REGIONAL INPUT-OUTPUT WITH ENDOGENOUS INTERNAL AND EXTERNAL NETWORK FLOWS

John R. Roy and Geoffrey J.D. Hewings

REAL 05-T-9 August, 2005

Regional Input-Output With Endogenous Internal and External Network Flows

by John R. Roy[1] and Geoffrey J.D. Hewings[2]

[1]ETUDES, PO Box 96, Mallacoota, Vic., 3892, Australia
[2]REAL, University of Illinois, 607 South Mathews Ave., Urbana, Illinois, 61801-3671, USA

Abstract: A major characteristic of dynamic regions or sets of dynamic regions is that external exports and imports are very significant in comparison with internally generated flows. Yet, it is precisely these flows of external exports and imports that are not evaluated in conventional regional input-output (I-O) analysis. The conventional approach merely determines internal flows of intermediate inputs. However, the total observable flows also contain internally generated flows to final demand, as well as external export flows, and external import flows to supply both intermediate inputs and final demand, implying five component flows. In most past models, the latter sets of flows have not been identified, with merely their totals at each region being assumed known. Although in recent work by the first author, internally generated regional flows to final demand were determined, certain of the regional external import and export flows were assumed to be available exogenously. This is quite an unrealistic expectation. In the following, all five sets of component flows are jointly determined, including transhipments of external exports and imports. In addition, rather than assuming just a single path between each set of regions, feasible multiple paths are assumed. A link/path transformation can then be made to obtain total link flows for each category, allowing a future consideration of congestion.

INTRODUCTION

Regional I-O analysis has a long history, including seminal works by Chenery (1953) and Leontief and Strout (1963), with the following analysis seen as a disaggregation of the Leontief-Strout (L-S) approach. Other significant contributors include Isard (1960), Polenske (1980), Hewings (1985), Miller and Blair (1985) and Oosterhaven (1988). An overview is provided in Roy (2004a). As more and more regional survey data became available, such analysis was approached with more confidence. In fact, regional I-O has become one of the most widely practised techniques in the field of regional science. Before proceeding further, we need to clarify terminology. For this, we turn to Isard et al (1998). The first class of regional model which they define is the *interregional* model, where both the flows and the I-O coefficients have four indices, that is, the flow of sector i into sector j from region r to region s. As it is extremely difficult to implement a full interregional model, most developments have concentrated on devising *multi-regional* models with less stringent data requirements. Although the dimensionality of these approaches reduces from four to three, different indices are absorbed in the flows compared to the I-O coefficients. The flows relate to the total flow of sector *i* as input to all other sectors between regions *r* and *s*, with the aggregation over destination sectors *j*. These flows are more likely to be available within freight statistics. The I-O coefficients relate to the amount of the sector *i* product being supplied as intermediate inputs to sector *j* in region *s* per unit of output of sector *j* in region *s*, aggregated over the different regions *r* which supply the inputs. It is precisely this different nature of the aggregation of the flows versus that over the I-O coefficients which creates the main challenge to development of sound multi-regional methods.

In conventional regional input output analysis, merely total flows are considered, without any attempt at disaggregation. In Roy (2004a,b), two advances were made (i) internally generated flows direct to final demand were determined and (ii) regional flows of external exports, as well as external imports to satisfy both internal intermediate and final demand were recognised in the flow relations, but not yet determined endogenously.[1] Although this was a useful advance on the classical analysis, it was not yet very practical, as multi-regional intermediate input flows arising from external imports needed to be available as data. This anomaly was recognised by the second author, who realised that a truly useful approach needed to yield endogenous multi-regional flows disaggregated into all of their five components (i) flows of internal intermediate inputs (as in the classical analysis), (ii) regional flows of internally generated final demand (as in Roy (2004a,b)), (iii) regional flows of external exports, (iv) regional flows of external imports to provide internal intermediate inputs and (v) regional flows of external imports to satisfy internal final demand. Part of the motivation for the need for this disaggregation may be found in an analysis of the Japanese interregional system over time. The results revealed that changes in interregional components were far more important that changes in technology in accounting for changes in output over time (see Hitomi *et al*., 2000). A further improvement is to identify each of the flows on possible multiple paths between each set of regions. Once such disaggregated flows are available, a path-link transformation, can evaluate

[1] We remain indebted to an anonymous reviewer for stimulating the recognition of internal final demand flows, as well as to Suahasil Nazara, formerly at Regional Economics Applications Laboratory (REAL), for suggesting the inclusion of the external import and export flows in the flow totals.

the flows on each link of the network, allowing a future consideration of congestion (not included in this paper; on this issue, see Sohn *et al*., 2004 and Kim and Hewings, 2004).

