The commodity technology revisited

Theoretical basis and an application to error location in the make–use framework

Albert E. Steenge

In this paper we provide conditions for the existence of non-negative homogeneous matrices of dimensions commodity × commodity or industry × industry in the SNA framework. These matrices are shown to satisfy all conditions underlying the commodity technology model. We point out that because the matrices exist even in the case of pure byproducts the identification of their columns as commodity specific limitational production functions is not straightforward and will require additional research on the role of fixed coefficients assumptions. In the second part of the paper we discuss a method to actually calculate these non-negative matrices via a systematic search of possible sources of inaccuracy in the basic make and use tables. The method is applied to aggregated tables for the USA, where a non-negative matrix is generated in a small number of steps. We show that the failure of present day statistical optimization methods may be due to a priori reliability estimates being off by hundreds of percent. Our main conclusion is that search methods attuned to the (labour intensive) way in which most input–output tables are compiled are most promising in error location in this field.

Keywords: Input–output analysis; Commodity technology; Error location

Basic to Leontief's input–output model is the simplifying assumption that each industry produces only one product and that each product is produced by only one industry. This immediately establishes a one–one relationship between industries and products (commodities): industries are characterized by their (single) products, while each product can immediately be associated with a specific industry. Consequently, Leontief's table of interindustry flows is square and 'homogeneous' ie it can be interpreted as being either of a commodity × commodity $(c \times c)$ or industry × industry $(i \times i)$ format. Last but not least Leontief's assumption implies that the mathematics to be used for further analysis of the table (eg for obtaining output multipliers) are readily available and not very complicated.

Statistical implementation of the scheme requires that industrial activity be broken down into a number of categories, the number being determined by the level of detail asked for. Individual establishments (ie individual business units) are assigned to an industry category according to their main or primary product.

The total output of a particular industry is then recorded as the sum of the outputs of all the establishments assigned to that industry. Clearly, the possible presence of secondary production in the individual establishments (ie production of a type considered primary to another industry) is a major complicating factor in implementing this scheme. In the early implementations adherence to the Leontief scheme required substantial reallocations: that is, an establishment's secondary production, together with its associated inputs, was reallocated or redefined to those industries where it was the main product. If the amount of secondary production was considered minor, several other procedures could be followed. Major problems arose if a significant amount of byproducts (products technically related to the main product) were present or if joint production was the case ie if byproducts were produced for which no industry was considered a primary producer. This solution often resulted in unrealistic and highly artificial constructs. (For a discussion of the various types of secondary production being distinguished, see Miller and Blair [7] Chapter 5.)

Rapid industrial diversification during the 1950s and 1960s caused further complications. Joint production had to be recognized as the rule rather than the exception. 'Disentangling' existing processes via a complex system of redefinitions and reallocations became extremely cumbersome and distortive. In fact insistence on the old habit of constructing tables based on monoproduct industries increasingly meant that artificial and arbitrary elements were introduced, often giving misleading results.

To cope with the problem the United Nations (UN), in its System of National Accounts (SNA) [13], proposed a break with the Leontief approach. The new methodology required the implementation of *two* matrices: a use matrix, of dimensions commodity × industry ($c \times i$), and a make matrix of dimensions industry × commodity ($i \times c$); see the next section for exact definitions. The use matrix can be viewed as an extension of the traditional input coefficients matrix and the make matrix as an extension of the output matrix in standard Leontief theory (the unit matrix, by definition). This new framework solved many problems of a descriptive nature. New problems arose, however. Leontief input–output economics derive their significance largely from the fact that output multipliers measuring the combined effects of the direct and indirect repercussions of a change in final demand were readily calculated. It was essential for the calculation of the multipliers that the matrix of input coefficients was always homogeneous. In the SNA framework these multipliers cannot be straightforwardly obtained because the basic make and use tables, being of the

$i \times c$ or $c \times i$ format, are *not* homogeneous. Therefore, calculating multipliers straightforwardly from make and use tables is not possible. This naturally seriously hampered the further use of the new framework, so a solution had to be found.

In its 1968 study [13] the UN discusses several possibilities. Two well known suggestions were to assume a so-called commodity technology (which required commodity specific inputs, in whichever industry the commodity was produced) or an industry technology, which required fixed output shares. Of the various alternatives the commodity technology model is generally acknowledged as a pure and theoretically superior method for the construction of homogeneous input–output tables in the SNA framework (ten Raa and van der Ploeg [12]). Mathematically this model simply postulates the existence of a non-negative matrix connecting the make and use matrices (for details see the next section). The columns of this matrix are to be interpreted as the sought after columns of commodity specific input requirements.

The empirical relevance of the commodity technology model has been tested repeatedly. The tests, however, were unfavourable. The problem was that negative elements invariably kept appearing in the calculated matrices. Because these negatives have no interpretation in an input–output framework, this implied that the commodity technology model did not produce meaningful results.

Attempts have recently been made to give a statistical explanation of the occurrence of the negatives. Statistical methods usually postulate a tolerance interval for each entry. However, approaches based on the acknowledgement of the presence of errors in the make and use tables invariably also resulted in rejection of the commodity technology model as an appropriate description of reality. (For the general methodology and data for the UK see eg Barker *et al* [1]; for recent contributions, see ten Raa *et al* [11], ten Raa and van der Ploeg [12].) In practice, therefore, mixed forms have been introduced regularly to obtain tables with the desired homogeneity (see eg the discussion in Inter-industry Economics Division [5]; see also UNSO [14] [15] and Rainer [8] for other comments). Mixed forms, however, have the drawback of a lack of theoretical underpinning and may therefore obscure the insight into the essential properties and characteristics of the economy.

