11

Merchandise trade data

Mark Gehlhar

11.1 Introduction

Development of merchandise trade data poses a somewhat different challenge from other components of the GTAP data base. The problem is less a matter of finding scarce data and more a matter of resolving inconsistent data. The reason is that one country's exports are its partners' imports but imports and exports are reported by both partners. This reporting arrangement produces two trade records representing the same trade flow. Large discrepancies can be found when comparing a country's export flows with its corresponding partner's import flows. Discrepancies in bilateral trade statistics pose a problem for use in the GTAP data base.

The GTAP model and data base link countries through bilateral trade. An accounting identity exists whereby the value of imports at *cif* prices minus the value of transportation services equals the value of exports at *fob* prices. For a given bilateral transaction, the reported import value can be substantially less than, or several times greater than the reported export value. Thus reported trade statistics in their "raw" form are not suitable for the GTAP data base. Various methods can be employed to make the import and export value consistent with each other. But satisfying this consistency condition alone does not solve the important problem of establishing reliable bilateral trade flows. There is a dual objective in the development of trade data for GTAP.

The reliability of trade data affects the credibility of model results. The direction and composition of trade, often referred to as the *structure of trade*, has a direct bearing on the final results of any simulation exercise. Trade structure determines how the impacts of a given policy shock are distributed across different economies in the world. But the fact that large discrepancies exist in reported trade statistics would suggest there is some degree of uncertainty attached to the

trade structure. Results of liberalization in those countries or sectors having high levels of protection can be particularly sensitive to alternative trade shares. Uncertainty associated with any component of the initial base data is troubling given that it is a permanent fixture in the GTAP model. There is only one solution to the problem and that is to remove the uncertainty associated with bilateral trade data. It requires generating the most reliable set of bilateral trade flows possible from the existing reported trade data. This is the objective for developing the bilateral merchandise trade data for GTAP.

Much effort goes into seeking the above objective which requires patience and open mindedness. Since the first version of the GTAP data base was released in 1993, much time has been spent performing analyses at a detailed level. Various reconciliation methods have been put to test. The results from these methods are evaluated and scrutinized. As better procedures are found they are incorporated into a standard procedure for GTAP trade data reconciliation. The trade data for version 3 of the GTAP data base were produced with a different procedure from previous versions. This document describes the procedure and provides evidence of its effectiveness in generating reliable trade data.

Just as a good recipe is not a substitute for good quality ingredients, a good method is not a substitute for good quality data. Even the best method employed in reconciling trade data cannot produce reliable results unless reliable statistics are reported. Reconciliation can only help to distinguish the *less* reliable records from *more* reliable trade records. The first objective in this documentation is to assess the quality of the raw ingredients found in bilateral trade statistics. The second objective is to show how the methodology is used to extract the most reliable information found in reported trade data. The overall quality of the bilateral commodity trade data depends both on the initial quality and on the ability of the method to yield the best statistics from the raw data.

The development of the merchandise trade data for GTAP like other undertakings in this project is an ongoing process. Not all problems in trade data can be solved at once. The most important problems receive highest priority and lesser problems are solved subsequently. Future work on merchandise trade data will focus on specific country and commodity problems and less on the general procedures which are now established here.

11.2 Comparisons of International Sources of Trade Statistics

The reliability of any data source is called into question when there is a lack of comparability. TheUN COMTRADE data base is the primary source of trade data for the construction of the merchandise trade data for GTAP. It seems appropriate to provide comparisons of the COMTRADE trade data with other sources of trade data.

The World Bank,[1] the International Monetary Fund (*IMF*),[2] and the Food and Agriculture Organization (*FAO*)[3] of the UN publish and disseminate international trade statistics on an annual basis. Using these sources, comparisons are made with the COMTRADE data base. Comparisons are possible if there is a common level of aggregation across sources. All four sources provide merchandise totals by individual country. The World Bank and the IMF publish some non-agriculture aggregates, not comparable with those found in COMTRADE. FAO publishes aggregate agricultural sectors and individual agriculture commodities. FAO's FAOSTAT data base contains individual agricultural commodities some of which are comparable to the COMTRADE categories. This permits us to compare COMTRADE and FAO individual agricultural commodity totals. FAO, the World Bank, and the IMF unfortunately do not provide bilateral trade flows which are comparable to COMTRADE. So comparisons are made only with merchandise trade totals for all sources, and for agricultural commodity totals with FAO as the only source.

The COMTRADE data base contains source-destination trade flows of individual commodities classified in accordance with the Standard International Trade Classification (*SITC*). The SITC system provides a 5-tier system of aggregation (1-5 digit levels). For generating merchandise trade totals, COMTRADE's bilateral-commodity trade must be aggregated. This is done starting at the 4-digit SITC level with bilateral trade.