Whereas Leontief and Strout (1963) developed balance relations to incorporate technology into the flow determination, Chenery (1953) defined trade coefficients to reflect the influence of the transport network (that was never explicitly modelled). However, in principle, the influence on the pattern of flows of the transport network and of the technology should be jointly determined. In order to achieve this, as well as to induce the model to be consistent with key base period observations, Wilson (1970) introduced uncertainty into the analysis via entropy, with his model being constrained in estimation to reproduce base period values of both a transport cost constraint and the right-hand sides of the L-S balance relations. In other words, the flows were co-determined by the transport cost information together with the technological information embodied in the L-S representation of input-output. The entropy approach represents one useful procedure to estimate models to determine flows which simultaneously satisfy certain observed base period quantity and price/cost relations, whilst at the same time implicitly accounting for variability in the behaviour of the individual agents within the market segments being modelled. It is a generic statistical technique possessing very useful asymptotic properties in the presence of large ‘populations’, defined and formalized by Smith (1990) as Most-Probable-State Analysis. It must be interpreted differently depending on the particular field of application. Many of these interpretations have been made in the field of regional science by the first author, as illustrated in Roy (2004a). Another alternative is to replace the entropy framework by one from information theory based on historical trade patterns (Snickars and Weibull, 1977), yielding models such as in Batten (1983).

In any short run model of regional supply, it is desirable to include output capacity constraints. Of course, an obvious way to achieve this is to introduce ≤ inequality constraints on production with respect to regional capacity for each sector. However, such constraints, when inactive, maintain separability and have no influence whatsoever on the flows - they only factor into the analysis once they become active. Intuitively, this is not very plausible. In most sectors, a 'vintage' distribution over regional capacity exists, and when the capacity limits of a sector are being hard pressed in a certain region, one would expect some spillovers into adjacent regions. In fact, the generic logistic forms of regional supply functions motivated by Hotelling (1932) and discussed by Johansson (1991) demonstrate this property. In this paper, an additional entropy term recognising heterogeneity within the available capacity is shown to yield a logistic supply function. This generates a further enhancement to the Wilson framework, as already included in Roy (2004a,b). In addition, the special information theory method of Roy (1987) allows the model here not only to perfectly reproduce the base period flows (as for conventional information theory models), but to be simultaneously responsive to future changes in the input-output coefficients, regional output capacities, freight prices and the transport network itself, with the technological balance relations being imposed anew in the projection time period.

MULTI-REGIONAL I-O MODEL WITH ENDOGENOUS INTERNAL AND EXTERNAL COMPONENT FLOWS

In Roy (2004a,b), a deterministic interregional I-O model was developed, with the flows including both internal and external components. The surprising result from this analysis was that the conventional interregional I-O model (the 'Isard' model) cannot be made consistent with externally provided final demand, thus undermining its special equilibrium assumptions. However, when we turned to a multi-regional analysis, the fundamental balance relations turned out to be generally consistent with those in the L-S approach, reinforcing the foundations of the latter. In the following development, L-S ideas are used again to develop the balance relations, but the flows are now endogenously determined into their five internal and external components, rather than as merely the aggregation of these five components, as implied by L-S.

Basic Definitions and I-O Balance Relations

Firstly, we define total aggregated flows $\overline{x}_{im}^{rs}$ of sector i on route m (strictly route m_{ij}) between regions r and s in terms of their internal and external components as

$$\overline{x}_{im}^{rs} = \sum_j (x_{ijm}{}^{rs} + i_{ijm}{}^{rs}) + y_{im}{}^{rs} + e_{im}{}^{rs} + i_{im}{}^{rs} \quad (1)$$

These *pool* the intermediate internal flows $x_{ijm}{}^{rs}$ and external (import) flows $i_{ijm}{}^{rs}$ of product i on route m between regions r and s to all sectors j, plus the internal flows $y_{im}{}^{rs}$ to final demand, plus the external export flows $e_{im}{}^{rs}$ and import flows direct to final demand $i_{im}{}^{rs}$ for product i on route m between regions r and s, which represent *transhipments* when regions r and s are different.

Note that, in this form, the sectoral aggregation over the destination sectors of the flows of a good between each pair of regions is a more general form of the *supply pool* assumption of Leontief-Strout. Then, we represent the overall *output technology* by defining multi-regional coefficients a_{ij}^{s}, denoting the total number of units $x_{ij}^{s} = \sum_{rm} x_{ijm}^{rs}$ and $i_{ij}^{s} = \sum_{rm} i_{ijm}^{rs}$ of internal and external intermediate inputs of sector *i* going into sector *j* along routes *m* from all regions *r* to region *s* required to produce a unit of sector *j* output in the same region *s*, given as

$$a_{ij}^{s} = [\textstyle\sum_{rm} (x_{ijm}^{rs} + i_{ijm}^{rs})] / X_j^{s} \qquad (2)$$

In order for the intermediate flows to be consistent with multi-regional I-O, they are defined as $\tilde{x}_{im}^{rs}$ and $\tilde{i}_{im}^{rs}$, being aggregated over all destination sectors *j*, via

$$\tilde{x}_{im}^{rs} = \textstyle\sum_j x_{ijm}^{rs} \quad ; \quad \tilde{i}_{im}^{rs} = \sum_j i_{ijm}^{rs} \qquad (3)$$

If these are now substituted into (1), we have the total flows $\overline{x}_{im}^{rs}$ in terms of all the component flows which are to be endogenously determined

$$\overline{x}_{im}^{rs} = \tilde{x}_{im}^{rs} + \tilde{i}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs} + i_{im}^{rs} \qquad (4)$$

The next step is to express the *usage* relations of the output X_i^{r} of sector *i* in region *r* in the form

$$X_i^{r} = \textstyle\sum_{sm} (\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \qquad (5)$$

Here, the usage of the output goes to provide internal intermediate inputs, internal flows direct to final demand and flows to external exports.

