the assumption in Table 2 of CES preferences and a fixed saving rate, the household's decision results from

$$\underset{\{C_{hs}, T_h^l\}}{\text{Max}} \left\{ \beta_h \left[T_h^l\right]^{-\rho_h^C} + \sum_s \alpha_{hs}^C \left[C_{hs}\right]^{-\rho_h^C} \right\}^{-\frac{1}{\rho_h^C}}$$

subject to the budget constraint

$$\sum_s (1+\tau_s^Z) p_s^Z C_{hs} = (1-\mu)(1-\tau^{Inc}) \left[w_L F_{Lh}^{sup} + \sum_{f \neq L} w_f \overline{F_{fh}^{sup}} \right]$$

and the time constraint

$$T_h^l + F_{Lh}^{sup} + \sum_s t_{hs} C_{hs} = \overline{T_h}.$$

Using the latter constraint to substitute out F_{Lh}^{sup} from the budget equation, we get:

$$\sum_s (1+\tau_s^Z) p_s^Z C_{hs} = (1-\mu)(1-\tau^{Inc})\left(-w_L T_h^l - w_L \sum_s t_{hs} C_{hs} + w_L \overline{T_h} + \sum_{f \neq L} w_f \overline{F_{fh}^{sup}}\right).$$

This can be rewritten as

$$\omega T_h^l + \sum_s \pi_{hs} C_{hs} = (1-\mu)(1-\tau^{Inc})\left(w_L \overline{T_h} + \sum_{f \neq L} w_f \overline{F_{fh}^{sup}}\right)$$

where the net wage (net of saving and taxes) $\omega = (1-\mu)(1-\tau^{Inc}) w_L$ is the consumer's valuation of leisure time, $\pi_{hs} = (1+\tau_s^Z) p_s^Z + \omega t_{hs}$ is the cost to consumer per unit of consumption good *s*. If *s* denotes travel, this is usually referred to by transport economists as the "generalised cost per unit of travel". Solving the maximization problem yields:

$$p_h^{Con} Con_h = (1-\mu)(1-\tau^{Inc})\left[w_L \overline{T_h} + \sum_{f \neq L} w_f \overline{F_{fh}^{sup}}\right]$$

$$\pi_{hs} C_{hs} = \left[\alpha_{hs}^C\right]^{\sigma_h^C} \left[\frac{p_h^{Con}}{\pi_{hs}}\right]^{\sigma_h^C - 1} p_h^{Con} Con_h \qquad \forall s$$

$$\omega T_h^l = \left[\beta_h\right]^{\sigma_h^C} \left[\frac{p_h^{Con}}{\omega}\right]^{\sigma_h^C - 1} p_h^{Con} Con_h$$

$$\left[p_h^{Con}\right]^{1-\sigma_h^C} = \left[\beta_h\right]^{\sigma_h^C} \omega^{1-\sigma_h^C} + \sum_s \left[\alpha_{hs}^C\right]^{\sigma_h^C} \pi_{hs}^{1-\sigma_h^C}$$

Labour supply F_{Lh}^{sup} is then determined using the time constraint. The share parameters α_{hs}^C and β_h can be calibrated from base-year data as before, provided information is available on the household's time endowment, leisure and travel time.

Though straightforward, this approach has two drawbacks. The first is that, in its valuation of travel time, the household takes into account only that part of lost labour income that goes into consumption, neglecting the one that goes into saving. This is due to our restrictive assumption of a constant savings rate. To relax this assumption would require an intertemporal approach beyond the scope of this chapter. The second drawback is that econometric estimates of valuations of travel time savings (VTTS) are typically considerably smaller than the wage rate, even if corrected for income taxes and the saving rate.

