Contents

[Trade Misinvoicing Analysis (Mirror Analysis) 2](#_Toc9601786)

[Introduction 2](#_Toc9601787)

[Acknowledgement 2](#_Toc9601788)

[General Process 3](#_Toc9601789)

[Data Sources: 5](#_Toc9601790)

[Brief of Methodology 6](#_Toc9601791)

[Preliminary Data Treatments: 6](#_Toc9601792)

[FOB Equivalents: 6](#_Toc9601793)

[Weighted Trade Gaps: 7](#_Toc9601794)

[References 8](#_Toc9601795)

# Trade Misinvoicing Analysis (Mirror Analysis)

## Introduction

This project supplies trade misinvoicing data to GFI’s national and global IFFs reports. For national reports, tariff loss is also estimated. Powered by cloud computing, GFI’s trade misinvoicing analysis utilizes global trade data at 6-digit HS code level for each pair of countries for the past 16 years in an efficient and robust manner. Averaging 15 million transactions per year, the Comtrade database of United Nations is our main source of data. Data from various other sources is incorporated into our analysis to address re-exportation and reporting differences.

GFI is looking for financial and technical support for the development of our analysis. Please contact our data analyst at [szhao@gfintegrity.org](mailto:szhao@gfintegrity.org) if you are interested in supporting us.

Related reports:

## Acknowledgement

As an ever evolving process, GFI’s trade misinvoicing analysis keeps improving towards perfection thanks to every contributor. We hereby present our gratitude to the following individuals who contributed before the creation of this GitHub repository. Any change of methodology and mechanism from version 1.0.0 onwards doesn’t reflect the believing of previous contributors.

## General Process[[1]](#footnote-1)

Download and Clean 16-year Global Comtrade Data

Pair Trade Data and Separate Matched, Orphaned, and Lost

Prepare Data for Modeling

Build Model to Estimate CIF FOB Difference

Calculate Trade Misinvoicing

(branching global and national)

Calculate Tariff Loss

(only for national analysis)

##### Downloading and Cleaning 16-year Global Comtrade Data

GFI used the [bulk API of UN Comtrade](https://comtrade.un.org/data/Doc/api/bulk) database to download global trade data year by year to ensure that every necessary transaction was covered. Only transactions of 6-digit level HS code were included in our analysis. The following contries/regions were removed.

|  |  |
| --- | --- |
| Country/Region | Note |
| World |  |
| France excluding Monaco & overseas | UN data for France includes Monaco and overseas areas |
| France including Monaco excluding overseas | UN data for France includes Monaco and overseas areas |
| part of Europe EU, nes |  |
| Gaudeloupe | included in UN data for France |
| Martinique | included in UN data for France |
| Reunion | included in UN data for France |
| French Guiana | included in UN data for France |
| Switzerland excluding Liechtenstein | UN data for Switzerland includes Liechtenstein |
| Liechtenstein | UN data for Switzerland includes Liechtenstein |
| USA excluding Puerto Rico and Virgin Islands | UN data for USA should include PR & USVI |
| USA including Puerto Rico | UN data for USA should include PR & USVI |
| US Virgin Islands | UN data for USA should include PR & USVI |
| Puerto Rico | UN data for USA should include PR & USVI |
| US miscellaneous Pacific Islands | not in USA but likely redundant with US Minor Outlying Parnters which is kept |
| Svalbard & Jan Mayen Islands | UN data for Norway includes Svalbard & Jan Mayen Islands |
| Norway excluding Bouvet | Svalbard & Jan Mayen Islands |
| Belgium-Luxembourg aggregate 1 | UN data include Belgium and Luxembourg as separate countries |
| Belgium-Luxembourg aggregate 2 | UN data include Belgium and Luxembourg as separate countries |
| Netherlands Antilles and Aruba | UN data includes the two as separate countries |
| Saint Kitts | Nevise & Anguilla |
| British Virgin Islands & Montserrat | UN data have BVI and Montserrate reported separately by all |
| non-geographic: Neutral zones |  |
| non-geographic: Bunkers |  |
| non-geographic: Free zones |  |
| non-geographic: Special categories |  |
| Africa, nes |  |
| Rest of America, nes |  |
| Europe, nes |  |
| North & Central America, nes |  |
| Oceania, nes |  |
| Areas, nes |  |
| South America, nes |  |
| Western Asia, nes |  |

Swiss foreign trade in gold, silver and coins was added for 2001 to 2011. Trade for Slovakia and Montenegro was adjusted.