The UN maintains a data set for *reported exports* and a separate data set for *reported imports*. The two data sets are reported and maintained independently of each other. A country trade total

---

1. The World Bank publishes its own trade statistics in the World Tables. According to the 1995 World Tables, the primary source of foreign trade data is the UNCTAD data base. It is supplemented with data from the UN COMTRADE data base and the IMF's International Financial Statistics. Balance of payments data are from the IMF's balance of payments statistics.

2. The IMF maintains that its source of trade data is in general customs statistics reported under the general trade system according to the recommendations of the UN.

3. FAO maintains trade statistics in the FAOSTAT data base and claims its trade data come from governments through magnetic tapes, national publications and FAO questionnaires. For EU members FAO trade data are obtained from EUROSTAT.

can be calculated in two ways. An export total can be calculated as the sum of a country's reported exports across all of its partners. The data for this total are from reported exports and it is referred to as a *country-reported* total because it is reported by the country itself. Another calculation representing the same country total is made by taking the sum of all partners' reported imports from the given exporting country. This total is referred to as a *partner-reported* total. Trade totals published by international organizations are generally country-reported totals.

From the COMTRADE data base country-reported merchandise totals are obtained by aggregating across commodities and partners. Here we denote the reported export flows as $X_{i,r,s}^{rp}$ where *i* is a commodity exported from region *r* to region *s*. Likewise we denote the reported import flows as $M_{i,r,s}^{rp}$, where *i* is a commodity exported from region *r* to region *s*. Country-reported export totals are shown below as

$$X_r^{rp} = \sum_i \sum_s X_{i,r,s}^{rp} \quad .$$

Similarly, country- reported import totals are calculated as

$$M_s^{rp} = \sum_i \sum_r M_{i,r,s}^{rp} \quad .$$

Table 1 shows merchandise export totals from FAO, the World Bank, the IMF and UN COMTRADE. Countries for this comparison include the top 25 exporting countries in the world. For most of these countries there is little difference in totals across sources with the exception of Mexico. In the Mexican case both the UN and the World Bank have comparable figures of $46.2 billion and $46.3 billion, respectively. FAO and the IMF also have comparable figures of $27.5 billion and $27.6 billion, respectively. The reason for the difference is that FAO and the IMF apparently disregard Mexico's reported export value and use the *balance of payments* level of trade. The same can be said for Mexican imports shown in table 2. FAO and the IMF use the level of imports shown for the balance of payments of approximately $48 billion rather than the Mexican reported import value of $62 billion which the World Bank publishes.

FAO merchandise totals are generally higher for both imports and exports. The reason for this is that the country totals are not strictly country-reported totals. FAO substitutes partner-reported trade flows when a country fails to report its own trade flows. This is done at the detailed agriculture commodity level. By filling in holes in country-reported trade by partner-reported trade higher levels of trade are obtained in the aggregate.

Overall we find there is significant comparability in total merchandise trade across different sources. The main reason for this is that there is actually only one source of trade data. That source is the customs trade files from individual countries which are submitted to the UN and shared by all international organizations. International organizations aggregate data for their own purposes and disseminate it as such.

We now turn to specific agriculture commodities, using the FAO as the source for comparisons. Exports of wheat, bovine meat, and bananas are used in this comparison. Table 3 shows reported exports for FAO and COMTRADE. In the last column the ratio of FAO to

COMTRADE is given. There is much similarity in the commodity totals. The only exception is the case for Australia. The reason for the difference is that FAO does not use a calendar year for Australian agriculture trade. Agricultural commodity data are provided to the UN and FAO on a monthly basis but FAO places Australia on a fiscal year beginning June 30 of the year in question.

We turn now to major agricultural importers using the commodities wheat, bovine meat, coffee, and bananas. The comparison of agricultural import totals is shown in table 4. Again as with exports the totals are nearly identical for FAO and COMTRADE. The exceptions are Brazilian imports of wheat and Italian imports of bananas. The exact reasons for the differences for Brazilian wheat not known. But it is known that partner data would give a total closer to the FAO level than the COMTRADE level which strictly used Brazil's reported imports. For the case of imports of bananas for Italy, FAO has made adjustments for transhipments from Italy. This will reduce the level of imports from what is officially reported. Transhipments pose a problem for data reconciliation (especially for Hong Kong).