Before proceeding further, we provide some basic definitions. The total exports E_i^{r} of sector *i* abroad out of region *r* and total imports I_i^{r} of sector *i* into region *r* from abroad are expressed in terms of the component flows as

$$E_i^r \;=\; \Sigma_{sm}\; e_{im}{}^{sr} \quad ; \quad I_i^r \;=\; \Sigma_{sm}\,(\tilde{i}_{im}^{rs} + i_{im}{}^{rs}) \qquad (6)$$

These denote transshipments[2] to ports r from producing regions s for the exports, and transhipments from the ports r to consuming regions s for the imports, which include components direct to final demand and those to supply intermediate inputs. Also, the final demand Y_i^r satisfied for sector i in region r, including that of its own region and that flowing in from all other regions s, is

$$Y_i^r \;=\; \Sigma_{sm}\,(\,y_{im}{}^{sr} + i_{im}{}^{sr}\,) \qquad (7)$$

These contain both the internal and externally imported components.

Because the demand-driven relations (5) relate to *usage* of the output in terms of *outflows*, we follow L-S to develop a viable way to eliminate the unknown outputs from the analysis by transposing the basic I-O relations (2), reversing the r and s indices and summing over j to yield the following *inflow* relations of goods i to all sectors j in region r via

$$\Sigma_{jsm}\,(x_{ijm}{}^{sr} + i_{ijm}{}^{sr}) \;=\; \Sigma_j\, a_{ij}{}^r\, X_j^r \qquad (8)$$

In terms of the multi-regional analysis quantities in (2), this yields

$$\Sigma_{sm}\,(\,\tilde{x}_{im}^{sr} + \tilde{i}_{im}^{sr}\,) \;=\; \Sigma_j\, a_{ij}{}^r\, X_j^r \qquad (9)$$

The next step is the elimination of the outputs X_j^r via (5), simply reversing i and j

$$\Sigma_{sm}\,(\,\tilde{x}_{im}^{sr} + \tilde{i}_{im}^{sr}\,) \;=\; \Sigma_j\, a_{ij}{}^r\,[\Sigma_{sm}\,(\,\tilde{x}_{jm}^{rs} + y_{jm}{}^{rs} + e_{jm}{}^{rs})] \qquad (10)$$

These represent the balance relations in terms of endogenous quantities needing to be satisfied to represent the technology. Although (10) is derived from the L-S philosophy, the implementation of this analysis in terms of the five component quantities rather than the aggregated flows, yields

[2] We could augment the conventional I-O structure to have such entities that do not enter the local production system: however, they do incur transport costs in moving through the region (see later for solution).

a form where all terms are endogenous. As such, the relations (10) here represent consistency conditions, rather than equations with exogenous right-hand sides, as in L-S.

Formulation of the Objective Functions

Armed with the above balance relations (10), the regional export and import definitions (6) and the final demand relations (7), all in terms of various of the five unknown sets of quantities, we can now proceed further with development of a probabilistic model. A considerable enhancement of the Wilson (1970) approach is made by ensuring that the flows reflect the current technology by disaggregating the analysis, inserting the enhanced technological balance relations (10) and setting logistic constraints on regional output capacity. Whereas the fundamental form of the conventional entropy model in the absence of economic information would yield equal sectoral production in each region, the enhanced model yields equal *relative capacity utilisation*, a more plausible hypothesis. Also, a strong distinction is made between the estimation form of the model and a transformed version for projection, as illustrated copiously in Roy (2004a).

The first major point is that we will need two separate but linked entropy models. These are necessary because the capacity usage quantities (5) contain no import terms. Thus, the first entropy term relates to division of the available capacity into that which is utilised and that which is non-utilised, with the former then being subdivided into the individual component flows. The simpler second entropy relates to expressing the total regional imports into their component flows. The key linkage constraints are (7) on final demand (also containing the import flows i_i^{sr} direct to final demand) and the balance constraints (10) (containing the import flows $\tilde{i}_{im}^{sr}$ of sector i which are supplied as intermediate inputs to all sectors). These constraints provide a rich linkage between the two models. Note that, an alternative procedure would be to maximize a weighted sum of the two entropies, with the weights being determined endogenously, as illustrated by Roy and Lesse (1985). However, we consider it more expedient to iterate between the two models to obtain consistent flows.

Entropy related to the productive capacity Consider that we have base period data, both on the total output capacity $\tilde{X}_{ir}^{0}$ of sector i to supply intermediate inputs, final demand and external exports from region r, as well as (optionally) the total in-flow $\overline{X}^{s0}$ of all sectors into regions s.

