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.A). 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 6 Data Base, many countries' input-output 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 I-O tables using any reasonable aggregation of this classification, (Huff, McDougall and Walmsley, 2000). Where the contributed table uses an aggregated classification, we must disaggregate it to the standard 57 sectors in-house.

Altogether, in GTAP 6, 17 regions require sectoral disaggregation. Of these five require disaggregation only in non-agricultural sectors, while four 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* representing a kind of average of the I-O structures for the regions where we have fully disaggregated contributed I-O 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 accordingly we use country-specific data, namely the agricultural I-O data compiled by Everett Peterson (chapter 12). 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 eleven of the twelve 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 use 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.

We perform the agricultural disaggregation before the non-agricultural disaggregation. 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.

13.2 Agricultural Disaggregation using Agricultural I-O Data

In this section we document the disaggregation of the agricultural sectors in the contributed I-O tables, using outside information on the I-O structure of the agricultural sector. We apply this procedure to twelve regions listed in table 13.1. At the end of this stage, the contributed tables are fully disaggregated in agriculture, but are in their original classification outside agriculture. Briefly, 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 table is non-zero. Such zeros would othewise 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 pro-rating.

— Clean up any minor problems in the disaggregated 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*) of those 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 table and in the agricultural I-O tables; and it is the *finest* (that is, the most detailed) such classification. In this context, the FCA classification matches the contributed table sectoral classification in agriculture, and the agricultural I-O classification outside agriculture.

In the central step in the procedure, we rebalance the agricultural I-O table so that it is consistent with the contributed table at the finest common aggregation. That is, if you aggregate both the rebalanced I-O table and the contributed table to the FCA classification, the results are the same. Since the FCA classification has exactly as much agricultural detail as the contributed I-O table, this means that that if you aggregate the agricultural sectors in the rebalanced agricultural I-O table to the same sectoral classification as the the contributed table, they agree with the contributed I-O table. In other words, the rebalanced agricultural I-O table is a disaggregation of the contributed I-O table in 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, 1c*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.

Just as we use block scaling targets to achieve the block target totals, we use row and 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 13.3), we may consider extending it 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.

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

GTAP Regions	Entropy Discrepancy
Hong Kong	1.73
India	1.59
Morocco	1.14
Malaysia	0.70
Tunisia	0.54
Uganda	0.39
Philippines	0.33
Viet Nam	0.30
Indonesia	0.23
China	0.21
Thailand	0.11
Chile	0.11

The region with the largest discrepancies is a small economy with small agricultural sectors, Hong Kong. Taking account of the size of the agricultural sector, the most significant discrepancies are those for India, China, and Viet Nam. Table 13.2 shows the factor cost shares for these regions.

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

	Contributed Table	Agricultural Table
India		
Paddy rice	14.8	16.8
Wheat	8.2	3.0
Other grains	2.7	1.6
Vegetables and fruits	16.5	10.5
Oilseeds	11.5	4.5
Sugar cane and beet	5.4	2.0
Plant-based fibres	3.5	2.8
Other crops	10.5	3.1
Cattle and sheep	0.0	0.3
Other livestock; wool	6.9	0.6
Milk	14.3	5.1
Other food products	2.3	27.6
Other meat	0.0	0.3
Vegetable oils	0.9	0.9
Sugar	0.9	1.4
Beverages and tobacco	1.7	19.4
China		
Crops	58.4	56.9
Livestock	24.7	13.0
Meat and dairy products	1.3	4.9
Other food products	11.1	12.2
Beverages and tobacco products	4.5	12.9
Viet Nam		
Paddy rice	36.8	45.2
Other crops	26.2	26.8
Sugar cane and beet	2.8	0.5
Livestock	15.1	6.3
Meat products	0.5	3.9
Vegetable oils	0.5	0.1
Dairy products	0.5	0.0
Other food products	11.6	6.4
Sugar	0.8	0.2
Beverages and tobacco products	5.3	10.7

While the differences between the two sources are quite substantial at certain points, on the whole they are not extreme. As expected, of the three countries examined in detail, India shows the largest discrepancies, with for example "other food products" getting a factor usage share of 2 per cent in the contributed table but 28 per cent in the AFP table. One feature common to all three listed countries is that the AFP tables show higher shares for beverages and tobacco products than do the contributed tables.