Because of the importance of the matter, and because in our view the issue has not been dealt with satisfactorily, in this paper we shall concentrate again on the commodity technology model. We shall show that *two* issues are actually involved. The first is the question of under exactly which conditions a non-negative matrix exists, connecting the make and use

matrix. (The present literature does not address this at all.) We shall see that this issue can be settled if the economy satisfies a certain mild condition in its production processes. As we shall see, the basic mathematical result here – based on successive application of a series of theorems of the alternative – has been provided by Mangasarian [6]. (Theorems of the alternative for linear inequalities provide an elegant and direct method for deriving equivalences of the same type as Proposition 1.) Mangasarian proved the result for a general type of rectangular matrix. However, as we shall show, in the case of square make and use matrices (the case dealt with in practice), an approach simpler than Mangasarian's quite complicated one – and much more in line with conditions underlying the building of an input–output model – is available.

The second issue concerns the question of how to look at past experiences with the model. As already mentioned, statistical methods postulating a tolerance interval for each cell in the make and use tables invariably rejected the commodity technology model. A main argument in this paper is to show that the econometric model used in the earlier studies is not appropriate. We shall show that, particularly if certain row or column sums are known with reasonable accuracy, circulating accuracy estimates representing the precision of individual entries in the input and output coefficients matrices may be hundreds of percent wrong. This implies that the commodity technology model can be rejected in practice simply because of incorrect dealing with cell accuracy estimates.

But, given that inaccuracies are present in the make and use tables, how can we locate them? Addressing this question is important because, as we shall see, negatives can often be linked straightforwardly to specific elements in the make or use matrices. Therefore, as an alternative to the methods based on *a priori* estimates of cell accuracies, we suggest a procedure to actually locate possible sources of error, a problem whose solution has been elusive up to now. Steenge [9] discussed the preliminary results of a vice versa inquiry ie research on the effects of changes in certain selected entries in the make–use tables in the elements of the derived homogeneous tables. Because the elements that were investigated in the make–use tables were *a priori* choices (although based on arguments concerning the plausibility of the size of certain entries to do with subsidiary production), a systematic inquiry into the causes of *all negatives* in the derived homogeneous tables was not possible. The present paper seeks to remedy this. The empirical part of the paper ends with an application to the US economy, using the aggregated 14 × 14 tables for the year 1977, as published in Young [16]. Estimates of inaccuracies in *big* entries are shown

to be within the intervals suggested in Barker *et al* [1]; on the other hand, error variances in the smaller entries can be infinite. A discussion of the relevance of our results in the light of by- and joint production is included.

The model

Let us first introduce our accounting framework (Table 1), where we employ the now familiar notation (UN [13]). In this table, U stands for the (square) commodity by industry use or absorption matrix, V for the (square) industry by commodity make matrix, e for the vector of final demands and y' for the (row) vector of primary inputs; the prime denotes vector transposition. Thus q is the vector of commodities being produced, and g the vector of industrial outputs. With i being the one vector, we have successively for each row and column:

Row-wise

$$\begin{cases} Ui + e = q \\ \quad Vi = g \end{cases} \tag{1}$$

Column-wise

$$\begin{cases} \quad i'V = q' \\ i'U + y' = g' \end{cases} \tag{2}$$

In the redefinition techniques that we shall discuss, the following assumptions regarding the existence of constant coefficient matrices play a role. First we have:

$$U = B\hat{g} \tag{3}$$

where B is a matrix of constant use or absorption coefficients. Furthermore, we encounter:

$$V' = C\hat{g} \tag{4}$$

where C is a matrix of fixed output coefficients.

The UN's *commodity–technology* model states that each commodity is characterized by its own – non-negative – input coefficients. This immediately leads to an input–output matrix, either in commodity × commodity $(c \times c)$ or industry × industry format $(i \times i)$, see UN [13] and Stone [10]. Denoting the $c \times c$ input–output matrix by M^c, we have[1]

$$M^c \equiv BC^{-1} = U(V')^{-1} \tag{5}$$

(From $BC^{-1} = U(\hat{g})^{-1}[V'(\hat{g})^{-1}]^{-1} = U(\hat{g})^{-1}(\hat{g})(V')^{-1}$ $= U(V')^{-1}$.) Denoting the $i \times i$ variant M^i, we have

$$M^i \equiv C^{-1}B = (\hat{g})(V')^{-1}U(\hat{g})^{-1} \tag{6}$$

Table 1. Accounting framework.

	Commodities	Industries	Final outputs	Totals
Commodities		U	e	q
Industries	V			g
Primary inputs		y'		
Totals	q'	g'		

The commodity technology model

From the above we can see that if the commodity model were valid, matrices B and C would be connected by a non-negative matrix of input proportions. However, as already mentioned, in practice problems arise: calculations using empirical material invariably reveal negative elements in the matrices BC^{-1} or $C^{-1}B$, usually in the order of a few percent of the total number of coefficients. Partly because of the lack of generally accepted techniques to eliminate or to circumvent the problem of the negatives many countries have turned to the industry technology approach (although application of this method requires fixed market shares), or have adopted one of many alternative mixed forms (UNSO [15]), thereby sacrificing transparency.

