Issues in the implementation of interregional commodity by industry input-output models

By

Randall W. Jackson and **Walter R. Schwarm**

RESEARCH PAPER 2007-12

Randall W. Jackson
Regional Research Institute
3040 University Avenue, Ste 3010
West Virginia University
Morgantown, WV 26506

Walter R. Schwarm
California Dept of Finance
Demographic Research
915 L St, 8th Floor
Sacramento, CA 95814

North American Regional Science Association
Savannah, GA
November 8, 2007

Abstract: A number of procedures for generating interregional commodity by industry accounts have been developed recently (Canning and Zhi, 2005; Robinson and Liu, 2006; Jackson et al, 2006; Lindall, Olsen and Alward 2006). While each approach shares a common organizational framework, very little attention has been devoted to the use of these accounts in impacts assessment application. This paper presents the common organizational framework for the data and introduces relevant issues by briefly reviewing its lineage. The paper then addresses a number of issues surrounding many-region modeling applications and demonstrates the implications of adopting different assumptions and approaches.

Issues in the implementation of interregional commodity by industry input-output models

Introduction

Regional and interregional input-output models have long been central research themes within regional science and cognate disciplines. From inception, IO modeling at the regional level has been dominated by a focus on industry-based analysis. This has been the case especially in the United States, despite the 1972 shift from industry-based to commodity by industry-based data reporting at the national level. The understandable reluctance to shift emphasis on the part of US regional analysts is based in large part on the preponderance of sub-national industry data on employment, income, hours worked, etc., and the paucity of regional level commodity-based data. Nevertheless, analysts faced with the need to construct regional IO tables rarely if ever rely on primary data and resort instead to regionalizing national accounts via one method or another. Hence, working with the national industry and commodity data becomes a practical necessity.

One option in dealing with the national commodity by industry accounts is to first assume either commodity- or industry-based technology, construct a national industry by industry table from the Make and Use tables, then regionalize using industry-based regional data and either location quotients, supply-demand pooling, regional purchase coefficients (Stevens, Treyz et al. 1983; Kuehn JA 1985 May; Stevens, Treyz et al. 1988), GRIT (West 1990) or similar methods. There is ample treatment of these options in the literature. The alternative is to use region-specific data to generate regionalized versions of the national Make and Use tables, then construct the desired commodity by industry, industry by industry, or other single region account format. Jackson (Jackson 1998) presented a comprehensive method of this type for US researchers, to which Lahr (Lahr 2001) subsequently offered a series of qualifications and refinements.

Lacking from the literature, however, is an enumeration and elaboration of approaches to constructing interregional input-output accounts from the commodity by industry foundation framework once the necessary Make and Use data are in hand. To our knowledge, there is little in the literature to guide the analyst in the construction of such models, either in the basic format and layout or the extended implications of decisions and assumptions leading to the final framework of the interregional model constructed. While Canning and Wang (Canning and Wang 2005) presented a method for generating interregional input-output data, and Jackson et al. (Jackson, Schwarm et al. 2006) present the basis for estimation of commodity flows, (Lindall, Olson et al. 2006) discuss multi-region models in the IMPLAN framework, and Schwarm et al.

(Schwarm, Jackson et al. 2006) and Robinson and Liu (Robinson and Liu 2006) provide comparisons of the results of various techniques to published flow data and to one another, no works to date focus directly on conceptual implications of modeling decisions and assumptions in the combined context of the interregional input-output and the commodity by industry format of the U.S. national benchmark accounts (U.S. Department of Commerce 1991).

The purpose of this paper, therefore, is to initiate a discussion of the explicit treatment and use of national commodity by industry data in the construction and use of interregional input-output models. Rather than focus on methods for estimating the raw interregional interaction data, this paper will instead confront conceptual issues in model construction and application that arise in selecting from organizational and implementation alternatives. The goal is not only to guide applications practitioners, but also to move the discussion forward toward an eventual consensus, which can in turn guide those developing the base data.