Furthermore, VTTS vary significantly over travel purposes, being smaller for leisure trips than for commuting. This is indirect evidence that people prefer spending time travelling over spending time working (see Hensher, this volume, Ch. 19). The simplest way to take this into account is by adding a preference term $\sum_s \gamma_{hs} t_{hs} C_{hs}$ to the household's objective function, with γ_{hs} denoting the utility per unit of time spent on travel item *s*. The household's demand system becomes:

Commentaire [u2] : Editor, see above.

$$C_{hs} = \zeta_h \left[\alpha^C_{hs}\right]^{\sigma^C_h} \left[\lambda_h \tilde{\pi}_{hs}\right]^{-\sigma^C_h}$$
$$T^l_h = \zeta_h \left[\beta_h\right]^{\sigma^C_h} \left[\lambda_h \omega\right]^{-\sigma^C_h}$$

with

$$\tilde{\pi}_{hs} = \pi_{hs} - (\gamma_{hs} / \lambda_h) t_{hs} = (1 + \tau^Z_s) p^Z_s + \tilde{\omega}_{hs} t_{hs}$$
$$\tilde{\omega}_{hs} = \omega - \gamma_{hs} / \lambda_h .$$

The two new unknowns ζ_h and λ_h are obtained by the budget constraint and the restriction

$$1 = \left[\beta_h\right]^{\sigma^C_h} \left[\lambda_h \omega\right]^{1-\sigma^C_h} + \sum\nolimits_s \left[\alpha^C_h\right]^{\sigma^C_h} \left[\lambda_h \tilde{\pi}_{hs}\right]^{1-\sigma^C_h} .$$

λ_h is the Lagrangian multiplier associated with the budget constraint, i.e. marginal utility of income. If *s* denotes travel, $\tilde{\pi}_{hs}$ is now the generalised cost per unit of travel, with VTTS $\tilde{\omega}_{hs}$. The VTTS is the net wage, corrected for the preference term γ_{hs} / λ_h representing the marginal utility of spending time with travel type *s*, translated into monetary units by the term $1/\lambda_h$. Note that the VTTS now not only depends on the wage rate, but also on all prices, travel time, income and the time endowment. The larger γ_{hs} is, the more the VTTS is reduced compared to the specification without travel time in the utility function. As before, all parameters except the elasticity of substitution can be calibrated from observed benchmark data. The additional information needed is the VTTS for each item of travel demand necessary for the calibration of γ_{hs}. In what precedes, we have specified all preferences as one-level CES: obviously, everything can be extended to nested CES or other functional forms.

3.1.2 Commuting

An interesting application of CGE models in transport is to look at the interaction between commuting costs and the labour market. In many countries commuting costs are deducted from the tax base. This reduces distortions in job choice, but may distort residential location choices (Wrede, 2003). In order to quantify these distortions an explicit modelling of commuting costs is needed.

Assume commuters do not care whether they spend time working or travelling to their job. They will choose commuting modes so as to maximise hourly wage, net of commuting cost, where an hour covers working plus commuting time. Let *m* index travel modes and assume for simplicity that industry disaggregation is such that each travel mode is a specific sector. Then, C_{hm} is demand by *h* of a specific commodity with market price $(1+\tau_m^Z)p_m^Z$. Denote unit travel time-cost by t_{hm} and let T_h^W be gross working time, including commuting time. The household wants to maximize the net wage w_{Lh}^{net} per hour worked, subject to a "commuting production function" that combines travel quantities by modes as inputs to produce an aggregate "travel to work" service. The household thus obtains the net wage from solving

$$T_h^W w_{Lh}^{net} = \underset{\{C_{hm}\}}{\mathrm{Max}} \left\{ w_L \left(T_h^W - \sum\nolimits_m t_{hm} C_{hm} \right) - \sum\nolimits_m (1+\tau_m^Z) p_m^Z C_{hm} \right\}$$

$$\text{s.t.} \quad \left[\sum\nolimits_m \delta_{hm} C_{hm}^{-\rho_h^M} \right]^{-\frac{1}{\rho_h^M}} = T_h^W$$

where we have assumed a CES commuting production function. (Any level of nesting could of course be introduced here.) Note that the amount of the travel service needed is assumed proportional to work time. This implies that work time is varied in terms of person-days per week, not time per day. Solving the problem is straightforward and left to the reader. As before, share parameters δ_{mh} can be calibrated from observed commuting data by mode, while substitution elasticities $\sigma_h^M = 1/\left(1+\rho_h^M\right)$ have to be imported from econometric studies on responses of mode choice on generalised costs. In the household's decision problem we just have to replace F_{Lh}^{sup} in the time constraint with gross working time T_h^W, and w_L with the net wage w_{Lh}^{net}.

