Ann Reg Sci (1998) 32:367–387

The Annals of
Regional Science


# Operational spatial computable general equilibrium modeling

**Johannes Bröcker**

Technical University Dresden, Faculty of Traffic Sciences, D-01062 Dresden, Germany
(e-mail: broecker@rcs.urz.tu-dresden.de)

**Abstract.** A prototype spatial computable general equilibrium model is developed and illustrated by a numerical example. The theoretical basis is a complete Arrow-Debreu equilibrium under perfect competition. The leading principle of model design is parsimony: The specification restricts the number of parameters in a way allowing for a model calibration relying on a limited data base, which is readily available in a country with a well developed statistical service. No "data generating" first stage, using entropy maximisation or other methods not in line with the philosophy of microeconomic equilibrium analysis, is required.

## 1. Introduction

Multiregional input-output (MIO) analysis is one of the standard methods in the tool box of empirical spatial economics. One of its strengths is its ability to take fully account of interregional interindustry interdependencies. Whatsoever, among others, the following three objections have been raised against MIO models.

*First*, MIO models are not flexible due to the fixed-coefficients assumption, which is particularly inconvenient with respect to trade coefficients.
*Second*, MIO models do not sufficiently take account of income-expenditure interdependencies.
*Third*, MIO models are one-sided demand driven, such that effects coming from the supply side, like cost and capacity variations, can not be modelled appropriately.

In recent years, computable general equilibrium models are coming up [17], which are already available as multiregional multiindustry variants [9, 10, 12, 18]. They do not suffer from any of the three aforementioned drawbacks of MIO models, while preserving all its modelling capacities.

Therefore spatial computable general equilibrium (SCGE) models seem to be the natural candidates for a new generation of applied interregional interindustry models. Why are they not yet a popular tool in applied work? This is probably because most regional scientists would subscribe to Oosterhaven's statement [13, p. 147] that general equilibrium models "are not operational at all". They are supposed to be difficult to understand and to be awfully costly in terms of data requirements and computational effort.

It is the aim of this paper, by presenting a simple prototype SCGE model, to demonstrate that this needs not be true. Of course, there exists no upper bound to the cost of such a model, if we want to build all details of the real world into it. But if we are content on making plausible assumptions about things which we can not observe for acceptable costs, and if we are satisfied by calibration instead of econometric estimation, SCGE can be cheap and still highly satisfying from a methodological point of view.

## 2. A SCGE-model with NCES functional forms

### *2.1 Basic ideas*

We develop a model for a closed economy with $I$ sectors, $i=1,\ldots,I$ and $R$ regions, $r=1,\ldots,R$. There are three types of activities, production, which is done by a number of representative firms, transport, which is carried out by a number of "transport agents", and final demand, which is the activity of a number of representative households, who earn their income by selling to firms primary factors, certain given amounts of which are their property. For the sake of simplicity, there is no public sector in this economy, and final demand is not subdivided into components like consumption, investment etc.

In order to keep data requirements low, we apply the so-called pooling concept in interregional trade, which was introduced by Moses and Chenery. According to this concept, all commodities produced by sector $i$, say, in various regions and delivered to region $s$, say, for intermediate or final use, are first merged into a pool of commodity $i$ in region $s$, from where they are delivered to intermediate or final users (see [3]). Thus, no direct link exists between producers and customers. Conceptually, within each sector goods in the place of production ("outputs") have to be distinguished from goods in the pool of the region of consumption ("pool goods"), and output as well as pool goods are to be distinguished by region. Thus, in each sector, there are $2R$ distinct economic goods and the same number of distinct prices.

Each region shelters $I$ representative firms, one representative household, and $I$ transport agents. Firm $i$ in region $r$ produces the output of sector $i$ in region $r$ by a linear-homogeneous production function, taking pool goods of all kinds, $i=1,\ldots,I$, from the regional pool and primary factors of all kinds, $k=1,\ldots,K$, as inputs.

Transport agent $i$ in region $s$ is responsible for transforming outputs of sector $i$ in all regions, $r = 1, \ldots, R$, including $s$ itself, into pool goods of kind $i$ available in $s$. Transport agents also have a linear-homogeneous technology at their command for transporting commodities to the region of destination and merging them into a pool.

Finally, households earn their income by selling the factors, which they own, to the firms, and expend it completely for commodities in the pools of the region where they reside. The amount of factors owned by the households is exogeneous.

We assume perfect competition. Firms, transport agents and households are well informed about all prices and take them as given. Firms and transport agents maximise profits, which in view of linear-homogeneity implies that in equilibrium prices equal minimal unit cost and no profits are left. Households maximise utility under their budget constraint.

The only thing which is left to be done now is to specify functional forms of technologies and preferences. We could deliberately choose from the full menu of functional forms offered by the econometric literature. One has to choose with care, however, in order to limit the costs.

### *2.2 Formal structure*

We begin by listing all endogeneous variables. Numbers in parenthesis point to the equation of the equilibrium system, which is "responsible" for the determination of the respective variable. Superscripts $i, j = 1, \ldots, I$ refer to sectors, superscripts $k = 1, \ldots, K$ to primary factors, subscripts $r, s = 1, \ldots, R$ to regions.

*Quantities:*

| | | |
|---|---|---|
| $x_r^i$ | Output of sector $i$ in region $r$ | (10) |
| $d_s^j$ | Final demand of $j$-goods in region $s$ | (7) |

*Prices:*

| | | |
|---|---|---|
| $p_r^i$ | Price of one output unit of sector $i$ in region $r$ | (1) |
| $q_s^i$ | Price of one unit of the pool good of sector $i$ in region $s$ | (4) |
| $w_r^k$ | Price of one unit of factor $k$ in region $r$ | (9) |

*IO coefficients:*

| | | |
|---|---|---|
| $a_s^{ij}$ | Input of pool goods of sector $i$ per unit output in sector $j$ in region $s$ | (2) |
| $c_s^{kj}$ | Input of factor $k$ per unit output in sector $j$ in region $s$ | (3) |
| $t_{rs}^i$ | Delivery of output of sector $i$ in region $r$ per unit of pool-good in region $s$. | (5) |

*Income and utility:*

| | | |
|---|---|---|
| $y_r$ | Household's income in region $r$ | (8) |
| $u_r$ | Household's level of utility in region $r$ | (6) |

In contrast to MIO analysis, IO coefficients and final demand vectors are endogeneous, the former depending on prices, the latter on prices and incomes. We first describe the behaviour of firms in order to derive $a$'s and $c$'s from cost minimisation, then that of transport agents in order to derive $t$'s from cost minimisation, and finally that of households in order to derive $d$'s from utility maximisation.