Recently a number of studies have appeared exploring a statistical approach to the problem of the negatives (ten Raa *et al* [11], and ten Raa and van der Ploeg [12]). However, despite various model formulations, negatives were persistent. For example, ten Raa and van der Ploeg employed 9 × 9 make and use tables for the United Kingdom. The observed data u_{ij}^0 and v_{jk}^0 are considered as the sum of true values u_{ij} and v_{jk} and error terms δ_{ij} and e_{jk}:

$$\begin{cases} u_{ij}^0 = u_{ij} + \delta_{ij} \\ v_{jk}^0 = v_{jk} + e_{jk} \end{cases} \tag{7}$$

[1] The UN's *industry–technology* model states that the inputs into commodities are in proportion to the value of these commodities in the total output of their industry. Again we obtain two matrices, one in commodity × commodity format (*S*), and the other in industry × industry format (*T*). For the first type we have:

$S = BD$

and for the second:

$T = DB$

Mixed forms often involve an eclectic choice of elements of both methods. For further details see eg Stone [10] or Miller and Blair [7].

The error terms were assumed to be normally distributed with zero means and known standard deviations σ_{ij} and τ_{jk}. Estimates of the standard deviations were based on subjective information obtained using 1975 statistics. The problem addressed is to minimize

$$f(U, V) \text{ subject to } U(V')^{-1} \geqslant 0 \tag{8}$$

where

$$f(U, V) = \sum_{i,j} \sigma_{ij}^{-2}(u_{ij} - u_{ij}^0)^2 + \sum_{j,k} \tau_{jk}^{-2}(v_{jk} - v_{jk}^0)^2 \tag{9}$$

The study shows that, in this formulation, the commodity technology model must be rejected at the 5% level. We should realize, of course, that if the accuracy estimates had been based on a different principle, the result might have been quite different. We come back to this later.

The existence of a third, non-negative matrix

The first issue we mentioned – the existence of a non-negative matrix connecting B and C – can be given a theoretical basis if the economy possesses certain elementary characteristics. To see this, let us first consider a traditional single product input–output model with non-singular coefficient matrix M. Now for any output vector

$$x > 0 \tag{10}$$

we have

$$Mx > 0 \tag{11}$$

In fact, the specification of M is such that always

$$x > 0 \Rightarrow Mx > 0 \tag{12}$$

That is, non-negative outputs are guaranteed only if non-negative inputs are available. Analogously, a

non-negative price vector implies non-negative input costs:

$$p > 0 \Rightarrow pM > 0 \qquad (13)$$

We may look upon (12) and (13) as fundamental properties of single product models, justifying separate attention.

In the single product case x is both the vector of sectoral intensities and the total output vector. In joint production models this equivalence is lost and we have to make a distinction between a vector of commodities, a vector of industrial outputs and a vector of production levels. Given the matrices of input and output coefficients introduced above, we may wish to consider joint production variants such as

$$Cx > 0 \Rightarrow Bx > 0 \qquad (14)$$

Note that x need not be non-negative here. To see this, let x_1 and x_2 stand for two vectors of production levels corresponding, say, to two different demand bundles. The corresponding output vectors are Cx_1 and Cx_2. Then $C(x_1 - x_2)$ is the vector of production increments. Writing $x = x_1 - x_2$, we obtain the expression appearing in (14).

Mangasarian [6] presented a number of results of the following kind. Suppose B and C are rectangular, real $m \times n$ and $k \times n$ matrices, then (our notation):

Proposition 1 (Mangasarian):

$$(Cx \geqq \Rightarrow Bx \geqq 0) \langle = \rangle$$

(There exists an $M^c \geqq 0$: $B = M^c C$), with M^c a real $m \times k$ matrix.)

Interest in these and related results (Proposition 1 is part of a more general theorem) is increasing. This is true for theoretical contributions in price and distribution theory (see eg Fujimoto and Krause [4]). However, Mangasarian's theorem may also be extremely helpful in empirical work as it may provide a basis for the commodity technology assumption. Unfortunately, although the backward implication is trivial to prove, proving the full theorem is quite complicated. The forward implication, for example, requires the successive application of several theorems of the alternative.

Because the result may be important for other applications as well, we shall show that in our case a quite different and much simpler approach may be followed. In our input–output model matrices B and C are square by construction. If, in addition, we assume that matrix C is non-singular (a condition which will normally be satisfied), the matrix equation $B = M^c C$

will be consistent. (Consistency is not guaranteed if B or C are rectangular.) Then we have:

Proposition 2:

Let B, C and M^c be $n \times n$ matrices such that $B = M^c C$, and let C have rank n. Then $(Cx \geqq 0 \Rightarrow Bx \geqq 0) \langle = \rangle$ $M^c \geqq 0$.

Proof:

"$\Rightarrow$": $(Cx \geqq 0 \Rightarrow Bx \geqq 0)$

$\Rightarrow (Cx \geqq 0 \Rightarrow M^c Cx \geqq 0)$

$\Rightarrow (y \geqq 0 \Rightarrow M^c y \geqq 0)$ (with $y \equiv Cx$)

$\Rightarrow M^c \geqq 0$

"$\Leftarrow$": $Cx \geqq 0 \Rightarrow M^c Cx \geqq 0 \Rightarrow Bx \geqq 0$

Returning to the SNA system, if (our) matrices B and C are square and if C has full rank, M^c is easily seen to exist.