As with merchandise trade it is not surprising to find comparability between FAO and COMTRADE given that they share primary data sources. But comparability of trade totals across sources is not by itself evidence of reliable data. One major problem in reported trade is the misrepresentation of bilateral trade flows. This can be a result of the reporting country's misidentifying its partners. This type of reporting error would not affect the country-reported total. But it does suggest that country totals for export and import totals are not consistent with one another. We examine a specific example of why this is so.

We examine individual wheat import flows to China reported both by China and exporters. Table 5 provides this information. We disaggregate Chinese wheat imports by 5 partners. The partners are shown in table 5 as Canada, the United States, France, Poland, and other. The first column of values shows the reported values reported by these exporters. Canada reported exports to China of $1,020,897 thousand. But China reported imports from Canada of $847,553 thousand. The third column gives the ratio of reported imports to reported exports. For Canadian wheat exports it is 0.83. Because of this discrepancy the trade totals of Canadian wheat exports and Chinese wheat imports are not consistent. The China-reported total from the world for wheat imports is $1,503, 730 thousand. This value is the same for both COMTRADE and FAO. The Canada-reported total for its wheat exports is $3,870,900 thousand. Again this value is the same for both COMTRADE and FAO. These totals are the sums of the individual bilateral flows. But China cannot be importing wheat from Canada valued at $847,553 thousand while Canada is exporting wheat to China valued at $1,020,897 thousand. The acceptance of the total wheat imports reported by China and the acceptance of the total wheat exports reported by Canada are not consistent. Both totals cannot be correct. If we accept what is reported by the Chinese we must use the Chinese value in the calculation of the Canadian wheat import total. This is shown in the fifth column in table 5. Canada's total must drop at least by 6% to accept the reported Chinese values.

The above illustration highlights the problem in using totals for imports and exports which are not from the same bilateral trade flows. This problem applies to all international trade totals for

agriculture and nonagriculture commodities. Importer- and exporter-reported totals as they are published by FAO, the World Bank and the IMF are inconsistent. This is because reported bilateral trade flows have not been reconciled. Because of the inconsistency problem there is no reason to use published trade totals as *targeted* totals in the development of the GTAP trade data. The most reliable trade totals are a summation of the most reliable bilateral trade flows.

11.3 Assessment of the Reliability of Bilateral Trade Statistics

The source of uncertainty in trade data is linked directly to discrepancies in bilateral-commodity data. Discrepancies make country totals unreliable and lessen the overall integrity of the reported structure of world trade. Discrepant transactions have caused some to call into question the reliability of the entire UN COMTRADE system. It is not always recognized that the UN is not the source of reporting errors. Individual countries contribute reporting errors. One unreliable reporter can generate a multitude of discrepant trade transactions. And if one unreliable reporter trades with every country in the world in every commodity, it makes the entire world trade structure appear unreliable. Identification of the unreliable reporters is essential in data reconciliation. Reporting problems and their significance in COMTRADE can be better understood with specific examples. We now turn to specific examples found in the COMTRADE data base.

It is impractical to examine all bilateral commodity trade for all regions at a detailed level. However, it is instructive to provide selected examples illustrating typical problems found at the detailed level in COMTRADE. Most aggregated sectors in GTAP comprise a variety of goods from various industries. We start with the GTAP sector Chemicals, Rubber, and Plastic products (*CRP*) using Australian exports. There is no particular reason for us to believe that Australia would bias the reporting of CRP any more than other products. There is also no particular reason why products found *within* CRP should be reported with a common bias because they belong to the CRP sector. As far as reporting goes, products within sectors are reported independently without any association to the GTAP sectoral classification; for example, Australia reports individual CRP products at the individual SITC level without association to CPR as sector. Biases in reporting tend to be product-specific rather than sector-specific. Observations on sectors lead to less informative generalizations than observations on individual products. We are interested in observing the CRP products at the individual level rather than as an aggregated CRP sector.

Table 6 displays bilateral transactions of Australia's exports for selected CPR transactions. There are five examples of 4-digit SITC product classes and six importing-partner countries. Each of the importing countries reports a value for each product as imports from Australia. The ratio of reported imports to reported exports for each transaction is provided in the last column in tables 6.

A subtotal is given for both a sum of Australia's reported trade with the six partners and a sum of partners' reported imports from Australia. Except for Canada, all importing countries report imports on a *cif* basis. Australia reports the value of exports for each of the partners on an *fob* basis.

We start at the top of table 6 with the product category shown as 4-digit SITC 5136 with the description "inorganic bases". Of the 6 transactions there is only one non-discrepant transaction. This transaction is with the United Kingdom. Shown in the last column for the United Kingdom we see that the ratio of reported imports to reported exports is 1.10, meaning the reported import value is 10% greater than the reported export value. The 10% difference could accurately represent the transportation margin. We contend that the consistency found in this transaction is not a mere coincidence but is the result of reliable reporting by both countries. We have no reason not to believe that Australia exported $1,506 thousand of chemicals classified as inorganic bases to the United Kingdom in 1992. When there is comparability in reports there is confidence and a high degree of certainty associated with the reported trade flow.