Technologies and preferences will be specified by nested functions with constant elasticities of substitution (NCES-function). They are explained in the appendix (section A below). It is convenient to work with the cost function, which is dual to the production function. It assigns the minimal unit costs to the vector of input prices.

Any NCES-function is completely specified by its substitution structure and a vector of position parameters having the same dimension as the input vector. In our notation a NCES cost-function with $n$ inputs is written as a function of $2n$ arguments, $n$ input prices and $n$ position parameters. The respective substitution structure is regarded as given and not shown explicitly.

*2.2.1 Firms.* The firm $j$ in region $s$ produces the output $j$ with intermediate inputs $i=1,\dots,I$, taken from the pool in region $s$, and with primary inputs. There are $K$ of them, indexed by $k=1,\dots,K$. The firm's technology is completely specified by its NCES unit-cost function $\mathrm{cf}^j\ (q_s, w_s; \alpha^j, \gamma^j)$. $q_s := (q_s^1, \dots, q_s^I)$ and $w_s := (w_s^1, \dots, w_s^K)$ are the price vectors for pool goods and primary factors in region $s$, respectively. $\alpha^j := (\alpha^{1j}, \dots, \alpha^{Ij})$ and $\gamma^j := (\gamma^{1j}, \dots, \gamma^{Kj})$ are the corresponding vectors of position parameters. Note that neither the function nor the position parameter bear a regional index, which means that within each sector firms in all regions produce by the same technology, an assumption justified by nothing but the desire to keep data requirements low.

In equilibrium, the output price equals minimal unit costs (otherwise supply would be either infinity or zero), which gives the first equation:

$$p_s^j = \mathrm{cf}^j(q_s, w_s; \alpha^j, \gamma^j). \tag{1}$$

According to Shephard's lemma [20, p. 54], the cost minimising input coefficients are the derivatives of the cost function with respect to price, which gives:

$$a_s^{ij} = \frac{\partial \mathrm{cf}^j(q_s, w_s; \alpha^j, \gamma^j)}{\partial q_s^i}, \tag{2}$$

$$\mathrm{c}_s^{kj} = \frac{\partial \mathrm{cf}^j(q_s, w_s; \alpha^j, \gamma^j)}{\partial w_s^k}. \tag{3}$$

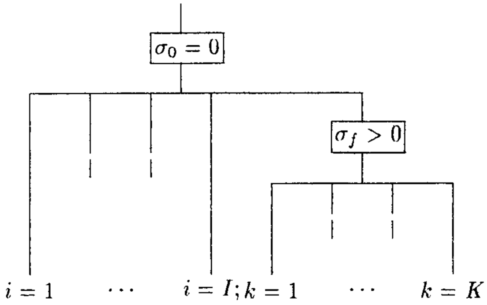


**Fig. 1.** Substitution tree

At this stage, the design of the tree structure of the NCES technology can be left open. It depends on intuition and the availability of information on the respective elasticities of substitution. A traditional design is shown in Fig. 1, allowing for substitution only between primary factors.

*2.2.2 Transport.* Next, we have to specify the transportation technology. Remember that there are $I$ transport agents per region. Transport agent $i$ in region $s$ uses up commodities produced in sector $i$ in all regions, $r=1,\ldots,R$, for generating the pool of commodity $i$ in region $s$. This transformation is also described by a NCES production function, transforming inputs from the various regions into an "output", namely a commodity of kind $i$ in the pool in region $s$. Using a NCES function in this context allows for taking account of product diversity. Purchases are not completely concentrated to the region with the lowest cif prices, because products from different regions are not regarded as identical. The degree of homogeneity within a sector is reflected by the elasticity of substitution (or the elasticities of substitution in case of nesting). This approach to handling interregional substitution is called Armington's assumption in international trade modelling (according to Armington [2]).

The specification of the transportation technology has to take into account that transportation uses up resources, the amount of which depends on economic distance. If we would take that literally, we had to assume transport agents to require transport services, produced by a special transport sector, as additional inputs, beyond commodities which have to be transported. Theoretically this is no problem, but it makes the model structure rather complicated in practice.

This is why we fall back on Samuelson's [16] iceberg model, which simply means that a certain percentage of the transported commodity itself is used up during transportation. This is plausible in case of oil transport, when the oil is partly used up as fuel, while a tanker crosses the Atlantic, whereas it seems rather strange in general. It could be argued, however, that pure costs of transportation anyway are only a small part of distance costs involved in interregional trade. The main part are costs for communication, services, storage, and others. They may be regarded as a joint product of the firm supplying the commodity. Thus, its output consists of the delivered commodities plus these services, which are paid by the customers and used up, while the commodity is delivered from the source to the destination.

Let the transport rate $\eta^i > 0$ be the share of commodity $i$ lost per unit of distance, and let $z_{rs}$ denote the distance from $r$ to $s$. Then the amount arriving in $s$, if one unit of output $i$ has been sent from $r$ to $s$, is $\exp(-\eta^i z_{rs})$, which is less than unity, if $z_{rs}$ is positive.