The incorporation of the output capacity in the entropy yields a logistic form of the flow function, consistent with Hotelling (1932), as suggested by Johansson (1991). Also, let $\bar{X}^0$ be the total value of all inputs in the system, $\bar{c}^0$ be the average transport cost (or distance) per value unit of commodity shipped and c_{im}^{rs0} be the internal transport cost[3] per value unit of sector *i* between regions *r* and *s* along path *m*. We must also include the extra transport costs c_i^{r0} to bring the external imports into their entry port *r* from outside our regional system and the extra transhipment costs c_i^{s0*} to transport the external exports from their exit region (port) *s* to their final destination outside our regional system. Note that, by normalizing with respect to prices, transport costs could be converted into *quantity* units. The number of microstates *Z* can now be given as the number of ways $\tilde{X}_{ir}^0$ distinguishable output capacity units may be divided into X_i^r which are utilised and $(\tilde{X}_i^{r0} - X_i^r)$ which remain unutilised, with the former then being allocated to the three categories of flows contained in the usage relations for X_i^r in (5)

$$Z = \pi_{ir} \{ \tilde{X}_{ir}^0 ! / [\{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \}! \, . \{ \pi_{sm} \; \tilde{x}_{im}^{rs} ! \; y_{im}^{rs}! \; e_{im}^{rs}! \}] \} \qquad (11)$$

As usual, π denotes the product sign. Setting the entropy *S* as the natural log of *Z* and applying the Stirling approximation, we obtain the entropy maximization objective

$$S = - \Sigma_{ir} \{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \} [\log \{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \} - 1] -$$

$$\Sigma_{irsm} \{ \tilde{x}_{im}^{rs} (\log \tilde{x}_{im}^{rs} - 1) + y_{im}^{rs}(\log y_{im}^{rs} - 1) + e_{im}^{rs}(\log e_{im}^{rs} - 1) \} \qquad (12)$$

In relation to the above objective, it is important to remain consistent with total regional destination in-flow information $\bar{X}^{s0}$ (if available), yielding the constraints via (4), with help of the export and final demand definitions in (15), in the form

$$\Sigma_{irm} (\tilde{x}_{im}^{rs} + \tilde{i}_{im}^{rs}) = \bar{X}^{s0} - \Sigma_i (E_i^{s0} + Y_i^{s0}) \qquad (13)$$

[3] In cases where transport costs are a relative small proportion of the total transaction costs, such as for a set of small densely-populated regions, or for cases where the other logistic costs of the transactions are quite high, transport costs and the associated constraint [see (14)] can be omitted.

The logistic capacity 'constraint' in the second entropy term in (11) implies that out-flow constraints on regional output should be omitted. However, we can enrich the model by applying the following output constraints, aggregating the regional outputs X_i^{r0} in (5) over both region and sector

$$\Sigma_{rsm} (\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) = X_i^0 \quad ; \quad \Sigma_{ism} (\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) = X^{r0} \tag{14}$$

Now, apply an average total internal and external transport/transhipment cost constraint over the entire set of flows as

$$\Sigma_{irsm} (\tilde{x}_{im}^{rs} + \tilde{i}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs} + i_{im}^{rs}) c_{im}^{rs0} + (\tilde{i}_{im}^{rs} + i_{im}^{rs}) c_i^{r0} + e_{im}^{rs} c_i^{s0*} = \bar{c}^0 \bar{X}^0 \tag{15}$$

Also, we apply the first of (6) as a constraint on base period exports E_i^{r0} of sector i out of region r and (7) for the base period final demand Y_i^{r0} for sector i in region r via

$$\Sigma_{sm} e_{im}^{sr} = E_i^{r0} \quad ; \quad \Sigma_{sm} (y_{im}^{sr} + i_{im}^{sr}) = Y_i^{r0} \tag{16}$$

The import terms are not included in the differentiation, as they are provided iteratively by the companion model which follows. Finally, the technological balance relations (10) are imposed, with a_{ij}^{r0} denoting the base period I-O coefficients. Applying Lagrangian theory, (12) is maximized in terms of $\tilde{x}_{im}^{rs}$, y_{im}^{rs} and e_{im}^{rs} under (13) with multipliers λ_{s1}, (14) with multipliers η_r and τ_i respectively, (15) with multiplier β_1, (16) with multipliers α_{ir} and γ_{ir1} respectively and (10) with multipliers ϕ_{ir1}. Differentiation with respect to $\tilde{x}_{im}^{rs}$, y_{im}^{rs} and e_{im}^{rs} to and equating to zero gives

$$\tilde{x}_{im}^{rs} = \{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \} \exp - \{ \lambda_{s1} + \eta_r + \tau_i + \beta_1 c_{im}^{rs0} + \phi_{ir1} (1 - \Sigma_j a_{ij}^{r0}) \}$$

$$y_{im}^{rs} = \{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \} \exp - \{ \beta_1 c_{im}^{rs0} + \eta_r + \tau_i + \gamma_{is1} - \phi_{ir1} (\Sigma_j a_{ij}^{r0}) \}$$

$$e_{im}^{rs} = \{ \tilde{X}_i^{r0} - \Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) \} \exp - \{ \beta_1 (c_{im}^{rs0} + c_i^{s0*}) + \eta_r + \tau_i + \alpha_{is}$$