It is worthwhile pausing for a moment to consider the implications of the above. The above propositions tell us that if B and C have the property that for any x we have $Cx \geqq 0 \Rightarrow Bx \geqq 0$, a column of a non-negative matrix M^c can be associated with each commodity. (The exact numerical configuration of (the columns of) M^c of course depends on the coefficients pattern of matrices B and C.) The columns can straightforwardly be interpreted as columns of commodity specific bundles of (imputed) input proportions. However, they do not necessarily represent *technologically interpretable* commodity specific limitational production functions, as in the standard Leontief model. Clearly if the set of commodities contains only primary and secondary products (products that are the main product of some other industry) we may expect that in a number of cases, perhaps even in most cases, they will allow an interpretation in terms of production functions. However, before any verdict can be given, much more testing by sector and industry specialists on empirical material is required to see if the columns of any particular matrix M^c can indeed be identified as production functions.

In the case of pure by- or joint production the interpretation of the columns of M^c is more complicated. Interpreting them as (technologically determined) production functions makes no sense here, because true primary processes are – by definition – absent. Nevertheless, they do represent imputed commodity specific input proportions, consistent with the overall make–use framework. Even though no production function interpretation is readily present the availability

of knowledge on such proportions is extremely useful in terms of price analysis, cost accounting motivations etc.

Above all it seems that we have to obtain more insight into the properties of the coefficients of M^c over time. The reasons for this are straightforward. For example, we know that in practice the subsidiary product is often produced only if some specific goal regarding the main product has been reached (or, contrarily, turns out to be out of reach). A change in emphasis on the subsidiary product may then result in a change in M^c as well. It can also happen that profits made on the subsidiary product are used to temporarily subsidize the main product, which otherwise would not be produced at all or only at a severe loss to the industry. This again may imply a changing M^c if a firm's financial policies are changed. Clearly if matrices B and C do not change, the same M^c matrices would be found each time. On the other hand, changes in the make and use tables and in the (derived) coefficients matrices B and C do not necessarily imply that M^c will change.

In any case research on these issues will require much further effort based on time series of input–output tables which are not yet available. At present it seems advisable to recognize that the columns of M^c, as well as typical commodity and industry specific characteristics, may also reflect characteristics of the entire economy. Similar remarks are valid concerning matrix M^i, *mutatis mutandis*.

The location of errors

It is well known that a substantial degree of arbitrariness is present in the balancing process of the commodity flow system which underlies all further exercises. We may think of the use of certain redistribution techniques, such as the RAS or Lagrange methods, problems in distinguishing stock or flow characteristics, the use of inter- or extrapolation techniques, errors in sampling and sampling methods, incorrect homogeneity assumptions, aggregation problems etc. The nature of many of the fundamental issues of classification that are involved is illustrated by the following remarks by Robert Eisner ([3], p 1614):

The identification of household and government purchases with final product, and goods and services bought by business as intermediate product, unless they are deemed 'capital' or accumulated in inventories, raises some critical questions and indeed a number of anomalies relating to the sometimes shifting identities of purchasers. Thus police services purchased by government are final product and included in GNP while the services of guards or watchmen purchased by a business are not; they are presumably resold as part of the business production in which they are

instrumental. The repair of state-owned highways is final product, while repair of private railroad tracks is intermediate. If the highways were private tollways and the railroads were nationalized, the classifications would be reversed. A great part of government activity is in fact clearly devoted to direct service to business as well as maintenance of the system and infrastructure essential to private production.

The amusement or entertainment services of movies watched in the theaters, on cable television sold by subscription, or on public (government) television are generally included in GNP. Similar services provided by regular commercial television in the United States are not included because they are paid for by purchases on current account by one business firm, in the form of advertising expenses from another.

But what about travel expenses in connection with work, including the commuting costs of getting from home to job? If these are paid directly by a business firm there will be intermediate product, included in the value of the 'final' output of the firm. If firms pay their employees higher wages and let them pay their own travel costs, those expenditures will be counted as personal consumption and their value will enter independently into gross national product in addition to the value of whatever output the firm is producing.

Conventions such as the above can easily lead to serious misallocation, including clear errors such as double counting. Another source of inaccuracy lies in the nature of the commodities being measured. For example, Dutch practice, in line with experience elsewhere, has shown that many column totals can be estimated quite accurately independently. That is, preliminary estimates rarely have to undergo substantial change to obtain overall consistency in the tables. (The popularity of the RAS method, for example, depends to a large extent on this stability phenomenon.) Furthermore it appears that among the various input categories the use of raw materials and certain semifinished goods, according to the same criterion, is normally estimated quite accurately. The same is true for many entries connected with typical manufacturing activities. On the other hand reliable estimates of other categories of inputs are sometimes much harder to obtain. This is especially true for those inputs that characterize an information economy, such as banking, advertising or administrative services. In practice, to obtain coefficients for these inputs, a number of keys often have to be used in the allocation process. This, of course, means that the quality of the obtained coefficients depends to a large extent on the quality of the keys that are being used. Furthermore, depending on the *type* of key used, estimates of the magnitudes of the entries on the affected input categories may fluctuate significantly. It then depends on the specifics of each particular input whether the direction of the fluctuation is systematically one sided or two sided or otherwise. (Compare here also the argumentation in Carter [2] on measurement issues in the information

sectors.) Another potential source of inaccuracy, naturally, is related to the source providing the raw data (see below for the way US input–output data are obtained).

