

# Trade, Production and Protection 1976-2004<sup>\*</sup>

Alessandro Nicita<sup>°</sup>  
Marcelo Olarreaga<sup>\*</sup>

October 2006

## Abstract

The database described herein provides researchers with a broad set of data covering trade, production and protection data for 28 manufacturing sectors corresponding to the 3-digit level International Standard Industrial Classification (ISIC), Revision 2. The database covers potentially 100 developing and developed countries over the period 1976-2004, but data availability varies by country and year. The Trade, Production and Protection database is available online and can be freely accessed through the World Bank trade website.<sup>⊗</sup>

---

<sup>\*</sup> We are grateful to three anonymous referees, the Editor of the World Bank Economic Review, and many users of the earlier release of the database for their comments and suggestions. We also thank the World Bank's Research Support Budget for funding.

<sup>°</sup> Development Economics Research Group (Trade), World Bank, Washington, D.C. 20433, USA; Tel. (202)473-4066; Fax: (202)522-1159; email: [anicita@worldbank.org](mailto:anicita@worldbank.org).

<sup>\*</sup> Office of the Chief Economist for Latin America, World Bank, Washington, D.C. 20433, USA, and CEPR, London, UK; Tel. (202)458-8021; Fax: (202)522-1159; email: [molarreaga@worldbank.org](mailto:molarreaga@worldbank.org)

<sup>⊗</sup> The database is available under the "Data & Statistics" section of the World Bank trade website ([www.worldbank.org/trade](http://www.worldbank.org/trade)).

The database described in this note includes annual data on trade flows (exports and imports), domestic production (output, value-added, employment, etc), and trade protection (tariffs and non-tariff barriers) for 100 countries over the period 1976 to 2004. The main contribution of this dataset is to have merged data from different sources in a common industry classification. The data is disaggregated into 28 manufacturing sectors, corresponding to the 3-digit level of the International Standard Industrial Classification (ISIC), Revision 2. The database is freely available for download through the World Bank trade website ([www.worldbank.org/trade](http://www.worldbank.org/trade)) under the “Data & Statistics” section.

The database updates the earlier release made available in Nicita and Olarreaga (2001). Besides the longer time coverage, the data contained in the database has been improved in a number of ways: the coverage has increase to cover 100 countries, the concordance table between SITC revision 2 and ISIC revision 2 has been updated, and more variables have been added (non-tariff barriers and their ad-valorem equivalents, elasticities of import demand, import price indices, import-weighted tariffs and their standard deviation).<sup>1</sup> A supplemental appendix to this article describes the different dimensions of the database, variable definition and data availability in more detail. It can be found at [wber.oxfordjournals.org](http://wber.oxfordjournals.org).

Researchers may find it useful to utilize this dataset jointly with other datasets available on the web. A particularly useful website for trade related data is Jon Haveman and Raymond Robertson’s webpage at [www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html](http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html). Another source of data are the web pages of the Centre d’Etudes Prospectives et d’Informations Internationales at [www.cepii.fr](http://www.cepii.fr), in particular the extension by Mayer and Zignago (2005) to the earlier release of the Trade and Production database, available at [www.cepii.fr/anglaisgraph/bdd/TradeProd.htm](http://www.cepii.fr/anglaisgraph/bdd/TradeProd.htm). Another wide collection of trade and protection data is available at the Center for International Data at UC Davis at [cid.econ.ucdavis.edu](http://cid.econ.ucdavis.edu).

---

<sup>1</sup> The database has also been made more compact by not including the ISIC 4 digit classification.

## **I. Production**

The source of domestic production related data is the United Nations Industrial Development Organization (UNIDO), which collects annual data from its member countries. This data is published yearly in the International Yearbook of Industrial Statistics using the 3-digit level of the ISIC, Revision 2 classification. The production related data of this database includes information on output, value added, gross fixed capital formation, index of real output, wage bill, number of establishments, number of employees and number of female employees for each of the 28 manufacturing sectors. Unfortunately, the country and time coverage of the production related data is limited. Some countries, especially in the developing world, only report data sporadically, and when they do they may only report some of the variables mentioned above. UNIDO makes a great effort to standardize the data and make it comparable across countries and years, however some issues still persist. Yamada (2005) provides detailed information on the UNIDO data and issues related to its use in research projects.