History of Many-Region IO

Two methods of handling many-region models are well entrenched in the literature. The first is the interregional model, first presented by (Isard 1951). The structure of this model is such that there is a complete enumeration of all flows among all sectors. Formally, transaction $z_{ij}^{LM} \in Z^{LM} \in Z$ represents a flow from sector i in region L to sector j in region M. So for a two-region IRIO, we have

$$Z = \begin{bmatrix} Z^{LL} & Z^{LM} \\ Z^{ML} & Z^{MM} \end{bmatrix}$$

$$A = Z\hat{X}^{-1}$$

$$(I - A)X = Y$$

$$A = \begin{bmatrix} A^{LL} & A^{LM} \\ A^{ML} & A^{MM} \end{bmatrix}$$

$$X = \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix} \quad Y = \begin{bmatrix} Y^{L} \\ Y^{M} \end{bmatrix}$$

and

$$\begin{bmatrix} Z^{LL} & Z^{LM} \\ Z^{ML} & Z^{MM} \end{bmatrix} + \begin{bmatrix} Y^{L} \\ Y^{M} \end{bmatrix} = \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix}$$

$$\begin{bmatrix} A^{LL} & A^{LM} \\ A^{ML} & A^{MM} \end{bmatrix} \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix} + \begin{bmatrix} Y^{L} \\ Y^{M} \end{bmatrix} = \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix}$$

$$\begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} - \begin{bmatrix} A^{LL} & A^{LM} \\ A^{ML} & A^{MM} \end{bmatrix} \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix} = \begin{bmatrix} Y^{L} \\ Y^{M} \end{bmatrix}$$

where X denotes output and Y final demand. So $\begin{bmatrix} I - A^{LL} & -A^{LM} \\ -A^{ML} & I - A^{MM} \end{bmatrix} \begin{bmatrix} X^{L} \\ X^{M} \end{bmatrix} = \begin{bmatrix} Y^{L} \\ Y^{M} \end{bmatrix}$ or $(I - A)X = Y$. Hence $(I - A)^{-1} Y = X$, and the standard impacts assessment solution is $(I - A)^{-1} \Delta Y = \Delta X$.

In the IRIO, the coefficients in the various A matrix quadrants are regional trade coefficients, not regional technical coefficients.

The second general class of many-region models is the multiregional input-output model, or MRIO. Often called the Chenery-Moses model, this formulation is attributed to (Chenery 1953) and (Moses 1955), who developed essentially the same structure independently. Polenske (Polenske 1980) and her colleagues later took on the ambitious task of implementing the MRIO for the 50 US states and the District of Columbia. The MRIO approach begins with a set of regional *technical* coefficients tables as the basic building blocks, as opposed to the regional input coefficients tables of the IRIO. To take advantage of the kinds of data likely to be available, a set of trade tables is developed. Trade flows in the multi-regional framework are estimated first by sector. For a particular sector, i, data are gathered on the flows of i from one region to all others, forming an interregional shipments table for each good, of the following form:

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & \cdots & z_{1M} \\ z_{21} & z_{22} & \cdots & \cdots & z_{2M} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ z_{M1} & z_{M2} & \cdots & \cdots & z_{MM} \end{bmatrix}$$

Total shipments of good i into region K are represented by a column sum of this table, or $T_i^K = z^{1K} + z^{2K} + ... + z^{MK}$. When each column in Z is divided by its column total, we obtain the proportion of all good i used in K that comes from each region, L, denoted $c_i^{LK} = z_i^{LK} / T_i^K \in C$. Next let $C_{LK} = \begin{bmatrix} c_1^{LK} \\ c_2^{LK} \\ \vdots \\ \vdots \\ c_n^{LK} \end{bmatrix}$, which shows the proportion of the total amount of each of n goods used in K that comes from region L. This vector shows the proportion of the total amount of each of n goods used in M that comes from region L. There will be one of these vectors for each region-region pair, including K-K and L-L, etc. The counterpart to the IRIO A^{LM} in the MRIO framework is $\hat{C}^{LM} A^M$. The counterpart to the IRIO A^{MM} in the MRIO framework is $\hat{C}^{MM} A^M$. For a two-region model, then, we will have the following matrices:

$$A = \left[\begin{array}{c|c} A^L & 0 \\ \hline 0 & A^M \end{array}\right] \qquad C = \left[\begin{array}{c|c} C^{LL} & C^{LM} \\ \hline C^{ML} & C^{MM} \end{array}\right] \quad X = \left[\begin{array}{c} X^L \\ \hline X^M \end{array}\right] \quad Y = \left[\begin{array}{c} Y^L \\ \hline Y^M \end{array}\right]$$