The above illustrates – given the way the tables are compiled in practice – that the *a priori* assignment of error margins to individual cell entries can be hazardous. Yet, in a number of studies, some of which we already referred to, such *a priori* estimates play a dominant role testing the validity of the various technology assumptions. For example, in ten Raa and van der Ploeg's [12] maximum likelihood approach, cell specific accuracy estimates are used ranging from a few percent for most activities to more than 20% for a number of entries involving services. Employing exogenously determined information regarding cell reliabilities in such a way means that we are relatively sure that we have located the error sources, and that we have some idea about their magnitudes. However, if we adopt this kind of approach various kinds of systematic sources of correlations between cell accuracies are not accounted for. It is therefore a matter of discussion whether such approaches, based on normality, are generally consistent with the way the tables have been put together. Consider, for example, the above mentioned well known fact that column or row totals are often given with a much higher degree of reliability than the entries in the cells. Assuming for simplicity that certain column totals in the use or make table are known correctly, a relatively small inaccuracy of, say, 5% in a large entry in a particular column may result in errors of *several hundred percent* in small entries in those columns. If such correlations are present it may be advisable, before using exogenous data on accuracies, to spend some time on the question of how to proceed if error margins can vary widely in such a way. Clearly the fact that cell variances may in fact range from zero up to infinity dismisses the maximum likelihood approach as a candidate for testing the relevance of the commodity technology model.

A systematic way to proceed, however, is to calculate the matrices

$$X_{ij} = \frac{\delta M^c_{ij}}{\delta C} = \frac{\delta (BC^{-1})_{ij}}{\delta C} = -(C^{-1})' B' E_{ij} (C^{-1})'$$

$$Y_{ij} = \frac{\delta M^c}{\delta C_{ij}} = \frac{\delta BC^{-1}}{\delta C_{ij}} = -BC^{-1} E_{ij} C^{-1}$$

and

$$Z_{ij} = \frac{\delta M^c_{ij}}{\delta B} = E_{ij} (C^{-1})'$$

where E_{ij} is an elementary matrix which has unity in its (i, j)th position while all other elements are zero (see below for proofs). If certain elements of, say, X_{ij} are large, this means that small errors in the corresponding elements of C will have a large effect on M^c_{ij}. It may therefore be worthwhile to investigate the accuracies of these elements in particular. Generally, this will imply a thorough reconsideration of the way these elements have been composed out of elementary data. The above matrices naturally only locate a potential source of a specific error. Sector specialists will have to give us insight into the way in which a specific entry has been obtained. This should then give us further indications regarding the reliability of the published figures. Above we discussed the commodity × commodity variant. A similar argument can, however, be put forward regarding the industry × industry variant.

Proof

The first and second formulae are related via a simple transformation. The proof of the second formula is as follows. Starting from

$$M^c (M^c)^{-1} = I$$

the product rule gives

$$M^c \frac{\delta (M^c)^{-1}}{\delta C_{ij}} + \frac{\delta M^c}{\delta C_{ij}} (M^c)^{-1} = 0$$

or

$$\frac{\delta M^c}{\delta C_{ij}} = -M^c \frac{\delta (M^c)^{-1}}{\delta C_{ij}} M^c$$

But

$$\frac{\delta (M^c)^{-1}}{\delta C_{ij}} = \frac{\delta CB^{-1}}{\delta C_{ij}} = E_{ij} B^{-1}$$

so

$$\frac{\delta M^c}{\delta C_{ij}} = \frac{\delta BC^{-1}}{\delta C_{ij}} = -BC^{-1} E_{ij} B^{-1} BC^{-1}$$

$$= -BC^{-1} E_{ij} C^{-1}$$

The first formula is obtained via a simple transformation:

$$\frac{\delta M^c_{ij}}{\delta C} = -(C^{-1})' B' E_{ij} (C^{-1})$$

The third formula is straightforward.

Table 2. USA (1977), matrix of input coefficients B

0.2460	0.0003	0.0026	0.0460	0.0001	0.0002	0.0016	0.0002	0.0054	0.0221	0.0027	0.0006	0.0000	0.0000
0.0015	0.0718	0.0076	0.0563	0.0005	0.1818	0.0000	0.0000	0.0001	0.0000	0.0001	0.0282	0.0000	0.0000
0.0108	0.0372	0.0011	0.0064	0.0352	0.0350	0.0068	0.0023	0.0721	0.0058	0.0136	0.1174	0.0000	0.0000
0.2012	0.0936	0.3731	0.3865	0.0944	0.0938	0.0377	0.0248	0.0054	0.3279	0.1083	0.0751	0.0000	0.0000
0.0177	0.0077	0.0261	0.0266	0.1013	0.0227	0.0323	0.0271	0.0025	0.0209	0.0280	0.0446	0.0000	0.0000
0.0123	0.0244	0.0030	0.0186	0.0085	0.2027	0.0195	0.0101	0.0072	0.0209	0.0170	0.0822	0.0000	0.0000
0.0432	0.0154	0.0885	0.0424	0.0155	0.0104	0.0130	0.0031	0.0025	0.0581	0.0229	0.0141	0.0000	0.0000
0.0139	0.0077	0.0087	0.0058	0.0123	0.0076	0.0146	0.1960	0.0254	0.0116	0.0100	0.0047	0.0000	0.0000
0.0547	0.0654	0.0023	0.0060	0.0133	0.0047	0.0388	0.0225	0.0645	0.0302	0.0409	0.0116	0.0000	0.0000
0.0008	0.0038	0.0008	0.0050	0.0069	0.0009	0.0169	0.0093	0.0025	0.0023	0.0105	0.0047	0.0000	0.0000
0.0200	0.0256	0.0632	0.0331	0.0565	0.0133	0.0978	0.0790	0.0203	0.0558	0.0888	0.0305	0.0000	0.0000
0.0008	0.0013	0.0008	0.0024	0.0027	0.0028	0.0049	0.0170	0.0022	0.0023	0.0056	0.0094	0.0000	0.0000
0.0002	0.0013	0.0001	0.0058	0.0203	0.0002	0.0010	0.0023	0.0001	0.0003	0.0010	0.0117	0.0000	0.0000
0.0000	0.0003	0.0001	0.0034	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0006	0.0000	0.0000