## **II. Trade**

The source of trade data is the COMTRADE database kept by the United Nations Statistic Division (UNSD). The trade data contained in the Trade, Production and Protection database is originally provided by COMTRADE using the Standard International Trade Classification (SITC) Revision 2 classification. This data is then converted into the ISIC Revision 2 classification using a concordance table. Trade data contains exports and imports information and is reported both at the aggregate and bilateral level. The country and time coverage is very complete. However, there are some missing observations, especially among developing countries. To fill missing observations, researchers often resort to mirrored data.<sup>2</sup> Mirrored data uses partner data. For example, export data of country *C* is calculated using imports from country *C* reported by its trading partners. The database provides mirrored data for both imports and exports. It is left to the final user to choose if, and how to use mirrored data.

---

<sup>2</sup> An example is Feenstra *et al.* (2005).

Trade data contains information on the value of shipments (in thousand of US dollars), but also physical quantities (in Kilograms). The database also contains unit values, measured in dollars per kilo. They are calculated as ratio of the value of shipments and physical quantities at the 3-digit level of the ISIC, Revision 2 classification.<sup>3</sup> They are provided for exports and imports at the aggregate and bilateral level. Important cautionary notes regarding the use of mirrored data and unit values are provided in Section V.

### **III. Protection**

The protection data in this database includes data on tariffs and non-tariff barriers (NTB). The main source of protection data is UNCTAD's TRAINS, but part of the data was collected by the authors through national statistical documents and websites. Protection data availability starts in 1988. This data originally comes at the 6-digit level of the Harmonized System (HS) 6 digit, and it is converted to the 3-digit level of the ISIC Revision 2 using a concordance table. The country and time coverage of the protection data is far from complete, especially for the NTB data.

The tariff data contains simple and import weighted average tariffs for the 28 manufacturing sectors. It also reports their standard deviation, and the maximum and minimum values at the 6-digit level of the HS within each ISIC code. These indicators are reported for two types of tariffs: Applied and Most Favored Nation (MFN). MFN rates are those granted to all WTO members to whom no preferential access is granted. Applied rates take into consideration the available data for preferential schemes, and therefore are calculated at the aggregate level using bilateral tariff and import data.<sup>4</sup>

Non-tariff barriers' (NTB) data is reported as a single category "Core NTB" which includes price-control measures, finance-control measures, and quantity-control

---

<sup>3</sup> For about 5 percent of products at the SITC 5 digit level, quantities are not reported at all or reported in units rather than using the metric system. The aggregate quantities at the ISIC level do not take into account these products.

<sup>4</sup> Bilateral protection data covers most preferential trade agreements, but not all. Researchers interested in the bilateral tariffs should refer directly to the UNCTAD TRAINS database, which is available through WITS ([wits.worldbank.org](http://wits.worldbank.org)). An alternative source of bilateral tariff data that includes a careful calculation of ad-valorem equivalents of specific tariffs is Bouet *et al.* (2005).

measures. NTB data is reported using coverage ratios (the percentage of import subjected to NTB) and frequency ratios (the percentage of tariff lines subjected to Core NTB). This database also provides simple and import-weighted AVE of NTB at the 3-digit level of the ISIC. The methodology to obtain the AVE of NTB is described in detail in Kee, Nicita and Olarreaga (2006).

#### **IV Other data**

Trade, production and protection data is often used in the context of gravity models. To facilitate researchers working with these types of models, this database includes a number of “gravity” type variables. The variables included are: geodesic distance between nations capitals, language, GDP, GDP per capita (PPP adjusted), and dummies for common language, shared border, landlocked, and island. Part of this data comes directly from the World Bank’s World Development Indicators, and part has been collected and constructed during the years by the authors. Researchers working with gravity models may find additional data at the CEPII’s web site:

<http://www.cepii.fr/anglaisgraph/bdd/distances.htm>, and at Andrew Rose’s web site: <http://faculty.haas.berkeley.edu/arose/RecRes.htm#Software>.

The trade, production and protection database includes also import demand elasticities for each country and each of the ISIC codes at one point in time. This data may be particularly useful for researchers interested in simulation exercises. The methodology used to estimate the import demand elasticities is described in detail in Kee, Nicita and Olarreaga (2004). Import demand elasticities are provided in a separate file.