$$- \phi_{ir1} (\textstyle\sum_j a_{ij}^{r0})\} \qquad (17)$$

Calling the exponential terms on the right of the three equations (17) A_{im}^{rs}, B_{im}^{rs} and D_{im}^{rs} respectively, summing these equations over *(sm)* and subtracting from $\tilde{X}_i^{r0}$, we finally obtain explicit relations for the unknowns as

$$\begin{aligned} \tilde{x}_{im}^{rs} &= \tilde{X}_i^{r0} A_{im}^{rs} / [1 + \textstyle\sum_{sm} (A_{im}^{rs} + B_{im}^{rs} + D_{im}^{rs})] \\ y_{im}^{rs} &= \tilde{X}_i^{r0} B_{im}^{rs} / [1 + \textstyle\sum_{sm} (A_{im}^{rs} + B_{im}^{rs} + D_{im}^{rs})] \qquad (18) \\ e_{im}^{rs} &= \tilde{X}_i^{r0} D_{im}^{rs} / [1 + \textstyle\sum_{sm} (A_{im}^{rs} + B_{im}^{rs} + D_{im}^{rs})] \end{aligned}$$

In many ways, it is tempting to estimate the non-linear equations obtained by substituting (18) into the constraints entirely by Newton-Raphson iteration. However, we may alternatively use Newton-Raphson for the multipliers β_1 and ϕ_{ir1} associated with the economic constraints and successive substitution for the pure summation constraints.

Entropy for External Imports As discussed earlier, the values of the intermediate and final demand external import flows need to be determined via a linked entropy model, with rich linkages existing within the constraints. Firstly, note that there is no natural capacity measure for regional external imports such as we had in the previous section for regional productive capacity. Thus, let us consider the number of ways Z' that the observed external imports I_i^{r0} may be allocated using the definition in (6) into the component flows $\tilde{i}_{im}^{rs}$ to supply intermediate inputs and those i_{im}^{rs} to supply final demand, yielding

$$Z' = \pi_{ir} \{ I_i^{r0}! / \pi_{sm} (\tilde{i}_{im}^{rs}! \,.\, i_{im}^{rs}!) \} \qquad (19)$$

Upon taking the natural log of both sides of (19), applying the Stirling approximation and removing constant terms, the external import entropy S' comes out as

$$S' = - \textstyle\sum_{imrs} [\tilde{i}_{im}^{rs} (\log \tilde{i}_{im}^{rs} - 1) + i_{im}^{rs} (\log i_{im}^{rs} - 1)] \qquad (20)$$

The entropy *S'* is to be maximised in terms of the import flows under various constraints. The first is the second of (6) expressed as regional import constraints

$$\textstyle\sum_{sm} (\tilde{i}_{im}^{rs} + i_{im}{}^{rs}) \;=\; I_i{}^{r0} \qquad (21)$$

These have multipliers ψ_{ir}. The inflow constraints (13) are re-applied with multipliers λ_{s2}, the travel cost/distance constraints (15) with multiplier β_2, the final demand constraints in (16) with multipliers γ_{ir2} and the inflow balance constraints (10) with multipliers ϕ_{ir2}. Upon differentiation of (20) and the nominated constraints with respect to the external import flows and equating to zero, we obtain

$$\tilde{i}_{im}^{rs} \;=\; exp - [\, \psi_{ir} + \lambda_{s2} + \beta_2 \,(\, c_{im}{}^{rs0} + c_i{}^{r0}) + \phi_{ir2} \,]$$

$$i_{im}{}^{rs} = exp - [\psi_{ir} + \beta_2 \,(\, c_{im}{}^{rs0} + c_i{}^{r0}) + \gamma_{ir2} \,] \qquad (22)$$

Upon elimination of ψ_{ir} via (21), the results are expressed as

$$\tilde{i}_{im}^{rs} \;=\; I_i{}^{r0}\; exp - [\, \lambda_{s2} + \beta_2 \,(c_{im}{}^{rs0} + c_i{}^{r0}) + \phi_{ir2} \,] \,/\, \textstyle\sum_{sm} \{\, exp - [\, \lambda_{s2} + \beta_2 (\, c_{im}{}^{rs0} + c_i{}^{r0})$$

$$+\; \phi_{ir2} \,] + \; exp - [\, \beta_2 (\, c_{im}{}^{rs0} + c_i{}^{r0}) + \gamma_{ir2} \,] \,\}$$

$$i_{im}{}^{rs} \;=\; I_i{}^{r0}\; exp - [\, \beta_2 \,(c_{im}{}^{rs0} + c_i{}^{r0}) + \gamma_{ir2} \,] \,/\, \textstyle\sum_{sm} \{\, exp - [\lambda_{s2} + \beta_2 \,(c_{im}{}^{rs0} + c_i{}^{r0}) + \phi_{ir2} \,]$$

$$+\; exp - [\, \beta_2 (\, c_{im}{}^{rs0} + c_i{}^{r0}) + \gamma_{ir2} \,] \,\} \qquad (23)$$

As before, (23) can be solved for the unknown Lagrange multipliers using Newton-Raphson iteration.

