Readme for "Do Export Support Programs affect Prices, Quality, Markups, and Marginal Costs? Evidence from a Natural Policy Experiment"

By Magnus T. Buus, Jakob R. Munch, Joel Rodrigue, and Georg Schaur

Data Availability Statement

Confidential data sources

The data set used in this paper is based on several Danish administrative registers. Physically these administrative micro data are stored on servers located in Statistics Denmark and may not be transferred to computers outside Statistics Denmark due to data security considerations. Researchers and their research assistants are allowed to use these data if their research project is approved by Statistics Denmark and if they are cooperating with a research institution approved by Statistics Denmark. Access to the data at Statistics Denmark is provided through the internet. Researchers or their assistants are allowed access to these data from research institutions in Denmark or through remote desktop connections to servers at Danish research institutions. If a researcher at a university or other research institution outside Denmark wishes to use these data, this may be accomplished by visiting a Danish research institution or by cooperating with researchers or research assistants working in Denmark. Statistics Denmark provides guidelines for data access here: https://www.dst.dk/en/TilSalg/Forskningsservice.

Further, we use data on the Danish Trade Council's export support program. These are not publicly available, and there are no general rules or guidelines on how to obtain access. Relevant contact information is found here: https://um.dk/en/the-trade-council.

Publicly available data sources

All data sets listed below are placed in the folder 1_data\input.

1. Concordance of EU's Combined Nomenclature (CN8) over time

   To account for changes in product categories over time, we use the algorithm proposed by Van Beveren et al. (2012), aggregating categories to the so-called CN8+ level.

   The do-file (1_data\code\1_7_CN8_over_time.do) and lists of product codes for 2002-2010 (CN_2002-CN_2010.dta) are downloaded from the author's website: https://sites.google.com/site/ilkevanbeveren/concordances.

   Lists of product codes for the period 2011-2015 (CN_2011-CN_2015.dta) are downloaded from Eurostat: https://ec.europa.eu/eurostat/ramon/nomenclatures.

   List of changes in product codes over time (CN_2018_update_of_codes_xls.xls) is downloaded from Eurostat: https://ec.europa.eu/eurostat/ramon/relations.

2. Input expenditure indices

   We deflate input expenditures before performing the De Loecker et al. (2016) estimation procedure.

Wages are deflated by the consumer price index (PRIS112.xlsx): https://www.statistikbanken.dk/PRIS112.

Capital expenditures are deflated by the gross fixed capital formation index (NABK69.xlsx): https://www.statistikbanken.dk/NABK69.

Material expenditures are deflated by the intermediate consumption index (NABP69.xlsx): https://www.statistikbanken.dk/NABP69.

NABK69 and NABP69 are based on the industry classification in the Danish National Accounts, the 69 grouping. These can be mapped to Dansk Branchekode (DB07), the Danish version of EU's nomenclature (NACE) in two steps. First, map to standard grouping 127 using the key (Dansk-Branchekode-2007-(DB07)-v3-2014.xlsx) from here: https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/nr-branche. Second, map to DB07 using the key (typed manually in 1_input\code\1_6_public.do) from here: https://www.dst.dk/da/Statistik/dokumentation/nomenklaturer/db07.

3. Bilateral trade flows
   Bilateral trade flows are collected by the UN in the Comtrade database and processed by CEPII in the BACI database.
   Data (BACI_HS92_Y2002_V202001- BACI_HS92_Y2015_V202001.csv) are downloaded here: http://www.cepii.fr/DATA_DOWNLOAD/baci/legacy/202001/trade_flows/BACI_HS92_V202001.zip. Due to their size, these files are not included.
   Country codes (country_codes_V202102.csv) are downloaded here: http://www.cepii.fr/DATA_DOWNLOAD/baci/nomenclatures/country_codes_V202102.csv.
4. Country income groups

   We use country income groups as defined by the World Bank (WITSCountryProfile-Country_Indicator_ProductMetada-en.xlsx) from here: https://wits.worldbank.org/countryprofile/metadata/en/country/all.

5. BEC product groups

   We use the Broad Economic Category (BEC) classification to distinguish between end-use product categories.

   A key between CN8 and BEC (CN_BEC_2021.xlsx) is downloaded from Eurostat: https://ec.europa.eu/eurostat/ramon/relations.

   We classify BEC groups as "Non-consumption" and "Consumption" (typed in manually in 4_reg_variables\code\4_2_productgroups.do) according to paragraph 9 here: https://ec.europa.eu/eurostat/ramon/other_documents/bec/BEC_Rev_4.pdf.

Documentation of programs

The analysis is divided into six parts. Each part has its own folder (listed below). Each folder has four subfolders: i) "input" where all relevant data must be placed; ii) "code" where all program files are placed; iii) "output" where all data sets relevant for subsequent folders as wells as final tables and figures will be located, and iv) "temp" where data sets not relevant for subsequent folders are placed.

Below is the list of folders as well as a brief description of the analysis performed therein. For details on each program file, please see the comments posted directly in the relevant file.

0. 0_letters. In this folder, we produce the paper's Figure 1. This is done separately from the remaining analysis because Figure 1, in contrast to all other results in the paper, is based on the raw non-anonymized data from the Trade Council that cannot be matched to Statistic Denmark's administrative data. One should place these raw non-anonymized data in the subfolder "input". Figure 1 will appear in the subfolder "output".
1. 1_data. In this folder, we load and collect the raw data. The subfolder "input" contains all publicly available raw data, and one should place all confidential raw data here, too.
2. 2_prepare. In this folder, we prepare data for the De Loecker et al. (2016) estimation procedure.
3. 3_lgkp. In this folder, we perform the De Loecker et al. (2016) estimation procedure. This folder is *not* organized as listed above, but follows the original structure. The original files are downloaded from: https://www.econometricsociety.org/publications/econometrica/2016/03/01/prices-markups-and-trade-reform. We have added notes to the original program files only to indicate where our implementation differs from the original program.
4. 4_reg_variables. In this folder, we construct samples and variables for the subsequent regression exercises.
5. 5_reg_analysis. In this folder, we perform all regression exercises. Each table and figure in the paper (main text as well as appendix) is associated with a .do file in the "code" subfolder as indicated by the file name (e.g. 5_2d_table_4 constructs Table 4 in the paper). All output is placed in the subfolder "output".

To execute the entire analysis: i) open 0a_global.do, change directory in line 1, and run; then ii) open 1_run.do and run. Alternatively, one can run each .do file separately in the order dictated by 1_run.do.