Online Data Appendix to:

Quantitative Trade Models in *Annual Review of Economics*

Timothy J. Kehoe, Pau S. Pujulàs, Jack Rossbach

February 2017

Structure of Appendix

This appendix accompanies an archive containing three folders, *Data, Output,* and *Programs*, containing files and a fourth folder *temp* which the STATA programs save temporary output to.

Below I describe the contents of the folders *Data* and *Programs*.

Data Files

This section describes the files in the folder "Data"

WDI_GDPs.csv: This file contains GDP (Current price, USD) for the United States, Australia, Chile, China, and New Zealand between 2001 and 2015. Downloaded from the Worldbank's World Development Indicators (WDI) database: http://databank.worldbank.org

NAFTA_GDPs.csv: This file contains GDP (Current price, USD) for the United States, Canada, and Mexico between 1987 and 2014. Downloaded from the Worldbank's World Development Indicators (WDI) database: http://databank.worldbank.org

NAFTAData_HS92.csv: This file contains bilateral trade data between the United States, Canada, and Mexico between 1989 and 2015. Data is at the 6-digit HS1988/1992 level, and is reported in 1000's USD in current prices. The reporter country is the exporter. Data downloaded from the World Integrate Trade Solution (WITS) database: http://wits.worldbank.org/

GTAPData_HS02.csv: This file contains bilateral trade data between the United States and Australia, the United States and Chile, China and Chile, and China and New Zealand between 2002 and 2015. Data is at the 6-digit HS2002 level, and is reported in 1000's USD in current prices. The reporter country is the exporter. Data downloaded from the World Integrate Trade Solution (WITS) database: http://wits.worldbank.org/

HS_GTAP_Concordance.csv: This file contains a concordance between the 6-digit HS2002 product codes and the GTAP industry codes. Concordance was produced by Thomas Hutcheson for GTAP: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1916

concordance_H0_to_I3.csv: This file contains a concordance between the 6-digit HS1988/1992 product codes and 4-digit ISIC Rev. 3 industry codes. This concordance is available for download from WITS: http://wits.worldbank.org/product_concordance.html

concordance_H0_to_I3_missing.dta: This file contains a concordance between the 6-digit HS1988/1992 product codes and 2-digit ISIC Rev. 3 industry codes for 21 products that were not included in the above concordance from WITS. This concordance was constructed by hand.

CP_concordance.csv: This file contains a concordance between 2-digit ISIC Rev. 3 industry codes and the industries used in Caliendo & Parro (2015). The concordance is from Table A.1 in their paper.

GTAP_Results.dta: This file contains the counterfactual results for the free trade agreements using the GTAP framework. Results were produced using RunGTAP v3.61 with the standard closure and Gragg solution method. The counterfactual involves setting tariffs equal to zero for all commodities exported between the two countries involved in each FTA by shocking the variable tms. The column *sim_growth* reports the predicted change, computed as percent change, of trade flows for each exporter-importer-industry pair, which is reported via the variable qxs in RunGTAP. The GTAP aggregations used to construct these results are available in the subfolder ***GTAP Aggregations***. RunGTAP is available at the link below:
https://www.gtap.agecon.purdue.edu/products/RunGTAP/default.asp

GTAP Aggregations (subfolder): This folder contains the GTAP aggregations used in the calibration and simulation of the GTAP model. The aggregations were produced using GTAPAgg using the GTAP9 database with a base year of 2004. The regional aggregation for each trade agreement is each of the partner countries and then a ROW aggregate. Sectors are left fully disaggregated. Factors are left at their default GTAPAgg aggregation. The naming of the zips are *(Country 1)_(Country 2)_(Base Year)_All.zip*, where the *All* stands for fully disaggregated sectors.

CP_Results.dta: This file contains the counterfactual results for the Caliendo & Parro (2015) framework. The column *CP_pair* lists the partner countries and years in the format (exporter iso3)-(importer iso3)-year. *CP_industry_code* contains the industry code the trade flow is for. *sim_growth* is the predicted change, in log differences, of trade flows given all post-NAFTA tariff changes. *sim_growth_oN* is the predicted change, in log differences, of trade flows given only NAFTA tariff changes. *sim_growth_oN_noIO* is the predicted change, in log differences, of trade flows given only NAFTA tariff changes for the model version with no input-output structure. These results were produced using the data appendix from Caliendo & Parro (2015), available on the author's website: http://faculty.som.yale.edu/lorenzocaliendo/Data_and_Codes_CP.zip

GTAP_Tariffs.dta: This file contains tariff data for trade between the United States and Australia, the United States and Chile, China and Chile, and China and New Zealand between 1989 and 2015. The reporter country is the importer. Data is from the United Nations' Trade Analysis Information System (TRAINS) and downloaded via WITS: http://wits.worldbank.org/

KPR2017_Level_Coefficients.xlsx: This file contains the estimated coefficients that we use to produce level predictions for the LTP methodology. The coefficients are estimated using equations (A.3)–(A.5) in section A.5 of the appendix. Tariffs are from TRAINS, γ is from Kehoe & Ruhl (2013) and the estimated tariff elasticity of –3.511 is from Kehoe, Rossbach, & Ruhl (2015). The estimated coefficients for NAFTA are in the sheet "NAFTA", and the estimated coefficients for the FTAs we use to evaluate the GTAP model are in the sheet "GTAP".

