Data and Code for "Exporter Heterogeneity and Price Discrimination: A Quantitative View"

Jae Wook Jung[*]

KIEP

Ina Simonovska[†]

University of California, Davis and NBER

Ariel Weinberger[‡]

University of Oklahoma

August 2018

1 Overview

This document describes the use of the dataset and provides instructions to implement the code used to generate all empirical results in "Exporter Heterogeneity and Price Discrimination: A Quantitative View". The overview is split into two main sections: section 2 describes the micro data and section 3 describes the estimation programs which are conducted in Matlab. For each section, we separate the "Firm Moments" (Chilean firm data) and "Price Moments" (cross-country price data). There are three main folders in the data folder: "Chile" includes the data and programs described in section 2 (mostly Stata). "Firm moments" and "Price moments" includes the Matlab code to replicate the programs described in section 3.

2 Data

2.1 Firm Moments: Sales and Markup Distributions

The data targets constructed for the "firm moments" in our paper are produced with a census of Chilean firms. The data is publicly available at:
http://historico.ine.cl/canales/chile_estadistico/estadisticas_economicas/industria/series_estadisticas/series_estadisticas_enia.php. On the INE website, these can be downloaded in Windows Access format. The data in our paper is in Stata .dta format, but reflects a subset of the data available on the public site.

---

[*]jwjung@kiep.go.kr

[†]inasimonovska@ucdavis.edu

[‡]ariel.weinberger@ou.edu

The following files produce Figure 1 in the text, as well as the Data Targets of the Firm Moments reported in the Panel B of Table 1.

- Do-file: firmmoments.do is the main driver file, which inputs data and produces Figure 1 as well as firm moments of the sales and markup distribution.

- Data: raw_chile.dta includes the variables of interest from the Chilean census data, from 1995-2007. Firms are coded with a unique ID over time. Each firm is listed within a 4 digit ISIC (Rev 3) code. The variables of interest include: total sales, value added, export value, total wages paid, and total value of gross fixed assets. It also calls sales_deflator.dta and input_deflator.dta which are deflators at the ISIC-year level used to deflate sales, etc. Finally, irate_wdi.dta is called to produce annual interest rates for all Chilean firms.

2.2   Price Moments: Cross-Country Price Data

Cross-country good-level price data is necessary to construct the Data Target for the Price Moments. The raw data is only available through a subscription to the Economist Intelligence Unit (EIU) at http://www.eiu.com/home.aspx#offer-ss-data. The data provides a matrix of prices for 110 goods to 71 destinations. This matrix of price data is used in the SMM estimation programs in order to construct the data targets for price moments. The Matlab driver file main_pricemmt.m takes in this data and produces the data targets reported in Panel B of Table 1.

3   Programs for the Simulated Method of Moments Estimation

In this section we describe how to replicate the SMM estimation. There are two separate estimation strategies, targeting firm and price moments respectively. The estimation procedures are described in the main text.

3.1   Gravity Estimation

As described in the main text, the first step to both estimation strategies includes a gravity estimation that computes bilateral trade costs (and wages from market clearing) for a guess of the Pareto parameter, θ. Each of the two main driver files (main_firmmmt.m and main_pricemmt.m), described in detail below, run the gravity estimation at the beginning of the code by calling est.m. The following data is included in both "Firm Moments" and "Price Moments" and necessary to run the gravity estimation:

- cepii_data.mat is a matrix of CEPII data that includes all of the regressors in the gravity equation (page 20 in the main text). For more details on this data, see the CSV file, cepii66.csv.

- data_firmmmt.mat includes the following data:

- A tradeshare matrix with values that correspond to λ_{ij} in the text. Columns (j) represent the destination and rows (i) represent the origin. This matrix cannot be provided publicly as bilateral trade flows and gross output are not public. To construct the tradeshares, we combine trade flows from 2004 at the SITC-4-digit level for each country-pair in the Feenstra et al. (NBER, 2005) dataset. Trade flows are aggregated to reflect aggregates of the manufacturing sector. The construction of this matrix is described in Simonovska and Waugh (JIE, 2014), with exact details provided here: https://github.com/mwaugh0328/Create-Trade-Share-Matrix-JIE-SW-2014.

- Gross output (Y) – source: UNIDO

- Total workforce (L) – source: Penn World Tables 8.0

3.2 Firm Moments

main_firmmmt.m is the main driver file that runs the estimation and produces parameter estimates plus standard errors using the procedure described in the main text. The results from running this file are the results provided in the left hand side of Table 1 (panels A and B). All the files necessary to run this estimation are under the folder "Firm moments." data_firmmmt.mat includes the publicly available data necessary to produce the gravity estimation, although it requires the tradeshare matrix.

Notice that within this benchmark estimation using firm moments, there are 2 separate estimations: an exactly identified one using only the sales advantage and markup advantage of exporters, plus an over-identified estimation that adds percentile-differences from the sales and markup distributions. Both of these estimations can be run using the same file, by choosing the appropriate options on line 15. The number of moments is equal to 2 for the exact identification, and equal to 8 for the over-identification. The optimal weighting matrix is only produced in the case where N.mmts==8.

gen_plots_firmmmt.m produces Figure 2, given the estimated parameters. gen_tables_firmmmt.m produces the model predictions displayed in Table 2 and 3 of the paper.

3.3 Price Moments

main_pricemmt.m is the main driver file that runs the estimation and produces parameter estimates plus standard errors using the alternative trade elasticity procedure described in the main text. The results from running this file are the results provided in the right hand side of Table 1 (panels A and B). All the files necessary to run this estimation are under the folder "Price moments." data_pricemmt.mat includes the publicly available data necessary to produce the gravity estimation, although it requires the tradeshare matrix.

Notice that within this benchmark estimation using price moments, there are 2 separate estimations: an exactly identified one with only price gaps, plus an over-identified estimation that adds further moments from the price gaps and the covariance of the gaps with destination wages. Both of these estimations can be run using the same file, by choosing the appropriate options on line 15. The data moments cannot be produced without the EIU price data. The optimal weighting matrix is only produced in the case where N.mmts==3.