3.1.3 Firms

Firms buy transport services (for both passengers and freight) as a production input. In a multiregional model, firms' transport demands are explicitly related to the interregional flows of goods. In single region models they are treated just as any other input with one change: firms bear costs not only in money but also in time. Monetary costs for transport can be observed in a sufficiently detailed input-output table, while time costs have to be imputed using travel time information and VTTS estimates for firms. For producing the input "freight services", say, one introduces a production function with a service as output and transport quantities by mode as inputs. A nested CES is again convenient here. To take account of time costs, a simple trick helps: the transport quantity by mode is itself regarded as a Leontief composite of a transport service (produced e.g. by the trucking industry) and a service called "travel time" (i.e. spending time with employees, equipment and goods "on the road"). This service is simply introduced as another commodity, produced by an industry, that may for example only use labour (representing employee's time) and capital (representing capital costs of goods, and equipment bound in transport).

3.1.4 Endogenous travel times and externalities

Travel times have been assumed fixed so far. Transport infrastructure is however a collective good with congestion, and thus travel times depend on both capacity and aggregate demand. To model this, t_{hm} is usually related to the ratio of aggregate travel demand D_m to capacity K_m as follows:

$$t_{hm} = t_{hm}^0 + \alpha_{hm}^t (D_m/K_m)^{\varepsilon_m}$$

where t_{hm}^0 denotes free-flow travel time. Agents in the economy, be they households or firms, are of course assumed too small to perceive any influence they could have on D_m so that t_{hm} remains fixed in their individual optimization problems. The elasticity ε_m is notoriously difficult to estimate, and its reliability is questionable because congestion varies a lot across different parts of the network, time of the day and day of the week and year. For welfare evaluations of policies affecting congestion such as road pricing, fuel taxes and infrastructure investment, imperfect as they are, endogenous travel times are nevertheless an indispensable model ingredient.

Another important element are externalities imposed by transport on the other parts of the economy. The least demanding way to take them into account is to neglect externalities on firms, and to assume separability for households meaning that no household decision is affected by externalities. Formally this means to specify household utility as a function of externalities and a sub-utility that only depends on the decision variables. In this case, the solution of the equilibrium can be done first, disregarding externalities, and the evaluation of the welfare impact is added afterwards. The main difficulty is to get a reliable parameterisation of the welfare impact of externalities. This includes, first, a measure of how traffic emissions translate into immissions affecting the household's wellbeing, and second, a measure of the impact of the immissions on utility. Under additive separability one would subtract a linear function of damage indicators from the mentioned sub-utility. The coefficients of this linear expression are calibrated such that one reproduces the willingness to pay (WTP) for damage reductions in the benchmark equilibrium. The WTPs must be imported from econometric estimates using revealed or stated preference data.

3.2. Transport in multi-region models

Regions can have any scale from parts of the world (Asia, Europe,…) down to residential zones in an urban area. In a multi-location setting for an urban area the focus is on shopping, commuting and other passenger trips, and on residential and commercial location choices. On a larger regional or national scale the focus is on long distance passenger traffic and on freight. We here concentrate on the latter and briefly deal with urban models in subsection 3.3.