And the equation system can be represented as $(I - CA)X = CY$. Given a change in final demand,

Note that the final demand vectors, Y, are not identical between the IRIO and MRIO specifications. For the IRIO approach, the partitions separate final demand for region L goods from final demand for region M goods. For the MRIO approach, Y^L and Y^M refer to total region L and total region M final demand. In essence, CY in MRIO approximates Y in IRIO.

Extensions of single-region IO assumptions

In the transition from closed nation to single-region to many-region IO, some of the assumptions necessary to obtaining a solution vector are extended and indeed take on new meaning. First, the assumption of fixed coefficients for a closed region implies linearity in production such that a doubling of outputs will require an exact doubling of each input. The coefficients reflect the technical relationships among inputs and outputs. When a nation is opened to trade, each technical coefficient is effectively split into two additive components: a regional coefficient and an import coefficient, or $a_{ij} = r_{ij} + m_{ij}$. Now, not only is the technical relationship fixed, but also the regional input coefficients are assumed to be fixed, the ratios of domestic to import supply for each coefficient also become fixed in the standard impacts assessment solution. This is a much stronger assumption, and one that has received attention in the literature (Beyers 1983). The final transition to the many-region context not only implies that total import coefficients are fixed, but so also the distribution of origins for imports. I.e., a doubling of output in an industry will require an exact doubling of purchases of all intermediate goods from all origin regions (and industries in IRIO) from which the purchases are made.

Commodity by industry single-region modeling issues

While the following section will be review to many, it is included to establish a basis for the ensuing discussion of commodity by industry data in interregional format. We first present the single-region framework, following closely the presentation in (Miller and Blair 1985), with minor notational differences. Diagram 1 presents a schematic of the basic layout of the commodity by industry framework.

Diagram 1. Single-region commodity by industry framework

Matrices U , V , W , and E are Use, Make, Value Added and Final Demand, respectively. The Use matrix depicts column industry use (purchases) of row commodity; the Make matrix depicts the column commodity output of each row industry; value added includes the payments sectors such as households, government (taxes and fees), and proprietors' income; Final Demand depicts row commodity final demand by column final demand activity, such as consumption, investment, government expenditures, and exports. For purposes of notational simplicity, we will assume in the discussion that follows that a) final demand columns have been aggregated to a single column and likewise that the rows of W have been aggregated to a single row, and b) the number of commodities is equal to the number of industries.

Given these definitions, we can enumerate a series of identities and establish a set of relationships that enable a set of solution counterparts to the interindustry impacts assessment framework. First, the identities are

$$1)\ \ U\mathbf{i} + E \equiv q$$
$$2)\ \ V\mathbf{i} \equiv g$$
$$3)\ \ V'\mathbf{i} \equiv q$$

Where *q* and *g* are total commodity and total industry output vectors. Now let

$$4)\ \ B = U\ddot{\varphi}^{-1}$$
$$5)\ \ U = B\ddot{\varphi}$$
$$6)\ \ q = Bg + E$$
$$7)\ \ D = V\ddot{\varphi}^{-1} \rightarrow d_{ij} = v_{ij} / q_j$$

Equation 7 is referred to as the industry-based technology assumption, and indicates that commodities are produced by industries in fixed proportion, such that as commodity production increases, each industry’s contribution to output of that commodity increases. Continuing,

$$8)\ \ V = D\ddot{\varphi}$$
$$9)\ \ D\ddot{\varphi}\mathbf{i} = g \text{ (from 2)}$$
$$10)\ \ g = Dq$$
$$11)\ \ q = BDq + E \quad \text{(from 6)}$$
$$12)\ \ (I - BD)^{-1} E = q$$