Table 3. USA (1977), matrix of output coefficients C.

0.9699	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.9141	0.0000	0.0003	0.0000	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0285	0.0564	0.0000	0.9841	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0023	0.0000	0.0000
0.0002	0.0000	0.0000	0.0000	0.9504	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0798	0.0000	0.0000
0.0000	0.0295	0.0000	0.0000	0.0064	0.9962	0.0000	0.0000	0.0000	0.0000	0.0000	0.3192	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0399	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9954	0.0000	0.0000	0.0000	0.0023	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0027	0.0516	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0423	0.0000	0.0000
0.0008	0.0000	0.0000	0.0139	0.0421	0.0000	0.0000	0.0054	0.0000	0.0000	0.9973	0.0117	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0009	0.0000	0.0000	0.0000	0.0000	0.0000	0.4484	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
0.0000	0.0000	0.0000	0.0016	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Application

We have seen that if (14) is accepted, the presence of negatives in M^c should rather be taken as indicative of the quality of the make and use tables. Of course, if we have confidence in the statistical office that built the tables, we should expect the number of serious errors to be small.

As an example of the error location method we have described, we shall present an application using the 14 × 14 summary version of the US make and use tables for 1977 (Young [16]). The following industries were distinguished: (i) agriculture, forestry and fishing; (ii) mining; (iii) construction; (iv) manufacturing; (v) transportation and communication; (vi) electric, gas and water services; (vii) wholesale and retail trade; (viii) finance and insurance; (ix) real estate; (x) eating and drinking places; (xi) services; (xii) government enterprises; (xiii) non-comparable imports; and (xiv) scrap and special industries. The 14 distinguished commodities bear the same name.

Tables 2 and 3 give matrices B and C as calculated from the make and use matrices. (Values recorded as less than $0.05 billion have been put at $0.025 billion.)

Table 4 contains matrix $M^c = BC^{-1}$. (The zeros in columns 13 and 14 need not concern us here.) As we see, a total of 11 elements are negative. By far the largest negative element is $M^c_{2,12} = -0.0612$. Looking at Table 3, we see that the secondary product (electric, gas and water services) of industry 12 is large compared to the primary product (government enterprises). Because mining is a large input in the industry producing electric etc services, the large size of the entry for this particular jointly produced good could be indicative of a potential source of error. Matrix $X_{2,12} = \delta M^c_{2,12}/\delta C$ is reproduced in Table 5. Not unexpectedly, we see that the elements with large influence are found in column 12. Thus if $C_{6,12}$ decreased by 0.1, $M^c_{2,12}$ would increase by about 0.041, *ceteris paribus*. Similar calculations, naturally, can be made regarding the other goods jointly produced by industry 12. Matrix $Y_{6,12} = \delta M^c/\delta C_{6,12}$ is reproduced as Table 6. Again we see that errors in $C_{6,12}$ have a substantial influence on the 12th column of M^c. (Matrix $Z_{2,12}$ is straightforwardly calculated, and is therefore not reproduced.)

From Table 7 we see that errors in the input coefficients matrix B have a large influence as well, the

Table 4. USA (1977), matrix M^r.