Our introduction of trade in subsection 2.4 assumed that the price charged by a firm exporting is the same as the price paid by the cross-border customer. Realism requires that, upon destination, goods be priced including "trade margins" (freight cost, wholesaling, storing, etc). Consider freight costs alone for expositional ease. The destination price p_{ijs}^{M} of good s exported by i to j is the mill price (possibly including local taxes though we neglect these hereafter) plus the freight cost f_{ijs}:

$$p_{ijs}^{M} = p_{is}^{E} + f_{ijs}.$$

f_{ijs} is of course paid to the industry producing the transport service from i to j. Industries producing a service driving a wedge between mill price and customer price are called margin industries. Wholesale and retail trade are other important margin industries. Transport policy affects the economy via its impact on f_{ijs}.

A Leontief technology is usually adopted to transform the good at the factory gate into the good at the location of the customer. Let θ_{ijs} denote transport service per unit of delivered good so that $\theta_{ijs}E_{ijs}$ is demand for the transport service associated with trade flow E_{ijs}. This service may be supplied by a firm located in the exporting region, or in the destination region.

In the latter case, $f_{ijs} = \theta_{ijs} p_{jt}^{Z}$ is the per unit transport margin with p_{jt}^{Z} the price of a unit transport service (sector t) in region j. Industries producing the transport service have monetary as well as time costs, which are taken account of in just the same way as explained for transport input of firms in subsection 3.1.3 above.

Obviously, introducing margin industries makes complex multiregional models even more complex. A popular alternative is the "iceberg" approach. Here, no transport service is produced: the exported good melts on its way from origin to destination, so that the exported quantities E_{ijs} differ from those that reach destination, denoted M_{ijs}. Let $0 < \psi_{ijs} < 1$ be the melting factor, so that $M_{ijs} = \psi_{ijs} E_{ijs}$. Assuming that transport is a competitive zero profit activity, the destination price is $p_{ijs}^{M} = p_{ijs}^{E} / \psi_{ijs}$, so that values at origin and destination are identical:

$$p_{ijs}^{M} M_{ijs} = p_{ijs}^{E} E_{ijs}$$

The parameter ψ_{ijs} has to be calibrated so that transport costs are a percentage share in trade value for the benchmark. Transport policies affecting trade costs can be evaluated by changing this parameter. Obviously, ψ_{ijs} can account for both monetary and time costs; furthermore, it may be made to depend on endogenous travel times.

3.3 Applications

3.3.1 Single-region models

Typical applications of single-region models to transport issues are the studies of Conrad (1997) and Conrad and Heng (2002) and a series of papers by Mayeres and Proost (e.g. 2001, 2004); see also the useful review by Munk (2003). Mayeres and Proost (2004) introduce a highly detailed structure of the transport market for passengers, distinguishing private versus business as well as different modes. They also introduce different types of households in order to identify distributional impacts of transport policies.

The applications have in common that they take account of congestion in some way. Conrad and Heng (2002) assume the effective stock of capital in the transport industry to be decreasing in capacity use. Their aim is to show whether a capacity increase in Germany is welfare improving. Given the calibration of the model, which is debatable regarding the congestion function, their answer is affirmative.

Models that cover private passenger flows take account of monetary as well as time costs determining demand decision, much in the same way as described in subsection 3.1.1 above. This leads to demand functions for transport with generalized costs substituted for prices. Most importantly, the VTTS becomes an endogenous variable that generally depends on all prices and income. Particularly, it depends on wages, thus introducing an interdependence between transport and the labor market (Berg, 2007). This approach allows for taking account of congestion in a more direct and less ad hoc way than in Conrad and Heng (2002). Relying on speed-flow relations from transport engineering one can make travel times depend on the volume of flows, given infrastructure capacities (see subsection 3.1.4 above). A policy affecting transport demand through taxes, fees or fuel prices, or a policy affecting capacity through infrastructure investment has a direct impact on travel times and possibly monetary travel costs; these in turn enter the demand decisions such that adjustments of congestion, travel times, prices, flows as well as transactions on all goods and factor markets eventually lead to a new equilibrium. In equilibrium all agents make their optimal choices, given prices as well as travel times determined by the equilibrium level of congestion. The most sophisticated brands of such models even take other externalities like noise, accident risks and air pollution into account (Mayeres and Proost, 2004). For the sake of simplicity preferences are usually assumed to be separable between utility from goods and travel on the one hand and environmental quality on the other, such that environmental externalities have an impact on utility, but not on decisions (Mayeres and Proost, 2004). It is thus neglected that people might for example travel more, if they move to the suburbs in order to escape from urban noise and air pollution.