Thus, BD forms the commodity by commodity requirements coefficients matrix counterpart to the industry by industry coefficients matrix. From equation 7, where $g = Dq$, we see that the commodity-standardized Make matrix provides a mechanism by which to move between industry and commodity space. Hence, commodity by industry total requirements using the industry-based technology assumption can be derived as

$$13)\quad Y = DE$$

$$14)\quad E = D^{-1}Y$$

$$15)\quad (I - BD)^{-1} D^{-1}Y = q$$

That is, the commodity by industry total requirements matrix is $(I - BD)^{-1} D^{-1}$. Industry by commodity total requirements using the industry-based technology assumption is obtained by

$$16)\quad (I - BD)^{-1} E = D^{-1}g$$

$$17)\quad D(I - BD)^{-1} E = g$$

where $D(I - BD)^{-1}$ is the expression for industry by commodity total requirements.

$$18)\quad E = (I - BD)D^{-1} \quad \text{(from 16)}$$

Finally,

$$19)\quad E = (D^{-1} - B)g$$

$$20)\quad DE = (I - DB)g$$

$$21)\quad (I - DB)^{-1} Y = g$$

yields DB and $(I - DB)^{-1}$ as the expressions for industry by industry direct and total requirements respectively, using the industry-based technology assumption. Of course, equations 12, 15, 17, and 21 can be expressed in an impacts assessment format.

It is with the assertion of behavioral assumptions that accounting frameworks are transformed into models of economic behavior. The first such assumption introduced above in equation 4 establishes that there is a fixed production function relationship; the ratios of commodities used per industry dollar output are constant. This is the counterpart to the fixed coefficient assumption in the single-region interindustry framework. The second assumption introduced in equation 7 defines the relationships among industry and commodity production. An alternative assumption, the commodity-based technology assumption, states that industries produce commodities in fixed proportion, or more formally

$$22)\quad C = V\hat{g}^{-1} \quad \text{or} \quad c_{ij} = v_{ij} / g_i$$

which indicates that as an industry increases its output, it produces the same commodity proportions. Interested readers can find the parallel development of the four total requirements matrices using the commodity-based technology assumption elsewhere (Miller and Blair 1985). For the purposes of this discussion, however, we focus more directly on the interpretive assumptions of the two technology assumptions. Likewise, there has been a good deal of debate in the literature concerning the appropriateness of one versus the other assumption in which we will not engage at this point, although what follows may eventually add to the basis for that discussion (de Mesnard 2004).

Commodity by industry interregional issues

To transition to the many-region model we first revisit the basic data layout providing a simple 2-region, 2-industry, 2-commodity numerical example. Begin, for simplicity, with a closed national economy with the relationships shown in Table 1. The assumption of a closed economy is only for expositional simplification. Extension to an open national economy and additional regions would be straightforward.

[redacted]

Table 1. Closed national economy

Splitting this system into two regions yields the representation in table 2 with a set of numerically plausible values.

[redacted]

Table 2. 2-region economy, with interregional Use relationships

Table 2 shows the regional sources for commodities used to satisfy industry and final demands in both regions. Likewise, a depiction of the regional and industry supply of commodities available to a region can be constructed, as shown in Table 3. Parallel to the behavioral considerations in the single-region commodity by industry framework, we are obliged now to consider whether the existing interregional disposition of output should determine future trade relationships or whether the existing interregional, industry purchasing patterns will be perpetuated in future system output production.

[redacted]

Table 3. 2-region Make relationships

To formalize these relationships using the industry-based technology assumption, define a matrix RU and a matrix RV as the commodity by industry partition of Table 2 and the industry by commodity partition of Table 3, let RB and RD be appropriately standardized versions of RU and RV, and define the corresponding “consolidated” Use and consolidated Make, U and V as shown below in Table 4, with B and D the corresponding standardized, consolidated U and V.

Table 4. Consolidated Use and Make Matrices

The consolidated and regionalized tables each have different functions and interpretations. The regionalized Use, RU, depicts a fully enumerated interregional Use table as described earlier, and in standardized form it depicts the region-specific commodity input – industry output regional direct requirements coefficients (and since this is a closed economy, also technical coefficients).   The regionalized Make, RV, depicts a fully enumerated interregional Make table as described earlier, describing the region- and industry-specific source of commodities supplied to each region. Note that the supply of commodities to regions includes not only supply to industries but also to regional and export final demand. In standardized form, RD depicts the region and industry-specific distribution of commodity outputs.

