Economic Systems Research, Vol. 10, No. 3, 1998 223

Regionalizing National Commodity-by-Industry Accounts

RANDALL W. JACKSON

(Received May 1997; revised February 1998)

ABSTRACT *Exended input–output (IO) models are increasingly prominent in regional economic analysis. Social accounting matrices and associated multiplier decompositions, IO econometric model hybrids and computable general equilibrium models are finding greater acceptance in contexts in which simple IO models once dominated. Although the extended regional models build primarily on the foundation of regional, interindustry accounting frameworks, the data from which these regional accounts are drawn are most commonly in the form of a national commodity-by-industry account. Despite this long-standing fact, the IO table adaptation literature has focused almost solely on methods of adapting national interindustry accounts to regional economies. This paper presents a method designed specifically to regionalize commodity-by-industry accounts, in the context of the US reporting system. The focus on commodity-by-industry data demands a confrontation with several important issues that otherwise might go unattended. Using a particular system and its accompanying classification scheme ensures a comprehensive and consistent regionalization method.*

KEYWORDS: *Regional accounts, input–output, commodity-by-industry*

Introduction

When the input–output (IO) model was becoming the framework of choice for regional analysis, there was great interest in regional IO table construction methods and associated issues (for reviews, see Hewings & Jensen, 1986; Miller & Blair, 1985). As the social accounting matrix, mixed econometric and IO and computable general equilibrium alternatives gained popularity, academic interest and attention shifted quite naturally to the theoretical underpinnings and development of these extensions. In the process, however, important aspects of the regional IO accounting foundations of these models have been neglected.

R. W. Jackson, Department of Geography, The Ohio State University, 1036 Derby Hall, 154 North Oval Mall, Columbus, OH 43210, USA. Tel: 1-614-292-7998; Fax: 1-614-292-6213; E-mail: jackson.21@osu.edu. The author appreciates the comments of Jan Oosterhaven and anonymous reviewers on earlier drafts of this paper. Both the method and the presentation of the paper have benefited. Responsibility for any remaining deficiencies, of course, is wholly that of the author.

0953-5314/98/030223-16

The modeling extensions build most commonly on a set of regional interindustry (institutional) accounts, virtually all of which are adapted from national accounts. However, the literature on adapting national to regional IO tables almost invariably uses industry-based national accounts as the point of departure (for a recent example, see Boomsma & Oosterhaven, 1992).[1] Indeed, this remains the case nearly three decades after the United Nations adopted the commodity-by-industry framework for national reporting (Gigantes, 1970; United Nations, 1968). The purpose of this paper is to renew interest in the theoretical and practical issues that accompany the transition from national to regional commodity-by-industry accounts.

Industries have been dominant units of analysis for as long as economies have been studied. The continued focus on industry-based IO accounts and modeling frameworks can be partly explained by simple inertia, and partly by the industry-based statistical reporting systems employed by government agencies. Another factor in explaining this continued focus may be the relevance and intuitive appeal of the industry as a meaningful agent in a regional economy. Furthermore, common impact assessments and a number of other IO applications can be carried out without estimating the parameters that lie outside the interindustry quadrant of traditional accounts. Many traditional and recent forms of key sector analysis (Hirschman, 1958; McGilvray, 1977; Rasmussen, 1956), industrial complex and clustering analyses (Czmanski & Czmanski, 1977; Howe & Stabler, 1992; O'hUallachain, 1984; Roepke *et al.*, 1974) and a wide variety of analyses of regional economic structural change (Aroche-Reyes, 1996; Dewhurst, 1993 Jackson & Hewing, 1984; Jackson *et al.*, 1989) and many other IO applications are largely independent of the other quadrants (final demand, payments and institutional).

The traditional focus on processing sectors, however, neglects some of the important attributes that distinguish regional economies. Final demand activities, such as state and local government spending on educational or non-educational activities, vary widely from region to region, for reasons that include the presence or absence of colleges and universities, prisons and state capitals (government itself). Failing to account for these final demand variations can lead to over-estimation or underestimation of regional industry's ability to supply goods or services for local demand. Hence, neglecting the final demand portion of the accounts can be a major cause of biased parameter (coefficient) estimates being obtained by applying standard, quotient-type IO table adaptation methods.

To illuminate the relevant issues and directly confront the format of the base data, this paper sets out a method for regionalizing commodity-by-industry-based, national IO accounts. Its objectives are to reduce the parameter bias that results from neglecting the non-processing sectors of regional accounts, and to take full advantage of the additional detail and flexibility in application provided by the national commodity-by-industry accounts. The outcome of this effort is a complete set of accounts, depicting the processing and the non-processing sectors of the economy. The detailed presentation of the procedures serves to identify certain important conceptual issues and to foster further discussion of the role of the commodity-by-industry framework in regional IO table construction.

The next section presents a brief review of the data tables, comprising the national accounting framework published for the US by the US Bureau of Economic Analysis. Although an officially sanctioned system of national accounts (SNA) exists (World Bank, 1993, p. 47), different nations compile and publish data that conform to varying degrees with the SNA guidelines. Regional analysts in each nation, therefore, will begin from somewhat different points of reference, and may need to make minor adjustments to their national accounts or to the procedure to be outlined here. The general presentation of the national accounts serves to introduce the necessary terminology and notation, focussing on those features of the national accounting framework that are particularly relevant to the construction of a regional counterpart. Although the US case identifies a single point of reference, it provides the important advantages of a concrete example. Section 3 presents a comprehensive and replicable method for regionalizing a national commodity-by-industry accounting framework. The paper concludes with a section that summarizes the regionalization method and highlights table construction issues.

