



BACI: INTERNATIONAL TRADE DATABASE AT THE PRODUCT-LEVEL THE 1994-2007 VERSION

Guillaume Gaulier & Soledad Zignago

NON-TECHNICAL SUMMARY

Empirical international trade analysis increasingly calls for accurate and disaggregated trade statistics. This working paper documents the construction of BACI, a detailed international trade database, which covers more than 200 countries and 5,000 products, between 1994 and 2007. New approaches have been developed to reconcile data reported by over 150 countries to the United Nations Statistics Division, which disseminates them via their COMTRADE database. When both exporting and importing countries report flows, we have two different figures for the same flow. In order to have a single consistent figure of a bilateral flow, we reconcile them using the procedure detailed below.

Firstly, to enable comparisons between import values, which are generally reported CIF (cost, insurance and freight), and export values, reported FOB (free on board), we need to remove transport costs from the reported imports. Detailed information on the product and transport components of CIF rates is not available. The CIF rate is therefore estimated using a gravity-type equation taking into account bilateral distance (in a non-linear manner), dummies for both contiguity and landlockness, year fixed-effects and the world median unit-value for each product category.

Secondly, we need criteria to average the FOB-FOB mirror numbers. We evaluate the reliability of each country's reporting by computing an indicator of the reporting distance among partners (the absolute value of the natural log of the ratio of mirror flows) and decompose it using a (weighted) variance analysis. The relative reliability of country reporting is then cleaned from the effects of its geographical and sectoral specialization. These adjusted qualities of reporting are finally used as weights in the averaging of mirror flows.

The three main advantages of BACI data, in comparison to other similar databases, are its product-level (more than 5,000), its geographical coverage (more than 200 countries) and its unit values, which are more reliable than the raw data, since the reconciliation of mirror figures tend to correct discrepancies. Since our methodology is purely statistical and does not require extensive additional data, the procedure can be applied exhaustively, even to cases in which knowledge on each country and product is very limited. Thus, the main aim of this work is to



N° 2010 – 23

Octobre

BACI: BASE POUR L'ANALYSE DU COMMERCE INTERNATIONAL VERSION 1994-2007

Guillaume Gaulier & Soledad Zignago

RESUME NON TECHNIQUE

L'analyse empirique du commerce international réclame de plus en plus de données désagrégées et fiables. Nous présentons ici la méthode de construction de BACI, notre base de données du commerce international qui couvre plus de 5 000 produits et pratiquement tous les pays du monde (plus de 200), annuellement de 1994 à 2007. Pour construire BACI, nous avons développé des méthodes originales d'harmonisation des données-sources de la base COMTRADE des Nations Unies qui compile les déclarations de quelque 150 pays. Il peut y avoir en effet, pour un même flux, deux sources d'information : le pays exportateur et le pays importateur, et les divergences entre ces deux flux miroir peuvent être importantes. Notre procédure d'harmonisation consiste à réconcilier ces deux sources, afin de disposer de données plus exhaustives et plus fiables, notamment en termes de valeurs unitaires (rapport des valeurs aux quantités).

Tout d'abord, les importations, généralement déclarées CAF (y compris les coûts, assurances et fret), sont calculées hors fret pour pouvoir être comparées aux déclarations FAB (franco à bord) des exportateurs. N'ayant pas d'information suffisamment détaillée sur les taux de ces coûts de transport, nous les estimons à partir d'une équation de type gravitationnel retenant comme variables explicatives la distance entre les partenaires (en tenant compte de l'existence d'une éventuelle frontière commune ou d'une situation d'enclavement) et la valeur unitaire médiane mondiale de chaque produit retenue comme indicateur de la plus ou moins grande facilité de son transport.

L'harmonisation consiste ensuite à calculer une moyenne des deux flux miroir, en accordant plus de poids à la déclaration considérée la plus fiable. La qualité de déclaration est estimée en considérant que la distance observée entre les deux déclarations d'un même flux comporte quatre composantes, attribuables respectivement à l'exportateur, à l'importateur, au produit et à l'année considérés ; de cette façon, il est possible d'isoler la qualité propre d'un déclarant, indépendamment de sa spécialisation géographique ou sectorielle.

Notre objectif est de permettre avec BACI des analyses très détaillées du commerce international pour le plus large échantillon de pays et la plus longue période possibles. De fait, BACI est utilisée dans des travaux de recherche pour trois raisons principales : sa dimension produit, son exhaustivité géographique et la fiabilité de ses valeurs unitaires. Notre méthodologie, purement statistique, permet de corriger de nombreuses sources d'erreur. Pour les utilisateurs de COMTRADE, BACI est librement disponible sur le site du CEPII (<http://www.cepii.fr/anglaisgraph/bdd/baci.htm>), en différentes nomenclatures produit.

Classification J.E.L. : F10, F14, F13, C80

Mots clés : Commerce international, Bases de données, Coûts au commerce, CAF/FAB, Harmonisation des flux de commerce international.

provide with an international trade database covering the largest number of countries at the highest degree of product-detail, for the longest period.

Users of Comtrade can freely download our BACI database in different available classifications (HS92, HS96, SITC) from <http://www.cepii.fr/anglaisgraph/bdd/baci.htm>.

J.E.L. Classification: F10, F14, F13, C80

Keywords: International Trade, Trade Costs, CIF/FOB, Trade data reconciliation.