The consolidated U, in contrast, depicts the regional industry use of commodities irrespective of region of origin of production. In standardized form, the block diagonals contain commodity input – industry output *technical* coefficients for each region's industries, since we have again assumed a closed national economy.   Likewise, the consolidated V depicts the total commodity supply for intermediate and total consumption, irrespective of the origin of production. Its standardized version therefore depicts the industry-specific distribution of commodity output, irrespective of region of output destination.

In developing the various versions of the total requirements matrices using these base data, a choice of which combinations of these matrices is most appropriate must be made.   At the outset, we restrict our focus to the commodity by commodity form of the solution using the industry-based technology assumption. Using B and D in combination would result in a matrix devoid of region-specific origin and destination detail, so is obviously excluded from consideration.   Conversely, using RB and RD would generate a nonsensical interregional commodity by commodity table whose values would have effectively been twice regionalized, resulting in overestimates of interregional and underestimates of intraregional values. Using B and RD will generate an interregional commodity by commodity table consistent with the region- and industry-specific commodity output distribution derived from the accounts. Using RB and D will generate an interregional commodity by commodity table consistent with the region- and industry-specific commodity use patterns derived from the accounts.

Although they will not be the focus of much discussion, the Leontief inverse tables from each of the four formulations using the above numerical example are presented in Table 5 to provide verification of the above assertions and a sense of the extent to which the alternatives can influence the results. As expected, the interregional partitions of BD-based inverse are zeros, and the RBRD-based values are correspondingly larger in interregional and smaller in intraregional partitions than the remaining two examples. The inverse based on the consolidated B matrix has consistently larger intraregional and consistently smaller interregional values than the consolidated D counterpart. This result, while not unexpected, is clearly dependent upon structure of production. The column multipliers from these same tables shown in Table 6 are strikingly similar.

[redacted]

Table 5. Leontief inverse tables based on the alternative direct coefficients formulations shown

[redacted]

Table 6 . Column multipliers from the four inverses based on the alternative formulations

Although of some interest, the above results are a function of the fictitious numerical example. The decision as to which of the formulations is appropriate should be made, rather, on conceptual and theoretical grounds. As noted, the BD and RBRD formulations generate either an undesired or nonsensical result, which narrows the choice to one between the fully interregionalized Use and the fully interregionalized Make formulations.   We leave mathematical proofs to others, and focus instead on the conceptual interpretations of systems defined according to either alternative.

The interregionalized Use formulation represents a system in which region-specific industrial production functions are the driving force behind the interregional frameworks generated. In a demand driven framework, it seems likely that establishments that have identified extra-regional sources of imports would indeed increase the size of their existing input orders according to increased production demands. The interregionalized Make formulation, in contrast, generates a system in which increases in an industry's total output will result in each region and each purchaser of its outputs will increase their consumption proportionately. The parallel in the single-region Make-Use framework is the commodity-based technology assumption, which de Menard (2004) asserts is itself sensible only in the context of the supply-driven input-output model. The interregionalized Make matrix appears to rest on heroic behavioral assumptions

However, there are potential problems associated with the use of the consolidated D matrix, which defines the aggregate region-specific industrial commodity output distribution (irrespective of destination) and applies it to regional industry production for use in all regions. For the two-region closed nation example, this is of little consequence, but could potentially take on greater importance – and hence introduce more error – as the number of regions and corresponding intervening distances increase.   It might well be the case, for example, that a large portion of an industry's primary commodity output is exported great distances, while its secondary commodities are produced and sold to a more localized market. Nevertheless, from the standpoint of rational economic behavior, the relationships in the interregionalized Use rest on the foundation of production relationships, and support it over the alternative.