2. The National Commodity–Industry Framework

To lay the foundations for the regionalization method, this section presents a review and discussion of the most prominent features of the US national accounting framework. The discussion results in a modified structure for the US accounting tables, which brings them more closely into line with the SNA guidelines and simplifies the regionalization procedure. A more detailed discussion of the data source can be found in US Department of Commerce, Bureau of Economic Analysis (BEA) (1997).

Table 1 presents a slightly modified version of the conventional layout of annual commodity and industry accounts.[2] Empty blocks in the table are devoid of values, while the relationship across the rows and down the columns form mathematical identities. Using | to denote the horizontal matrix augmentation operator, the commodity final demand block has been replaced by an equivalent expression of its component activities. The component terms in $\mathbf{E}=\mathbf{F}|\mathbf{x}|(-\mathbf{m})$ correspond to a matrix of domestic final demand, an exports column vector and a transposed row vector of imports with signs reversed. Commodity accounts are comprised of blocks **U**, **E** and **q**.

Element $u_{ij} \in \mathbf{U}$ denotes, in producers' prices, the value of commodity i used by industry j during a calendar year. **U** reports all commodities used, irrespective of their geographic production origins, and is refered to as either the 'use matrix' or the 'absorption matrix'.

Element $e_{ij} \in \mathbf{E}$ denotes the value of commodity i sold to final demand activity j, also in producers' prices. Final demand activities in this set-up include personal consumption expenditures, investment, inventory changes, federal expenditures, state and local government expenditures, exports (**x**) and imports (**m**). The ith

Table 1. Schematic layout of commodity and industry accounts with imports shown as negative entries in final demand

	Commodities	Industries	Final demand	Total output
Commodities		$\mathbf{U}$	$\mathbf{F}\|\mathbf{x}\|(-\mathbf{m})$	$\mathbf{q}$
Industries	$\mathbf{V}$			$\mathbf{g}$
Value added		$\mathbf{W}$		
Total inputs	$\mathbf{q}'$	$\mathbf{g}'$		

element in the imports vector, transposed and entered in the negative in **E**, corresponds to the value of commodity i imported by all domestic industries and other final demand activities during the year.

Element $q_i \in \mathbf{q}$ is the total dollar value of commodity i domestic output. The effect of the imports column in **E** is to offset the values of commodity imports in the 'use' and domestic final demand blocks. Hence, the accounting identity for commodity output, or the commodity balance equation, is

$$q_i \equiv \sum_{j=1}^{n} u_{ij} + \sum_{j=1}^{p} e_{ij}, \qquad \forall i, i = 1, \ldots, n \tag{1}$$

Letting **i** denote an appropriately dimensioned column summing vector, equation (1) can be rewritten as

$$\mathbf{Ui} + \mathbf{Ei} \equiv \mathbf{q} \tag{2}$$

The block labeled **V** in Table 1 represents production-oriented relationships and is commonly referred to as the 'make matrix'. A make matrix augmented by a totals vector and one or more imports vectors constitutes the 'supply matrix' in the terminology of the most recent SNA. All BEA benchmark IO accounts to date—indeed, including the recently released 1992 benchmark—have been constructed and published in the format described and used in this presentation. Thus, element $v_{ij} \in \mathbf{V}$ denotes, in producers' prices, the value of commodity j produced by domestic industry i during the year. The sum of the jth column of **V** is the total domestic commodity j output, i.e. q_j, and the sum of the ith row is the total domestic industry i output, i.e. g_i. Therefore, we have

$$\mathbf{Vi} \equiv \mathbf{g} \tag{3}$$

and

$$\mathbf{i'V} \equiv \mathbf{q'} \tag{4}$$

The block labeled **W** in Table 1 denotes the components of value added in the production of the output of corresponding column industries. Value-added components include wages and salaries, profit-type income and indirect business taxes, for example. Therefore, element w_{ij} denotes the dollars of component i value added to the economy during the year by production in industry j.

The final identity in the system is the commodity-by-industry input balance. The sum of the value of all commodity inputs used by an industry and the total value added by that industry is equal to its domestic output value. Formally, we have

$$\mathbf{i'U} + \mathbf{i'W} \equiv \mathbf{g'} \tag{5}$$

Together, equations (2)–(5) provide the basic identities of the national accounting framework.

The two most critical deviations of the US accounts from the SNA lie in the treatment of imports and in the valuation of entries in the use table. In the US accounts, the use table entries are in producers' prices, while margins have been collected and transferred to the corresponding use table rows. This deviation from the SNA simplifies the regionalization procedure, particularly where subnational data on margins are scarce or non-existent. The treatment of imports, however, not only deviates from the SNA but also complicates the transition from accounting to modeling frameworks. Modifying the national accounts prior to regionalization

Table 2. Schematic layout of commodity and industry accounts with imports shown as a commodities source

	Commodities	Industries	Final demand	Total output
Commodities		**U**	**F** | **x**	**q** + **m** = **s**
Industries	**V**			**g**
Imports	**m**′			**m**′**i**
Value added		**W**		
Total inputs	**s**′	**g**′		

is useful, since estimating regional imports and exports is a central task. It is critical to understand fully and emphasize imports accounting before constructing accounts for subnational regions the economies of which are generally much more open than those of their national counterparts.