Coordination of the two models In our combined model, (18) and (23) are estimated successively, passing the current estimated values of $\tilde{x}_{im}^{rs}$, $y_{im}{}^{rs}$ and $e_{im}{}^{rs}$ as temporary known values into the constraints associated with (23), estimating new values of $\tilde{i}_{im}^{rs}$ and $i_{im}{}^{rs}$ from re-estimation of (23) and substituting these values in turn back into the constraints associated with

(18), and so on. As both objective functions are strictly concave, convergence would be obtained if the data are consistent. Once all the sets of the five component flows are obtained, substitution into (4) yields the total observable multi-regional flows $\bar{x}_{im}^{rs}$. As we satisfy the balance relations (10) which insert the usage relations (5) into the fundamental input output relations (9), the regional outputs X_i^r can be evaluated directly from (5) and should automatically satisfy (9) when the X_j^r values are substituted, reversing indices *i* and *j*.

With the flows on alternative paths *m* between regions *r* and *s* being available above, we now examine link flows. Let us define a *{0,1}* matrix $\delta_{ma}{}^{rs}$, with entries equal to 1 if link *a* occurs on path *m* between regions *r* and *s* and zero otherwise. From this we can obtain the total flows for each link *a* on the network. For example, for the flows $\tilde{x}_{im}^{rs}$ of intermediate internal inputs in (17), the associated link flows $\tilde{x}_{ai}$ for sector *i* on any link *a* are simply given as

$$\tilde{x}_{ai} = \sum_{rsm} \delta_{ma}{}^{rs} \; \tilde{x}_{im}^{rs} \qquad (24)$$

This is useful output, even if the model does not yet consider congestion.

Use of models for projection

The main task in transforming the estimated models in (18) and (23) into a form suitable for projection is to choose the Lagrange multipliers which should be treated as parameters in projection, and those which must be evaluated anew. The formalism of adapting the Lagrangian procedure to handle this change is illustrated copiously in Roy (2004a). The main issue is which information should reasonably be treated as exogenous input and which treated as endogenous output. In terms of the structure of the model, it is considered that the exogenous input should include any quantity changes, such as changed regional output capacities $\tilde{X}_i^r$, changed external exports E_i^r, changed external imports I_i^r and changed final demand Y_i^r. If transport costs change to $c_{im}{}^{rs}$, the availability as a parameter of the gravity Lagrange multipliers β_1 and β_2 allows the influence of these changes to be assessed by the model. Also, not expecting the total inputs $\bar{X}^s$ to region *s* to be available in the projection period, its Lagrange multipliers λ_{s1} and λ_{s2} should also be treated as parameters. The same applies to the multipliers η_r and τ_i on the aggregated output constraints (14). As with the classical analysis, any new multi-regional input-output coefficients

$a_{ij}{}^{r}$ must be provided exogenously. Visually, the projection equations will be of identical form to the base period relations (18) and (23). The main differences are (i) that the multipliers η_r, τ_i, β_1, β_2, λ_{s1} and λ_{s1} are now knowns rather than unknowns, with their associated right-hand sides now becoming outputs rather than inputs, and (ii) we now enter potentially new values of the exogenous data, including output capacities $\tilde{X}_i^r$, final demand Y_i^r, external exports E_i^r, external imports I_i^r, unit travel costs $c_{im}{}^{rs}$, $c_i{}^{r}$ and $c_i{}^{s}$ and regional I-O coefficients $a_{ij}{}^{r}$.

An adjustment from information theory

As stressed by Batten (1983), the more general information theory approach may give improved prediction ability for flow models of this type. If the goodness of fit of the estimated models is not satisfactory, the projection model is likely to yield improved results if information bias terms are computed and inserted into the models, as demonstrated via a new information theory procedure in Roy (1987).

One interesting feature of the current formulation is that we are unlikely to routinely have survey values of the full five sets of component flows – the most we can expect is to have survey values $\bar{x}_{im}^{rs0}$ of the *total* flows for each sector along each path between each pair of regions. Although this limits the potential power of the procedure in Roy (1987), it still allows the bias terms to be accommodated in a more aggregate sense. For example, if the result of the base period estimation of the total flows in (3) is given as $\bar{x}_{im}^{rs'}$, we define bias terms $q_{im}{}^{rs}$ = $\bar{x}_{im}^{rs0}$ / $\bar{x}_{im}^{rs'}$ which are normalised and applied as prior probabilities to the entropies of (12) and (20). For each component flow, such as for the intermediate internal inputs $\tilde{x}_{im}^{rs}$, this would yield the first of (18) in the revised form

$$\tilde{x}_{im}^{rs''} = q_{im}{}^{rs}\ \tilde{X}_i^{r0} A_{im}{}^{rs} / [\, 1 + \textstyle\sum_{sm} q_{im}{}^{rs} (A_{im}{}^{rs} + B_{im}{}^{rs} + D_{im}{}^{rs})] \qquad (25)$$

If we make the same corrections to the other four components in (18) and (23), then sum the results to the total values $\bar{x}_{im}^{rs''}$ in (3), we would ensure that $\bar{x}_{im}^{rs''}$ = $\bar{x}_{im}^{rs0}$, that is, the *total* estimated flows are identical to the *total* observed flows in the base period. Such bias factors are then be applied to the models when used in projection.

