

Mackay, Euan A.

Article — Digitized Version

The accuracy of international trade data

Intereconomics

Suggested Citation: Mackay, Euan A. (1982) : The accuracy of international trade data, Intereconomics, ISSN 0020-5346, Verlag Weltarchiv, Hamburg, Vol. 17, Iss. 5, pp. 251-255, <http://dx.doi.org/10.1007/BF02928196>

This Version is available at:

<http://hdl.handle.net/10419/139826>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The Accuracy of International Trade Data

by Euan A. Mackay, Edinburgh*

International trade data are of great importance to planners in the Third World as well as to academic researchers. They are often considered to be the best economic series available. Just how accurate are these data really?

Awareness of the degree of data accuracy is important for any researcher involved in empirical work. The accuracy of observations used as an input to a model determines how much reliance can be placed on the results. When data is recognised as being unreliable, the researcher can acknowledge this in his interpretation of the results and thus indicate the level of confidence one can have in them. This is particularly important when results are used as an input to another model or decision-making process.

In many cases within economics the data used is not usually collected by the researcher but is more often accessed through published sources. Researchers collecting their own data should be able to indicate its limitations in describing the phenomenon under study. On the other hand published data is not often collected for the purpose to which the researcher wishes to put it, the methods used to collect the data may not be reported, and there is rarely any chance of repeating the "experiment" to obtain another observation.

The concept of accuracy is not easily defined, although there is an implicit notion of being close to the "true" value. If there is only one observation of an economic phenomenon there is no way of describing the true value other than in terms of the observed value. There may be *a priori* beliefs as to what the true value should be, and observations may be used to test the underlying theory. If the observation does not support the theory then the theoretical hypothesis is rejected, or the validity and accuracy of the observation is questioned, or the observation is rejected as not corresponding to the theoretical definition of the

variable. Data inaccuracy may derive from the methods used to produce the data, or in actual data errors. If we can be sure that the methods of producing the data are correct then either the theory or the data is wrong. Assuming we can obtain at least two observations of the same experiment and we can account for a part of the difference between the observations, we may describe the unaccounted difference as an error and so define the relative accuracy of the data.

International trade data is considered in many cases to be the best economic series available. These data are usually available for every country, and for some may be the only economic series produced. They are frequently used as an aid in describing the economies of developing countries when other data are unavailable or of dubious quality. Trade is generally an important aspect of the development process and trade data may be an essential input to development plans. Revenues from import and export duties are often the major source of finance for Third World governments, and future spending plans will be based on projections of future trade values. The accuracy of this data is therefore of great importance to the planners in the Third World as well as to academic researchers.

International trade data offers a rare example of an economic transaction being reported twice, once by the exporter and once by the importer. Because of this, if there is data missing from one partner's statistics, an alternative estimate can be provided from the other. It would appear easy therefore to test whether trade data is accurate given the classification systems used and a knowledge of the other causes of divergences between the data. Exports are usually reported on a free on board (f. o. b.) basis and imports on a c. i. f. basis including the

*University of Edinburgh.

costs of insurance and transportation. If the insurance and freight charges were known it would be possible to measure the residual error between the c. i. f. and f. o. b. valuations of a trade flow.

For some major international trade series, methods to estimate missing values imply a belief that on average the difference between the c. i. f. value reported by an importer and the f. o. b. value reported by the exporter is around 10 % and is accounted for by transport and insurance costs and so an assumption is made that transport and insurance will account for 10 % of the value of any trade flow. This process is described in the introduction to the IMF Direction of Trade Statistics Yearbook 1981:

"Some countries either do not compile or do not publish trade flows or compile them with considerable time lags. In such cases, data from the records of partner countries are used to derive the trade of the non reporting countries. For example, if country B is a non reporting country in a given period, A's reported f. o. b. exports to B, multiplied by a factor of 1.1 to allow for costs of freight and insurance, are entered in the table for B as an estimate of B's imports from A; and A's reported imports from B, divided by a factor of 1.1 if country A reports its imports on a c. i. f. basis are entered as an estimate of B's exports to A. For 1980 such derived data account for 15 per cent of world exports and 11 per cent of world imports."

This approach can be criticised on various points. It seems to assume that freight and insurance charges would represent the same percentage of the value of A's exports to B and B's exports to A, ignoring the types of products included in trade. If trade is not symmetrical in commodity type, then using the same conversion factor for trade in both directions is not necessarily justified. The same argument applies to trade between different pairs of countries: should the same freight and insurance charges in percentage terms be applied to trade between A and B as well as A and C? Is there any evidence that over time between any pair of countries insurance and freight charges remain a constant percentage?