The database also contains information on input-output tables based upon the Global Trade Analysis Project (GTAP) database version 4, which is based on data from the early '90s.<sup>5</sup> The GTAP 4 database aggregates some countries into regions; therefore, countries in the same GTAP region will have the same input-output table. To give more flexibility to the user of these tables, the data is broken down into two different tables.<sup>6</sup>

---

<sup>5</sup> The GTAP database is now in its sixth release. By using newer releases of the GTAP database the interested user can build updated input-output tables. For information on how to access the GTAP databases visit [www.gtap.org](http://www.gtap.org).

<sup>6</sup> The supplemental appendix provides more detail on the construction of these tables.

The first table reports the share of each manufacturing sector output that is sold as an input to the production of each sector. The second table reports the amount of each of the other sectors' output necessary to produce one unit of final output in each sector. Note that, because the GTAP industry disaggregation does not exactly match the ISIC 3-digit level industry disaggregation, the input-output data is provided at a higher level of aggregation than the 3-digit level of the ISIC, Revision 2.

## **V. Special considerations**

The data in the Trade, Production and Protection Database has been organized to facilitate its use for a large number of purposes. Its objective is not to produce quick answers, but rather to help researchers in the lengthy and cumbersome exercise of collecting and organizing data into a common classification. For a meaningful use of the database it is important that researchers are aware of its limitations.

Most importantly, the database is an unbalanced panel, as it contains a large number of missing observations. As such, it needs to be used with care; e.g., industry or country averages may not be very meaningful if they correspond to different time periods, or contain different countries in different years. Of course, it is always possible to interpolate missing data, but the decision is left to the final user to interpolate or not and how to do it.

A second issue relates to the use of mirrored data. In theory, export data is recorded as free on board (FOB), while import data is recorded as cost, insurance and freight (CIF). Therefore, it may be appealing to use the difference between the bilateral import value and its corresponding export value to impute trade costs. In practice, the differences between a trade flow (whether imports or exports) and its mirrored counterpart should not be considered a good measure of trade costs, especially for countries where the custom capacity is weak. In many cases the discrepancies are usually attributable to a series of reasons besides trade costs, such as different product classifications, existence of entrepôts, weak accounting methods, under-invoicing and custom corruption. Moreover, in the case of trade aggregates, discrepancies may be due

to missing reporting partners in the mirrored data (i.e., not all partners are reporting trade to COMTRADE).

Another issue when using trade data is the existence of *entrepôts* (countries where transits of trade take place but which do not constitute the origin or the final destination of these) which may again create accounting discrepancies between reported data and mirrored data (Hanson and Feenstra, 2001). In some cases the country of origin (O) mistakenly reports the *entrepôt* (T) as the destination of the shipment. Meanwhile, correctly, the *entrepôt* country (T) does not report the import, and the final importer (F) reports the original exporter (O) as the origin. When comparing bilateral imports and exports this creates discrepancies. In the example above, country (F) reports an import from (O), which is not reported (as an export to F) by country (O). The researcher should keep this in mind especially when analyzing bilateral trade flows that may interest *entrepôts* such as Hong Kong, Macao, Singapore and the Netherlands.

Note however, that in some cases, mirrored data may be considered of better quality. One reason for this is that the partner may have a much better customs administration relative to the reporter. See Yates (1995) for a discussion.

Another important consideration relates to a few cases (about 1 percent of observations) where the value of exports is larger than the sum of output plus imports. There are several explanations for this. First, there may be discrepancies between the year of production and trade flows; goods that were produced last year may be exported this year. Thus, declines in stocks or delay in exports can partly explain it. Second, there may be a misallocation of production data across ISIC categories. Finally, for some countries the reported production data may exclude a significant portion of industrial activity, either because the coverage of small-scale establishments may be incomplete, the data may refer only to a certain area of the country (urban), or may refer to only part of the manufacturing sector (i.e. excluding any informal sector). Thus, measurement error should be a concern.