The transport agent's activity can now be thought of as being separated into two parts: One is transporting the outputs from all regions of origin to the region of destination, whereby they are partly used up; the other is merging the amounts left, after all commodities arrived in the region of destination, into the pool. The latter is regarded as a transformation process subject to a NCES technology characterised by the unit-cost function $\mathrm{ct}^i(v_s^i; \vartheta^i)$. $v_s^i := (v_{1s}^i, \ldots, v_{Rs}^i)$ is the vector of prices per unit of commodities arriving in $s$ from regions $r = 1, \ldots R$. As there are no costs of transportation other than the loss of a certain percentage of the respective commodity, we have $v_{rs}^i = p_r^i \exp(\eta^i z_{rs})$. Notice that the NCES functions, according to which commodities from different regions are pooled, vary over sectors, but not over regions of destination. Cost minimisation of transport agents then yields

$$q_s^i = \mathrm{c}t^i(v_s^i;\, \vartheta^i),$$

$$t_{rs}^i = \frac{\partial \mathrm{c}t^i(v_s^i, \vartheta^i)}{\partial v_{rs}^i}\; exp\;(\eta^i z_{rs}).$$

The latter equation takes account of the fact that the quantity sent off from $r$ exceeds that arriving in $s$ by the factor $\exp(\eta^i z_{rs})$. Note that $t_{rs}^i$ is the delivery from $r$ in terms of output units produced in $r$ per unit of pool good in $s$. With the definition

$$\mathrm{c}\vec{t}_s^{\,i}(p^i; \vartheta^i) := \mathrm{c}t^i(v_s^i; \vartheta^i)$$

the latter two equations are equivalently written as

$$q_s^i = \mathrm{c}\vec{t}_s^{\,i}(p^i;\, \vartheta^i), \tag{4}$$

$$t_{rs}^i = \frac{\partial \mathrm{c}\vec{t}_s^{\,i}(p^i; \vartheta^i)}{\partial p_r^i}. \tag{5}$$

*2.2.3 Households.* It is well known in duality theory of modern microeconomics that a household's preferences are completely specified by his expenditure function $\mathrm{eh}_s\,(q_s, u_s)$, assigning to the price vector $q_s$ the minimal expenditure required for attaining utility level $u_s$. It is non-decreasing, linear-homogeneous, concave in $q_s$ and increasing in the utility level $u_s$ [20, p. 122–123].

The equilibrium level of utility is obtained from the equality

$$y_s = eh_s(q_s, u_s), \tag{6}$$

where $y_s$ is the household's income. Demand is obtained as

$$d_s^j = \frac{\partial\, eh\,(q_s, u_s)}{\partial q_s^j} \tag{7}$$

by Hotelling's theorem, which is the same as Shephard's lemma, but stated in a different context.

Again we could choose from the full menu of functional forms in the econometric literature like CD, LES, CES, Translog, and AIDS [7, 11]. In order to keep in line with what we have assumed hitherto we choose a NCES form, though it is somewhat inconvenient in so far as it implies unitary income elasticities for all goods, contradicting empirical evidence. A different choice, however, would require additional information in the calibration process.

The NCES expenditure function has the form

$$eh(q_s, u_s) = u_s\, \mathrm{c}h\,(q_s; \delta)$$

with NCES cost function ch and vector of position parameters $\delta = (\delta^1, \ldots, \delta^I)$. All households are assumed to have identical preferences. thus, neither ch nor $\delta$ has a regional index.

Equations (1) to (7) determine output prices and pool prices for all commodities in all regions, the IO coefficients, final demand and utility in all regions, given the factor prices and incomes in all regions. Model parameters are the vectors of position parameters in the cost functions of firms, transport agents, and households, respectively, as well as distances and transport rates.

As already mentioned, households earn their income by selling exogeneously fixed amounts of factors to the firms. Let $f_{rs}^k$ denote the amount of factor $k$ in $r$, owned by the household in $s$. Selling the factor services the household earns an income

$$y_s = \sum_{r,k} \mathrm{f}_{rs}^k w_r^k\,. \tag{8}$$

Finally, factor markets have to clear:

$$\sum_s \mathrm{f}_{rs}^k = \sum_i \mathrm{c}_r^{ki} x_r^i\,. \tag{9}$$

Outputs $x_r^i$ are obtained from the standard linear system of IO equations

$$x_r^i = \sum_s t_{rs}^i \left( d_s^i + \sum_j a_s^{ij} x_s^j \right). \qquad (10)$$

Equations (1) to (10) determine the 10 groups of endogeneous variables listed above (see p. 3), given all NCES functions with their respective vectors of position parameters, and given the amounts of production factors.

## 3. Solving the SCGE-model

The system of Eqs. (1) to (10) is highly non-linear. General existence theorems, however, imply that a solution always exists if $\sum_s \mathrm{f}_{rs}^k > 0$ for all $k, r$. The solution is determined only in terms of relative prices, which means that in any solution prices and incomes may be multiplied by an arbitrary positive scalar to obtain another solution. Uniqueness can not be guaranteed, but the weaker property of regularity in terms of relative prices holds, which means that generically no further solution exists in a sufficient small neighbourhood of any solution [6]. In other words, generically solutions are isolated points in price space. The word "generically" means that incidentally isolation may not hold, but this can only happen on a set of measure zero in parameter space. An arbitrary small change of parameters would recover isolation in this case.

The problem of numerical solution can be reduced to the problem of finding the vector $w$ of relative factor prices clearing the factor market, which means solving the $KR$ equations

$$\sum_s \mathrm{f}_{rs}^k = \mathrm{f}d\,{}_r^k(w)$$

for the $KR$-vector $w := (w_1, \ldots, w_R)$, $\mathrm{fd}_r^k$ denoting the factor demand function for factor $k$ in region $r$. There are globally convergent fixed-point algorithms doing the job. They are usually based on the homotopy principle. The algorithms follow the homotopy path, the start point of which is known and the end point of which is the unknown solution. Path following is possible either by simplicial approximation [19] or by a continuation method like the prediction-correction method [1, 5]. In practice, however, a globally converging technique is rarely required, and a standard Newton method – if it converges – is much faster.

The values of the factor demand function for a given vector of factor prices are obtained as follows:

1. Insert $w_s$ into (1).
2. Solve the $2IR$ Eqs. (1) and (4) for the two $IR$-vectors of output prices $p$ and pool prices $q$.
3. Insert prices into (2), (3), and (5) and calculate coefficients.
4. Calculate incomes by (8) and then final demand vector $d$ by (6) and (7).