What if we do not have sector output capacity data?

Data on the net sector output capacities $\tilde{X}_{im}^{r0}$ for each sector in each region may be quite difficult to acquire. Also, the concept of capacity may be 'fuzzy' in situations of high demand, where some dormant or outmoded plants may well be called into emergency production. In such cases, base period constraints should be introduced on the value X_i^{r0} of the observed output of sector *i* from each origin region *r*, writing (5) in the form

$$\Sigma_{sm}(\tilde{x}_{im}^{rs} + y_{im}^{rs} + e_{im}^{rs}) = X_i^{r0} \qquad (26)$$

Clearly, these regional sectoral output constraints make the aggregated output constraints on both total sectoral production and total regional production in (14) redundant. In addition, the capacity entropy is omitted in the first term in (12), which now deals just with the allocation of utilized capacity, and (26) is attached with multipliers η_{ir}, yielding the results

$$\begin{aligned}
\tilde{x}_{im}^{rs} &= exp - \{\lambda_{s1} + \eta_{ir} + \beta_1\, c_{im}^{rs0} + \phi_{ir1}\,(1 - \Sigma_j\, a_{ij}^{r0})\} \\
y_{im}^{rs} &= exp - \{\eta_{ir} + \beta_1\, c_{im}^{rs0} + \gamma_{is1} - \phi_{ir1}\,(\Sigma_j\, a_{ij}^{r0})\} \\
e_{im}^{rs} &= exp - \{\eta_{ir} + \beta_1(c_{im}^{rs0} + c_i^{s0*}) + \alpha_{is} - \phi_{ir1}\,(\Sigma_j\, a_{ij}^{r0})\} \qquad (27)
\end{aligned}$$

This model is estimated rather similarly to the capacity-constrained model in (18) and can be solved via a combination of the Newton-Raphson method and iterative adjustments. With the total inflow $\bar{X}^{s0}$ to region *s* being not available as input in the projection period, its Lagrange multiplier λ_{s1} in the first of (27) must be treated as a parameter in projection. However, constraints (16) are applied anew in the projection period, with new export totals E_i^r and final demand quantities Y_i^r. Thus, their associated Lagrange multipliers are unknowns, and can be eliminated from the second and third relations in (27), yielding

$$y_{im}^{rs} = (Y_i^s - \Sigma_{rm}\, \tilde{i}_{im}^{rs})\; exp - \{\eta_{ir} + \beta_1\, c_{im}^{rs} - \phi_{ir1*}\,(\Sigma_j\, a_{ij}^{r})\} \,/\, [\Sigma_{rm}\, exp - \{\eta_{ir} + \beta_1\, c_{im}^{rs} - \phi_{ir1*}\,(\Sigma_j\, a_{ij}^{r})\}]$$

$$e_{im}^{rs} = E_i^s\, F_i^s\; exp - \{\eta_{ir} + \beta_1\, c_{im}^{rs} - \phi_{ir1*}\,(\Sigma_j\, a_{ij}^{r})\} \,/\, [F_i^s\{\Sigma_{rm}\, exp - \{\eta_{ir} + \beta_1\, c_{im}^{rs}$$

$$- \phi_{ir1*} (\Sigma_j a_{ij}^r)\}] \qquad (28)$$

in which $F_i^s = exp - c_i^{s*}$ and ϕ_{ir1*} are the new Lagrange multipliers on the L-S balance constraints (10) when estimated with revised I-O coefficients in the projection period and the intermediate import flows $\tilde{i}_{im}^{rs}$ are passed up iteratively from the linked import model in (23).

Although the necessity of (27) to satisfy the balance relations (10) plus the export and final demand constraints (16) in the projection period induces it to remain non-separable, the logistic form (18) has a more complex interdependency structure and should be used where reasonable capacity data can be found. It's important to realise that the multipliers η_{ir} are treated as known parameters in projection, allowing the final output X_i^r to emerge endogenously via direct substitution of the new flows into (5). This biases the flows to take the observed base period output into account in projecting final output via (5). Thus, this class of projection model shares some of the properties of the general information theory approach of Batten (1983), whilst allowing the transport costs to be explicitly represented and changed for the projection period. Note that, as (26) contains no unknown external import terms, our companion import model (23) remains unaffected in structure. Of course, as with the previous model with logistic capacity constraints, it still needs to be estimated iteratively with the new form in (27) and (28).

CONCLUSIONS

The approaches presented above integrate technology change, output capacity change, changes in regional final demand, endogenous determination of external export and external import flows, as well as transport network and cost changes, in the evaluation of multi-regional flows. Thus, they are well suited to analysis of dynamic regions, where external flows are important. As probabilistic models, they fit parameters to observations, rather than relying on deterministic optimization. The aggregation from an interregional approach adopts the same pooling assumptions as Leontief-Strout. The logistic form of the supply relationships promotes spillovers into adjacent regions when there is both a vintage distribution of capacity and a high pressure on this capacity in a given region. These models do not attempt to project changes in technology. It is the user's responsibility to provide any changed I-O coefficients a_{ij}^r.