There are some further issues regarding protection data. First, while the applied tariff data takes into account preferential schemes of developed countries, some of the smaller agreements of developing countries may not be included. This issue is usually more relevant for early years. Second, preferential schemes are not always fully utilized

(partly due to the inability to meet origin requirements), but there is no systematic information on preference utilization rates and therefore the data assumes that these are fully utilized. Third, tariff data only includes the ad-valorem component of tariff schedule, i.e., it does not provide ad-valorem equivalents of specific duties. This is not a major omission as the database focuses on manufacturing where specific duties are rare. Fourth, the database provides simple and import-weighted average tariffs, however none of these has a sound theoretical basis as measures of trade restrictiveness. For example, in calculating import weighed average tariffs, goods subject to prohibitively high tariffs have zero weight, underestimating trade restrictiveness. Similarly, when computing simple average tariffs, very low tariffs on economically meaningless goods would downward bias this measure of trade restrictiveness. For a detailed discussion and some solutions to these problems see Kee, Nicita and Olarreaga (2006).

There is also an important caveat regarding the use of unit values. Unit values are useful to analyze many international trade issues, in particular price competitiveness. However, they should be used with caution as they are a noisy proxy for prices.<sup>7</sup> This is particularly true for large product aggregates, as the ones provided in this database. The main reason are changes in product quality or product mix as these can affect average unit values. There is no straightforward solution to all these problems, rather a need for awareness and caution when using the data. In econometric work, one way out of some of these issues is to instrument unit prices to get rid of measurement error.

## **V. Technical information**

The database is free of charge and available by accessing the World Bank Trade website ([www.worldbank.org/trade](http://www.worldbank.org/trade)) under “Data & Statistics”. Most of the data is stored as ASCII files and can be read with any text editor or statistical software. A supplemental appendix to this article describes the different dimensions of the database, variable definition and data availability in more detail. It can be found at [wber.oxfordjournals.org](http://wber.oxfordjournals.org) or at World Bank Trade website.

---

<sup>7</sup> In practice, sudden jumps in the time series of unit values, when not substantiated by other data, are usually an indication of a change in the accounting/recording methods.

## References

- Bouet, Antoine, Yvan Decreux, Lionel Fontagné, Sébastien Jean, and David Laborde (2004), “A consistent, ad-valorem equivalent measure of applied protection across the world: The MacMap-HS6 database”, CEPII Working Paper No. 2004-22.
- Feenstra, Robert, Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo, (2005) “World Trade Flows: 1962-2000”. NBER Working Paper No. 11040.
- Hanson, Gordon H. and Robert C. Feenstra, (2001), “Intermediaries in entrepôts trade: Hong Kong re-exports of Chinese goods”. NBER Working Paper Series No. 8088.
- Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga (2004), “Import demand elasticities and trade distortions”, World Bank Policy Research Working Paper No. 3452.
- Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga (2006), “Estimating trade restrictiveness indices”, World Bank Policy Research Working Paper No. 3840.
- Mayer, Thierry and Soledad Zignago (2005), “Market access in global and regional trade”, CEPII Working Paper No. 2005-02.
- Nicita, Alessandro and Marcelo Olarreaga, (2001), “Trade and Production: 1976-1999,” World Bank Policy Research Working Paper No. 2701.
- Yamada Tetsuo (2005) “Relevance and Applicability of the UNIDO Industrial Statistics Database for Research Purposes”, UNIDO ESA/STAT/AC.105/21.
- Yeats, Alexander (1995), “Are partner country statistics useful for estimating "missing" trade data?”, World Bank Policy Research Working Paper # 1501.

## **Appendix**

This contains information on each of the files that are made available, a table on data availability and variable names.

### **Files Description**

#### **TPP\_Data.ZIP**

The database at the 3 digit ISIC level by country and year (zip archive, comma separated file).

#### **COVERAGE.xls**

Excel tables on data availability, variable coverage, and variables description for the Trade, Production and Protection Database (MS Excel) that go beyond what is provided in this supplemental appendix.

#### **BILATERAL.ZIP**

Bilateral trade data (zip archive, comma separated file).

#### **GravityData.ZIP**

Other data. Contains the following files (zip archive, comma separated file)

TPP\_ISICnames – ISIC rev2 codes and product description.

TPP\_countrynames – Country codes and full names.

TPP\_region\_inclv1 – geographic region and income level by country.

TPP\_SITCtoISIC – Concordance SITC rev2 to ISIC rev2.