What can be said for the UK transaction unfortunately cannot be said for all other transactions. Brazil reports imports of $5,443 thousand from Australia, while Australia reports exports of $41 thousand to Brazil. Reported imports are 134 times larger than reported exports. Because this discrepancy is unreasonably large by any measure, it is deemed *unreliable* and there is uncertainty attached to it. The information from this transaction alone does not however suggest that both reporters are unreliable, or which reporter is unreliable. It shows only that reported exports are small compared to reported imports. More information is necessary to link reliability to a specific reporter.

Other transactions for inorganic bases show a similar pattern to the Brazilian case where the import value is much larger than Australia's export value. Canada, Germany, New Zealand, and the United States all report import values several times larger than Australia's export value. The largest trading partner, the United States, reports an import value of 73.7 times larger than Australia's export value. For all six importers shown in the subtotal, reported imports are 83.7 times larger than Australia's reported exports of inorganic bases. Is Australia an unreliable reporter ? The answer is still uncertain without further information.

The other examples of CRP categories include medicaments, pharmaceutical goods, rubber tires, and articles of plastic. Transactions involving these categories do not exhibit the same reporting pattern as inorganic bases. In fact, several transactions appear quite reliable. For medicaments, New Zealand reported imports of $76,184 thousand while Australia reported exports of $64,201 thousand, giving a ratio of imports to exports of 1.2. All other transactions involving New Zealand, other than inorganic bases, appear reliable.

The reporting of other transactions is rather mixed. Australia reports export values substantially higher than Germany's import values for medicaments and pharmaceutical goods. But Australia reports values only slightly less for rubber and articles of plastic than what Germany reports. For UK imports of medicaments, Australia reports $25,530 thousand, consistent with the

UK report of $23,776 thousand. For US imports of articles of plastic Australia reports $4,448 thousand, while the United States reports $4,448 thousand: an exact match.

For the entire aggregated CRP sector (2,873 transactions), Australia's partners reported an import value that is twice the value of what Australia reports as exports to these partners. Does this mean Australia is an unreliable reporter ? The information shown here for Australia-reported exports of CRP is inconclusive. There are several examples of accurate transactions. We cannot generalize and say Australia systematically under-reports its exports of CRP. There are cases where reported exports were significantly higher than reported imports. But clearly, the most obvious problem for Australia-reported trade is in the reporting of inorganic bases which happens to be a large component of CRP exports for Australia. But from what was observed in the data we can still only say there is probable cause to *suspect* that Australia severely and systematically under-reports exports of inorganic bases. We therefore must investigate further to come to a conclusion.

A question that comes to mind is whether the reporting problem lies with Australia or with its partners. Perhaps the importers are systematically over-reporting ? That assertion can be challenged. To do this we examine exports of the same products to the same importers but exclude Australia as an exporter from the aggregate partner. Table 7 shows reported data in a similar format to table 6. The exporters include all non-Australian exporters. We do not observe large discrepancies for inorganic bases as we did in table 6. Brazil reports imports of $128,596 thousand and the aggregate exporter reports $110,027 thousand giving a ratio of reported import to reported exports of 1.2 . This ratio would not indicate systematic over-reporting by Brazil. The United States, which reported an import value 74 times greater than Australia's export value, reports imports only 10 percent greater than reported exports from its aggregate partner. Importers' reports together are only 10% greater than their non-Australian exporters' as shown in the subtotal for inorganic bases.

We view all the above information as evidence that Australia systematically and severely under-reports exports of inorganic bases. The exact reason why Australia reports what it does is not known nor is it necessary to know for our purpose at hand. Throughout the COMTRADE data base there are numerous examples like this. Some transactions have extreme discrepancies and some have less extreme discrepancies, and some are non-discrepant transactions.

Part of the problem in reported trade is knowing when there is a problem or not a problem. We now turn to another GTAP sector and country reporter. For this case we focus on the GTAP sector Non-Grain Crops (*NGC*). Like the CPR sector, NGC contains a variety of SITC categories. Five selected 4-digit SITC commodities belonging to the NGC sector are examined. These include bananas, fresh fruit n.e.s., manila fibre, live plants, and vegetable matter n.e.s. We focus on Philippine exports of NGC.