Program Files

This section describes the files in the folder "Programs"

Prior to running all Stata programs you must edit the location of the base_directory in each do file so that the paths to the Data, temp, and Output folders are correct. Requires the Stata package estout. Code is commented; feel free to contact the authors if anything requires clarification. Also see the data appendix for Kehoe, Rossbach, Ruhl (2015), available for download from the author's website: http://www.econ.umn.edu/~tkehoe/papers/NewProductsDataAppendix.zip

KPR_Appendix_GTAP.do: This file contains Stata code for producing the output for Table 1 (correlations only), and Appendix Tables A.1, A.4, A.5, and A.11. As inputs it loads the files *GTAPData_HS02.csv*, *HS_GTAP_Concordance.csv*, *WDI_GDPs.csv*, and *GTAP_Results.dta,* and produces the files *GTAP_LTP_Results.dta* and *KPR2017_Table1.txt* in the folder *Output*. It also produces the files *KPR2017_TableA1.txt, KPR2017_TableA4.txt, KPR2017_TableA5.txt, & KPR2017_TableA11.txt* for the Robustness Appendix.

KPR_Appendix_Tariffs.do: This file contains Stata code for producing the tariffs reported in Table 1 of the paper. As inputs it loads the file *GTAPtariffs.dta* and produces the file *KPR2017_Table1_Tariffs.txt* in the folder *Output*. This program does not involve any calculations, it simply reformats the data and reports the simple average tariff from the input data.

KPR_Appendix_NAFTA.do: This file contains Stata code for producing the results in Table 2, and Appendix Tables A.2, A.3, A.5, A.6, A.7, and A.12. As inputs it loads the files *NAFTAData_HS92.csv*, *CP_Results.dta, NAFTA_GDPs.csv*, *concordance_H0_to_I3.csv concordance_H0_to_I3_missing.dta,* and *CP_Concordance.csv,* and produces the files *CP_LTP_Results.dta* and *KPR2017_Table2.txt* in the folder *Output*. It also produces the files *KPR2017_TableA2*.txt, *KPR2017_TableA3*.txt, *KPR2017_TableA5*.txt, *KPR2017_TableA6*.txt, *KPR2017_TableA7*.txt, and *KPR2017_TableA12*.txt for the Robustness Appendix.

KPR_MatlabCodeforCP.m: This file contains snippets of matlab code and instructions for producing the counterfactual results for Caliendo & Parro (2015). This file is meant to accompany the data appendix for Caliendo-Parro, and should only be run after running their programs. Their appendix is available here: http://faculty.som.yale.edu/lorenzocaliendo/Data_and_Codes_CP.zip

KPR_Appendix_GTAP_diff_base_years.do: This file is identical to *KPR_Appendix_GTAP.do* in terms of its structure and inputs, only changing the base year. As inputs it loads the files *GTAPData_HS02.csv*, *HS_GTAP_Concordance.csv*, *WDI_GDPs.csv*, and *GTAP_Results.dta,* and produces the file *KPR2017_TableA9.txt* in the folder *Output* for Table A.9 in the Appendix.

KPR_Appendix_NAFTA_diff_base_years.do: This file is identical to *KPR_Appendix_NAFTA.do* in terms of its structure and inputs, only changing the base year. As inputs it loads the files *NAFTAData_HS92.csv*, *CP_Results.dta, NAFTA_GDPs.csv*, *concordance_H0_to_I3.csv concordance_H0_to_I3_missing.dta,* and *CP_Concordance.csv,* and produces the file *KPR2017_TableA10.txt* in the folder *Output* for Table A.10 in the Appendix.

Output Files

This section describes the files in the folder "Output". Text files for tables are not described below, as each text file simply corresponds to the labeled table in the paper/robustness appendix, where they are described in full detail.

CP_LTP_Results.dta: This file contains the output file used to construct the NAFTA and CP evaluation tables after running the program "KPR_Appendix_NAFTA.do". The variables are described as follows: *cp_industry_code* is the industry code from Caliendo & Parro (2015) [missing values indicate trade flows for products that are not included in any of the industries studied by Caliendo and Parro], *exporter* is the 3-digit ISO3 code of the exporting country, *importer* is the 3-digit ISO3 code of the importing country, *pre_val* is the trade flows for the given industry in the base year (1991), *pre_val_avg* is the average value of trade over the first three years for each given industry (1991–1993), *post_val* is the trade flow in the final period (2006). The variable *least_val_pre* is the value of trade flows for LTP in the base year, *least_val_post* is the value of trade flows for LTP is the final period; *notleast_val_pre* is the value of trade flows for non-LTP in the base period, and *notleast_val_post* the value of trade flows for non-LTP in the final period; *least* is the number of products in the industry that are considered LTP, *notleast* is the number of products in the industry that are consider non-LTP. *GDP_pre* is the GDP of the exporter country in the base year, *GDP_post* is the GDP of the exporter GDP in the final period. The variable *perleast* indicates the share of trade in each industry for the base year made up of LTP. The variable *growth* indicates the percentage change in trade flows between the base year and final period for each industry, where growth is deflated by exporter GDP. The growth rate is compared against the growth rates from the CP paper; *sim_growth* is the default predicted growth rate for trade flows for each industry from Caliendo & Parro (2015), *sim_growth_oN* is the predicted growth rate when only changes in tariffs resulting from NAFTA are considered, and *sim_growth_oN_noIO* is the predicted growth rate when only changes in tariffs from NAFTA are considered and there are no input-output linkages across industries in the model.

GTAP_LTP_Results.dta: This file contains the output file used to construct the GTAP evaluation tables after running the program "KPR_Appendix_GTAP.do" All variables are the same as for *CP_LTP_Results.dta*; except for the following exceptions: *gtap* is the industry code for the GTAP model, the base year is 2002, the final period is 2015, and the growth rate is compared against the growth rates from the GTAP model; *sim_growth* is the predicted growth rate for each industry using the GTAP model and setting tariffs equal to zero.