Models of this brand seem to be an ideal framework for analyzing the impact of transport policies on a wide range of interesting variables such as transport quantities, congestion, incomes and prices. Even more important is their ability to assess welfare effects, for the aggregate economy and/or for different household types. They thus extend the classical welfare-theoretical cost-benefit analysis to a general equilibrium framework. There are however also drawbacks. One is the notorious uncertainty about elasticities, that is of course a general problem of CGE applications. The prior choice of functional forms that is usually left untouched in sensitivity analysis, might even be more problematic. Another drawback is that the macro style of these models averages out a lot of details, that could be decisive for the policy conclusions. A case in point is the macroeconomic congestion function. Congestion greatly varies by region, time of day, day of the week and from link to link. For calibrating the macro congestion function one must fix a point on the macro speed-flow schedule for the whole economy, which must be understood as some kind of average. But obviously, speed as a function of average flow can be very different from average speed, when the average is taken over speeds as functions of flows under a lot of different conditions regarding link, time of day et cetera. One can of course try to differentiate to any degree, but the lack of spatial detail remains a problem.

3.3.2 Multi-region models

Multiregional models aim at quantifying regional effects of transport policy, particularly of infrastructure investment. Typically, they introduce trade costs that are reduced by investing into certain transport links. An early contribution starting this literature is Buckley (1992). His model is a standard perfect competition approach with three regions and five industries. Interregional trade follows an Armington approach. Cost and expenditure functions are nests of either Leontief or CD functions. Transport is a Leontief complement of interregional flows. It is assumed to be produced at the place of origin. Buckley's experiment is to increase labour productivity in one region's transport sector. The results show how the welfare gain is distributed across regions.

Venables and Gasiorek (1998) improve upon this idea by allowing for more regions and industries, more flexible functional forms and – most importantly – by applying the Dixit-Stiglitz approach to monopolistic competition in the production sector. This brings scale effects into the impact analysis, which are not existent in the traditional perfect competition framework. Cost reductions lead to expansion of output; this in turn makes producers move down the average cost curve. This gives rise to effects that the SACTRA report (Department for Transport, 1999) has called "wider economic effects" of transport cost reductions. In a perfect competition framework without externalities such effects cannot exist: the welfare gain in monetary terms, generated by a marginal transport cost reduction, is just this marginal transport cost reduction, no less, no more. With economies of scale this is different: the marginal welfare gain tends to exceed the marginal cost reduction. The ratio of the former over the latter, called the "total benefit multiplier", is in the order of 1.4 in the authors' numerical experiments. One should be aware that this multiplier may not only blow up gains, but possible losses as well: regions losing due to other regions moving closer to one another can lose more with increasing than with constant returns, because they move up rather than down the average cost curve.

In a series of research projects for the European Commission Bröcker and co-authors (Bröcker et al., forthcoming) have applied a similar approach with a smaller number of industries (just one tradable and one non-tradable sector in most cases), but a very large number of regions, such that the spatial distribution of welfare effects generated e.g. by the commission's TEN-T infrastructure program can be monitored in much detail (see also Bröcker, 2001a, 2001b, 2002).

Kim and Hewings (2003) and Kim, Hewings and Hong (2004) follow a different line of argument for identifying regional impacts of transport infrastructure improvements. They let firms use transport infrastructure as a production input that is provided for free. The level of service of the transport infrastructure is measured as a Harris (1954) type potential indicator of accessibility. The authors find a positive network effect of infrastructure policy, meaning that the welfare gain of an entire network of new projects exceeds the sum of the effects, if all projects are evaluated separately.