Table 8 displays the reported values of imports and Philippine reported value of exports. For almost all transactions in Table 8 the import value is greater than the export value. Looking at the subtotals by commodity, for bananas the import value is 3.2 times greater than the export value and for fresh fruit it is twice as large. For other NGC commodities the import value is less than 20% greater than the reported export value which seems acceptable. As with the case with Australian

exports it appears that there could be a problem of systematic under-reporting of exports. But what is different here is that the discrepancies are less severe and more uniform than for the case of Australian exports of inorganic bases. Although reported imports are 2-3 times larger than reported exports, the uniformity in the discrepancy is curious.

It is commonly known that fresh fruit and in particular bananas are typically more *bulky*, i.e., the price per unit of weight or volume is comparatively low. For such commodities transportation costs adds a proportionately higher value to the *fob* export value than it does for other nonbulk goods. Transportation cost can sometimes help explain discrepancies between the *cif* value and the *fob* value. It is instructive to examine transportation cost in some detail to gain a better understanding of what is represented in reported trade values.

A comprehensive examination of transport cost requires data on reported values, quantity, and unit values. Table 9 provides these data. We focus on Philippine banana exports to Japan and Korea. In order to gain a better perspective of the role of transportation cost over time we examine a 10-year time-series (1985-1994). Table 9 contains four columns of reported trade data. The first column is the quantity in metric tons as reported by the importer. The second column is the quantity in metric tons as reported by the Philippines. The third and fourth columns contain the value reported by the importer and the value reported by the Philippines, respectively.

The fifth column in table 9 is the ratio of the quantity of reported imports to reported exports. What is clear from this ratio is that reported trade of the quantity of bananas is consistent for the most part. There is no evidence of the Philippines' under-reporting banana exports. On average, the Philippines reports a higher export quantity than the importers. A slightly higher quantity of exports over import quantity is typical for fresh fruits given that some spoilage occurs in transit.

The sixth column in table 9 shows the ratio of the reported value of imports to exports. In 1994, Korea reported twice the value of bananas that the Philippines reported but reported exactly the same quantity of bananas as the Philippines. How can this be explained ? Transportation cost is a major factor here. It is explained by the difference between the export unit value reported by the Philippines and the import unit value reported by Korea. Import and export unit values are shown in the seventh and eighth columns of table 9. The ratio of the import unit value to the export unit value is given in the tenth column in table 9. The difference between the import and export unit value is the unit transportation cost. In 1994, the unit transportation cost was $174 per metric tonne for shipping bananas from the Philippines to Korea. There were 117.8 thousand metric tonnes shipped giving a transportation cost of $20,440 thousand. This explains the difference in the *cif* and *fob* reported values.

The discrepancy in the reported value of bananas varies by the ratio of unit value of imports to unit value of exports. For Korea the largest discrepancy was in 1985 with an import:export value of 4.10 (sixth column). This corresponds to the largest unit price ratio of 4.55 (ninth column). For Japan, the largest discrepancy was in 1988 with an import:export value of 3.18 and a unit price ratio of 3.39.

Table 9 also shows that transportation costs are route-specific. The importer unit price (seventh column) differs between Japan and Korea.

We now turn to a few summary statistics giving an overall assessment of the reliability of reported bilateral trade. Thus far specific examples have been provided showing reported bilateral trade. These specific examples are helpful in understanding the nature and probable source of the discrepancies. For assessing the overall reliability of bilateral trade it is useful to show summary statistics that provide information on the extent of the problem of discrepant trade data. We know there can be extremely large discrepancies as was shown in the case of Australian exports. A question worth pursuing is how large and how pervasive are such discrepancies in the COMTRADE system. Statistics on the size and frequency of discrepancies are informative in this regard.

Table 10 displays summary statistics on the discrepancies found in COMTRADE at the 4-digit SITC level of aggregation. Transactions are classified by 5 sizes. The reported import value is used as a measure of size for transactions. For each size class four summary statistics are given. The first column is the average discrepancy for the class of transactions. It is calculated as a simple average. We see that for the class of transactions under $1 million the average discrepancy is 7794%. Clearly this indicates there are wide disparities in reported imports and exports giving much uncertainty to trade flows in this class. The size of transactions varies within each class and thus it is useful to calculate a value-weighted average. This is provided in the second column. For the “less than 1 million” class the weighted average is 146% compared to 7773% for the simple average. In comparing the simple average with the weighted average for other classes we see that the weighted average is lower than the simple average. Larger discrepancies tend to be with smaller transactions. This can be seen within classes and between classes.