0.2522	-0.0025	0.0026	0.0467	0.0000	0.0002	0.0016	0.0002	0.0054	0.0221	0.0027	-0.0020	0.0000	0.0000
-0.0001	0.0691	0.0076	0.0572	-0.0007	0.1824	0.0000	0.0000	0.0001	0.0000	0.0001	-0.0672	0.0000	0.0000
0.0109	0.0392	0.0011	0.0063	0.0362	0.0349	0.0068	0.0023	0.0721	0.0058	0.0135	0.2206	0.0000	0.0000
0.1960	0.0755	0.3731	0.3871	0.0939	0.0933	0.0377	0.0243	0.0054	0.3279	0.1086	0.0445	0.0000	0.0000
0.0175	0.0060	0.0261	0.0267	0.1052	0.0227	0.0323	0.0271	0.0025	0.0209	0.0281	0.0584	0.0000	0.0000
0.0122	0.0189	0.0030	0.0186	0.0069	0.2033	0.0195	0.0100	0.0072	0.0209	0.0171	0.0321	0.0000	0.0000
0.0432	0.0139	0.0885	0.0428	0.0152	0.0103	0.0130	0.0030	0.0025	0.0581	0.0229	0.0316	0.0000	0.0000
0.0141	0.0078	0.0087	0.0058	0.0124	0.0076	0.0146	0.1968	0.0254	0.0116	0.0099	-0.0038	0.0000	0.0000
0.0562	0.0710	0.0023	0.0055	0.0122	0.0047	0.0388	0.0223	0.0645	0.0302	0.0408	0.0057	0.0000	0.0000
0.0006	0.0039	0.0008	0.0049	0.0068	0.0009	0.0169	0.0093	0.0025	0.0023	0.0105	0.0062	0.0000	0.0000
0.0196	0.0256	0.0632	0.0324	0.0554	0.0132	0.0978	0.0789	0.0203	0.0558	0.0890	0.0296	0.0000	0.0000
0.0007	0.0012	0.0008	0.0023	0.0025	0.0028	0.0049	0.0171	0.0022	0.0023	0.0056	0.0173	0.0000	0.0000
0.0000	0.0010	0.0001	0.0059	0.0213	0.0002	0.0010	0.0023	0.0001	0.0003	0.0010	0.0220	0.0000	0.0000
-0.0001	0.0001	0.0001	0.0034	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0013	0.0000	0.0000

Table 5. USA (1977), matrix $X_{2,12}$.

0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000
-0.0001	0.0000	0.0003	0.0130	0.0494	0.0062	0.0080	0.0004	0.0065	0.0013	-0.1543	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0014	0.0054	0.0007	0.0009	0.0007	0.0007	0.0001	-0.0169	0.0000	0.0000	0.0000
0.0000	0.0000	0.0002	0.0107	0.0408	0.0051	0.0066	0.0051	0.0054	0.0010	-0.1277	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	-0.0001	-0.0005	-0.0001	-0.0001	-0.0001	-0.0001	0.0000	0.0015	0.0000	0.0000	0.0000
-0.0001	0.0000	0.0007	0.0342	0.1302	0.0162	0.0210	0.0162	0.0172	0.0033	-0.4071	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0002	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0002	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000
0.0000	0.0000	-0.0003	-0.0126	-0.0480	-0.0060	-0.0077	-0.0004	-0.0063	-0.0012	0.1501	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0000	0.0000	0.0000	0.0000

Table 6. USA (1977), matrix $Y_{6,12}$.

0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0003
0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.4071
0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0779
0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.2081
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0507
0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.4538
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0231
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0169
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0104
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0021
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0294
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0063
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0005
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 7. USA (1977), matrix C^{-1}.

1.0310	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.0000	1.0940	0.0000	-0.0003	0.0000	-0.0010	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007
0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-0.0299	-0.0627	0.0000	1.0162	0.0000	-0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0040
-0.0002	0.0000	0.0000	0.0000	1.0522	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.1874
0.0000	-0.0324	0.0000	0.0000	-0.0068	1.0045	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.7141
0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0891
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0053
0.0000	0.0000	0.0000	0.0000	0.0001	0.0001	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	-0.1152
0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	1.0000	0.0000	-0.0027	0.0000	-0.0943
-0.0004	0.0009	0.0000	-0.0142	-0.0445	0.0000	-0.0055	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	-0.0183
0.0000	0.0001	0.0000	0.0000	0.0000	-0.0021	0.0000	0.0000	0.0000	0.0000	0.0000	1.0027	0.0000	2.2319
0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
0.0000	0.0001	0.0000	-0.0016	-0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 8. Changes in selected elements of the original make and use matrices required to generate a non-negative matrix M^c.

(1) Coefficient in original coefficient matrices	(2) Absolute value	(3) Calculated absolute change	(4) Required percentage change in (2)	(5) Size of corresponding element in original make or use table (in billions of US$)
$B^0_{2,1}$	0.0015	0.0001	7	0.2
$B^0_{14,1}$	0.0000	0.0001	∞	0.0
$B^0_{1,2}$	0.0003	0.0027	900	0.025[a]
$B^0_{2,5}$	0.0005	0.0008	160	0.1
$B^0_{1,12}$	0.0006	0.0013	217	0.025[a]
$B^0_{2,12}$	0.0282	0.0168	60	1.2
$B^0_{8,12}$	0.0047	0.0013	28	0.2
$C^0_{6,12}$	0.3192	−0.0750	−23	0.1
$C^0_{12,12}$	0.4484	0.0750	17	19.1

[a] The recorded value in the original use table was given as 'less than 0.05 billion'.

diagonal elements of C^{-1} normally being larger than one. To show that only minor changes in the make and use tables may be required to obtain a non-negative matrix M^c, we have calculated, using as a guide the matrices X_{ij}, Y_{ij} and Z_{ij}, calculated for various appropriate i and j, a set of coefficient changes in B or C that would be sufficient to obtain a non-negative corresponding matrix M^c (see Table 8).

First of all we observe that only a very limited number of entries needs correction. Furthermore, we see that $C^0_{12,12}$ is the only large entry that would require substantial (absolute) correction. (This correction, however, appears well within the margins for this kind of product as given by ten Raa and van der Ploeg ([12], Table 4), who refer to earlier work by Barker *et al* [1]. They were within the range of inaccuracy that seems appropriate for the US data; see also further below.) The other entries referred to in Table 8 are all small, for which the required *percentage* change may be very large. For example, looking at the elements of the first column of the use matrix, we see that misallocations in this column (which may be found just as well in the value-added sphere), accounting for a share of merely 0.0002 (= 0.0001 + 0.0001) of the input value in this input column, may imply a (small) error of about 7% in $B^0_{2,1}$ and an infinitely large error in $B^0_{14,1}$.