3.3.3 Urban models

A more recent branch of CGE applications in transport looks at urban passenger transport, focusing on the transport - land use nexus. Anas and collaborators (Anas and Hyok-Joo, 2006; Anas and Kim, 1996; Anas and Liu, 2007; Anas and Xu, 1999) lead the field. These authors succeeded in modelling, in a general equilibrium framework, location decisions of households and firms, travel decisions for shopping and commuting, goods and services production decisions of firms and goods and services consumption decisions of households. Households' consumption and travel decisions are micro based: households maximize utility subject to a budget as well as a time constraint. Travel times are obtained from a stochastic user equilibrium (Sheffi, 1985) in a congested network. In equilibrium, markets for land, labour, goods and services clear, and travel times are expected minimal times given equilibrium flows through the network. Due to the congestion externality the equilibrium allocation is not Pareto-efficient. In a recent extension there are also housing, construction and demolition sectors in order to model the dynamics of the housing stock.

An important methodological innovation in this work is merging the continuous demand approach of traditional CGE models with the discrete choice concept. If one took all households as homogenous, a rather unrealistic equilibrium pattern with strictly separated land use zones would emerge, and bang-bang type responses of households' location decisions to shocks would be observed.

The utility U_{ij} of a household residing at i and working at j is assumed to have three additive components, $U_{ij} = \tilde{U}_{ij} + A_{ij} + u_{ij}$:

- The systematic component $\tilde{U}_{ij}$ is a function of continuously measured quantities of goods and service consumption, as usual. It must be defined in a way that makes it dimensionless. For homothetic preferences it is $\tilde{U}_{ij} = \log V_{ij}$, if V_{ij} is a linear-homogeneous representation of preferences.
- A_{ij} is the inherent attractiveness of the residence-work place pair *ij*. It delivers the degree of freedom needed to reproduce any observed distribution of the population across such pairs in a benchmark data set.
- u_{ij} is an idiosyncratic component varying across individuals of the *ij*-population, which otherwise are taken to be identical. u_{ij} is assumed to be independent identically Gumbel distributed. This implies that the share of the total population choosing the *ij*-pair is described by a logit model.

This framework is about to replace the so-called LUTI models in urban simulation (Waddell, 2000; Wegener, 2004), that follow a tradition initiated by Lowry (1964). Models of the latter kind do a good job in simulating land use implications of urban transport policies, but due to the lack of micro-foundation they are unable to quantify welfare effects. Furthermore, understating the price-mechanism in these models leads to ad-hoc mechanisms equilibrating markets that are not very convincing. In both respects Anas and co-authors did a big step forward, offering a framework for simulating a wide range of policies such as infrastructure provision, subsidizing certain modes, road pricing, cordon pricing, supply of parking lots and more. For any such policy one can not only simulate price and quantity impacts, but also welfare impacts by residential zone, type of household and income group. These are the issues that debates about urban transport policies typically focus on.

4. Conclusions

During the last twenty years, computable general equilibrium (CGE) models have become standard tools of quantitative policy assessment. Their appeal has built on their rigorous grounding in economic theory: individual agents' decision-making behaviour is derived from explicit optimization under strictly specified technological or budget constraints, given market signals that ensure global consistency. These theoretical foundations have made CGE models appear particularly useful for ex-ante evaluations of policy reforms. In this chapter, we have discussed how the standard CGE framework can be extended to include most –if not all–the elements that are the focus of transportation policy analysis.

Powerful as it is, the whole apparatus relies on the concept of "representative agent" despite unclear aggregation procedures to link these aggregate optimizing decision-makers to the numerous individual agents whose behaviour they are meant to capture. Yet, large and detailed micro data-sets on individual behaviour in their full heterogeneity are increasingly being made available, and for many issues, working with myriads of actual economic agents rather than with a few hypothetical ones is extremely appealing as it makes possible to precisely identify the winners and the losers of a reform – obviously a major concern to policy-makers. One can therefore conjecture that in the future, CGE modelers will devise explicit aggregation procedures in order to be able to keep track, in their general equilibrium models, of the full heterogeneity in individual behaviours provided by the micro data-sets. See Magnani and Mercenier (2009) for an effort in that direction.