The frequency and the value of transactions belonging to each class are also given in the third and fourth columns of table 10. Frequency of transactions by class is expressed as a share of the total number of transactions in the data set. The smallest size class claims 73.1 percent of total transactions. Because this class represents the bulk of transactions one could easily come to the conclusion that most transactions in COMTRADE are unreliable. This is true. But what is often not recognized, as shown in table 10, is that although 73% of the number of bilateral flows could be viewed as unreliable these transactions contribute less than 2% of the total value of trade. Over 75% of the value of individual transactions are reported with less than a 25% discrepancy even without accounting for the *cif*:*fob* difference.

Table 11 provides the same statistics as table 10 except that the data set is aggregated to the GTAP sector level. Even after aggregating, the large discrepancies associated with small transactions remain. Aggregation redistributes the number of transactions to larger size classes. At the GTAP commodity level of aggregation small transactions of less than $1 million represent 32% of the total number of transactions, but represent a minuscule 0.04% of the total value of trade. Aggregation concentrates most of the value of trade into transactions above $75 million, making up 96% of total trade.

The above statistics show that the reliability of total value trade is much less of a problem than the reliability of small individual transactions. But the degree of importance should not be necessarily correlated to the value of trade. Small transactions might seem small from the total world perspective but they are actually large transactions for small countries. Trade flows among developing countries are just as important as trade flows among industrialized countries. The reliability of all transactions large or small is important for the GTAP model. The examples and summary statistics in this section reveal some characteristics of the COMTRADE data set. The methodology of reconciling bilateral data must be tailored to accommodate the specific problems found in the reported data.

11.4 Methodology and Rationale

The large share of discrepant transactions found in COMTRADE makes the task of reconciling bilateral commodity trade seem almost impossible. Partners rarely communicate with each other with regards to what each one is reporting.[4] There are no rules coordinating the reporting process. And so reported exports are reported independently of reported imports. Tackling the task requires some optimistic forethought. It is useful to draw an analogy for trade data reconciliation.

What comes to mind after observing partner trade is a ballroom filled with people performing couples dances.[5] Consider if there was no communication between dance partners. Each dancer attempts to lead a dance regardless of whether their partner knows the dance. It would be only when *both* partners know the dance very well that the couple is *in step*. Those couples who know the dance would gracefully move across the floor in step. Even if one partner knows the correct dance steps the couple painfully stumbles over one another *out of step*. And so it is for trade reporting.

What can be done with a noncommunicative group of pathetic ballroom dancers ? Rules must be imposed. One rule is to have all women follow exactly what the men do, analogous to choosing reported imports over reported exports. This would at least help to keep partners in step with each other, but it gives no assurance that the correct dance is being performed. Not all men know the dance steps for every dance. And the same goes for women. Another solution is to work out a compromise. Women could lead the first half of a given dance and men lead the second half. This is analogous to taking an average of the two reported values of trade. So, if Brazil reports 5,442 and Australia reports 41 we end up with 2,742. But there is no reason to believe that the average is the

4. The United States and Canada share reported trade information making them consistent reporters.

5. While ballroom dancing and trade data reconciliation are worlds apart, this type of analogy is useful in understanding the rational for the methodology.

actual trade flow. The objective in ballroom dancing is not simply to have everyone in step with each other, but to perform the correct dance in step. The same goes for bilateral trade where the correct trade flow is most important. It only seems logical that the dancer who knows the dance should lead regardless of gender. This approach makes use of all dance talent found in the entire group. This is the approach taken in reconciling trade data.

Misrepresentation of reported bilateral trade by reporter countries can either be intentional or unintentional. There are numerous explanations for discrepant transactions but the reasons for reporting behavior are not important. All that is important for this exercise is that countries which systematically misrepresent their reported bilateral trade are identified. Identification of unreliable reporters requires empirical evidence in the form of an indicator reliability.

An appropriate measure of reliability is of the utmost importance for a successful method to yield quality results. There are numerous statistics that might be thought of as representing reliability. A measure of the difference between country-reported totals and partner-reported totals can be used to measure biased reporting. But totals can be misleading and are of limited value in measuring the reliability of reporters' bilateral trade. The reason for this is that countries can incorrectly report all of their bilateral trade flows but still report correct totals. The most common reporting problem is misidentifying partners and misclassifying commodities. Some countries cannot or do not identify all of their partners and commodities. Recognizing this problem the UN reporting system provides residual categories for both unidentified partners and commodities. These are generally referred to as "not elsewhere specified" (*n.e.s.*) categories. Countries can under-report trade with specific country partners but can make up for it by reporting to the n.e.s. partner category. In that case using the reported trade total for the world would not reveal a bilateral reporting problem. Unreliable reporters can correctly report totals without the use of an n.e.s. category as well. This is done by confusing partners. But misidentifying partners leads to under-reporting for one partner and over-reporting for another without affecting total trade. And as before the reported trade total would not reveal a bilateral reporting problem. For these reasons country- reported totals of are of little help in revealing unreliable reporter countries.