##### Pairing Trade Data and Separating Matched, Orphaned, and Lost

This step paired reporter countries and their partner countries record by record (the other 2 pairing criteria are same year and same 6-digit commodity code). Hong Kong re-export was adjusted after pairing. From paired data, GFI separated matched, orphaned, and lost transactions by the following claims for each transaction from its reporter and its partner.

|  |  |  |
| --- | --- | --- |
|  | **reporter** | **partner** |
| **matched** | Yes | Yes |
| **orphaned** | Yes | No |
| **lost** | No | Yes |

##### Preparing Data for Modeling

Data used from this step onward is a subset of annual UN Comtrade data where our preliminary requirements hold.

Preliminary Requirements:

(1) non-zero values for the trade are reported by both the reporting country and its partner;

(2) non-zero volumes for the trade are reported by both the reporting country and its partner;

(3) volumes are reported in the same physical units for both the reporting country and its partner.

GFI added regional trade agreement type[[2]](#footnote-2), geographical information[[3]](#footnote-3), and development Level[[4]](#footnote-4) data to the subset and removed every incomplete record.

##### Building Model to Estimate CIF FOB Difference

A regression model (refer to FOB equivalents for details) was built from a subset that satisfied the following criteria: (1) the mirror trade volumes differed by more than 2.5 percent, and (2) the ratio of the (implied) unit values of imports to exports was greater than 1.8 or less than 0.8.

GFI is looking for financial and technical support to utilize actual transportation and insurance cost data to upgrade this model. Please contact our data analyst at [szhao@gfintegrity.org](mailto:szhao@gfintegrity.org) if you are interested in supporting us.

##### Calculating Trade Misinvoicing

The difference between reporters’ reporting value and their partners’ reporting value on FOB bases is the gross value of trade misinvoicing. The gross value of trade misinvoicing was then weighted by a weighting scheme detailed in the weighted trade gaps section.

Advanced economies were used as a benchmark against which potential trade misinvoicing is estimated. Not every developing country reported bilateral trade with advanced economies in every year published in the DOTS and Comtrade databases.

##### Calculating Tariff Loss

For national reports, GFI used the WITS database of World Bank for tariff rates, which were based on weighted effectively applied tariff rates in TRAINS and WTO-IDB. In case of duplication, TRAINS data was used.

Tariff rates were then applied to weighted values of trade misinvoicing at 6-digit HS code level for every partner country to calculate tariff loss.

## Data Sources:

##### UN Comtrade

Illicit financial flows were estimated from a subset of annual UN Comtrade data where our preliminary requirements hold. The used time span is 16 years and only countries listed in Appendix 1 were included.

##### Census and Statistics Department of the Government of Hong Kong Special Administrative Region

Hong Kong re-export data for adjusting effect of re-export passing Hong Kong in UN Comtrade database.

##### NSF-Kellogg Institute for International Studies Data Base on Economic Integration Agreements (April 2017)

Data of types of regional trade agreement between countries from 1950 to 2012. GFI used 2012 data for every subsequent year. Types of regional trade agreement was used in a regression model to estimate gaps between CIF and FOB reporting, in other words, the cost of transportation and insurance.

##### Centre d’Études Prospectives et d’Informations Internationales (CEPII)

Geographical information used in the regression model to estimate gaps between CIF and FOB reporting.

##### Swiss Customs Administration

Swiss foreign trade in gold, silver and coins by trading partner, 1982-2013

##### World Bank WITS[[5]](#footnote-5)

Tariff rates were based on weighted effectively applied tariff rates in TRAINS and WTO-IDB. In case of duplication, TRAINS data was used.

## Brief of Methodology

### Preliminary Data Treatments:

##### Non-reporting of Swiss trade flows of precious metals in Comtrade for years prior to 2012

GFI added trade data in gold, nonmonetary, unwrought others (other than powder) from Swiss Customs Administration for 2001 to 2011.

##### Entrepôt trade through Hong Kong

GFI followed OECD’s practice in BALANCE INTERNATIONAL MERCHANDISE TRADE DATA: VERSION 1.