This paper addresses the question of whether there is any empirical justification in using a factor of 1.10 to convert trade data from a f. o. b. to a c. i. f. basis. It needs to be established whether on average there is a 10 % difference between c. i. f. and f. o. b. valuations for exports and imports. Unfortunately it is not possible here to determine how important freight and insurance charges are in the trade values. We are also faced with the question, if the cif-fob ratio is significantly different

from 1.10 what are the causes? Has there been an error in recording the data or do the insurance and freight charges diverge significantly from the average 10 % of trade value?

Causes of Discrepances

In a recent study, Yeats¹ used data that gave c. i. f. and f. o. b. values recorded by the USA for USA imports. The c. i. f. value measured the value of imports at the first port of entry in the United States and included freight, insurance and other charges excluding import duties. By comparing this data to that of partner f. o. b. export valuations he was able to isolate the transport and insurance cost component from the difference between c. i. f. and f. o. b. based valuations. These results confirm that there was a wide variation in the cif-fob ratio over a cross-section of countries exporting to the USA. For example, Kuwait overvalued exports to the USA by 259 % and Sierra Leone underestimated exports to the USA by 87 %. Large cif-fob ratios were not always associated with inaccurate data but in some cases were accounted for by high freight charges. For Surinam an observed differential of 21.4 % was largely accounted for by freight charges as a result of Surinam's export concentration on bulky, low-value items. However, out of the fifty-five countries in his sample, twenty had an unexplained residual of over 20 % over and above freight and insurance costs on exports to the USA.

Differences between c. i. f. and f. o. b. valuations, over and above transport and insurance costs, may be attributable to the varied systems used in recording trade as well as to actual errors in the data. These disparities may be the consequence of different recording definitions, notably the way in which warehoused and re-exported goods are recorded. For example *special imports* are a combined total of imports for direct consumption and withdrawals from bonded warehouses or free zones for domestic consumption, whereas the *general system* includes both imports for direct domestic consumption and imports into bonded warehouses or free zones. Therefore depending on how much of the trade is into bonded warehouses, goods exported to a country using the *special system* will appear as a different import value compared to that recorded by another country using the *general system*.

Differences will also arise depending on the valuation and classification procedures used as well as the effects of multiple exchange rates, and the time-lags between

¹ A. J. Yeats: On the Accuracy of Partner Country Trade Statistics, in: Oxford Bulletin of Economics and Statistics, Nov. 1978.

exports being recorded and delivery of the goods to the final destination. In some cases, especially for oil and petroleum products, there may be diversion of the cargo en route. Although Kuwait may record an export of oil to the USA, if the cargo is diverted en route it will be recorded as an import by a third country. Differences will also arise from errors in the recording process. These include errors in counting and recording transactions, in some cases due to difficulties in data collection. We also have to take into account deliberate falsification of trade data, bribery of customs officers and, especially in poorly policed areas, smuggling, which may be an important component of trade when there are high tariffs or total bans on certain imports or exports. Although it may not be possible to obtain a direct measure of the contribution of some or all of these items, we are not justified in assuming they do not exist, nor can it be assumed that any difference over and above 10 % (accounted for by transport and insurance) is attributable to these possible sources of error.

The Example of the West Indies

Although we are aware that these problems exist for international trade data, we are unable to make the necessary corrections until it is established for all pairs of countries over time what the average cif-fob ratios are and whether a conversion ratio of 1.10 is suitable in all cases. Yeats' study was confined to USA imports, although another study by Naya and Morgan² in 1969 examined the case of Southeast Asia. An analysis of all world trade is necessary, but studies for various regions will contribute to an understanding of the problems facing users of international trade data.

To evaluate the accuracy of international trade data and to test whether a factor of 1.10 is reasonable as an estimate of the cif-fob difference, the ratio $R_{ij} = M_{ij} / X_{ij}$ has been calculated, where M_{ij} is j 's imports from i (as reported by j) and X_{ij} is i 's exports to j (as reported by i). If the standard assumption made is correct then we would

expect that on average between any pair of countries over time, R_{ij} would not be significantly different from 1.10.

The sample used covers trade between four West Indian countries: Barbados, Jamaica, Guyana and Trinidad-Tobago; and six industrial countries: Canada, UK, France, Germany, the USA and Japan. The ratios R_{ij} for trade between these countries are calculated for the years 1969 to 1978. By including both developing and industrial countries we are able to observe how these ratios perform for intra-West Indian as well as intra-industrial country trade and trade between these two areas. Although the sample is small the results will give an indication of how the ratios can vary for trade between countries at the same, and at differing, levels of development.