TPP\_HStoISIC – Concordance HS88 to ISIC rev2.

TPP\_GDPdata – Country GDP data time series.

TPP\_languagesvector – language by country.

TPP\_island\_landlocked – Dummies for island and landlocked status, by country.

TPP\_border – Shared border, by country pair.

TPP\_distance – Bilateral distance, by country pair (kilometers).

#### **IDE\_ISIC.CSV (comma separated file).**

Contains import demand elasticities (ID\_elast), by country and ISIC sector, and their standard error (ID\_elast\_se)

#### **IOTABLES.zip**

TPP\_IOTables – Input-output tables. (MS excel file).

## Data Availability

Country Name	Code	Production	TRADE	Protection (tariff)	Protection (NTB)
Algeria	DZA	76-03	76-04	93,97,98,01-03	97,01
Argentina	ARG	76-02	80-04	92-04(94)	99,01
Armenia	ARM	86-03	97-04(98,01)	01	
Australia	AUS	76-01	79-04	91-04(02,94-95)	99
Austria	AUT	76-00	78-04	90,95-03	99
Azerbaijan	AZE	90-02	96-04	02	
Bangladesh	BGD	76-03	77-04(99)	89,94,99-04(01)	00
Belgium-Luxemburg	BLX	76-03	78-04	88-03	99
Benin	BEN	76-99 (94-95)	92-02	01-04	
Bolivia	BOL	76-02	77-04	93-04(03)	99,01
Botswana	BWA	76-02	00-01	01,04	
Brazil	BRA	76-03	83-04	89-04	99,01
Bulgaria	BGR	76-03	92,96-04	01-04(02)	
Cameroon	CMR	76-02	76-04 (81,83-85,88,91-94)	94,95,01,02	97
Canada	CAN	76-01	78-04	89,93-04(94)	00
Chile	CHL	76-02	83-04	92-04(96,03)	99,01
China	CHN	77-97	85-04	92-04(95,02)	97,01
Colombia	COL	76-00	78-04	91-04(93,98,03)	97,01
Costa Rica	CRI	76-02	86-04	95,99-04(03)	98
Cote D'Ivoire	CIV	76-02	76-03(80,84,86-94,01)	93,96(01-04)	01
Cyprus	CYP	76-03	76-04	02	
Czech Republic	CZE	80-03(94)	93-04	96,99,02,03	99
Denmark	DNK	76-98	76-04	88-03	99
Ecuador	ECU	76-02	78-04	93-04(00,01,03)	97,01
Egypt	EGY	76-02	81-04	95,98,02	97,01
El Salvador	SLV	76-03	86-04	95,97,98,00,01,02,04	97
Ethiopia	ETH	76-02	93-03(94,96)	95,01,02	95
Finland	FIN	76-00	76-04	88,95-03	99
France	FRA	76-00	78-04	88-03	99
Gabon	GAB	76-02	76-04(84-92,95)	95,98,01,02	94
Germany (76-90 West)	DEU	91-03	78-04	88-03	99
Ghana	GHA	76-00	92-04(93-95)	93,00,04	95
Greece	GRC	76-03	76-04	88-03	99
Guatemala	GTM	76-98	86-04	95-04(96,99,03)	88
Honduras	HND	76-03	86-04	95-04(96-98,03)	98
Hong Kong	HKG	76-03	78-04	88-03	94
Hungary	HUN	76-03	76-04	91,93,96,97,02	99
Iceland	ISL	76-96	77-04	93,96,01,03	96
India	IND	76-01	78-04	90,92,97,99,01	97
Indonesia	IDN	76-02	80-04	89,90,93,95,96,99-04	99
Iran	IRN	76-02	97-04	00,03,04	
Ireland	IRL	76-02	76-04	88-03	99
Israel	ISR	76-02	81-04	93,04	
Italy	ITA	76-00	77-04	88-03	99
Japan	JPN	76-03	76-04(93)	88-04	96,01
