Chapter 13

Disaggregation of Input-Output Tables

Robert A. McDougall

13.1 Overview

In this chapter we document the disaggregation of the input-output (*I-O*) tables. This is the major step undertaken in our in-house processing of the tables (for an overview of the processing, see chapter 11). It takes place after some initial clean-ups and adjustments, and before the synthesis of the composite region tables (chapter 14), further minor adjustments, and the fitting of the I-O tables to external data (chapter 19).

While we use a 57-sector classification for the GTAP data base, many countries' I-O source statistics do not support such detail. In particular, few I-O sources match the relatively ambitious requirements for agriculture and food processing, where GTAP distinguishes 20 sectors, while some sources distinguish as few as two (agriculture and food processing separately). Accordingly, we accept contributed tables using any reasonable aggregation of this classification, (Huff, McDougall, and Walmsley, 1999). Where the contributed table uses an aggregated classification, we must disaggregate it to the standard 57 sectors in-house.

Altogether, in GTAP 5, 24 regions require sectoral disaggregation. Of these, seven require disaggregation only in non-agricultural sectors, while 17 require disaggregation in agriculture. For non-agricultural sectors we use a disaggregation procedure that is relatively undemanding in data, basing the I-O structure of the disaggregate industries on a *representative table* which represents a kind of average of the I-O structures for the regions where we have fully disaggregated contributed tables (chapter 14). Obviously this is less than ideal, but to the extent that industry structures reflect patterns of consumption and technology common across countries, we may hope that it generates reasonable results. For agriculture however, geographic factors such as climate and soil influence the industry structure heavily; here we use country-specific data, namely the agricultural I-O data compiled by Everett Peterson (chapter 13). This comprises a collection of country I-O tables using a sectoral classification heavily oriented toward agriculture.

More precisely, we use the agricultural I-O data set for 16 of the 17 tables requiring agricultural disaggregation. For Sri Lanka, however, we do not use the agricultural I-O data, but instead use other estimates of agricultural imports, production, and exports. This is not because of any gaps in the agricultural I-O data set, but for technical reasons related to the Sri Lankan aggregation mapping. This differs from the mappings for other regions, in combining some agricultural sectors together with non-agricultural sectors in a single aggregate sector. Our current main line procedure cannot handle this situation. In principle we could revise it to handle it, but in the interests of expediency we simply used an old procedure from GTAP 4. The old procedure was designed to exploit import, production, and export data for regions lacking agricultural I-O data.

The major inputs into the procedure are the contributed I-O tables, the agricultural I-O data set, and a *representative* I-O table. The contributed tables are not in their original state, but have already undergone initial adjustments. The representative I-O table is an average I-O structure based on the regions whose source tables are fully disaggregated (see chapter 14 for further details). We use it in disaggregating non-agricultural sectors (section 13.4). We perform the agricultural disaggregation before the non-agricultural.

13.2 Agricultural Disaggregation using Agricultural I-O Data

In this section, we document the disaggregation of the agricultural sectors in the contributed tables, using outside information on the I-O structure of the agricultural sector. We apply this procedure to sixteen regions, listed in table 13.1 below. At the end of this stage, the contributed I-O tables are fully disaggregated in agriculture, but not in their original classification outside agriculture. Briefly, for each region, we:

- Make small adjustments to the agricultural I-O table, designed to replace zeros with small non-zero values in cases where the corresponding value in the contributed I-O table is non-zero. Such zeros would otherwise render the next step infeasible.
- Rebalance the agricultural table to make it consistent with the contributed table, using a scaling factor approach.

— Mutually disaggregate the rebalanced agricultural table and the contributed table, by prorating.
— Clean up any minor problems in the disaggregated I-O table.

At the beginning of this procedure we have two I-O tables: the contributed I-O table and an agricultural I-O table. The contributed table is somewhat aggregate in agriculture, and perhaps also outside agriculture, but typically contains many non-agricultural sectors. The agricultural table is drawn from the agricultural I-O data set; it is fully disaggregated in agriculture, but highly aggregated outside agriculture.