For each annual trade flow covered, an R_{ij} ratio was computed; where data was not available for one of the partners the ratio was excluded. For each pair of trading partners an average R_{ij} over the ten year period was calculated. These averages are reported in the table. The averages were tested to see whether, within a 95 % confidence interval, they could have come from a population with a mean value of 1.10. In all cases where there was a significant difference between the observed average and the assumed average of 1.10 it would appear that there is little empirical evidence supporting a conversion factor of 1.10.

Overall, for the eighty-seven averages computed, only thirty-two of them (36 % of the total) were accepted as coming from a population with a mean of 1.10. The other fifty-five (63 % of the total sample) were significantly different from 1.10. Of those rejected, twenty-four came from a population with mean less than 1.10 and thirty-one from a population with mean greater than 1.10. Although there seems to be a reasonably

² S. Naya, T. Morgan: The Accuracy of International Trade Data: The Case of Southeast Asian Countries, in: Journal of the American Statistical Association, June 1969.

KONJUNKTUR VON MORGEN

The short report on domestic and world business trends and raw material markets published every fortnight by HWWA-Institut für Wirtschaftsforschung – Hamburg

Annual subscription rate DM 120,-

ISSN 0023-3439

VERLAG WELTARCHIV GMBH - HAMBURG

even number of averages belonging to populations with means above and below 1.10, this is not the case when the sample is subdivided into intra-West Indian and intra-industrial country trade.

Intra-West Indian Trade

Of the twelve averages for intra-West Indian trade, only four were accepted as coming from a population with mean of 1.10. The remaining eight were rejected, all of them coming from a population with a mean less than 1.10, their value lying within the range 0.84 and 1.05. Of the four accepted values only one was greater than 1.10. There would therefore appear to be a downward bias with respect to the expectation of an average of 1.10. With the relatively underdeveloped transport system within the West Indies one would expect an upward bias in cif-fob ratios reflecting high transport costs. The countries included in this sample are generally considered to be the more developed economies with the facilities to accommodate large ships. The smaller, less developed islands are not usually serviced by the main shipping lines, most of their trade coming via one of the more developed countries. Their import figures would reflect the much higher costs involved in off-loading and reloading onto smaller vessels, as well as higher freight rates because of the smaller size of their cargoes.

It is worthwhile to point out that low cif-fob ratios do not necessarily imply accurate data. In the case of intra-West Indian trade it is possible that negative errors have cancelled out high positive transport costs. This highlights the problems faced when data on actual transport and insurance costs are not available.

Imports from Industrial Countries

From twenty-three averages of West Indian imports from the industrial countries, eleven were accepted as coming from a population of mean 1.10. Twelve averages were rejected, all coming from populations with means greater than 1.10 and values lying between 1.14 and 1.46. In this sub-sample there appears to be an upward bias, which in many cases may be accounted for by high transport costs, but there is probably also a large proportion of error in the West Indian import data. Two of the particularly high averages, for US and Japanese exports to Trinidad are accounted for by one of the years in the sample having an uncharacteristically high value.

Exports to Industrial Countries

Out of twenty-three observations for West Indian exports to the industrial countries, thirteen were accepted as being from a population with an average of 1.10. Of the ten rejected, eight were greater than 1.10. Although it is dangerous to generalise from such a small sample we can see that for trade between the West Indies and the six industrial nations there is a bias towards average R_{ij} ratios greater than 1.10 when Canada's imports are excluded.

In some cases this upward bias may be associated with high transport costs. For example, one may expect that in the cases of Guyana and Jamaica, exports of bauxite would be subject to high freight charges. Yeats, in his study of American imports³, found transport costs for the four West Indian countries in 1974 to be 12.6 % for Jamaica, 5.5 % for Trinidad, 11 % for Guyana and 6.2 % for Barbados. Unexplained variation was -39.2 %

³ A. J. Yeats, *op. cit.*

Average R_{ij} (1969-1978)

Importer \ Exporter	Barbados	Guyana	Jamaica	Trinidad	Canada	United Kingdom	France	Germany	Japan	USA
Barbados	—	0.91	0.90	0.88	0.93	1.22	3.45 ×	11.99 ×	—	1.19 ×
Guyana	1.03 ×	—	1.12 ×	0.84	1.31 ×	1.14 ×	1.50	1.85 ×	2.36	1.21
Jamaica	1.00	1.03	—	0.97	0.93	1.10 ×	3.27	1.53	1.16 ×	1.12 ×
Trinidad	1.05	1.09 ×	1.04 ×	—	1.10 ×	1.05 ×	1.38 ×	3.79	5.74	1.01 ×
Canada	1.28	1.23	1.09 ×	1.25	—	1.14	1.45	1.34	1.19	1.05
United Kingdom	1.26	1.18	1.10 ×	1.14	1.01	—	1.02	1.01	1.15	0.99
France	1.13 ×	1.19	1.26	1.14 ×	0.98	1.12 ×	—	1.05	1.26	—
Germany	1.26 ×	1.46	1.36	1.46	1.02	1.02	0.96	—	1.14	0.98
Japan	—	0.97 ×	1.26	3.31 ×	1.01	0.84	1.50	1.60 ×	—	0.89
USA	1.33 ×	1.57 ×	1.04 ×	11.19 ×	0.98 ×	1.11 ×	1.40	1.31	1.17	—

× — Accepted within a 95 % confidence interval that R_{ij} is from a population with mean = 1.10.