References

Anas, A. and R. Hyok-Joo (2006), 'Curbing urban sprawl with congestion tolls and urban boundaries', *Regional Science and Urban Economics*, 36, 510–541.

Anas, A. and I. Kim, (1996), 'General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration', *Journal of Urban Economics*, 40, 232–256.

Anas, A. and Y. Liu (2007), 'A regional economy, land use, and transportation model (RELU-TRAN): Formulation, algorithm design and testing', *Journal of Regional Science*, 47(3), 415–455.

Anas, A. and R. Xu (1999), 'Congestion, land use and job dispersion: A general equilibrium model', *Journal of Urban Economics*, 45, 451–473.

Armington, P.S. (1969), 'Geographic pattern of trade and the effects of price changes', *IMF Staff Papers*, 16, 176–199.

Berg, C. (2007), 'Household transport demand in a CGE-framework', *Environmental & Resource Economics*, 37, 573–597.

Bröcker, J. (2001a), 'Spatial effects of transport infrastructure: The role of market structure', in J. Roy and W. Schulz, eds., *Theories of Regional Competition*, Baden-Baden: Nomos.

Bröcker, J. (2001b), 'Trans-European effects of 'Trans-European Networks'', in F. Bolle and M. Carlberg, eds., *Advances in Behavioral Economics: Essays in Honor of Horst Todt*, Heidelberg: Physica.

Bröcker, J. (2002), 'Spatial effects of European transport policy: A CGE approach' in G.J.D. Hewings, M. Sonis, D. Boyce, eds., *Trade, Networks and Hierachies – Modelling Regional and Interregional Economies*, Berlin, Heidelberg and New York: Springer.

Bröcker, J. (2004), 'Computable general equilibrium analysis in transportation economics', in D.A. Hensher, K.J. Button, K.E. Haynes and P.R. Stopher, eds., *Handbook of Transport Geography and Spatial Systems*, Amsterdam: Elsevier.

Bröcker, J., A. Korzhenevych and C. Schürmann (forthcoming), 'Assessing spatial equity and efficiency impacts of transport infrastructure projects', *Transportation Research B: Methodological*, doi:10.1016/j.trb.2009.12.008.

Buckley, P.H. (1992), 'A transportation-oriented interregional computable general equilibrium model of the United States', *The Annals of Regional Science*, 26, 331-348.

Conrad K. (1997), 'Traffic, transportation, infrastructure and externalities – A theoretical framework for CGE-analysis', *The Annals of Regional Science*, 31, 369–389.

Conrad, K. and S. Heng (2002), 'Financing road infrastructure by savings in congestion costs: A CGE analysis', *The Annals of Regional Science*, 36, 107–122.

Department for Transport (1999), *Transport and the Economy*, Report by Standing Advisory Committee on Trunk Road Assessment (SACTRA). http://www.dft.gov.uk/pgr/economics/sactra/

Dixit A. and J. Stiglitz (1977), 'Monopolistic Competition and Optimum Product Diversity' *American Economic Review*, vol. 67(3), 297-308.

Ethier W. (1982), 'National and International Returns to Scale in the Modern Theory of International Trade', *American Economic Review*, 73, 389-405

Fossati A. and W. Wiegard, eds. (2002), *Policy Evaluation with Computable General Equilibrium Models*, Routledge.

Gorman W. (1959), 'Separable Utility and Aggregation', *Econometrica*, 27, 469-81.

Ginsburgh V. and M. Keyzer (1997), *The Structure of Applied General Equilibrium Models*, Cambridge, MA: MIT Press.

Harris, C. (1954), 'The Market as a Factor in the Localization of Industry in the United States", *Annals of the Association of American Geographers*, 64, 315–348.