Measuring the consistency found in individual bilateral trade flows is most useful for establishing reporter reliability. But this is not as straightforward as one would hope. What exactly should be measured in regards to bilateral trade requires some thought. One could simple calculate an average discrepancy across bilateral trade flows for each reporter. But as shown in the last section, the average size of discrepancy can vary widely and the size of discrepancy is not correlated with reliability. Suppose there are 2 reporters: reporter *A* and reporter *B*. Reporter *A* has large and small trading partners. Reporter *B* has fairly uniformly sized partners. Both reporters make the same mistake by misidentifying their partners. Reporter *A* confuses only two of its partners. But reporter *B* confuses all of its partners. But the mistake of *A* would result in a much larger average discrepancy than *B*'s average discrepancy even with misrepresentation of all trade flows. The larger average discrepancy is because reporter *A* has both large and small partners and when *A* mismatches just 2 partners it produces very large discrepancies. For reporter *B*, even mismatching all partners

does not result in a large average discrepancy since the partners are more uniform in their trade flows. But the size of *A*'s average discrepancies does not make *A* a less reliable reporter than *B*. If anything, reporter *B* is less reliable because it has mismatched all of its bilateral trade flows.

Average size of discrepancies is not a good indicator of reliability. It is like judging a dance partner on just how bad their worst partners' dances were. Conventional wisdom would suggest that it is the rate of successful dances that reflects knowledge of dance talent. A better measure of bilateral reliability is that which measures the rate of successful transactions.

The sum of the total value of *accurate* partner matches as a share of total reported trade is a measure well suited for measuring reliability. We now formalize this as a quantified measurement. First, what is an accurate partner match ? Rarely does a reported export value match perfectly with a corresponding reported import value. Although there are no perfect matches some discrepancies are small enough to be considered accurate matches. A threshold level must be established. It is the difference as a percentage between reported exports and reported imports. This level as been established at 20%.

There are a few preliminary steps that must take place prior to checking for accuracy in reporting. As we saw in the last section reported imports can differ from each other because of the *cif*:*fob* transport margin. Some margins can be over 100% as was shown for the case of bananas. Before determining whether a transaction is accurate imports when reported on a *cif* basis must be converted to *fob*. This is done using a *cif*:*fob* conversion factor at the SITC commodity level.[6]

Once imports are converted to *fob* the accuracy level $AL_{i,r,s}$ of each transaction is calculated. This is shown as

$$AL_{i,r,s}^{fob} = \frac{\left| M_{i,r,s}^{fob} - X_{i,r,s}^{fob} \right|}{M_{i,r,s}^{fob}}.$$

The value of trade reported by the importer for commodity *i* exported from region *r* to region *s* is denoted by $M_{i,r,s}^{fob}$ and likewise the value reported by the exporter for commodity *i* from region *r* to region *s* is denoted by $X_{i,r,s}^{fob}$. Whether exports exceed imports or vice versa is not relevant, so accuracy is not sign-specific . The magnitude of the difference only matters. The goal is to measure reliability indices which are importer-commodity-specific or exporter-commodity-specific. The reliability index is rather simple to construct but very useful in the information it provides.

We start with construction of the importer-commodity-specific index. First we calculate total trade $M_{i,s}^{T}$ reported by the importer *s* for commodity *i*. This is shown as

$$M_{i,s}^{T} = \sum_{r} M_{i,r,s}^{fob} \qquad \text{for all } r.$$

---

6. The *cif*:*fob* conversion factor by individual commodities is first estimated using all transactions. Once a subset of most reliable reporters are found it is then re-estimated using only transactions reported by most reliable reporters. This conversion factor is not the same as the bilateral margins found in the GTAP data base and documented elsewhere.