5. Solve the linear system (10) for the output vector $x$.
6. Insert into the RHS of (9) to obtain the factor demand.

Only step 2 needs further elaboration. After inserting Eqs. (4) for $q_s$ into Eq. (1), the Eqs. (1) are an interdependent system in $p$, given $w$ (as well as the vectors of position parameters). Write $p = g(p, w)$ for this system, for short. Due to linear-homogeneity one has

$$g(p, w) = I_p p + I_w w, \tag{11}$$

with $I_p$ and $I_w$ denoting $g$'s Jacobians with respect to $p$ and $w$. The $n$th step of a Newton procedure for the system is the solution of

$$p(n) = g(p(n-1), w) + I_p(p(n) - p(n-1))$$

for $p(n)$, which by Eq. (11) reduces to

$$p(n) = I_p^{n-1} p(n) + I_w^{n-1} w.$$

Superscript $n$–1 means that the respective Jacobians are calculated with prices from the last step, $p(n–1)$. Using the derivative properties (2), (3) and (5) this means explicitly

$$p_s^j(n) = \sum_{ir} t_{rs}^i a_s^{ij} p_r^i(n) + \sum_k \mathrm{c}_s^{kj} w_s^k,$$

with coefficients calculated by inserting prices from the previous step. In other words, one simply iterates Leontief's price system, adjusting coefficients in each iteration to changes in prices.

## 4. Calibrating the SCGE-model

Calibrating a CGE-model means to fix its parameters such that certain benchmark data are exactly reproduced in the equilibrium solution. We say that a model is "just identified" by a given set of benchmark data if, in the family of models emerging from different choices of free parameters, there is one and only one member reproducing these data. It's somewhat an art to design a CGE model such that it is just identified by a set of readily available benchmark data.

The structure of the described model is made up of the following components:

1. The substitution structure in the NCES functions cf, ct, and ch;
2. the interregional distances;
3. the transport rates;
4. the vectors of position parameters in these three types of NCES functions;

5. the amounts of the various factors in the regions owned by the households.

If these things are known, an equilibrium solution can be computed. Unfortunately, one benchmark does not suffice for calibrating the model with respect to all five aspects. The substitution structures have to be given. Usually they are obtained by gathering estimates of elasticities in the literature and designing substitution trees more or less following tradition and intuition.

Interregional distances are assumed to be given as well. For the moment[1] we also assume that the transport rates $\eta^i$ for each sector are known in form of data about transportation cost per kilometre as percentages of the respective commodity values. Then position parameters and factor amounts remain to be calibrated.

The position parameters are the vectors $\alpha^i$, $\gamma^i$, $i=1,\ldots,I$, making up an $((I+K)\times I)$-matrix, the vectors $\vartheta^i, i=1,\ldots,I$, making up an $(R\times I)$-matrix, and the $I$-vector $\delta$. National IO data can be used for calibrating the $\alpha$'s, $\gamma$'s, and $\delta$'s, and regional employment data for calibrating the $\vartheta$'s.

Concerning point 5, an $(R\times R\times K)$-array with elements $f^k_{rs}$ has to be calibrated. Unfortunately there is usually no information on interregional income flows. Thus, for the sake of simplicity, let us assume that households own factors only in their own region, such that $f^k_{rs}=0$ for $k=1,\ldots,K$, if $r\neq s$. Then only an $(R\times K)$-matrix of factor stocks is left to be calibrated.[2]

We assume that there are data on regional factor prices, which allow for a calibration with respect to the $f$'s.

Summarising this, calibration of the model means to find $\alpha$'s, $\gamma$'s, $\delta$'s, $\vartheta$'s, and $f$'s such that the equilibrium solution coincides with benchmark observations with respect to

- national IO-data in value terms,
- regional employment data by industry, and
- regional factor prices.

Let $k=1$ denote labour and $L^i_r$ benchmark observations on regional employment by sector. Furthermore, let $\check{w}^k_r$ be the benchmark observations on regional factor prices, and let us introduce the notation for national benchmark IO data given in Table 1. The $A$'s denote interindustry flows, the $D$'s final demand, and the $C$'s primary inputs, all in value terms.

For calibrating the model, the following equations are introduced in addition to Eqs. (1) to (10):

$$A^{ij} = \sum_s q^i_s a^{ij}_s x^j_s, \qquad i,j=1,\ldots,I \tag{12}$$

$$D^i = \sum_s q^i_s d^i_s, \qquad i=1,\ldots,I \tag{13}$$

[1] For more on this see footnote 4.
[2] In the following, we write $f^k_r$ instead of $f^k_{rr}$ for short.

**Table 1.** National IO-data

| | | | | | | |
|---|---|---|---|---|---|---|
| $A^{11}$ | $\cdots$ | $A^{1j}$ | $\cdots$ | $A^{1I}$ | | $D^{1}$ |
| $\vdots$ | | $\vdots$ | | $\vdots$ | | $\vdots$ |
| $A^{i1}$ | $\cdots$ | $A^{ij}$ | $\cdots$ | $A^{iI}$ | | $D^{i}$ |
| $\vdots$ | | $\vdots$ | | $\vdots$ | | $\vdots$ |
| $A^{I1}$ | $\cdots$ | $A^{Ij}$ | $\cdots$ | $A^{II}$ | | $D^{I}$ |
| $C^{11}$ | $\cdots$ | $C^{1j}$ | $\cdots$ | $C^{1I}$ | | |
| $\vdots$ | | $\vdots$ | | $\vdots$ | | |
| $C^{K1}$ | $\cdots$ | $C^{Kj}$ | $\cdots$ | $C^{KI}$ | | |

$$C^{kj} = \sum_s w_s^k \mathrm{c}_s^{kj} x_s^j, \quad k = 1, \ldots, K, \quad j = 1, \ldots, I, \tag{14}$$

$$\check{w}_r^k = w_r^k, \qquad k = 1, \ldots, K, \quad r = 1, \ldots, R, \tag{15}$$

$$L_r^i = \mathrm{c}_r^{1i} x_r^i, \qquad i = 1, \ldots, I, \quad r = 1, \ldots, R. \tag{16}$$

It is suggested that the model is just identified by these equations. Unfortunately, we are not yet able to give a rigorous proof of this suggestion. As a first hint to its validity we count equations and unknowns: There are $I^2$ Eqs. (12) and $I^2$ unknown $\alpha$'s, $I$ Eqs. (13) and $I$ unknown $\delta$'s, $KI$ Eqs. (14) and $KI$ unknown $\gamma$'s, $KR$ Eqs. (15) and $KR$ unknown $f$'s, and finally $IR$ Eqs. (16) and $IR$ unknown $\vartheta$'s. Some additional reflections reveal, however, that among these equations $2I$+1 equations are redundant, i.e. they are implied by all other equations. This leaves $2I$+1 degrees of freedom which can be closed by choosing units for the $I$ outputs, $I$ pool goods and the level of utility. This choice is arbitrary and will not affect any comparative static result obtained by the calibrated model. We close these degrees of freedom by imposing the additional restrictions

$$\sum_r x_r^i = \sum_r p_r^i x_r^i, \quad i = 1, \ldots, I, \tag{17}$$

$$\sum_r \vartheta_r^i = 1, \qquad i = 1, \ldots, I, \tag{18}$$

$$\sum_i \delta^i = 1. \tag{19}$$

Restriction (17) fixes units of measurement for outputs such that the average benchmark price in each sector equals unity. Restriction (18) fixes units of measurement for pool goods. The rationale underlying (18) is to let pool prices equal output prices, if output prices do not vary over regions and if transport costs vanish.