Table 2 is a modified version of the schematic layout in Table 1. The imports entries in the commodity final demand block have been removed, and **m**′ has been added as a new, 'imports' row. The row sum **m′i** is simply the sum of all imports. Summing across a commodity row now yields total commodity use **s**, or

$$q_i + m_i = s_i \in \mathbf{s} \tag{6}$$

Likewise, we have

$$\sum_i v_{ij} + m_j = s_j \tag{7}$$

The vector of gross industry output **g** remains unaltered. The need for these modifications will become apparent in the Section 2.1.

2.1. From Accounts Construction to Modeling

IO accounts form the basis of a variety of modeling frameworks. Traditional IO models can be framed in a variety of functional and institutional formats, with buyers and sellers classified by commodity and institution. Given the historical dominance of the interindustry format for regional application, this section describes the transition from the commodity-by-industry accounts to the institutional modeling framework.

A first step in this transition is the construction of commodity-by-industry direct requirements coefficients, by defining a fixed relationship between commodity input and industry output values. This relationship is directly analogous to the traditional expression of industry technical coefficients, where both the numerator and the denominator are in industry value units. The formal expression is

$$\mathbf{B} = \mathbf{U}\hat{\mathbf{g}}^{-1} \tag{8}$$

By imposing one of two possible production assumptions, several forms of this basic relationship—and of the corresponding total requirements matrices—can be derived. The choice of assumption is influenced by the units in which it is desired or necessary to work, and by the preference for a model that reflects commodity-based or industry-based technology.[3]

Under the commodity-based technology assumption, industries produce commodities in fixed proportions. When an industry's output increases, its production of each commodity increases proportionately. Conversely, constant industry contributions to total commodity output characterize the industry-based technology assumption. Increases in the output of a commodity under the industry-based technology assumption indicate proportionate increases in the production of that commodity in each industry in which it is produced; therefore, they also indicate changes in the commodity mix of industry output. Because the industry-based technology assumption has been applied most often, particularly in the US, the discussion that follows also used the industry-based technology assumption, captured by the expression[4]

$$\mathbf{D} = \mathbf{V}\hat{\mathbf{q}}^{-1} \tag{9}$$

Most regional analysis has been and—particularly in the US—continues to be industry based. Again, this can be explained partly by the emphasis on industries as common and meaningful units for collecting and reporting associated data (such as on employment, income, investment, etc.), and partly by inertia within the community of researchers whose work began with interindustry accounting frameworks. Although other models can be developed in a similar fashion, the institutional model is the focus of this presentation.

From equation (9), we have

$$\mathbf{D}\hat{\mathbf{q}} = \mathbf{V} \tag{10}$$

Substituting the left-hand side of equation (10) for **V** in equation (3) yields

$$\mathbf{D}\mathbf{q} = \mathbf{g} \tag{11}$$

Because **V** embodies the relationship between commodity and industry output, **D** provides a general transformation from commodity to industry space and, given equal numbers of commodity and industry sectors, the inverse of **D** transforms from industry to commodity space. These transformations can be used to advantage. By premultiplying both sides of equation (2) by **D**, an industry-based output balance equation is obtained, i.e.

$$\mathbf{DUi} + \mathbf{DEi} = \mathbf{g} \tag{12}$$

However, this transformation is peculiar in several ways. First, because **U** includes the value of imported commodities, elements in a column of **DU** will continue to be comprised partially of imports, so will not be accurate representations of flows among domestic industries as expected. Elements in the columns of **F** will be similarly misrepresented. Assuming that exports are produced domestically, the elements of **Dx** will be correct, while $-\mathbf{Dm}$ implies that foreign and domestic commodity compositions of industry output are identical. The following method, which extends the concept of total-based make table coefficients (US Department of Commerce, BEA, 1991), is designed to overcome these often-overlooked inconsistencies and to avoid the implicit assumption about the nature of foreign production.

Begin with the notation of Table 2. A representation of commodities exported by each industry is obtained by premultiplying the diagonalized exports vector by **D**, as defined in equation (9). The result, i.e. $\mathbf{D}\hat{\mathbf{x}}$, is a matrix of the same dimension as **V** that shows domestic industry production of commodities for export—commodities that are, therefore, unavailable for domestic use. Subtracting

$\mathbf{D\hat{x}}$ from $\mathbf{V}$ then yields domestic commodity supply by industry, and the total supply for domestic demand is

$$\mathbf{H} = \begin{pmatrix} \mathbf{V} - \mathbf{D\hat{x}} \\ \mathbf{m}' \end{pmatrix} \tag{13}$$

The values in $\mathbf{H}$ represent commodities available to fill intermediate and final domestic and export demand. The column sums of $\mathbf{H}$ equal the domestic commodity supply minus exports plus imports, i.e.

$$\mathbf{l} = \mathbf{iH} = \mathbf{q} - \mathbf{x} + \mathbf{m} \tag{14}$$

Standardizing $\mathbf{H}$ yields a matrix that shows the composition of industry and import sources of commodity production for domestic demand, i.e.