To bring these two tables into contact with each other, we use a sectoral classification that constitutes the *finest common aggregation* (*FCA*) for the two initial I-O tables. This is a *common* aggregation, because it is an aggregation of both the sectoral classifications used in the contributed I-O table and in the agricultural I-O table; and it is the *finest* (that is, the most detailed) such classification. In this context, the FCA classification matches the contributed table's sectoral classification in agriculture, and the agricultural I-O classification outside agriculture

Following an entropy-theoretic approach, the rebalancing is undertaken using scaling factors. We define a *block scaling factor* for each *block* of cells in the agricultural table corresponding to a single cell in the FCA. The block may be 1×1, 1×n, m×1, or m×n, depending on whether the row and column containing the cell in the FCA are aggregated relative to the agricultural I-O classification. Suppose for example that the only disaggregation required is to separate "wool; ruminant meat" in the contributed table into two sectors, "wool" and "ruminant meat". Then the agricultural I-O table contains a 2×2 block corresponding to intra-industry usage of "wool; ruminant meat" in the FCA table, a 2×1 block for each element of the"wool; ruminant meat" sales structure in the FCA table, and a 1×2 blocks for each element of the cost structure of the "wool; ruminant meat" industry in the FCA table.

We balance the table by matching each cell to a block and scaling it by its block scaling factor. We set the block scaling factors to achieve certain block target totals; we get the block target totals by aggregating the contributed table to the FCA. Thus after balancing, the agricultural I-O table is a disaggregation of the aggregated contributed table.

At the same time, we impose targets on individual row and column totals in the agricultural I-O table. Each row total represents total sales of some input: a domestically produced commodity, an imported commodity, or a primary factor. Aggregated to the FCA, these must agree with the contributed table; but within each aggregate commodity, we maintain the individual commodity shares of the original agricultural I-O table. So in the example previously given, if the ratio of total sales of wool to total sales of ruminant meat is 30:70 in the original agricultural I-O table, and total sales of "wool; ruminant meat" in the contributed table are \$6 billion, then the total sales targets for "wool" and "ruminant meat" in the balanced table is 30 per cent of \$6 billion, or \$1.8 billion.

Table 13.1 Discrepancy in Industry Shares in Agricultural Factor Costs between Agricultural and Contributed I-O Tables

GTAP Region	Entropy Discrepancy
Rest of Central European Associates	1.1297
Malaysia	1.0717
Hong Kong	0.9985
Singapore	0.8172
Morocco	0.7837
Vietnam	0.3853
Former Soviet Union	0.3621
China	0.3080
Mexico	0.2870
Brazil	0.2436
Argentina	0.2201
Uganda	0.2013
Canada	0.1564
Philippines	0.1272
Turkey	0.1102
Chile	0.0666

Just as we use block scaling targets to achieve the block target totals, we use row and column scaling factors to achieve the row and column target totals. Thus each cell in the agricultural I-O table is scaled by three factors: one row-specific, one column-specific, and one block-specific. There are some further complications, relating to data categories with mixed sign data (changes in stocks, production taxes) and to avoiding singularities in the equation system (unacceptable in GEMPACK, liable to occur because of linear dependencies between the target totals), but in essence the balancing program just sets and applies the scaling factors.

It is debatable whether targeting the row and column totals is beneficial. The decision to apply them was made largely on grounds of conservatism. We targeted the row and column totals in the previous data base release. Since the disaggregation procedure seemed to work reasonably well then, we thought it safer to follow a similar approach with the current release. But since rebalancing without row and column target totals has worked well outside agriculture in the current release (section 13.4), we may consider extending that to agriculture in future releases.

Comparing the agricultural with the contributed I-O tables would be a very extensive project; we make only a beginning here. We compare industry shares in agricultural factor costs across the two tables. Table 13.1 uses the *entropy discrepancy* to indicate the degree of discrepancy between the two structures. To compare the two I-O tables we need to aggregate them to the FCA; therefore comparing results across countries may be misleading, since the smaller the number of agricultural sectors in the contributed table, the lower the discrepancy is likely to be. In the extreme, if the contributed table has just one agricultural sector, then the discrepancy is necessarily zero.

Two of the regions with the largest discrepancies are small economies with small agricultural sectors, Hong Kong and Singapore. Table 13.2 shows the factor cost shares for the regions with large discrepancies and significant agricultural production — Rest of Central European Associates, Malaysia, and Morocco.