Admittedly the coincidence between the number of equations and unknowns is no convincing argument. Another hint to the validity of the suggestion is that a simple iterative procedure delivers a solution not depending on the starting point, according to practical experience. The procedure starts with a set of position parameters, which are adjusted in each iteration until convergence, taking the following steps (factor prices are set equal to $\check{w}_r^k$ from the beginning):

1. Solve (1) and (4) for prices.
2. Calculate IO coefficients according to (2), (3) and (5).
3. Calculate output according to (16)

$$x_r^i = L_r^i / \mathrm{c}_r^{1i}, \quad i = 1, \ldots, I, \;\; r = 1, \ldots, R.$$

4. Calculate $f_r^k$ according to

$$\mathrm{f}_r^k = \sum_i \mathrm{c}^{ki} x_r^i, k = 2, \ldots, K, \quad r = 1, \ldots, R$$

(see (9) with $f_{rs}^k = 0$ if $r \neq s$).

5. Calculate incomes according to (8),

$$y_r = \sum_k \mathrm{f}_r^k \check{w}_r^k, \quad r = 1, \ldots, R\,.$$

6. Calculate final demand according to (6) and (7).
7. Adjust $\alpha^{ij}$ until (12) is fulfilled, $i,j = 1,\ldots,I$.
8. Adjust $\delta^i$ until (13) is fulfilled, $i = 1,\ldots,I$.
9. Adjust $\gamma^{kj}$ until (14) is fulfilled, $k = 1,\ldots K$, $j = 1,\ldots I$.
10. Adjust $\vartheta_r^i$ until the RHS in (10) equals $x_r^i$, $i = 1,\ldots,I$, $r = 1,\ldots,R$.
11. If position parameters remained almost constant, as compared to the preceding iterate, stop; else go to 1.

## 5. A numerical example

As described above, four kinds of data are required for implementing the model empirically:

- National IO data in value terms,
- regional employment data by industry,
- regional factor prices, and
- interregional distances.

In addition, all substitution structures (including the respective elasticities of substitution) as well as transport rates $\eta^i$ – one for each sector – must be known.

In this section we show a little numerical experiment demonstrating the tractability of the approach. It is based on a purely hypothetical data set for

**Table 2.** National IO data

| From | to $j=$ 1 | 2 | 3 | 4 | $D$ | $\sum$ |
|---|---|---|---|---|---|---|
| $i=1$ | 4 | 3 | 2 | 5 | 18 | 32 |
| 2 | 16 | 6 | 5 | 9 | 4 | 40 |
| 3 | 2 | 2 | 6 | 6 | 10 | 26 |
| 4 | 0 | 10 | 2 | 8 | 14 | 34 |
| $k=1$ | 6.8 | 3.8 | 2.8 | 1.1 | | |
| 2 | 1.8 | 12.7 | 3.4 | 1.5 | | |
| 3 | 1.4 | 2.5 | 4.8 | 3.4 | | |
| $\sum$ | 32 | 40 | 26 | 34 | | |

**Table 3.** Employment by region and sector

| $r$ | $i$ 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.204 | 0.834 | 0.843 | 0.161 |
| 2 | 1.635 | 0.167 | 0.211 | 0.161 |
| 3 | 0.409 | 0.834 | 0.702 | 0.321 |
| 4 | 3.269 | 1.167 | 0.351 | 0.321 |
| 5 | 0.817 | 0.333 | 0.351 | 0.040 |

a closed economy, made up by 5 regions, 4 sectors, and 3 primary factors of production (land, labour, and capital, say).

Table 2 shows national IO data in value terms for one year. This table is like Table 1, but filled with numbers.

Employment by region and sector is given in Table 3 and factor prices by region and factor in Table 4.[3] The distance matrix is obtained by calculating Euclidean distances between points located as shown in Fig. 2. Transport rates $\eta$ by sector are given in Table 5. Note that a fairly strong distance impact is assumed in the example. In sector 2 (the same is true for sector 3) more than half of a delivery is lost during transport from region 2 to 3, for example.

[3] Note that for each factor only ratios between factor prices of different regions have to be known, because the choice of units for measuring factor quantities is arbitrary. The factor price in one arbitrary region can be set equal to one for each factor (or the average over regions can be set equal to one) for fixing units of measurement.

**Table 4.** Factor prices by region and factor

| $r$ | $k$ | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 1.19 | 0.62 | 1.14 |
| 2 | 0.81 | 0.61 | 0.52 |
| 3 | 1.07 | 1.43 | 1.14 |
| 4 | 1.11 | 0.97 | 1.34 |
| 5 | 1.43 | 1.03 | 0.57 |

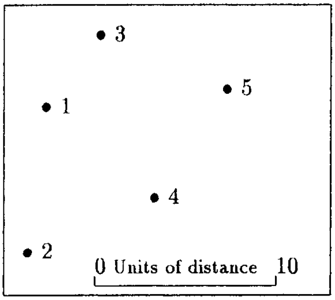


**Fig. 2.** Assumed location of regions

Finally, substitution structures have to be specified. The cost functions ct for transportation are assumed to be non-nested CES with elasticities of substitution by sector as given in Table 5 ($\sigma$-trans).[4] For production the 2-level NCES structure shown in Fig. 1 is chosen, with a Leontief form on the upper level (i.e. $\sigma_0=0$) and sector specific elasticities of substitution $\sigma_f$ between factors as given in Table 5 ($\sigma$-fact). The household's cost function ch is non-nested CES with elasticity of substitution equal to 0.8.