$$\tilde{\mathbf{D}} = \mathbf{H}(\hat{\mathbf{l}})^{-1} \tag{15}$$

with column sums of one. Postmultiplying $\tilde{\mathbf{D}}$ by $(\mathbf{U}|\mathbf{F})$ transforms the $(\mathbf{U}|\mathbf{F})$ from commodity space to industry and import sector space, i.e. the last row will now show domestic uses of imports. The value-added matrix $\mathbf{W}$ is consistent with the commodity-by-industry accounts or with the industry-by-industry accounts. These matrices constitute the components necessary to generate the institutional accounts in a format familiar to most regional IO modelers.[5] The accounting framework can be shown concisely as

$$\left(\begin{array}{cc|c} \tilde{\mathbf{D}}\mathbf{U} & \tilde{\mathbf{D}}\mathbf{F} & \mathbf{Dx} \\ \cline{3-3} & & 0 \\ \hline \mathbf{W} & & \end{array}\right) \tag{16}$$

3. Regionalizing the National Accounts

To avoid excess abstraction, and to ensure a comprehensive treatment, this section develops a regionalization procedure for the 85-sector US commodity-by-industry accounts published by the US Department of Commerce, BEA. With the goal of establishing a general and adaptable procedure, the approach described can accommodate the introduction of superior data, but does not depend on them for implementation. Generality is retained at the expense of sophistication in some areas, but the typical regional analyst faced with a lack of region-specific data will inevitably have few alternatives to the use of various simple proportional estimation (Kendricks–Jaycox) methods as employed here. Following the regionalization, the commodity-by-industry framework can be transformed into the institutional format, or any other modeling format, as desired. A description of the BEA data tables precedes the presentation of the regionalization method.

3.1. The BEA Data, Classification Scheme and Special Sectors

The IO structure of the US is presented as a series of data tables that include and build on the make and use table data. The BEA's Table 1 (BEA-1) is entitled 'Use of commodities by industries'. It is comprised of intermediate and final uses $(\mathbf{U}|\mathbf{E})$, along with a single value-added row $(\mathbf{W})$. The BEA's Table 2 is the make table, i.e. $\mathbf{V}$, as shown in Tables 1 and 2. The remaining BEA tables are primarily direct

Table 3. Special sectors

BEA sector number	Revised sector number	Sector name	BEA 1	Supply table
80	86	Non-comparable imports	Row	
81	84	Scrap	Row	Column
82	80	Government industry	Column, row	Column, row
83	81	Rest of world	Column, row	Column, row
84	82	Household industry	Column, row	Column, row
85	83	Inventory valuation adjustment	Column, row	Column, row

Table 4. Final demand distribution for special sectors

	PCE	GPFI	CBI	Exports	Imports	Federal govt. defense	Federal govt. other	State and local govt.	State and local govt.
BEA 80 NCI	×	×	×	×	×	×	×	×	×
BEA 81 scrap	×	×	×	×	×	×	×	×	×
BEA 82 GI						×	×	×	×
BEA 83 ROW	×			×	×	×	×		
BEA 84 HI	×								
BEA 85 IVA			×						

Notes: PCE, personal consumption expenditure; GPFI, gross private fixed investment; CBI, changes in business inventory; NCI, non-comparible imports; GI, government industry; ROW, rest of world; HI, household industry; IVA, inventory valuation adjustment.

and total requirements tables in various permutations of institutional and functional formats.

The BEA classification comprises sectors 1–79, the industry sector names of which match their primary product commodity classification, and special accounting sectors 80–85, as shown in Table 3. Sectors 82–85 have only value-added entries in their BEA-1 columns, in amounts that exactly equal their diagonal entries in the make table. Final demands for these sectors equal the corresponding sectoral outputs, which, therefore, do not enter into the processing sectors as inputs.

Table 4 shows the cell locations of non-zero final demand entries for the special sectors. No processing sectors use these outputs, and there is a one-to-one correspondence between industry and commodity output. The challenge for regionalizing sectors 82–85 lies in estimating total commodity output values and their distributions across final demand activities. These procedures will be presented, following the regionalization of final demands. The treatment of sectors 80 and 81, i.e. non-comparable imports (NCI) and scrap, is described next.

3.1.1. Non-comparable imports. NCI, i.e. sector 80, has a use table row but no use table column, since there is no domestic industry for this commodity sector. Also for this reason, there will be no associated row or column in the make matrix. From the standpoint of forming **D** (or computing its inverse), this presents no problem, but it does create an imbalance in the number of commodity sectors in BEA-1 and in the make table. For convenience and consistency in sequential numbering, the NCI row of BEA-1 is moved into the payments quadrant, effectively reducing the number of 'processing' sectors to 84. NCI becomes row 86, value added (VA) becomes row 85, sectors 82–85 become sectors 80–83, and scrap (discussed next) becomes sector 84.

3.1.2. Scrap. This commodity sector also has no corresponding industry. It has a commodity row but no industry column in BEA-1, and has a commodity column but no industry row in the make table. For convenience and consistency, the sector sequence is rearranged such that the scrap sector becomes the 84th row of BEA-1. Sectors 1–83 will now have corresponding rows and columns in BEA-1 and in the make table. The revised order of the sectors is indicated in Table 3.

3.2. Regionalizing the Accounts

The procedure begins by establishing a regional counterpart to the BEA make table. Although there is an interdependence between certain cells in the two tables, the most critical cells of the make table can be estimated without referring to BEA-1.

3.2.1. Output. The central task in any method designed to regionalize national accounts is estimating the ability of the region to supply its own needs. This information ultimately enables the estimation of regional imports and exports. To generate a regional commodity balance equation, industry output and final demand must be scaled to reflect regional activity levels.