In view of the way in which the basic data for the US make and use tables are assembled, these outcomes do not seem particularly large. (The tables are compiled using data collected by the Census Bureau. Unfortunately, the census procedures and classification schemes do not always concide with input–output convention. For example, regarding the differences between related activities such as construction, reconstruction and renewal (of houses, plant and so on), large differences can exist. The same is true for the classification of many services. This means that in the process of translation of the census data into input–output data

a multitude of keys has to be used. This also means that systematic errors of the type we have described may be a regular phenomenon and that especially small entries may be very far off percentage-wise.) Of course, before any definite verdict can be given, industry and sector specialists will have to verify every suggested correction. Tables 5, 6 and 7 may suggest that in practice, due to the sparseness of B and, especially, C, the number of realistic options is limited.

Further comments and conclusion

In this paper we have focused on the problem of the negatives in joint production input–output analysis. We have pointed out that under a mild condition on the economy's production characteristics, the make and use matrices are indeed connected by a third non-negative matrix.

Clearly, if the basic matrices B and C were error free, (14), or equivalent implications, could be verified straightforwardly by calculating M^c, provided B were non-singular. However, in practice, (14) cannot be verified in this way, as the basic tables may contain errors. We have shown that by focusing on a particular negative, the entries in the make and use tables can be divided into those that may be influential in explaining the negative and those that are not. The calculated derivatives provide us with a preliminary estimate of the potential error. The next step will then be to see if the implied correction is acceptable for the sector and industry specialists. The example we have presented (using the 14 × 14 tables for the USA for 1977) illustrates that only very few corrections may be required to obtain (corrected) make and use matrices such that the implied third matrix does not contain any negatives. However, further research on much larger tables is required here before a definite verdict can be reached.

Given the present status of the debate it seems likely that any future statistical optimization models (if that line of research is pursued) will have to incorporate inter- (or, possibly, even intrasectoral) correlations. This will mean additional quality requirements concerning the basic data, because, particularly if row or column sums are known with reasonable accuracy, present day circulating accuracy estimates representing the precision of individual entries in the input and output coefficients matrices, may be hundreds of percent wrong. In fact, accuracy estimates are dependent on the way the tables have been compiled. It may therefore be good practice to obtain insights into the impact of variations in any particular element by systematically investigating the magnitudes of the corresponding derivatives. The final assessment will have to be made by sector specialists, especially if systematic problems of classification and methodology are present.

References

1 T. Barker, F. van der Ploeg and M. Weale, 'A balanced system of national accounts for the United Kingdom', *Review of Income and Wealth*, Vol 30, 1984, pp 461–485.

2 A. Carter, 'Input–output recipes in an information economy', *Economic Systems Research*, Vol 1, 1989, pp 27–43.

3 R. Eisner, 'Extended accounts for national income and product', *Journal of Economic Literature*, Vol 24, 1988, pp 1611–1684.

4 T. Fujimoto and U. Krause, 'More theorems on joint production', *Journal of Economics*, Vol 48, 1988, pp 189–196.

5 Interindustry Economics Division, 'The input–output structure of the US economy, 1977', *Survey of Current Business*, May 1984, pp 42–84.

6 O.L. Mangasarian, 'Perron–Frobenius properties of $Ax = \lambda Bx$', *Journal of Mathematical Analysis and Applications*, Vol 36, 1971, pp 86–102.

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

8 N. Rainer, 'Descriptive versus analytical make–use systems: some Austrian experience', Paper presented at the Eighth International Conference on Input-Output Techniques, Sapporo, Japan, 28 July–2 August 1986.

9 A.E. Steenge, 'Second thoughts on the commodity technology and the industry technology approaches', in A. Franz and N. Rainer, eds, *Compilation of Input–Output Data*, Orac, Vienna, 1989.

10 R. Stone, 'Accounting matrices in economics and demography', in *Mathematical Models in Economics*, Wiley, Chichester, 1984.

11 Th. ten Raa, D. Chakraborty and J.A. Small, 'An alternative treatment of secondary products in input–output analysis', *Review of Economics and Statistics*, Vol 66, 1984, pp 88–97.

12 Th. ten Raa and R. van der Ploeg, 'A statistical approach to the problem of negatives in input–output analysis', *Economic Modelling*, Vol 6, 1989, pp 2–19.

13 United Nations, *A System of National Accounts: Studies in Methods*, Series F, No 2, New York, 1968.

14 United Nations Statistical Office, 'Input–output standards in the SNA framework', in A. Franz and N. Rainer, eds, *Problems of Compilation of Input–Output Tables*, Orac, Vienna, 1986.

15 United Nations Statistical Office, *Study of the Country Practices in Implementing the SNA Input–Output Framework in the 70s*, Report to the Working Party on National Accounts and Balances, 3–7 March 1986, CESWP.22/82/Add.1.

16 P.C. Young, 'The US input-output experience – present status and future prospects', in A. Franz and N. Rainer, eds, *Problems of Compilation of Input-Output Tables*, Orac, Vienna, 1986.