Next we calculate the reported imports that were accurately matched with partners' reported export values, denoted by $M_{i,s}^{A}$. This is calculated as

$$M_{i,s}^{A} = \sum_{r} M_{i,r,s}^{fob} \qquad \text{for all } r \text{ such that} \qquad AL_{i,r,s}^{fob} \leq 0.20.$$

We can now calculate the importer-commodity-specific reliability index as the share of accurate transactions, denoted by $RIM_{i,s}$:

$$RIM_{i,s} = \frac{M_{i,s}^{A}}{M_{i,s}^{T}} * 100.$$

Likewise we calculate the exporter-commodity-specific reliability index:

$$X_{i,r}^{T} = \sum_{s} X_{i,r,s}^{fob} \qquad \text{for all } s,$$

$$X_{i,r}^{A} = \sum_{s} X_{i,r,s}^{fob} \qquad \text{for all } s \text{ such that} \qquad AL_{i,r,s}^{fob} \leq 0.20,$$

$$RIX_{i,r} = \frac{X_{i,r}^{A}}{X_{i,r}^{T}}.$$

There is one important detail left out of the above. Each reporter is given an opportunity to disregard the value reported by its *worst partner*. The reason for this is to give some "saving grace" to reliable reporters who happen to trade with a large unreliable partner. Just as a good dancer can have a bad experience with a particular dance partner this it will not blemish the good dancer's record given the option to disregard it. Equal treatment must be given to all dancers. So all dancers good or bad disregard their worst dance experience.

A quantitative measure is needed to identify the worst partner. This is measured in a rather straightforward manner as the value-weighted accuracy level. This is calculated for the importer and exporter as follows:

$$WAL^{M_{i,r,s}} = \frac{M_{i,r,s}^{fob}}{M_{i,r,s}^{T}} \; AL_{i,r,s},$$

$$WAL_{i,r,s}^{X} = \frac{X_{i,r,s}^{fob}}{X_{i,r,s}^{T}} \; AL_{i,r,s}.$$

For each importer and exporter for a given commodity the partner which produces the largest *WAL* is dropped from its set of transactions. By dropping a large and less accurate transaction it raises the reliability index for all countries. But it has a greater positive effect on better reporting countries than it does for poor reporting countries. This provides more accurate information for the reconciliation process. Once reliability indices are generated, the actual reconciliation is nothing more than accepting the reported trade flows of the more reliable partners. No adjustments are made to reported bilateral trade flows.

We now turn to specific examples showing the calculated RIM and RIX and how they are used in reconciling individual bilateral trade flows. Recall that in the previous section we noted a problem with Australian exports in the CRP sector. We now see exactly how this problem was resolved. Table 12 displays selected transactions giving the reported values of importer and exporter, the RIM and RIX, and the most reliable reporter for each given transaction. Not surprising Australia received a very low RIX of 5.3 for inorganic bases. This means that only 5.3% of the value of inorganic bases was reported accurately by Australia. The Brazilian RIM is 54. Since this is larger than the Australian RIX of 5.3 we reject the value Australia reported of $41 thousand and accept the value Brazil reported of $5,443 thousand. Of the exporters listed for exports of inorganic bases, Canada received the highest reliability rating of 97.5.

The worst partner for Australia as an importer of inorganic bases was the United States. But even excluding the United States from Australia's total exports of inorganic bases did not seem to help the RIX much. The worst partner for the United States as an exporter of inorganic bases was Australia. Removing Australia as an exporter for inorganic bases to the United States gave a boost to the RIM for the United States. It is not all bad news for Australia's reporting record in the examples listed. Australia's reliability for medicaments was 42.2 which was greater than New Zealand's reliability of index of 27.2. Again, Australia is better as a reporting exporter of pharmaceutical goods than the United Kingdom and the United States are as reporting importers of these goods.

We noted before that there was a discrepancy between Canadian wheat exports to China and Chinese imports from Canada. From table 12 we can see that China has a higher reliability index as an importer of wheat than Canada has an exporter of wheat. Therefore we accept the value reported by China and reject the value reported by Canada. At the same time, we must reject FAO's total for Canadian wheat exports since we reject that Canada exported $1,020,890 thousand to China.

In some cases the reliability index can be 100. This means that all of the individual trade flows were deemed accurate. This was the case for Japanese and Syrian wheat imports. This means all Japanese wheat imports of wheat would be reported by Japan in the reconciled data for GTAP. The same goes for Syria. In some cases the reliability index can be zero. This means none of the trade flows were deemed accurate enough for use in the calculation. In the examples listed this occurred for Sri Lankan wheat imports. Therefore all imports of Sri Lankan wheat imports would be reported by the exporter in the reconciled data for GTAP.

What is important to recognize in this method is that it is based solely on the evidence of reliable reporting on a case by case level. Each reported trade flow goes to trial and is judged. Generalizations and inferences are not made. A country's reporting ability as an exporter can differ from its ability as an importer. A country's individual reporting ability varies by commodity. The method is tailored to account for the specific problems encountered. The process is very much a *bottom-up* approach where the totals are the sum of the most reliable bilateral commodity trade flows. When each and every trade flow is reconciled as done here the level of uncertainty attached

to a given trade flow is removed. We have no reason to believe that world trade is anything different from what is found in the GTAP trade data base.