Output data are rare for most subnational regions at any significant level of disaggregation. Industry-based employment data are much more commonly supported at the subnational level, and often form the majority of region-specific data available to the regional modeler. Given national and regional employment estimates, corresponding national output estimates and a willingness to assume productivity invariance among subnational regions within each industry, the regional output for industry j is estimated from the region's share of national industry j employment, denoted by $\varepsilon_j \in \boldsymbol{\varepsilon}$. We have

$$g_j^{\mathrm{R}} = \varepsilon_j g_j^{\mathrm{N}}, \qquad \forall j, j = 1, \ldots, n \tag{17}$$

If there are reliable estimates of regional-to-national productivity ratios by industry, $\rho_j \in \rho$, then they can be incorporated as

$$g_j^{\mathrm{R}} = \varepsilon_j \rho_j g_j^{\mathrm{N}}, \qquad \forall j, j = 1, \ldots, n \tag{18}$$

The expression $\varepsilon_j \rho_j = \tau_j \in \tau$ now indicates the relationship between regional and national industry output. Therefore, the rows of the national make table can be scaled to regional output levels using

$$\mathbf{V}^{\mathrm{R}} = \hat{\tau}\mathbf{V} \tag{19}$$

The sums of rows and columns 1–79 of $\mathbf{V}^{\mathrm{R}}$ are regional industry and commodity output estimates. Sectors 80–83 have only diagonal entries in $\mathbf{V}^{\mathrm{R}}$. The values for these entries will be derived from later steps.

3.2.2. The use table. The national use matrix can be scaled to a regional use matrix in a similar fashion, i.e.

$$\mathbf{U}^{\mathrm{R}} = \mathbf{U}\hat{\tau} \tag{20}$$

This equation implies that regional and national industry technology are identical. To enhance the accuracy of the results, the scaled use table can be directly edited to replace first estimates with any superior region-specific data, prior to subsequent steps. Analysts with access to superior data are always expected to use them to enhance the general procedure described here.

3.2.3. Final demand. The national final demand quadrant, now defined as $\mathbf{F}|\mathbf{x}$ and excluding $\mathbf{m}$, must be modified on an activity-by-activity basis to approximate its regional counterpart. Some of the elements in a regional final demand table are primarily functions of the size of the corresponding regional production sector and some are more directly functions of the size of the final demand activity sector within the region. A region's proportion of rest-of-the-world export, for example, can be assumed to be related to the contribution of regional industries to national industrial output. Given a (typical) lack of region-specific data, arguments might be made for treating inventory adjustments similarly. The second kind of final demand activity is perhaps best represented by personal consumption expenditures. Regional demand for food, for example, is more directly a function of total regional personal income than it is a function of the size of the food processing industry in a region. Similar arguments can be put forward for the treatment of various categories of government expenditure.

Let us consider local supply-dependent activities, i.e. those for which commodity final demand is determined primarily by the relative size of the local production sector. In the extreme, if commodity i were produced only in region A, then all final demands for that commodity would be met by region A. Exports are the obvious example of this type of final demand activity. A region's rest-of-the-world exports and commodity i can be estimated, for example, as the product of national commodity i exports and regional share of national commodity i output.

The most difficult final demand activity to characterize is investment. Although the bulk of region investment is typically construction expenditure, investment demand is by no means limited to construction. A capital flows table, describing investment purchase patterns by industry, would offer a most appealing set of data to use for regional investment estimation. For the US, however, the most recent capital flows tables are nearly 20 years old. Although more sophisticated methods of estimating investment final demand by commodity can be justified and are encouraged, investment and export demand estimates are treated the same as any other local supply-dependent final demand activity. Gross private fixed investment (GPFI), changes in business inventory (CBI) and exports are estimated as local supply-dependent final demand activities.

To scale any local supply-dependent final demand activity column, apply the regional shares of national output to the corresponding commodity final demand values. Let $\gamma_i \in \gamma = (q_i^{\mathrm{R}}/q_i^{\mathrm{N}})$ be the regional share of commodity i production. Elements of γ that correspond to government industry (GI), rest-of-world industry (ROW), household industry (HI), inventory valuation adjustment (IVA) and NCI are assigned zero values for convenience, since these commodities are not produced by conventional industries. For supply-dependent activities, $\mathbf{E}_j^{\mathrm{R}} = \hat{\gamma}\mathbf{E}_j^{\mathrm{N}}$.

The IVA, ROW and NCI values for GPFI, CBI and exports can be estimated by multiplying the ratio of the national final demand value to the national processing sector final demand total by the regional processing sector final demand total. For example, the general equation for the NCI values is

$$E_{\mathrm{NCI},j}^{\mathrm{R}} = \left(\frac{E_{\mathrm{NCI},j}^{\mathrm{N}}}{\sum_{i=1}^{79} E_{ij}^{\mathrm{N}}} \right) \sum_{i=1}^{79} E_{ij}^{\mathrm{R}} \tag{21}$$

Values for GI and HI for these activities retain their zero values (see Table 2).