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 the GTAP 4 Data

Base 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 data releases is both feasible and desirable.

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 in their original classification outside agriculture; at the end, they are fully disaggregated across in all sectors. The procedure applies to fifteen regions (of the total seventeen regions subject to disaggregation, only China's and India's are 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 need row scaling factors anymore. 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).

Out of all the total non-agricultural industries disaggregated in fifteen regions, table 13.3 shows the ten cases with the greatest entropy discrepancy. Uganda and Sri Lanka have multiple entries in the top ten (three and two respectively). More strikingly, four of the top ten involve minerals (either fossil fuels only, or other minerals also), and 3 involve “other private services”. Minerals tend to be involved because of the great differences in their sales dispositions, “other private services” simply because it represents a very large part of the economy. 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.

As an example of how sales dispositions can affect the aggregation shares, we take oil and gas in the Philippines. In the representative table, oil and gas are of almost equal size (ratio 52:48), but in the Philippines, after balancing, the oil sector is much larger than the gas sector (ratio 94:6). The reason is that, according to the contributed table, oil and gas in the Philippines are used almost entirely by the petroleum refining sector (95 per cent of total sales); but according to the representative table, oil and gas inputs into petroleum refining consist almost entirely of oil (input ratio 99:1).

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

	Representative Table	Disaggregated Regional Table
Sri Lanka, minerals: 1.04		
Coal	10.9	13.0
Oil	39.4	37.8
Gas	36.9	0.3
Other minerals	12.8	48.9
Philippines, oil and gas: 0.58		
Oil	51.7	94.5
Gas	48.3	5.5
Uganda, metals and prods: 0.24		
Ferrous metals	50.7	34.1
Other metals	28.9	13.2
Metal products	20.4	52.7
Uganda, minerals: 0.22		
Coal	10.9	8.4
Oil	39.4	25.6
Gas	36.9	25.0
Other minerals	12.8	41.0
Sri Lanka, other private services: 0.20		
Communication	12.5	13.0
Financial services	13.6	7.1
Insurance	4.4	13.9
Business services	39.9	18.0
Recreational services	29.6	48.0
Thailand, oil and gas: 0.16		
Oil	51.7	24.3
Gas	48.3	75.7
Uganda, machinery and equipment: 0.11		
Motor vehicles	40.6	58.8
Other transport equipment	12.2	16.8
Uganda, machinery and equipment: 0.11 (Contd)		
Electronic equipment	16.4	8.0
Other machinery and equipment	30.7	16.3
Argentina, other private services: 0.11		
Communication	12.5	12.5
Financial services	13.6	15.9
Insurance	4.4	15.9
Business services	39.9	23.9
Recreatl services	29.6	31.8
Viet Nam, ferrous metals; metal products: 0.09		
Ferrous metals	71.3	50.2
Metal products	28.7	49.8
Uruguay, other private services: 0.08		
Communication	12.5	10.0
Financial services	13.6	27.0
Insurance	4.4	8.9
Business services	39.9	29.0
Recreatl services	29.6	25.1

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

References

Huff, K., McDougall, R. And Walmsley, T. 2000. "Contributing Input-Output Tables to the GTAP Data Base," GTAP Technical Paper No. 1, Center for Global Trade Analysis, Purdue University.

Liu, J. and McDougall, R. 1998. "Disaggregation of Input-Output Tables," Chapter 13. in McDougall, R., Elbehri, A., and Truong, T.P. *Global Trade, Assistance, and Protection: The GTAP 4 Data Base*, center for Global Trade Analysis, Purdue University.