Source: Basic data from United Nations Yearbook of International Trade Statistics.

for Jamaica, -7.4 % for Trinidad, 16.1 % for Guyana and 28.9 % for Barbados. Although both Jamaican and Guyanese exports were subject to higher transport costs than those facing Trinidad and Barbados, the cif-fob ratios were largely determined by the size and direction of the unexplained residual.

Intra-Industrial Country Trade

The most surprising results were for intra-industrial country trade, where only four out of twenty-nine averages were accepted as belonging to a population with mean of 1.10. It is normally assumed that trade data for industrial countries is more accurate than that from the developing world. It has to be remembered, though, that this data is only considered inaccurate because it does not conform to a population with a pre-specified mean value. All that is asserted is that these averages do not come from a population with a mean value of 1.10. Within this sub-sample, over half of the values with means less than 1.10 can be attributed to Canadian and USA imports.

As both Canada and the USA report their imports on a f.o.b. basis, insurance and freight charges will not be included. It is reasonable to assume therefore that the ratios for Canada and the USA would on average be close to one, as exports f.o.b. should equal imports f.o.b. apart from errors in recording. When tested against a population mean of one, all of the Canadian and three out of four of the US averages were accepted. Of the twenty trade flows remaining only three were accepted as coming from a population with a mean of 1.10, those rejected lying between 0.84 and 1.50.

Problems Facing Users of Trade Data

The results presented here support those of Yeats and Naya/Morgan in highlighting the nature of the problems facing users of international trade data series. Differences between c.i.f. and f.o.b. valuations vary as a result of transport and insurance costs as well as unexplained errors in the data, and there is certainly little empirical support for assuming that cif-fob ratios are on average 1.10. If it were shown that freight and insurance did account for around 10 % of the value of trade, it would then be reasonable to use 1.10 as a conversion factor in creating missing trade data. But, as Yeats has reported for USA imports, transport costs accounted for between 3.5 and 10.6 % of the value of trade for developed countries, and between 0.9 and 21.4 % of the value for the developing countries.

Until the relative importance of transport costs is isolated, data accuracy cannot be measured. For example, for Trinidadian exports to the USA in 1974,

although the total variation between c.i.f. and f.o.b. values was only -1.8 %, the unexplained variation was -7.4 %. Ratios close to 1.10 are not necessarily accurate and this may just reflect negative errors outweighing transport costs. Removing this error may produce larger cif-fob ratios, but the data could be considered more accurate.

The accuracy of trade data has obvious implications for its use in trade models. The cif-fob ratio has been used before as a proxy for transport costs. In a recent paper by Conlon⁴, nominal transport costs were defined as "the ratio of the difference between the c.i.f. and the f.o.b. values of imports, to the f.o.b. value of imports". Conlon then notes that "the insurance and other charges which are included in the c.i.f. value of imports are considered to be unavoidable elements in the total cost of international transportation and as such are treated as transport charges". If a large proportion of the cif-fob ratio is data error, this proxy will result in negative or positive biases in the measure of transport costs. When estimating import or export demand and supply equations, which of the partners' data should be used? Elasticities derived from these equations will be biased if the data is not accurate, and policy decisions based on the results may have disastrous consequences. In the world of international trade negotiations, agreements to allow third world countries access to industrial country markets have been determined in the past by the rate of growth of exports. Past rates of growth may not truly reflect what is actually happening if data is subject to large errors. These examples indicate the type of mistakes that can be made if the researcher or policy-maker is not aware of the extent of errors in trade data.

To help improve trade data a number of approaches may be suggested. A study of cif-fob ratios for all world trade should be undertaken to assess the extent to which the ratios diverge from the 1.10 rule of thumb. Detailed examination of the classification and valuation procedures for different countries could provide information on systematic biases in the data and it may then be possible to correct data taking into account the procedures used to compile the basic statistics.

In the end, measuring the accuracy of trade data is going to require detailed information on the transport and insurance costs for trade between countries at the aggregate and commodity level. Until this information is available researchers must be made aware of the implicit assumptions they may be making when using this data.

⁴ R. M. Conlon: Transport Cost and Tariff Protection of Australian and Canadian Manufacturing: A Comparative Study, in: Canadian Journal of Economics, Nov. 1981.