If we now consider local demand-dependent activities, entries in several columns of **E** can be expressed as a function of the associated aggregate regional activity level. These columns include personal consumption expenditures (PCE), and federal, state and local government expenditures. Scaling these activities is relatively straightforward. Elements in the regionalized columns are the product of the national final demand values and the region's share of national total, i.e.

$$e_{ij}^{R} = \left(\frac{\text{Regional control total}_j}{\text{National control total}_j}\right) e_{ij}^{N} \tag{22}$$

The regional control totals might be measures of total personal income, total state and local government education expenditures, etc. This scaling procedure can be applied to all column values, including those for special sectors. The modified elements will continue to be expressed in commodity value units.

Using actual data in place of any of the scaled estimates is always preferred. Where reliable regional personal consumption data or government expenditures distributions are available, for example, their use will only enhance the overall accuracy of the regional accounts. For these activities as well as the local supply-dependent activities, the final demand entries can now be directly edited prior to the estimation of total imports.

3.3.3. Imports. Rest-of-world imports to a regional economy will be, in part, a function of the intra-regional demand for each commodity. Although it can be assumed that a region would share in rest-of-world imports of a commodity in proportion to its share of non-export demand for this commodity, local commodity production levels must also be incorporated. Demand and local production levels also enter into the estimation of rest-of-nation imports.

An alternative is to treat national and rest-of-nation imports as a single, or 'hybrid' final demand activity sector, estimated as residuals from output balance equations. For two main reasons, this 'supply–demand pool' approach is used here, despite the mixed results of comparisons of adaptation approaches in the literature (McMenamin & Haring, 1974; Morrison & Smith, 1974; Sawyer & Miller, 1983). First, although simpler quotient methods have often been ranked higher for their ability to replicate survey-based regional IO coefficients, those comparisons were based on a national interindustry coefficients table as a starting point, and focused only on replicating the processing sector parameters. For the regionalization method to provide the foundation for social accounting matrices, econometric IO and computable general equilibrium models, it must generate a consistent estimate of the entire set of regional accounts.

Second, the supply–demand pool approach can be argued to be theoretically superior to methods based on location quotients, which do not account for variations in the final demand structure. The most strongly theory-driven alternative is the regional purchase coefficient (RPC) of Stevens *et al.* (1983). Although it might be possible to incorporate that formulation of RPCs into a method similar to the method presented here, the supply–demand pool approach is chosen for simplicity and conceptual clarity.

Net imports of a commodity to a region are estimated as the difference between total regional industry supply and total regional demand. Let $f_{kj} \in \mathbf{F}^{R}$ denote the matrix of p domestic final demand activities that have been estimated independently (i.e. personal consumption expenditures, investment, inventory changes and government expenditures). Using $\mathbf{x}^{\mathrm{RW}}$ to denote rest-of-world exports, net imports of commodity k can be estimated as

$$m_k = \left(\sum_{j=1}^{79} u_{kj}^{\mathrm{R}} + \sum_{j=1}^{p} f_{kj}^{\mathrm{R}} + x_k^{\mathrm{RW}} \right) - \sum_{i=1}^{n} v_{ik}^{\mathrm{R}}, \qquad \forall k, k = 1, \ldots, 79 \tag{23}$$

The v_{ik}^{R} elements in equation (23) are those from a national make table that has been scaled analogously to the use table, as shown in equation (20). The elements $m_k \in \mathbf{m}$ correspond to demand less supply, so can be positive or negative. Negative values indicate less demand than supply and, therefore, a net regional surplus of commodity k, while positive values indicate a net regional supply deficit (bearing in mind that the estimated regional contribution to rest-of-world exports is now taken as fixed and treated as though it were a typical final demand activity).

To accommodate negative values, introduce a new commodity final demand activity to represent regional exports to the rest of the nation, i.e. $\mathbf{x}^{\mathrm{RN}}$. Negative values in $\mathbf{m}$ are transferred to this new column, reversed in sign (they are now net exports rather than net imports), and replaced by zero values in $\mathbf{m}$, i.e. the estimate of net imports from all origins. Commodity k can have a non-zero imports entry and a positive entry in the national exports sector, indicating that the region exports to other countries commodities that it imports from all origins.

Although it might be desirable to differentiate between rest-of-world and rest-of-nation imports, the data do not support such a distinction. It would be possible to use the framework of equation (21) to estimate an international imports share of total output (an IO coefficient) but the fact that these are competitive imports suggests that regions with strong concentrations of production for a given commodity would have smaller coefficients than would those that have little or no production of that commodity. Hence, net imports are estimated as shown in equation (23). Further note that whether to estimate rest-of-world and rest-of-nation exports as two separate vectors or as a hybrid is also a choice that is left to the analyst. Estimating exports as a hybrid vector would simply require eliminating the x_k^{RW} term from equation (23).

The treatment of imports described, however, still implies that no cross-hauling takes place either among the region and other countries or within subnational regions. If there were no cross-hauling, then the estimate of rest-of-nation exports would be gross rather than net, as would the estimates in the new imports column. Assuming no cross-hauling is naive, of course, and underestimates gross imports and exports from other regions. Some mechanism to account for cross-hauling is needed.

Cross-hauling values by commodity, from superior data or from estimates, can be added to $\mathbf{x}^{\mathrm{RN}}$ and $\mathbf{m}$. These adjustments could be made on an *ad hoc* basis or can be formally incorporated into the procedure by specifying a vector, say $\mathbf{k}$, of values expressed as the proportions of regional commodity outputs that are cross-hauled. For commodity i, the value to be added to $\mathbf{x}^{\mathrm{RN}}$ and subtracted from $\mathbf{m}$ will be $k_i q_i^{\mathrm{R}}$. Other forms of superior data can also be used. For example, known values for specific commodity imports can be used in place of mechanical estimates. Direct adjustments to imports (exports), however, must be offset by adjustments to exports (imports) to maintain the commodity balance.