The key advance with respect to Roy (2004a,b) is the identification and evaluation of the five sets of component flows, including regional flows to satisfy internal intermediate inputs, internal regional flows direct to final demand, as well as transhipment flows, including external export flows, external import flows to provide internal intermediate inputs and external import flows to go directly to internal final demand. As discussed by the second author in his critique of the earlier papers, for any route between regions, we can now evaluate the mix of the five component flows which are present. This can be further specialised to evaluation of the component flows on all individual links on the network. For example, we can see which links are especially vulnerable in a situation of increasing external imports or exports. In many ways, the above method may appear to resemble a commodity flow model. The main differences are that (i) technology is included by use of the L-S balance relations (10) and (ii) we follow the I-O method in which final demand is provided exogenously and we evaluate the values of all regional outputs plus the associated component flows.

Much empirical work remains to be done to demonstrate the relevance of the proposed model. In addition, a probabilistic RAS type approach should be devised to determine the full set of component flows, extending the analysis in Roy (2004a), Sec. 7.2, where changes in technology are inferred from changes in the regional outputs. A further challenge, if data is available, is to include the above ideas within a commodity by sector framework. This is especially important for the above models, where transport costs are included explicitly. Within a sector producing several commodities, the unit transport costs for each commodity can vary quite markedly, and are just averaged out in the sector by sector framework above.

References

Batten, D. F. (1983). *Spatial Analysis of Interacting Economies.* Kluwer, Boston.

Chenery, H. (1953). 'Regional analysis', in *The Structure and Growth of the Italian Economy.* eds. Chenery, H., Clark, P., Pinna, V., Rome, US Mutual Security Agency.

Hewings, G.J.D. (1985). *Regional Input-Output Analysis*, Sage Publications Inc., Beverly Hills, CA.

Hitomi, K., Okuyama, Y., Hewings, G.J.D.and Sonis, M. (2000). 'The role of interregional trade in generating change in the interregional Leontief inverse in Japan, 1980-1990', *Economic Systems Research*, 12, 515-537.

Hotelling, H. (1932). 'Edgeworth's taxation paradox and the nature of supply and demand functions', *Journal of Political Economy,* 40, 577-616.

Isard, W. et.al. (1960). 'Interregional and regional input-output techniques', Chapter 8 in *Methods of Regional Analysis,* John Wiley, New York.

Isard, W. et.al. (1998). *Methods of Interregional and Regional Analysis.* Ashgate.
Johansson, B. (1991). 'Regional industrial analysis and vintage dynamics', *Annals of Regional Science*, 23, 1-18
Kim, E., Hewings, G.J.D. and Hong, C. (2004). 'An application of an integrated transport network – multiregional CGE Model I: A framework for economic analysis of highway projects', *Economic Systems Research* 16, 235-258.
Leontief, W. and Strout, A. (1963). 'Multi-regional input-output analysis', in *Structural Interdependence and Economic Development.* ed. Barna, T., London, Macmillan.
Miller, R.E. and Blair, P.D. (1985). *Input-Output Analysis : Foundations and Extensions.* Prentice-Hall, Englewood Cliffs, NJ.
Ooosterhaven, J. (1988). 'On the plausibility of the supply-driven input-output model', *Journal of Regional Science*, 28, 203-217.
Polenske, K.R. (1980). *The U.S. Multiregional Input-Output Accounts and Model.* Lexington Books, Lexington, MA.
Roy, J.R. and Lesse, P.F. (1985). 'Entropy models with entropy constraints on aggregated events', *Environment and Planning A*, 17, 1669-1674.
Roy, J.R. (1987). 'An alternative information theory approach for modelling spatial interaction', *Environment and Planning A*, 19, 385-394.
Roy, J.R. (2004a). *Spatial Interaction Modelling: A Regional Science Context.* Advances in Spatial Science Series, Springer Verlag, Heidelberg.
Roy, J.R. (2004b). 'Regional input-output analysis and uncertainty', *Annals of Regional Science*, 38, 397-412
Smith, T.E. (1990). 'Most-probable-state analysis: A method for testing probabilistic theories of population behaviour, In: M. Chatterji, R.E. Kuenne (Eds.), New Frontiers in Regional Science, MacMillan, 75-94.
Snickars, F. and Weibull, J. (1977). 'A minimum information principle : theory and practice', *Regional Science and Urban Economics*, 7, 137-168.
Sohn, J. Hewings, G.J.D. Kim, T.J, Lee J.S and Jang, S-G. (2004). 'Analysis of economic impacts of an earthquake on transportation networks', In Y. Okuyama and S. Chang (Eds.), *Modeling Spatial and Economic Impacts of Disasters,*
Wilson, A. G. (1970). *Entropy in Urban and Regional Modelling.* Pion, London.