

15.A

Construction of the Trade Data

Robert McDougall

We describe here the initial construction of the GTAP trade data set. This is close to but not exactly the trade data that appear in the GTAP data base. It differs from it in two ways: it does not include exactly the same arrays, and it does not incorporate certain adjustments.

The trade data in the GTAP data base include six arrays. There are four bilateral trade arrays, showing value of trade at different prices. Two of these four arrays, the *fob* and *cif* value arrays, are included in the trade data set. The other two, showing trade at market prices in exporting and importing countries, use both the trade and protection data sets (chapter 16). The GTAP data base also contains two arrays of data on international trade margins, that is, the services used or costs incurred in moving goods from point of export to point of import. Both of these are included in the trade data set. Thus the trade data set contains four arrays corresponding to arrays in the GTAP data base.

After the initial construction described here, some adjustments are made to the trade data in the course of preparing the energy data set (chapter 17). These adjustments are initially made at the country level rather than at the GTAP region level. To allow this, the initial trade data set contains country-level as well as region-level bilateral trade data.

The major part of international trade is covered in the bilateral trade arrays. These have three dimensions: commodity, source region, and destination region. A major source for bilateral trade is the United Nations COMTRADE data (chapter 15.B). However this covers only merchandise trade (trade in goods not services, but including electricity). Accordingly, we need another data source for services. In GTAP 5, for the first time, we use the IMF balance of payments statistics (chapter 15.E).

Margin services are an especially difficult area. A full margin services data array would have five dimensions: the margin service or *mode of transport* (air transport, sea transport, etc.), the

country of origin of the service, the merchandise commodity to which the service is applied, the country of origin of the merchandise, and the country of destination. So for example, it would specify the value of *water transport services* supplied by *Norway* for the export of *motor vehicles* from *Germany* to *Canada*. (There are two countries of origin, one for the margin service and one for the merchandise, but only one country of destination, because the destination country imports the merchandise and the margin service together.)

For the GTAP data base we are less ambitious, and instead define two smaller arrays: a margin supply array, indexed by margin commodity and country of origin, and a margin usage array, indexed by margin commodity, merchandise commodity, country of origin of the merchandise, and country of destination. So for the example above, the data base specifies the value of *water transport services* supplied by *Norway*, and the value of *water transport services* used in exporting *motor vehicles* from *Germany* to *Canada*.

Margin services are considered exports of the country that supplies the service, and imports of the country that receives the merchandise to which they are applied to. Accordingly, they are included in the services trade statistics. However, the services trade statistics do not show which merchandise trade flows the margins apply to. So, to estimate the margin usage array, we combine them with further data, including margins data incorporated in the merchandise trade data set, and estimates of modal shares. Chapter 15.D documents these additional data and chapter 15.E describes how we combine it with the services trade data.

Another special case is *travelers' expenditures*. Purchases in one country by residents of another country are considered exports from the first country to the second. This includes tourism, but also such things as expenditures incurred in short-term employment overseas. The services trade statistics treat travelers' expenditures as a distinct commodity, but in the GTAP data structure, they are counted as trade in the goods and services actually purchased. Chapter 15.E documents the required data transformation.

The final special case involves *re-export services*. A large part of China's trade passes through Hong Kong, which earns substantial revenue from the difference between import and re-export prices. We account for this revenue as an export of trade services from Hong Kong to the countries of destination of the merchandise. Chapter 15.C describes the estimation of these flows.

To summarize, the source data used in preparing the trade data set include:

- merchandise trade statistics (chapters 15.B, 15.D),
- services trade statistics (15.E),
- modal shares statistics (15.D), and
- statistics for Hong Kong re-export services (15C).

The merchandise trade data and the modal shares data, along with the services trade statistics, are also inputs into the construction of various services trade data arrays (chapter 15.E). The resultant services trade data set includes arrays for ordinary services trade (bilateral), travelers'

expenditures, margin usage, and margin supply. All these data sets, both source and intermediate, are at the country rather than the GTAP region level.

After the data construction covered in chapters 15.B-15.E, there remains the final assembly of the bilateral trade arrays. For the bilateral trade arrays, we use data for:

- merchandise trade (chapter 15.B),
- ordinary services trade (15.E),
- travelers' expenditures (15.E),
- Hong Kong re-export services (15.C).

Finally, we aggregate from countries to GTAP regions.