There are three special sector values in the imports column of BEA-1. The NCI value for imports is defined to be equal to the negative of the sum of other

NCI. The scrap value can be computed as the difference between total scrap production, derived as the sum of the scrap column of the regionalized make matrix, and the sum of the regionalized scrap row of BEA-1 (obviously excluding the imports column). In the absence of superior data, there is little guidance as to the appropriate imports column ROW value. One approach is to assume that the regional imports share of total ROW mirrors the corresponding national share, and compute the value accordingly.[6]

The importance of accurately estimating commodity imports cannot be overemphasized. Recall that this column of imports values in accounts that parallel Table 1 becomes a row vector in Table 2, which is the system used in the generation of the modeling framework in Section 2.1. Thus, neglecting cross-hauling, for example, will result in overestimates of regional supply and correspondingly inflated output multipliers when models based on these accounts are used as the basis for impacts assessment.

3.4. *Value added*

The final step in regionalizing BEA-1 is adjusting the value-added block. Assuming that value added is proportionate to industry size, and that value added per dollar output within industries is invariant across subnational regions, values for the regional value-added block can be computed as

$$\mathbf{W}^{R} = \mathbf{W}\hat{\tau} \tag{24}$$

3.5. *From Accounts Construction to Modeling*

These steps generate a regional version of the accounts presented in Tables 1 and 2. The same sets of procedures used to construct national modeling frameworks, such as that presented in Section 2.1, can now be applied to the regional accounts to construct the basis of a regional modeling framework. The only additional consideration is that there are now two exports columns: rest-of-world and rest-of nation exports.

4. Conclusions

The method described reveals the richness of detail of the commodity-by-industry accounts for regionalization procedures. Equally importantly, however, it reveals a wide array of roles for professional judgement and the need for careful attention to the consistency of the procedure. The outcome of a procedure that regionalizes the entire accounting framework is fundamentally different from—and, presumably, more accurate and reliable than—the outcome of a procedure that regionalizes only the interindustry coefficients. As a foundation for extended models, of course, it is essential to generate a consistent and comprehensive set of accounts.

The transition from national commodity-by-industry accounts to regional accounts identifies a number of important issues to be confronted in regional accounts construction. First, even if quotient-based methods were applied to implement the regionalization, there is a critical difference between interindustry table derivations based on conventional versus modified make matrices, such as the matrix **H** derived here. For example, quotient methods are generally predicated on the assumption that the reference region (usually the nation) interindustry coefficients reflect 'self-sufficiency'. Thus, given a location quotient of unity, $a^{R}_{ij} = a^{N}_{ij}$. This would be appropriate only if the national coefficients were derived using the total-based make table. Using the standard make table would generate positively biased regional coefficients, leading to overestimates of industry and system mutlipliers.

Second, standard descriptions of the methods used to move from commodity-by-industry accounts to modeling frameworks (see, for example, Miller & Blair, 1985) overlook the procedural modifications that must be introduced to generate a use table that is consistent with the exports component of final demand. In the absence of these adjustments, imports by local industries will be underestimated and multipliers will be overestimated.

Third, explicitly incorporating cross-hauling ultimately reduces the size of the regional transactions estimates (and corresponding IO coefficients) and, therefore, reduces the individual industry and system multipliers.

The method described leaves room for extension and enhancement, particularly in the treatment of investment final demand and in greater use of superior data. Explicit means of incorporating superior data in this method include (1) the direct editing of the scaled make and use tables prior to the estimation of imports; (2) the inclusion of relative regional productivity estimates; and (3) the adjustment of net exports and imports estimates to account for cross-hauling.

This paper has only been directed towards the construction of accounts for a single region. Additional issues, such as aggregate consistency of comprehensive sets of subnational accounts, remain to be addressed. To this end, Oosterhaven (1984) has formalized a family of square and rectangular interregional IO models, and Eding and Oosterhaven (1996) are confronting these issues for accounts in The Netherlands—a different geographical and statistical reporting context.

A focus on specific national contexts is accompanied by an inescapable loss of generality. As a result, the methods presented here are centered on the US experience. This is a recurring, general and widespread problem, however, that will continue to plague IO analysts and research in related areas, unless and until all countries adopt identical data collection and reporting frameworks. Perhaps the World Bank (1993) advice to countries 'to start to compile accounts utilizing the revised *System* as soon as possible' will be taken seriously and alleviate these problems. With or without such standardization, however, advancements in national accounting practices continue and, with renewed attention, can lead to corresponding advances in regional accounts construction and economic systems research.

Notes

1. Eding and Oosterhaven (1996) also have begun to address issues involved in constructing regional IO tables from commodity-by-industry accounts. Their work is driven partly by a need to construct a consistent and exhaustive set of 14 regional tables for The Netherlands. Similarly, this paper stems from the increase in US interest and activity in the construction of extended IO models.
2. Similarly to many others, the US accounts deviate from the SNA. It is assumed that analysts will be able to rearrange their data, literally or conceptually, to match the format described here.
3. Kop Jansen and ten Raa (1990) have suggested that the choice of one technology assumption over the other seems to be a 'matter of judgement or taste' (p. 214). They provide some theoretical justification for using the commodity technology assumption but note that, in practice, problems of negative coefficients and violations of material balance constraints arise. Although Konijn and Steenge (1995) more recently introduced an alternative approach to creating IO accounts based on

an 'activity technology assumption', their approach 'requires very detailed (usually rectangular) *make* and *use* matrices, as well as the use of exogenous information on input structures of activities' (p. 33).