Table 13.2 Shares in Agricultural Factor Costs for Selected Countries: Comparison between I-O Tables (percent)

	Contributed I-O Table	Agricultural I-O Table
Rest of Central European Associates		
Paddy rice	0.2	0.1
Wheat	3.2	16.0
Other grains	4.9	17.6
Non-grain crops	28.4	39.6
Other livestock products	22.7	2.3
Wool	0.5	0.0
Meat products	5.3	1.8
Other food products	17.7	11.9
Dairy products	2.3	2.7
Processed rice	5.2	0.1
Beverages and tobacco products	9.6	7.8
Malaysia		
Paddy rice	32.1	8.0
Wheat	0.0	0.0
Other grains	0.1	0.1
Non-grain crops	42.3	25.9
Other livestock products	7.2	16.4
Wool	0.0	0.0
Meat products	1.2	15.8
Other food products	10.9	23.3
Dairy products	0.6	0.2
Processed rice	0.8	6.4
Beverages and tobacco products	4.8	3.8

Contd

Table 13.2 Shares in Agricultural Factor Costs for Selected Countries: Comparison between I-O Tables (percent) (Continued)

	Contributed I-O Table	Agricultural I-O Table
Morocco		
Paddy rice	0.0	0.1
Wheat	15.8	8.8
Other grains	7.2	4.9
Vegetables and fruits	21.2	36.7
Oilseeds	1.4	0.7
Sugar cane, sugar beet	1.4	1.5
Plant-based fibres	2.3	0.1
Other crops	1.8	1.6
Livestock	33.3	22.2
Wool	2.8	0.3
Meat products	0.0	8.1
Vegetable oils and fats	0.3	0.4
Dairy products	1.0	2.7
Other food products	7.6	8.8
Sugar	2.2	0.5
Beverages and tobacco products	1.5	2.5

We are not in a position to assess the relative merits of the two sets of estimates comprehensively. We do however observe that in each of the regions selected for table 13.2, the contributed table displays some curious features: for example, the substantial processed rice production in the Rest of Central European Associates; the combination of a high share for paddy rice and a low share for processed rice in Malaysia; and the combination of a high share for livestock and a low share for meat products in Morocco.

13.3 Agricultural Disaggregation using Summary Data

As noted above (section 13.1), for technical reasons, we use a special procedure to disaggregate agricultural sectors in Sri Lanka. The procedure is generally similar to that used in GTAP release 4 for “category 4” regions (Liu and McDougall, 1998). We estimate total production and usage of agricultural commodities using FAO food balances data sets. We adjust to allow for minor commodities liable to be found in heterogenous groupings such as “other food products” that may not be covered in the FAO statistics, reconcile against the contributed table, and apply to the representative table, rescaling it to meet the reconciled sector targets and so as to constitute a disaggregation of the contributed table. The somewhat baroque details are not presented here.

Elimination of this special procedure in future releases is both feasible and desirable. We look forward to omitting this section in the next edition of the data guide.

13.4 Non-agricultural Disaggregation

In this section we document the disaggregation of the non-agricultural sectors. At the beginning of this stage, the contributed tables are fully disaggregated in agriculture, but not in their original classification outside agriculture. At the end, they are fully disaggregated across in all sectors. The procedure applies to twenty-three regions (of the total 24 regions subject to disaggregation, only China's is confined to agriculture). The procedure is similar to that for the agricultural disaggregation (section 13.2), with some changes:

— The disaggregation is based on the representative I-O table (section 13.1), not a region-specific agricultural I-O table.

— We don't need to construct an FCA sectoral classification (section 13.2). Since the representative table uses the full GTAP sectoral classification, the FCA is just the classification used in the incoming contributed table (*i.e.*, fully disaggregated in agriculture but not more or less aggregated outside agriculture).

— In rebalancing the representative table to match the contributed table, we do not set row or column target totals for disaggregate sectors.

— Since the rebalanced representative table is already fully disaggregated, we do not need a mutual disaggregation step.

The decision not to target individual row or column totals in the non-agricultural disaggregation contrasts with the treatment in agriculture. For agriculture we could use region-specific agricultural I-O tables to set the target totals; for the non-agricultural sectors however we do not have a corresponding region-specific information source.