Jordan	JOR	76-03	81-04(96)	00-03	01
Kenya	KEN	76-03	80-04(89)	94,00,01,04	93
Korea	KOR	76-02	76-04	88-90,92,94,95,99,02,04	96
Kuwait	KWT	76-02	87-01	02	
Kyrgyzstan	KGZ	87-02	95-04(97)	02,03	
Latvia	LVA	86-03	94-04	96,97,01	96
Lithuania	LTU	90-03(91)	92-04(93)	95,96,01	99
Macau	MAC	76-02	76-04	88-03	
Malawi	MWI	76-03	77-04(92-93)	94,96-98,01	
Malaysia	MYS	76-02	78-04	88,91,93,96,97,01-03	96,01
Malta	MLT	76-01	90-04	98,01,02,03	
Mauritius	MUS	76-03	80-04	95,97,98,02	95
Mexico	MEX	76-02	89-04	91,95,97-03	99,01
Moldova	MDA	86-02	94-04	96,00,01	
Mongolia	MNG	76-02	96-03		
Morocco	MAR	76-03	76-04	93,97,00-03	99,01
Mozambique	MOZ	76-00	94-02(98)	01-03	
Myanmar	MMR	76-03	93	01-04	
Nepal	NPL	76-94,96,02	82-03(00-01)	93,98-00,02-04	
Netherlands	NLD	76-03	78-04	88-03	99
New Zealand	NZL	76-03	79-04	92,93,96-00,02-04	99
Nigeria	NGA	76-02	91-03(92-95)	88-90,92,95-02	01
Norway	NOR	76-01	76-04	88,93,95,96,98,00-03	96
Oman	OMN	90-03(91-92)	81-04	92,97,02	99
Pakistan	PAK	76-02	82-04(94)	95,98,01-04	
Panama	PAN	76-02	86-04	97-01(99)	
Peru	PER	79-96(93)	76-04	93,95,98-00,04	99,01
Philippines	PHL	76-97	77-04	88-03(91,96,97)	98,01
Poland	POL	76-00	80-04	91,92,95-03	99
Portugal	PRT	76-00	79-04	88-03	99
Qatar	QAT	80-02	81-04(86-87,97)	02	
Romania	ROM	76-03	89-04	91,99,01,04	99
Russian Federation	RUS	90-02	96-04	93,94,96,97,01,02	97
Senegal	SEN	76-03	77-04(82-85,88)	01-04	01
Singapore	SGP	76-03	79-04	89-03	92,01
Slovakia	SVK	90-99 (95-96)	94-04	02	
Slovenia	SVN	84-03	92-04	99,01-03	99
South Africa	ZAF	76-03	76-04(86-91)	88,90,91,93,96,97,99,01,04	99
Spain	ESP	76-00	78-04	88,03	99
Sri Lanka	LKA	76-01	79-04(95-98,00)	90,93,94,97,00,01,04	
Sweden	SWE	76-03	76-04	88-89,95-03	99
Switzerland	CHE	76-03	76-04		96
Taiwan	TWN	76-01	89-04	89,92,96,99-03	99,01
Tanzania	TZA	76-02	95-04	93,97,98,00,03	01
Thailand	THA	76-02	76-03(02)	89,91,93,95,00,01,03	94,01
Trinidad and Tobago	TTO	76-02	79-03	91,92,96,99,01-03	
Tunisia	TUN	76-03	80-04	90,92,95,98,02-04	99
Turkey	TUR	76-00	85-04	93,95,97,99,03	97
Uganda	UGA	76-02	94-04	94,00-04	
Ukraine	UKR	87-01	96-03	95,97,02	97
United Kingdom	GBR	76-03	78-04	88-03	99
United States	USA	76-01(98)	78-04	89-04(94)	99
Uruguay	URY	76-02	83-04	92,95-04(03)	99,01
Venezuela	VEN	76-01	82-04	92,95,97-00,02,04	99,01
Yemen	YEM	76-01	76-04(77,82-90,92-94)	00	

Note: The symbol "--" divides the first and last year of a time series, the comma divides single years. For the years in parentheses the data is not available. More detailed information is made available in the file coverage.xls.