The two-region closed system provided an additional simplification that should be noted. Because there were no foreign imports in the simple example, the coefficients in the Use tables were indeed technical coefficients. When the system is opened to foreign imports, competitive foreign imports must not be included in the Use tables, unless there is a corresponding Rest-of-World Industry in the Make table. Otherwise, all supply would be met by domestic industry. This implication is consistent with Dietzenbacher's (Dietzenbacher 2005)critique of the use of US-type Make-Use systems with embedded imports. Likewise, were the regionalized D and consolidated B approach chosen, the B matrix would need to represent regional technical coefficients, while the D matrix would need to include a Rest-of-World row industry, consistent with Jackson's (Jackson 1998) regionalization approach.

Summary

This paper has provided an initial discussion of unaddressed issues concerning the construction of many input-output tables founded on the Make-Use data framework. The primary focus is the choice between using the fully interregionalized Use data or the fully interregionalized Make data, since the two sets of information cannot be used in the same interregional table formulation. In the process, we define a “consolidated” form of the two tables, which either represents technical coefficients in the case of the Use, and which by including a Rest-of-World industry, represents total supply in the case of the Make matrix.

The discussion comes down on the side of using the combination of interregional Use and consolidated Make matrix approach. The preference is based on the foundation of production behavior consistent with the demand-driven input-output model rather than market share behavior, which appears to be more consistent with a supply-side input output model. The paper succeeds in laying out an array of relevant issues and implications of alternative approaches to the construction of interregional models, and provides an initial set of mechanisms for resolving those issues.

References

Beyers, W. B. (1983). "The interregional structure of the U.S. economy." International regional science review **8**(3): 213-232.

Canning, P. and Z. Wang (2005). "A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts." Journal of regional science **45**(3): 539-563.

Chenery, H. B. (1953). Regional Analysis. The Structure and Growth of the Italian Economy. H. B. Chenery, P. G. Clark and V. C. Pinna. Rome, U.S. Mutual Security Agency**:** 97-129.

de Mesnard, L. (2004). "Understanding the shortcomings of commodity-based technology in input-output models: An economic-circuit approach." Journal of Regional Science **44**(1): 125-141.

Dietzenbacher, E. (2005). The Fallacy of Using US-Type Input-Output Tables. 15th International Input-Output Conference. Beijing.

Isard, W. (1951). "Interregional and Regional Input-Output Analysis: A Model of a Space Economy." Review of Economics and Statistics **33**(4): 318-28.

Jackson, R. W. (1998). "Regionalizing National Commodity-by-industry Accounts." Economic Systems Research **10**(3): 15-39.

Jackson, R. W., W. R. Schwarm, et al. (2006). "A method for constructing commodity by industry flow matrices." Annals of Regional Science **40**(4): 909-920.

Lahr, M. L. (2001). "Reconciling Domestication Techniques, the Notion of Re-exports and Some Comments on Regional Accounting." Economic Systems Research **13**(2): 165 - 179.

Lindall, S., D. Olson, et al. (2006). "Deriving Multi-Regional Models Using the IMPLAN National Trade Flows Model." Journal of Regional Analysis & Policy **36**(1): 76-83.

Miller, R. E. and P. D. Blair (1985). Input-Output Analysis: Foundations and Extensions. Englewood Cliffs, New Jersey, Prentice-Hall.

Moses, L. N. (1955). "The Stability of Interregional Trading Patterns and Input-Output Analysis." American Economic Review **45**(5): 803-832.

Polenske, K. R. (1980). The U.S. multiregional input-output accounts and model. Lexington, Mass., Lexington Books.

Robinson, D. P. and Z. Liu (2006). "The Effects of Interregional Trade Flow Estimating Procedures on Multiregional Social Accounting Matrix Multipliers." Journal of Regional Analysis & Policy **36**(1): 94-104.

Schwarm, W. R., R. W. Jackson, et al. (2006). "An Evaluation of Method for Construction Commodity by Industry Flow Matrices." Journal of Regional Analysis & Policy **36**(1): 84-93.

U.S. Department of Commerce, B. o. E. A. (1991). The 1982 Benchmark Input-Output Accounts of the United States. Washington DC, US Department of Commerce.

West, G. R. (1990). "Regional Trade Estimation - a Hybrid Approach." International regional science review **13**(1-2): 103-118.