The model's calibration yields estimates of output prices (rows are sectors, columns regions)

$$p = \begin{pmatrix} 0.97 & 0.97 & 1.03 & 0.99 & 1.07 \\ 0.90 & 0.88 & 1.18 & 1.04 & 1.05 \\ 0.99 & 0.95 & 1.13 & 1.12 & 0.96 \\ 0.97 & 0.99 & 1.05 & 1.03 & 0.96 \end{pmatrix}$$

and prices for pool goods

[4] It can be shown that this specification implies trade flows fulfilling a gravity equation with distance function $\exp(-\lambda^i z_{rs})$. $\lambda^i$ is easily estimated from observed trade flows. Knowing $\lambda^i$ allows for inferring on either $\eta^i$ or $\sigma^i$, if $\sigma^i$ or $\eta^i$ are known from other sources.

**Table 5.** Parameters

| | *i* | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| $\eta$ | 0.028 | 0.047 | 0.062 | 0.016 |
| $\sigma$-trans | 4.3 | 0.5 | 2.5 | 1.8 |
| $\sigma$-fact | 0.8 | 1.5 | 3.0 | 2.5 |

$$q = \begin{pmatrix} 1.24 & 1.16 & 1.28 & 1.10 & 1.23 \\ 1.28 & 1.55 & 1.36 & 1.31 & 1.44 \\ 1.35 & 1.79 & 1.44 & 1.55 & 1.36 \\ 1.10 & 1.13 & 1.12 & 1.11 & 1.14 \end{pmatrix}.$$

Note that the price level in each sector is arbitrary, because the choice of units for quantities is arbitrary. Only interregional price ratios within each sector convey information. As already mentioned, scaling rule (17) makes weighted average output prices equal unity in each sector for the benchmark. Due to scaling rule (18) the pool prices' excess over unity reflects the transport costs included in these prices.

From calibrating the model we also get estimates of regional factor inputs for the two factors not given in the data (not shown here) as well as estimates of regional incomes obtained by multiplying regional factor inputs by regional factor prices:

$$(y_1, \ldots, y_5) = (11.37 \quad 4.32 \quad 7.43 \quad 14.03 \quad 8.84).$$

Finally, an interesting information is the value of the household's cost function $\mathrm{ch}(q_s;\delta)$. It is the natural regional price index in this model. The percentage deviation of this index for regions $2,\ldots,5$ from the respective index of region 1 is (6.20 3.83 –1.22 1.79)%. This index may also be interpreted as a complex measure of peripherality, taking the spatial distribution of all factors as well as preferences and technologies into account. In our example region 2 is the most peripheral (i.e. the one with the highest cost of living), region 5 is the most central one according to the index.

As an exercise in comparative statics we calculate effects from a reduction of distance costs – due to infrastructure investment, for example. Let us assume that the distance for transports from region 1 to region 5 is halved, while all other distances (including the one from 5 to 1) remain constant. As a consequence, all prices, outputs, trade flows, and incomes change. We will not present all the figures, but confine ourselves to the most interesting result, namely the welfare impact of the infrastructure investment. It is conveniently measured by the relative equivalent variation (REV). It is defined as the percentage increase of pre-investment income a household would need, in order to get the after-investment utility under pre-investment prices. More precisely, it is

$$REV_s := 100\left(\frac{eh(q_s,\hat{u}_s)}{eh(q_s,u_s)} - 1\right) = 100\left(\frac{\hat{y}_s/ch(\hat{q}_s;\delta)}{y_s/ch(q_s;\delta)} - 1\right).$$

$q_s$, $y_s$, and $u_s$ denote pre-investment, $\hat{q}_s, \hat{y}_s$, and $\hat{u}_s$ after-investment prices, incomes, and utilities, respectively. Due to the homotheticity of preferences, REV equals the percentage change of real income, with the cost function taken as price index. Our calculations give

$$REV = (6.57\ \ -1.10\ \ -1.09\ \ -0.15\ \ 4.67)\%.$$

As expected, regions 1 and 5 directly affected from the distance reduction are gaining. The other regions are loosing slightly because they suffer from higher factor prices in regions 1 and 5.

## 6. Conclusions

Though we have given only a small numerical example, it should be clear that the approach is tractable also for applications with more sectors and regions, different types of households, etc. Trade flows with the rest of the world are easily included. The fundamental advantage of the specification is that the model can be calibrated using readily available data only. No "data generating" process, which is the common practice starting point in MIO analysis, is required. Usually the theoretical basis of this kind of "data generating" is obscure and inconsistent with the spirit of CGE modelling. Instead, the philosophy of our approach is simple: Don't put more parameters into the model as you have independent observations (econometricians usually want us even to put much less parameters than observations into the model), and follow this principle from the very beginning in designing your model.

It is obvious that such a model leaves many things open. A straightforward extension is to introduce factor mobility with capital reacting on factor price and households reacting on utility. Another is to introduce dynamics through endogeneous investments. The formal structure, however, becomes much more complicated if rational expectations are assumed for investment and saving behaviour – as it is usually done in neoclassical model building.[5]

A final extension is introducing endogeneous technical chance. This would make the formal structure even more complicated, leading to applied models in the spirit of recent developments in the theory of endogeneous growth (see [8, 15]). Progress in this direction may be possible, if the model is kept very small in terms of numbers of regions and sectors included.

[5] See [4] for a dynamic 1-sector 2-regions model with migration and capital mobility.

## A The NCES (nested constant elasticity of substitution) system

Let $x{:}=(x_1,\ldots,x_I)>0$ be a vector of input quantities and $p{:}=(p_1,\ldots,p_I)>0$ the respective vector of input prices. Inputs are distinguished by industry or region of origin. (We drop the convention of pointing to industries by superscripts and to regions by subscripts.) With any concave linear-homogeneous production function $F^*(x)$ is associated a (unit) cost function

$$F(p) = \inf_x\{p \cdot x | F^*(x) \geq 1\}, \text{ [6]}$$

another concave linear-homogenous function showing the minimal costs for one unit of output, taking prices $p$ as given in the market. $F$ and $F^*$ form a polar pair of concave linear-homogenous functions [14, p. 136–139]. It is well known that $F$ is differentiable, if $F^*$ has strictly concave level sets, and that $F_p$, the gradient of $F$, is the vector of cost-minimising inputs per unit of output (Shephard's lemma).