##### Slovakia and Montenegro

Beginning in 2006, adjust Slovakia reports to reflect Montenegro independently from Serbia-Montegro.

### FOB Equivalents:

For those countries that report import values to the United Nations on a CIF basis only, FOB equivalents are calculated using margins predicted by a regression model that utilizes geographical information, regional trade agreements, infrastructure quality (proxied by level of development based on IMF), global median price (based on reported export prices at 6-digit level), commodity classification methods, and year of trade. The training set was a subset of 2001 to 2016 Comtrade data that has been applied preliminary treatments, filtered by our preliminary requirements, and satisfied the following criteria: (1) the mirror trade volumes differed by more than 2.5 percent, and (2) the ratio of the (implied) unit values of imports to exports was greater than 1.8 or less than 0.8.

Table Independent Variables in Regression for Gaps between CIF and FOB

|  |  |  |  |
| --- | --- | --- | --- |
| **Mnemonic** | **Variation** |  | **Description (Source)** |
| ln\_distw | {i,j} |  | log of the weighted distance between i & j (CEPII) |
| ln\_distw\_squared | {i,j} |  | ln\_distw x ln\_distw |
| log(uvmdn) | {t,k} |  | log of the median "world" price of good k at time t |
| d\_contig | {i,j} |  | dummy = 1 if i and j share a border (CEPII) |
| d\_conti | {i,j} |  | dummy = 1 if i and j are on the same continent (CEPII) |
| d\_rta | {t,i,j} |  | categorical variable for type of regional trade agreement between i and j in year t (Bergstrand & Baier (2015), extended to 2015 by GFI) |
| d\_landlocked\_i | {i} |  | dummy = 1 if i is landlocked (CEPII) |
| d\_landlocked\_j | {j} |  | dummy = 1 if j is landlocked (CEPII) |
| d\_dev\_i | {i} |  | dummy = 1 if i is a developing economy (IMF) |
| d\_dev\_j | {j} |  | dummy = 1 if j is a developing economy (IMF) |
| d\_hs\_diff | {t,i,j} |  | dummy = 1 if i and j used different HS classification in year t |
| d\_2001 ... d\_2015 | {t} |  | dummy = 1 for corresponding years |

## Weighted Trade Gaps:

Discrepancies in the volumes reported for each mirrored trade are then used to weight the raw trade gaps: a higher weight is applied to a given discrepancy in value the smaller is the associated volume discrepancy. The use of weighted measures (rather than the raw trade gaps) in the Comtrade estimates is intended to improve the reliability of the trade misinvoicing estimates. The weighting scheme is described in formal terms as follows: let QD and QA denote, respectively, the reported volume of trade (of a particular good in a particular year) between a developing country reporter (D) and its advanced-country trade partner (A). The weight applied to the trade gap in value terms was specified as



## References

Census and Statistics Department of the Government of Hong Kong Special Administrative Region. (n.d.). Entrepôt trade through Hong Kong. Hong Kong, Hong Kong, China. Retrieved 5 2017

Centre d’Études Prospectives et d’Informations Internationales. (n.d.). Geographical Information.

Swiss Customs Administration. (n.d.). *Swiss foreign trade in gold, silver and coins by trading partner, 1982-2013.* Retrieved 3 2015, from Swiss Customs Administration: https://www.ezv.admin.ch/dam/ezv/en/dokumente/abgaben/Aussenhandelstatistik/Diffusion/Daten/Waren/2\_3\_Gold\_LD.xlsx.download.xlsx/2\_3\_GOLD\_LD\_EXPIMP\_en.xlsx

United Nations Statistics Division. (n.d.). *UN Comtrade Database*. Retrieved 2 2019, from UN Comtrade Database: https://comtrade.un.org/

University of Notre Dame, Kellogg Institute For International Studies. (n.d.). *NSF-Kellogg Institute Data Base on Economic Integration Agreements.* Retrieved from https://kellogg.nd.edu/nsf-kellogg-institute-data-base-economic-integration-agreements

World Bank. (n.d.). *Tariff and Trade Analysis.* Retrieved from World Integrated Trade Solution: https://wits.worldbank.org/WITS/WITS/Restricted/Login.aspx

