<u>Availability of exporter-level customs transactions datasets to external researchers</u>

The calculation of the measures in the Exporter Dynamics Database version 2.0 required the use of an exporter-level dataset for each country. While we collected exporter-level datasets for 58 countries, for the other 12 countries the exporter-level data was not available for confidentiality reasons and therefore we asked counterparts with access to the exporter-level data to calculate the measures for us.

Given the strong confidentiality concerns that apply to exporter-level data, we asked permission from the data providers in those 58 countries to share this information with external researchers. The final list of countries for which we have been authorized to share the exporter-level datasets with external researchers (under the conditions specified below) is:

	Country	**Period**
1	Albania	2004-2012
2	Bulgaria	2001-2006
3	Burkina Faso	2005-2010
4	Guatemala	2003-2010
5	Jordan	2003-2010
6	Malawi	2006-2008
7	Mexico	2000-2009
8	Peru	1997-2008
9	Senegal	2000-2010
10	Uruguay	2001-2012
11	Yemen	2008-2012

For the other 47 countries, access to their exporter-level data by external researchers is not permitted.

The conditions to have access to any of the exporter-level datasets listed above are the following:
1. To present a one-page proposal indicating briefly the purpose of the research for which the datasets are requested and listing the countries for which the datasets are required.
2. To commit to share with the World Bank team the output of the research for which the data is requested in the form of working papers, book or report chapters, etc.
3. To commit to treat the exporter-level datasets in a confidential manner (not circulating/distributing these files without prior consent from our data providers). A notification specifying the extent of this commitment needs to be signed by the requestor of the datasets and returned to the World Bank before s/he receives the files with the information.
4. To cite the following paper as the source of the datasets:

Fernandes, A., Freund, C. and M. Pierola (2016). "Exporter Behavior, Country Size and Stage of Development: Evidence from the Exporter Dynamics Database" Journal of Development Economics vol. 119, pp. 121–137.

> Please fill in and sign the downloadable PDF form "Exporter Dynamics Database - Exporter-Level Data Request Form.pdf" and follow the instructions to request the data.

IMPORTANT INFORMATION FOR USERS

- The datasets shared include six main variables: a) Year of Exports; b) Exporter ID; c) HS 6-digit code; d) Value of Exports in USD; e) Net Weight in Kgs; and f) Destination (please see Annex below for the sample format).

- The exporter-level data that we are sharing includes artificial codes identifying exporters but no real tax ids nor names of firms – these were conditions discussed with data providers to be able to share the data.

- The exporter-level datasets contain the information on the products exported by each exporter in each country according to the HS classification at 6 digit level. This is the maximum level of disaggregation used in international comparisons. No further level of disaggregation will be made available in any of the country exporter-level datasets.

- The HS 6-digit product codes included in each of the datasets that we sent you are those provided by the countries and may be at different revision levels depending on the year. The 4 revision levels covering our datasets' sample period are HS 1996, HS 2002, HS 2007 or HS 2012. One obvious thing that does not happen in our datasets is that export data for 2005 cannot be provided at the HS 2007 level because that would be an anachronism. But other than that we do not know which specific HS revision is used for each country in each year (because countries did not provide us with that information). What we do know in that some developing countries take several years to begin to implement a given HS revision so it is possible to have export data for 2008 still provided at the HS 2002 level (instead of being provided at the more proximate in time HS 2007 level). If you want to assess which specific HS revision is used for each country in each year you can compare the list of product codes in the country's dataset in a year to the lists of product codes included in HS 1996, HS 2002, HS2007 and HS 2012 classifications as provided in WITS/COMTRADE.

- To minimize issues with this HS product classification, you may want to use the consolidated HS 6-digit codes that we used for the Exporter Dynamics Database. The consolidation of the 1996, 2002, 2007 and 2012 HS classifications into a single HS classification is described in documents available at http://go.worldbank.org/2F87NP14T0:
    - in general terms in the companion technical note under the heading Database Construction
    - in detail in the technical note prepared by Tolga Cebeci under the heading Consolidation of HS Classifications. Also available is a user-friendly excel file with explanations on the consolidation codes used for the Exporter Dynamics Database. Please contact Tolga Cebeci directly at tcebeci@worldbank.org (cc-ing afernandes@worldbank.org) if you have questions about the consolidation.

- The exporter-level data was provided by customs agencies and therefore it **covers ONLY export transactions**. This means that the data does **NOT** include any other firm characteristics, in particular there is no information on firm location, age, ownership, size in terms of employment, sales in the domestic market.
- The data that we are sharing is for exporters in each country, we do **not** have equivalent information for importers in each country to share. Also, we do **not** have data at the tariff level for any of the countries.

Annex: Sample format for the exporter-level customs datasets

Year (y)	Firm Identifier (f)	Destination Code 3-Digit Code (d)	HS 6-Digit Code (hs)	Export Value in USD (v)	Quantity (q)
2007	001A	JPN	902680	8,053	1
2007	001A	PER	441520	700	950
2007	002A	CHL	391910	671	104
2007	003A	COL	902000	4,500	157
2008	002A	TTO	853590	525	20
2008	003A	ECU	391990	3,443	587
2008	003A	COL	392690	390	18
2009	001A	PER	330499	442	31
2009	002A	BOL	482190	644	49
2009	002A	PER	392620	7,453	759

The list of destinations codes and names is available at http://go.worldbank.org/2F87NP14T0
in the downloadable file under the heading List of Variables, Countries and Sectors.