NCES is a particularly convenient form of such a pair of functions. It is obtained by nesting a series of CES functions. It is completely specified by its substitution structure and an $I$-dimensional vector of position parameters. The substitution structure is described by a substitution tree and a vector of elasticities of substitution.

What is meant by a substitution tree should be sufficiently clear from the example in Fig. 3. It consists of a set of nodes, $i=1,\ldots,I,\ I+1,\ldots,J$, and a set of directed arcs. Nodes 1 to $I$, the ends ot the tree, represent the inputs, node $J$, the root, represents the output, and nodes $I+1$ to $J$–1 represent artificial intermediate goods. The output and intermediate goods are regarded as being produced according to a CES production-function, using the goods represented by their respective predecessor nodes as inputs. Thus, one CES function with its respective elasticity of substitution is associated with each node except the ends of the tree.

The NCES function is defined recursively as follows:

$$F^*(x) := \mathrm{f}^*_J(x)$$

with

$$\mathrm{f}^*_j(x) := \begin{cases} \left[\sum_{i \in \mathcal{N}_j} (\mathrm{f}^*_i(x)/\beta_{ij})^{\frac{\sigma_j - 1}{\sigma_j}}\right]^{\frac{\sigma_j}{\sigma_j - 1}} & \text{if } j>I, \\ x_j & \text{else}, \end{cases}$$

$\sigma_j>0$, $\sigma_j \neq 1$, is the *elasticity of substitution* associated with node $j>I$. As shown below, the limiting cases $\sigma_j \to 0$, $\sigma_j \to 1$, and $\sigma_j \to \infty$ can be included as well. $\mathcal{N}_j$ is the set of predecessor nodes of node $j>I$.[7] $\beta_{ij}$, $i \in \mathcal{N}_j$, are the *weight parameters* of the CES function.

[6] The operator "·" denotes the scalar product of two vectors, i.e. $p \cdot x{:}=p_1x_1+\ldots+p_Ix_I$.

[7] In Fig. 3 these sets are $\mathcal{N}_7=\{2, 3, 4\}$, $\mathcal{N}_8=\{5, 6\}$, $\mathcal{N}_9=\{1, 7\}$, $\mathcal{N}_{10}=\{8, 9\}$ for nodes 7–10. They are empty sets for the end nodes 1 to 6.

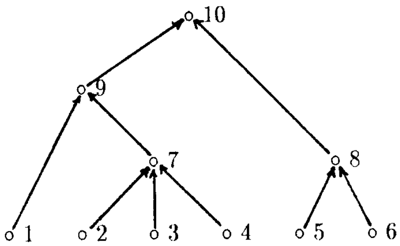


**Fig. 3.** Substitution tree

It's a matter of straightforward calculus not to be repeated here, to find the polar function of a CES function. Nesting these polar functions yields a recursive formula for $F$,

$$F(p) := \mathrm{f}_J(p)$$

with

$$\mathrm{f}_j(p) := \begin{cases} \left[\sum_{i\in\mathcal{N}_j} (\mathrm{f}_i(p)\beta_{ij})^{1-\sigma_j}\right]^{\frac{1}{1-\sigma_j}} & \text{if } j>I, \\ p_j & \text{else} \end{cases} \tag{20}$$

and another recursive formula for $F_p$

$$F_p(p) = (b_1, \ldots, b_I),$$

$$b_i := \begin{cases} \beta_{ij}^{1-\sigma_j}\left(\frac{\mathrm{f}_j(p)}{\mathrm{f}_i(p)}\right)^{\sigma_j} b_j & \text{if } i < J \\ 1 & \text{else} \end{cases} \tag{21}$$

In the last formula $j$ denotes the follower of $i$, i.e. $i\in\mathcal{N}_j$.

$F$ and $F_p$ are evaluated by first working through the substitution tree from the ends to the root for calculating $f_j, j>I$, and then backwards for calculating $b_i$.

For specifying a NCES function, the $\beta$-parameters are required beyond the substitution structure. Obviously, one such parameter is associated with each arc in the substitution tree, of which there are more than $I$ (except in the trivial non-nested case). As mentioned above, however, no more than $I$ parameters are required for a complete specification, given a certain substitution structure. This is so because different $\beta$-parameters yield identical CES-functions.

Consider the family of NCES functions obtained by choosing $\beta$'s, given a certain fixed substitution structure. Then it is easily seen that for any $\bar{p}>0$ and $a:=(a_1,\ldots,a_I)>0$ there is one and only one member $F$ in this family such that $F_p(\overline{p})=a$. Thus, given the substitution structure, the $I$-dimensional vector $a$, which $F_p$ shall attain at some price vector $\overline{p}$, is a valid parametrisation. The vector to be attained at $\overline{p}=(1,\ldots,1)$ could be taken as the parameter vector, for convenience. We call this vector the vector of *position parameters* of the NCES function.

In order to prove this statement, $F(p)$ is reformulated in terms of $a$ and $\overline{p}$, whereby the $\beta$'s are eliminated. Defining

$$g_i(p) := \frac{\mathrm{f}_i(p)}{\mathrm{f}_i(\bar{p})},\ i = 1,\ldots,J,$$

and dividing Eq. (20) by $f_j(\overline{p})$ yields

$$F(p) := g_J(p)\mathrm{f}_J(\bar{p})$$

with

$$g_j(p) := \begin{cases} \left[\sum_{i\in\mathcal{N}_j}\left(g_i(p)\frac{\mathrm{f}_i(\bar{p})}{\mathrm{f}_j(\bar{p})}\beta_{ij}\right)^{1-\sigma_j}\right]^{\frac{1}{1-\sigma_j}} & if\ j > I^{\text{c}} \\ p_j/\bar{p}_j & else \end{cases} \tag{22}$$

From Eq. (21) we obtain

$$\frac{\mathrm{f}_i(\bar{p})\bar{b}_i}{\mathrm{f}_j(\bar{p})\bar{b}_j} = \left(\bar{\beta}_{ij}\frac{\mathrm{f}_i(\bar{p})}{\mathrm{f}_j(\bar{p})}\right)^{1-\sigma_j},$$

with $\bar{b}$'s and $\bar{\beta}$'s such that (21) is fulfilled with $p=\overline{p}$ and $a=(\bar{b}_1,\ldots,\bar{b}_I)$.