4. The discussion of technology assumptions draws on a parallel presentation in Miller and Blair (1985), who also present the formulations of various functional and institutional forms of the IO model. Although this paper uses the industry-based technology assumption, modifying the regionalization method to be consistent with alternative technology assumptions should present no problems beyond those that equally concern national-level IO accounts.
5. For a numerical example of this procedure, please contact the author.
6. The ROW industry represents 'factor income by US residents from abroad, less payments of factor income by US residents to foreigners' (US Department of Commerce, BEA, 1991).

References

Aroche-Reyes, F. (1996) Important coefficients and structural change: a multi-layer approach, *Economic Systems Research*, 8, pp. 235–246.

Boomsma, P. & Oosterhaven, J. (1992) A double entry method for the construction of bi-regional input–output tables, *Journal of Regional Science*, 32, pp. 269–284.

Czmanski, D. Z. & Czamanski, S. (1977) Industrial complexes: their typology, structure and relation to regional development, *Papers of the Regional Science Association*, 38.

Dewhurst, J. D. L. (1993) Decomposition of changes in input–output tables, *Economic Systems Research*, 5, pp. 41–54.

Eding, G. J. & Oosterhaven, J. (1996) Towards a new, rectangular approach in the construction of regional input–output tables, paper presented to the North American Regional Science Association, Washington, DC.

Gigantes, T. (1970) The representation of technology in input–output systems, in: A. P. Carter & J. Brody (eds) *Contributions to Input–Output Analysis*, pp. 270–290 (Amsterdam, North-Holland).

Hewings, G. J. D. & Jensen, R. C. (1986) Regional, interregional and multiregional input–output analysis, in: P. Nijkamp (ed.) *Handbook of Regional and Urban Economics*, vol. 1, pp. 295–355 (Amsterdam, North-Holland).

Hirschman, A. O. (1958) *The Strategy of Economic Development* (New Haven, CT, Yale University Press).

Howe, E. C. & Stabler, J. C. (1992) The regional structure of the United States Economy, *Papers in Regional Science*, 71, pp. 175–191.

Jackson, R. W. & Hewings, G. J. D. (1984) Structural change in a regional economy: an entropy decomposition approach, *Modeling and Simulation*, 15, pp. 451–455.

Jackson, R. W., Hewings, G. J. D. & Sonis, M. (1989) Decomposition approaches to the identification of change in regional economies, *Economic Geography*, 65, pp. 216–231.

Konijn, P. J. A. & Steenge, A. E. (1995) Compilation of input–output data from the national accounts, *Economic Systems Research*, 7, pp. 31–45.

Kop Jansen, P. & ten Raa, T. (1990) The choice of model in the construction of input–output coefficients matrices, *International Economic Review*, 31, pp. 213–227.

McGilvray, J. W. (1977) *Linkages, Key Sectors, and Development Strategy* (British Association for the Advancement of Science, University of Lancaster: Cambridge University Press).

McMenamin, D. G. & Haring, J. V. (1974) An appraisal of non-survey techniques for estimating regional input–output models, *Journal of Regional Science*, 14, pp. 191–205.

Miller, R. E. & Blair, P. D. (1985) *Input–Output Analysis: Foundations and Extensions* (Englewood Cliffs, NJ, Prentice-Hall).

Morrison, W. I. & Smith, P. (1974) Non-survey input–output techniques at the small area level: an evaluation, *Journal of Regional Science*, 14, pp. 1–14.

O'hUallachain, B. (1984) The identification of industrial complexes, *Annals of the Association of American Geographers*, 74, pp. 420–436.

Oosterhaven, J. (1984) A family of square and rectangular interregional input–output tables and models, *Regional Science and Urban Economics*, 14, pp. 562–582.

Rasmussen, P. N. (1956) *Studies in Intersectoral Relations* (Amsterdam, North-Holland).

Roepke, H., Adams, D. & Wiseman, R. (1974) A new approach to the identification of industrial complexes using input–output data, *Journal of Regional Science*, 14, pp. 15–29.

Sawyer, C. & Miller, R. E. (1983) Experiments in regionalization of a national input–output table, *Environment and Planning A*, 15, pp. 1501–1520.

Stevens, B. H., Treyz, G. I., Ehrlich, D. J. & Bower, J. R. (1983) A new technique for the construction of non-survey regional input–output models and comparison with two survey-based models, *International Regional Science Review*, 8, pp. 271–286.

US Department of Commerce, Bureau of Economic Analysis (1991) *The 1982 Benchmark Input–Output Accounts of the US Economy: Make, Use and Supplementary Tables* (Washington, DC, US GPO).

US Department of Commerce, Bureau of Economic Analysis (1997) The benchmark input–output accounts of the US economy: make, use and supplementary tables, *Survey of Current Business*, 77.

United Nations (1968) *A System of National Accounts* (New York, United Nations).

World Bank (1993) *System of National Accounts 1993* (Washington, DC, World Bank).