## **Variables Description:**

### ***Production variables:***

**Value Added (value\_added)** for an industry is that portion of sales that is not accounted for by the use of inputs and supplies from other industries. *Gross Value added* for a particular industry represents its contribution to national GDP. It is sometimes referred to as GDP by industry. Value added is reported thousand dollars

**Industrial Production Index (ind\_prod\_index)** represents the physical output of a *fixed basket* of goods, measured in volume terms, such as tons or thousands of units. It is generally provided directly from national sources, but it is also sometimes estimated by UNIDO as a “commodity-based” indicator, in which output of specific commodities are counted and added together. The base year of the Industrial Production Index is different across countries, and it sometimes varies within countries and across sectors. Also, the disaggregation of the industrial production index is sometimes less disaggregated than the ISIC 3 digit revision 2, which implies that for some sectors at the 3 digit of the ISIC the series may be identical.

**Number of Establishments or Enterprises (n\_establ)** represents the number of entities operating in the sector. The data does not allow differentiating between establishments and enterprises. An “establishment ” is ideally a unit that engages, under a single ownership or control, in one, or predominantly one, kind of activity at a single location; for example, workshop or factory. An “enterprise” is a legal entity possessing the right to conduct business in its own name; for example, to enter into contracts, own property, incur liability for debts, and establish bank accounts. Number of Establishment or Enterprises is reported in units.

**Number of Employees (n\_employees)** is defined as the total number of persons who worked in or for the establishment during the reference year. The figures reported refer normally to the average number of persons engaged during the reference year. The number of employees is including all persons engaged other than working proprietors, active business partners and unpaid family workers. Number of Employees is reported in units.

**Number of Female Employees (n\_female\_empl)** is defined as the total number of female employees who worked in or for the establishment during the reference year. Number of Female Employees is reported in units.

**Wages and Salaries (wage\_bill)** include all payments in cash or in kind paid to "employees" during the reference year in relation to work done for the establishment. Payments include: (a) direct wages and salaries; (b) remuneration for time not worked; (c) bonuses and gratuities; (d) housing allowances and family allowances paid directly by the employer; and (e) payments in kind. Excluded are employers contributions in respect of their employees paid to social security, pension and insurance schemes, as well as the

benefits received by employees under these schemes and severance and termination pay. Wages and Salaries is reported thousand dollars.

**Output (output)** represents the value of goods produced in a year, whether sold or stocked. Output is reported thousand dollars.

**Gross Fixed Capital Formation (gr\_dix\_cap\_form)** refers to the value of purchases and own-account construction of fixed assets during the reference year less the value of corresponding sales. The fixed assets covered are those (whether new or used) with a productive life of one year or more. Gross Fixed Capital Formation is reported thousand dollars.

### *Trade variables*

**Import Values (imp\_tv)** represents the value imports entering the reporting country. Import values are reported in thousand dollars.

**Import Quantities (imp\_q)** represents the weight of the goods entering the reporting country. Import Quantities are reported in kilograms.

**Import Unit Values (imp\_uv)** represents the average unit value of the goods entering the reporting country. Import Unit Values are reported in dollars per kilogram.

**Export Values (exp\_tv)** represents the value of exports of the reporting country. Export values are reported in thousand dollars.

**Export Quantities (exp\_q)** represents the weight of the goods exported by the reporting country. Export Quantities are reported in kilograms.

**Export Unit Values (exp\_uv)** represents the average unit value of the goods exported by the reporting country. Export Unit Values are reported in dollars per kilogram.

**Mirrored Import Values (mir\_imp\_tv)** represents the value imports entering the reporting country observed as exports from partner countries. Mirrored Import values are reported in thousand dollars.

**Mirrored Import Quantities (mir\_imp\_q)** represents the weight of the goods entering the reporting country observed as exports from partner countries. Mirrored Import Quantities are reported in kilograms.

**Mirrored Import Unit Values (mir\_imp\_uv)** represents the average unit value of the goods entering the reporting country observed as exports from partner countries. Mirrored Import Unit Values are reported in dollars per kilogram.

**Mirrored Export Values (mir\_exp\_tv)** represents the value of exports of the reporting country observed as imports of partner countries. Mirrored Export values are reported in thousand dollars.

**Mirrored Export Quantities (mir\_exp\_q)** represents the weight of the goods exported by the reporting country observed as imports of partner countries. Mirrored Export Quantities are reported in kilograms.

**Mirrored Export Unit Values (mir\_exp\_uv)** represents the average unit value of the goods exported by the reporting country observed as imports of partner countries. Mirrored Export Unit Values are reported in dollars per kilogram.

