Description of data files for "Trade Liberalization, Quality, and Export Prices" by Haichao Fan, Yao Amber Li, and Stephen R. Yeaple (accepted by *Review of Economics and Statistics*)

****

DATA SOURCES**

****

Please see the online appendix for detailed instructions on data construction. The main data used in this paper is *proprietary* data, containing firm-level and firm-product information. The original data on manufacturing firm survey database and customs database were obtained in collaboration with NBSC (National Bureau of Statistics of China) and China's General Administration of Customs through their authorized vendor:

Beijing IVG Information Consulting Co., Ltd.
Website: www.china-ivg.com
Tel: 8610-63806335; 8610-83509152.
Fax: 8610-83509185.
Contact: Ms. SONG Jia.
Email: marketing@china-ivg.com
Address: CUV International Building, Wanliu Qiao, Fengtai District, Beijing, China.
Postal Code: 100070

If you use the following data sets and the corresponding programs, please cite:

Fan, Haichao, Yao Amber Li, and Stephen R. Yeaple "Trade Liberalization, Quality, and Export Prices" forthcoming *Review of Economics and Statistics*.

All the following data sets and programs are compiled using Stata/MP 13.1 for Windows (64-bit x86-64).

****

DATA SETS *****

****

Note: To save space for storage, all data files have been compressed using the free software 7-zip. After unzipping those data files, you will obtain the four master data files and one auxiliary data file.

Master data files:

1. Merge.final.fp.dta and Merge.final.fpc.dta: These two data sets are used to produce all tables and figures (except for robustness table 11 and Figure 3) in the paper. They have the similar data structure, with the former at firm-HS6 product level and the

latter at firm-HS6 product-country level. The data in these two datasets have the following layout:

variable name	variable explanation
FRDM	software-generated, numerical firm identifier in the NBSC firm survey data
year	year
coun_aim	export destination country (in Chinese) (only appears in firm-product-country level dataset Merge.final.fpc.dta)
HS6	numerical HS6 product code (Note: some products have 5-digits numerical codes if they start with 0)
price	(log) unit value export price
diff	a dummy variable denoting "Homogeneous" which is equal to 1 for homogeneous goods and 0 for differentiated goods according to Rauch's classification
HHI	Herfindahl index at 4-digit CIC level
l	ln(employment)
y	ln(deflated firm's value-added)
K_L	ln(capital/labor)
wage	ln(wage bill)
quant	quantity of exported product
value	export value (deflated)
price	ln(price)=ln(value/quant)
N_imp	The number of imported variety across product and country
duty_d	Weighted firm specific tariff reduction (Main Measure)
duty_un	Unweighted firm specific tariff reduction (Measure 1)
duty_exo_d	Tariff reduction measure by fixing the total number of imported varieties during the whole sample period (Measure 2)
duty_inter_d	Weighted firm-specific tariff reduction of only intermediate goods (Measure 3)
TFP	ln(TFP)
pairdummy	=group(FRDM HS6 coun_aim) in Merge.final.fpc.dta =group(FRDM HS6) in Merge.final.fp.dta
N_d	the change in the number of import varieties
K_L_d	Δln(capital/labor)
wage_d	Δln(wage)
l_d	Δln(employment)
price_d	Δln(price)=ln(price)2006 – ln(price)2001 at firm-product level in custom.final.fp.dta at firm-product-country level in custom.final.fpc.dta
TFP_d	Δln(TFP)
duty00	Firm specific tariff based on the 2001 tariff level
duty97	Firm specific tariff based on the 1997 tariff level
sigma_m	At HS2 digit level sigma based on Broda and Weinstein
var	Quality Variance at HS6-digit level
GK_ind	Gallop-Monahan Index
duty00s	initial tariff levels in 2001 facing by all other firms in the same 2-digit CIC industry
duty_ds	The tariff changed facing by all other firms in the same 2-digit CIC industry
share_d	Market share change
price_net5_d	the change in log quality-adjusted price when sigma=5

quality5_d	the change in effective quality when sigma=5
price_net10_d	the change in log quality-adjusted price when sigma=10
quality10_d	the change in effective quality when sigma=10
price_net_d	the change in log quality-adjusted price when sigma=sigma_i from Broda and Weinstein (2006)
quality_d	the change in effective quality when sigma=sigma_i from Broda and Weinstein (2006)
duty_in_d	the change in industry input tariff
duty_out_d	the change in industry output tariff
pi_d	the firm-level price change

2. custom.final.fp.dta and custom.final.fpc.dta: These two data sets are used to produce robustness table 11 and Figure 3. They have the similar data structure, with the former at firm-HS6 product level and the latter at firm-HS6 product-country level. The data in these two datasets have the following layout:

variable name	variable explanation
company	software-generated company identifier in the customs data
year	year
coun_aim	export destination country (in Chinese) (only appears in firm-product-country level dataset custom.final.fpc.dta)
HS6	numerical HS6 product code (Note: some products have 5-digits numerical codes if they start with 0)
price	(log) unit value export price
diff	a dummy variable denoting "Homogeneous" which is equal to 1 for homogeneous goods and 0 for differentiated goods according to Rauch's classification
duty_d	Weighted firm specific tariff reduction (Main Measure)
duty_un	Unweighted firm specific tariff reduction (Measure 1)
duty_exo_d	Tariff reduction measure by fixing the total number of imported varieties during the whole sample period (Measure 2)
duty_inter_d	Weighted firm-specific tariff reduction of only intermediate goods (Measure 3)
N_d	the change in the number of import varieties
price_d	$\Delta\ln(price)=\ln(price)_{2006} - \ln(price)_{2001}$ at firm-product level in custom.final.fp.dta at firm-product-country level in custom.final.fpc.dta
duty_in_d	the change in industry input tariff
duty_out_d	the change in industry output tariff

Auxiliary data sets:

tariff.MNF_applied_rate.HS6.dta: this data set provides tariff data to produce Table 3.

variable name	variable explanation
year	year
HS6	numerical HS6 product code (Note: some products have 5-digits numerical codes if they start with 0)
duty	MFN tariff rate

****

******* Programs****************

****

1_Sample_Construction.do

This do file constructs the samples contained in the four final data sets.

2_Stylized_Facts.do

This do file produces tables and figures in Section 3 "Stylized Facts" and Table 3 in Section 5.2 "The Measurement of Tariff Reductions".

3_Main_results.do

This do file produces all tables and figures in the main results of the paper (Section 6).

4_Robustness.do

This do file generates all tables and figures in Robustness (Section 7 and Section 8).