Although we do not target row or column totals, we still use column-specific scaling factors, at least for columns representing industries. We use them now to maintain sectoral balance, that is, to ensure that each industry's total costs is equal to its total sales. But we do not now need row scaling factors. Thus each cell is subject to two scaling factors: a block scaling factor, to ensure agreement with the aggregate table (as in section 13.2); and a column scaling factor, to ensure sectoral balance.

Under this approach, the shares of disaggregate sectors in total sales of an aggregate sector emerge from the rebalancing process. They are typically close to those in the representative table, but they can differ for compositional reasons. Table 13.3 reports the cases where the change is greatest (as measured by the entropy discrepancy).

Table 13.3 Aggregation Shares in Total Sales, Selected Industries (percent)

	Representative Table	Disaggregated Regional Table
Uganda, `machnry`: 0.33		
Motor vehicles	26.6	57.6
Transport equipment	6.5	14.5
Electronic equipt	14.7	3.8
Machinery, equipt	52.2	24.1
Sri Lanka, `osp`: 0.25		
Communication	9.9	13.0
Financial services	20.6	7.1
Insurance	7.6	13.9
Business services	37.9	18.0
Recreatl services	24.0	48.0
Vietnam, `olg`: 0.13		
Oil	90.7	99.5
Gas	9.3	0.5
Uganda, `mine`: 0.11		
Coal	16.3	9.7
Oil	39.0	32.1
Gas	4.0	0.3
Other minerals	40.6	57.9
Sri Lanka, `mv_ot`: 0.10		
Motor vehicles	80.3	59.8
Transport equipment	19.7	40.2
Uganda, `metals`: 0.10		
Ferrous metals	58.9	39.8
Other metals	12.9	10.4
Metal products	28.2	49.8
Vietnam, `isf`: 0.09		
Ferrous metals	67.6	47.3
Metal products	32.4	52.7
Argentina, `osp`: 0.08		
Communication	9.9	12.5
Financial services	20.6	15.9
Insurance	7.6	15.9
Business services	37.9	23.9
Recreatl services	24.0	31.8
Singapore, `t_t`: 0.08		
Trade	64.1	48.3
Other transport	26.1	28.9
Water transport	5.9	16.1
Air transport	3.9	6.7
Philippines, `osp`: 0.06		
Communication	9.9	12.6
Financial services	20.6	12.5
Insurance	7.6	10.5
Business services	37.9	28.8
Recreatl services	24.0	35.6

Header rows show: *Country*, `sector code`: *entropy discrepancy*.

Out of the total 95 non-agricultural industries disaggregated in 24 regions, table 13.3 shows the ten cases with the greatest entropy discrepancy. The seven largest discrepancies are concentrated in just three countries, Uganda, Sri Lanka, and Vietnam; even of these, only five exceed 0.10 (we show the entropy discrepancy at the right of each header row). Unlike the earlier table 13.2 in section 13.2, the revisions reported here are not a problem: those arose from inconsistencies between data sources, while these arise from the incorporation of region-specific information into a region-generic proxy data source.

On the other hand, while the revisions may be benign, they need not all be. Investigation of particular cases may sometimes lead us to problems or inconsistencies in the source data. We take for closer examination the last case reported in the table, the revisions in the composition of total sales of “other private services” in the Philippines. As measured by the entropy discrepancy (0.06), the revisions are quite moderate. The most important changes are a decrease in the share of “other financial services” from 21 to 12 percent, and an increase in the share of “recreational and other services” from 24 to 36 percent. Further investigation shows that the share of recreational services increases because (according to the representative table) it holds a high share of “other private services” exports; and because (according to the contributed table) exports account for a large share of “other private services” demand in the Philippines. This all seems quite plausible. On the other hand, the share of “other financial services” declines in part because it accounts for an especially large share of “other private services” usage by the “ownership of dwellings” sector; but, according to the contributed table, the Philippine “ownership of dwellings” sector (unlike other countries’) does not make any use of “other private services”. This may indicate a defect in the Philippine table, or at least an inconsistency between countries in the national accounting treatment of this sector.

References

Liu, J. and McDougall, R. 1998. “Disaggregation of Input-Output Tables,” chapter 3 in McDougall, R., Elbehri, A., and Truong, T.P. *Global Trade, Assistance, and Protection: The GTAP 4 Data Base*, Purdue University.