The LHS-term is the share of good $i$ in the value of good $j$ under $\overline{p}$, $i\in\mathcal{N}_j$. More precisely we have

$$\phi_{ij} := \frac{\mathrm{f}_i(\bar{p})\bar{b}_i}{\mathrm{f}_j(\bar{p})\bar{b}_j} = \frac{\sum_{k\in\mathcal{M}_i} a_k\bar{p}_k}{\sum_{k\in\mathcal{M}_j} a_k\bar{p}_k}.$$

For $i>I$ the set $\mathcal{M}_i$ denotes all ends, which are node $i$'s direct or indirect predecessors, i.e.

$$\mathcal{M}_i := \{k \mid 1\le k\le I \text{ and the path from } k \text{ to the root meets } i\}.$$ [8]

Inserting $\phi$ into (22) and taking into account that $\mathrm{f}_J(\bar{p}) = a\cdot\bar{p}$, a new recursive formula for $F$ is obtained, which only depends on $a$ and $\overline{p}$ instead of $\beta$:

$$F(p) := g_J(p)a\cdot\bar{p}$$

[8] In Fig. 3 we have, for example, $\mathcal{M}_{10}=\{1–6\}$ and $\mathcal{M}_9=\{1–4\}$.

with

$$g_j(p) := \begin{cases} \left[\sum_{i\in\mathcal{N}_j} (g_i(p))^{1-\sigma_j}\phi_{ij}\right]^{\frac{1}{1-\sigma_j}} & if\ j>I, \\ p_j/\bar{p}_j & else. \end{cases}$$

As to the limiting cases for the elasticity of substitution mentioned before, the CES reduces to the Leontief case for $\sigma_j \to 0$ with fixed quantities of the components per unit of the composite good. In this case we have

$$g_i(p) := \begin{cases} \sum_{i\in\mathcal{N}_j} g_i(p)\phi_{ij} & if\ j>I, \\ p_j/\bar{p}_j & else. \end{cases}$$

The case $\sigma_j \to \infty$ means complete substitutability. The composite good is simply a weighted average of its components, and we have

$$g_j(p) := \begin{cases} min_{i\in\mathcal{N}_j}\{g_i(p)\} & if\ j>I, \\ p_j/\bar{p}_j & else. \end{cases}$$

For $\sigma_j \to 1$ we obtain the Cobb-Douglas case with constant value shares. Evaluate $\log g_j$ for $j>I$: You get the undefined expression 0/0 for $\sigma_j=1$. According to L'Hospital's rule, however, we find the limit for $\sigma_j \to 1$ by inserting first derivatives with respect to $\sigma_j$ at $\sigma_j=1$ in the nominator and denominator. This yields

$$g_j(p) := \begin{cases} \prod_{i\in\mathcal{N}_j} g_i(p)^{\phi_{ij}} & if\ j>I, \\ p_j/\bar{p}_j & else. \end{cases}$$

## References


[1] Allgower E, Georg K (1980) Simplicial and continuation methods for approximating fixed points and solutions to systems of equations. SIAM Review 22:28–85

[2] Armington PS (1969) A theory of demand for products distinguished by place of production. International Monetary Staff Papers 16:159–176

[3] Batten DF, Boyce DE (1986) Spatial interaction, transportation, and interregional commodity flow models. In: Handbook of Regional and Urban Economics, Vol. 1, pp 357–401. North Holland, Amsterdam

[4] Bröcker J (1994) Kapitalakkumulation und Migration im vereinigten Deutschland: Ein dynamisches Zwei-Regionen-Modell. In: Todt H (ed) Beiträge zur Standortforschung, pp 9–26. Duncker & Humblot, Berlin

[5] Chow SN, Mallet-Paret J, Yorke JA (1978) Finding zeros of maps: homotopy methods that are constructive with probability one. Mathematics of Computing 32:887–899

[6] Dierker E (1982) Regular economies. In: Arrow KJ, Intriligator M (eds) Handbook of Mathematical Economics, Vol. 2, chapter 15, pp 795–830. North-Holland, Amsterdam

[7] Diewert WE, Wales TJ (1987) Flexible functional forms and global curvature conditions. Econometrica 55:43–68

[8] Grossman GM, Helpman E (1991) Innovation and Growth in the Global Economy. MIT-Press, Cambridge, Massachusetts

[9] Hewings GJD, Gazel R (1998) Regional and interregional impacts of the US-Canada free trade agreement. Annals of Regional Science 32 (in press)
[10] Hirte G, Genosko J (1988/89) Regionalisierte empirische allgemeine Gleichgewichtsanalyse: Eine Einführung mit einem einfachen Modell für die Bundesrepublik Deutschland. Jahrbuch für Regionalwissenschaft 9/10:32–56
[11] Lau L (1984) Functional forms in econometric model building. In: Griliches Z, Intriligator MD (eds) Handbook of Econometrics, Vol III, pp 1515–1566. North-Holland, Amsterdam
[12] Meagher GA, Parmenter BR (1998) Regional issues in the monash model. Annals of Regional Science 32 (in press)
[13] Oosterhaven J (1981) Interregional Input-Output Analysis and Dutch Regional Policy. Gower, Aldershot
[14] Rockafellar RT (1970) Convex Analysis. Princeton University Press, Princeton, NJ
[15] Romer PM (1990) Endogenous technological change. Journal of Political Economy 98:S 71–S 102
[16] Samuelson PA (1954) The transfer problem and transport cost, ii: analysis of effects of trade impediments. Economic Journal 64:264–289
[17] Shoven JB, Whalley J (1984) Applied general-equilibrium models of taxation and international trade: an introduction and survey. Journal of Economic Literature 22:1007–1051
[18] Trela I, Whalley J (1986) Regional Aspects of Confederation. Volume 68 for the Royal Commission on the Economic Union and Development Prospects for Canada. University of Toronto Press, Toronto
[19] Van der Laan G, Talman AJJ (1986) Simplicial algorithms for finding stationary points, a unifying description. Journal of Optimization Theory and Applications 50:165–182
[20] Varian HR (1984) Microeconomic Analysis, 2nd edn. Norton, New York