Appendix Countries Covered in GFI’s Analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Developing Economies (148)** | | | | | **Advanced Economies (36)** |
| **Sub-Saharan Africa (45)** | **Asia (25)** | **Developing Europe (23)** | **Middle East & N. Africa (22)** | **Western Hemisphere (33)** |
| Angola | Bangladesh | Albania | Afghanistan, Islamic Republic of | Antigua and Barbuda | Australia |
| Benin | Bhutan | Armenia, Republic of | Algeria | Argentina | Austria |
| Botswana | Brunei Darussalam | Azerbaijan, Republic of | Bahrain, Kingdom of | Aruba | Belgium |
| Burkina Faso | Cambodia | Belarus | Djibouti | Bahamas, The | Canada |
| Burundi | China, P.R.: Mainland | Bosnia and Herzegovina | Egypt | Barbados | Cyprus |
| Côte d'Ivoire | Fiji | Bulgaria | Iran, Islamic Republic of | Belize | Czech Republic |
| Cabo Verde | India | Croatia | Iraq | Bolivia | Denmark |
| Cameroon | Indonesia | Georgia | Jordan | Brazil | Estonia |
| Central African Republic | Kiribati | Hungary | Kuwait | Chile | Finland |
| Chad | Lao People's Democratic Republic | Kazakhstan | Lebanon | Colombia | France |
| Comoros | Malaysia | Kyrgyz Republic | Libya | Costa Rica | Germany |
| Congo, Democratic Republic of | Maldives | Macedonia, FYR | Mauritania | Dominica | Greece |
| Congo, Republic of | Mongolia | Moldova | Morocco | Dominican Republic | Hong Kong |
| Equatorial Guinea | Myanmar | Montenegro | Oman | Ecuador | Iceland |
| Eritrea | Nepal | Poland | Pakistan | El Salvador | Ireland |
| Eswatini, Kingdom of | Papua New Guinea | Romania | Qatar | Grenada | Israel |
| Ethiopia | Philippines | Russian Federation | Saudi Arabia | Guatemala | Italy |
| Gabon | Samoa | Serbia, Republic of | Sudan | Guyana | Japan |
| Gambia, The | Solomon Islands | Tajikistan | Syrian Arab Republic | Haiti | Korea, Republic of |
| Ghana | Sri Lanka | Turkey | Tunisia | Honduras | Latvia |
| Guinea | Thailand | Turkmenistan | United Arab Emirates | Jamaica | Lithuania |
| Guinea-Bissau | Timor-Leste, Dem. Rep. of | Ukraine | Yemen, Republic of | Mexico | Luxembourg |
| Kenya | Tonga | Uzbekistan |  | Nicaragua | Malta |
| Lesotho | Vanuatu |  |  | Panama | Netherlands |
| Liberia | Vietnam |  |  | Paraguay | New Zealand |
| Madagascar |  |  |  | Peru | Norway |
| Malawi |  |  |  | St. Kitts and Nevis | Portugal |
| Mali |  |  |  | St. Lucia | San Marino |
| Mauritius |  |  |  | St. Vincent and the Grenadines | Singapore |
| Mozambique |  |  |  | Suriname | Slovak Republic |
| Namibia |  |  |  | Trinidad and Tobago | Slovenia |
| Niger |  |  |  | Uruguay | Spain |
| Nigeria |  |  |  | Venezuela, Republica Bolivariana de | Sweden |
| Rwanda |  |  |  |  | Switzerland |
| Sao Tome and Principe |  |  |  |  | United Kingdom |
| Senegal |  |  |  |  | United States |
| Seychelles |  |  |  |  |  |
| Sierra Leone |  |  |  |  |  |
| Somalia |  |  |  |  |  |
| South Africa |  |  |  |  |  |
| Tanzania |  |  |  |  |  |
| Togo |  |  |  |  |  |
| Uganda |  |  |  |  |  |
| Zambia |  |  |  |  |  |
| Zimbabwe |  |  |  |  |  |
| Note: The designation of an economy as either "advanced" or "developing" is determined by the International Monetary Fund. | | | | | |

1. Refer to the master flow chart for detailed process from a technical aspect. [↑](#footnote-ref-1)
2. Source: NSF-Kellogg Institute for International Studies Data Base on Economic Integration Agreements, April 2017 version [↑](#footnote-ref-2)
3. Source: CEPII [↑](#footnote-ref-3)
4. Source: IMF [↑](#footnote-ref-4)
5. Only for national reports. [↑](#footnote-ref-5)