***Trade Protection Variables:***

**Simple Applied Tariff (tar\_savg\_AHS)** represents the simple average applied tariff rate applied on goods entering the country. Applied rates take into consideration the available (*however, not complete*) data for preferential schemes (i.e. the applied average tariff takes the tariff rates for each partner that export to the market country in constructing the average). Applied Tariff is reported in percentage points.

**Weighted Applied Tariff (tar\_iwAHS)** represents the import weighted average applied tariff rate applied on goods entering the country. Applied rates take into consideration the available (*however, not complete*) data for preferential schemes (i.e. the applied average tariff takes the tariff rates for each partner that export to the market country in constructing the average). Applied Tariff is reported in percentage points.

**Applied Tariff Standard Deviation (tar\_sdAHS)** represents the standard deviation of the applied tariff lines within each ISIC 3 digit product.

**Applied Tariff Minimum Tariff (tar\_minAHS)** represents the lowest tariff line among the applied tariff lines within each ISIC 3 digit product.

**Applied Tariff Maximum Tariff (tar\_maxAHS)** represents the highest tariff line among the applied tariff lines within each ISIC 3 digit product.

**Simple MFN Tariff (tar\_savg\_MFN)** is the simple average tariff rate that must be paid for the item at the border of the importing country. Most Favored Nation (MFN) tariff indicates a preferred status given to a nation's specially valued trading partners. MFN status is often reciprocal. In particular, all WTO members must offer MFN to other WTO member states. The MFN rate is normally discounted from the general tariff rate. MFN Tariff is reported in percentage points.

**Weighted MFN Tariff (tar\_iwMFN)** is the import weighed average tariff rate that must be paid for the item at the border of the importing country. Most Favored Nation (MFN) tariff indicates a preferred status given to a nation's specially valued trading partners. MFN status is often reciprocal. In particular, all WTO members must offer MFN to other

WTO member states. The MFN rate is normally discounted from the general tariff rate. MFN Tariff is reported in percentage points.

**MFN Tariff Standard Deviation (tar\_sdMFN)** represents the standard deviation of the MFN tariff lines within each ISIC 3 digit product.

**MFN Tariff Minimum Tariff (tar\_minMFN)** represents the lowest tariff line among the MFN tariff lines within each ISIC 3 digit product.

**MFN Tariff Maximum Tariff (tar\_maxMFN)** represents the highest tariff line among the MFN tariff lines within each ISIC 3 digit product.

**Core NTB Frequency Ratio (NTB\_core\_w)** represents the percentage of import subjected to Non-Tariff Measures that have an unfair protectionist impact. Core NTB includes price-control measures, finance-control measures, and quantity-control measures.

**Core NTB Coverage Ratio (NTB\_core\_s)** represents the percentage of tariff lines (within each ISIC product) subjected to Non-Tariff Measures that have an unfair protectionist impact. Core NTB includes price-control measures, finance-control measures, and quantity-control measures.

**Simple Ad-Valorem Equivalent of NTB (NTB\_AVE\_s)** represents simple average of ad-valorem equivalent of core NTB from HS 6 digit level data. Core NTB includes price-control measures, finance-control measures, and quantity-control measures.

**Weighted Ad-Valorem Equivalent of NTB (NTB\_AVE\_w)** represents import weighted average of ad-valorem equivalent of core NTB from HS 6 digit level data. Core NTB includes price-control measures, finance-control measures, and quantity-control measures.

**Number of HS lines (tar\_HS\_lines)** represents the number of HS lines in each ISIC 3 digit category. This is used when calculating simple tariff averages across industries.

### **Input-Output Tables**

The data for the input-output tables originates from the GTAP 4 database which is based on early 1990s data. Two tables are provided. The first table reports the value of input of good Y needed to produce a unit of good X. In GTAP term that is input from sector Y to sector X (imports + intermediates domestic purchases) divided total output of sector X (exports + intermediates domestic purchases + household purchases + government purchases). The second table reports the share of the domestic production of sector Y that is sold to sector X. In GTAP term that is: intermediates domestic purchases of sector X from sector Y divided by total output of sector X. Because the GTAP sector classification does not directly match the classification at the 3-digit level of the ISIC, these input-output tables are only provided for 18 sectors (the correspondence to the ISIC 3-digit classification is provided in each table).