Jara-Diaz, S.R. (2000), 'Allocation and value of travel time savings' in D.A. Hensher and K.J. Button, eds., *Handbooks in Transport, Vol 1: Handbook of Transport Modelling*, Amsterdam: Pergamon-Elsevier.

Kehoe T. (1991), 'Computation and Multiplicity of Equilibria', in Hildenbrand W. and H. Sonnenschein, eds., *Handbook of Mathematical Economics Vol IV*, pp. 2049-2143. Amsterdam: North-Holland.

Kim, E. and G.J.D. Hewings (2003), 'An application of integrated transport network – multiregional CGE framework II: Calibration of network effects of highway', *REAL Discussion Paper 03-T-24*, Urbana, Illinois.

Kim, E., G.J.D. Hewings and C. Hong (2004), 'An application of an integrated transport network-multiregional CGE model: A framework for the economic analysis of highway projects', *Economic Systems Research*, 16(3), 235–258.

Lowry, I.S. (1964), *A Model of Metropolis*, Memorandum RM-4035-RC, Santa Monica: RAND Corporation.

Magnani R. and J. Mercenier (2009), 'On Linking Microsimulation and Computable GE by Exact Aggregation of Heterogeneous Discrete-choice Making Agents', *Economic Modelling*, 26(3), 560-570.

Mayeres, I. and S. Proost (2001), 'Marginal tax reform, externalities and income distribution', *Journal of Public Economics*, 79(2), 343–363.

Mayeres, I. and S. Proost (2004), *Testing Alternative Transport Pricing Strategies: A CGE Analysis for Belgium*, Paper presented at the conference on 'Input-Output and General Equilibrium: Data, Modeling and Policy Analysis', Brussels, 2-4 September 2004.

Mercenier, J. (1995a), 'Can '1992' Reduce Unemployment in Europe? On Welfare and Employment Effects of Europe's Move to a Single Market', *Journal of Policy Modeling*, 17(1), 1-37.

Mercenier J. (1995b), 'Nonuniqueness of Solutions in Applied General Equilibrium Models with Scale Economies and Imperfect Competition', *Economic Theory*, 6, 161-177.

Mercenier J. (2002) 'An Applied Intertemporal General Equilibrium Model of Trade and Production with Scale Economies, Product Differentiation and Imperfect Competition', in Fossati A. and W. Wiegard, eds., *Policy Evaluation with Computable General Equilibrium Models*, pp. 85-104, Routledge.

Mercenier, J. and T.N. Srinivasan, eds.(1994), *Applied General Equilibrium and Economic Development : Present Achievements and Future Trends*, The University of Michigan Press.

Munk, K.J. (2003), *Computable General Equilibrium Models and their Use for Transport Policy Analysis*, Danmarks TransportForskning (DTF), Rapport 4, Lyngby, Denmark.

Sheffi, Y. (1985), *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Englewood Cliffs, New Jersey: Prentice-Hall.

Shoven J.B. and Whalley J. (1984), 'Applied general equilibrium models of taxation and international trade: An introduction and survey', *Journal of Economic Literature*, 22, 1007–1051.

Srinivasan T.N. and J. Whalley, eds. (1986), *General Equilibrium Trade Policy Modelling*, The M.I.T. Press.

Venables, A.J. and M. Gasiorek (1998), *The Welfare Implications of Transport Improvements in the Presence of Market Failure*, Report to Standing Advisory Committee on Truck Road Assessment (SACTRA), Department of the Environment, Transport and the Regions, UK.

Waddell, P. (2000), 'UrbanSim: modeling urban development for land use, transportation and environmental planning', *Journal of the American Planning Association*, 68, 297–314.

Wegener, M. (2004), 'Overview of land use and transport models', in D.A. Hensher, K.J. Button, K.E. Haynes and P.R. Stopher, eds., *Handbook of Transport Geography and Spatial Systems*, Amsterdam: Elsevier.

Wrede, M. (2003), '*Tax deductibility of commuting expenses and residential land use with more than one center*', CESifo Working Papers 972.