Paul Bergin, Reuven Glick, Jyh-Lin Wu. "The Micro-Macro Disconnect of Purchasing-Power Parity," *Review of Economics and Statistics*, Vol. 95 (3), 2013. MS #14419

Documentation

This file contains instructions for replicating Tables 1-9 and Figures 1-5 of the paper. All programs, datasets, and results are contained in the zipped file "BerginGlickWu Replication." Since the raw data on prices for individual goods and services in cities worldwide obtained from the Economist Intelligence Unit (EIU) are proprietary, we provide intermediate forms of the data series used. For more information on the raw data, go to the website http://eiu.enumerate.com/asp/wcol_HelpAboutEIU.

We proceed below in two steps: Step (1), generating intermediate datasets from input datasets, and Step (2), performing analysis and generating results from the intermediate datasets.

Other supporting files
"BGW Input and Intermediate Datasets.docx": lists the input (original) and intermediate (for analysis and output) datasets used in programs
"BGW Variable and Result Names.docx": provides descriptions of variable and result name labels used in the analysis
"BGW Code & Datasets List.xlsx": lists all programs and datasets, including the ado files called by each program

Helpful Notation for understanding the naming of data files:
Drop1 – Turkey is dropped
F1 – No filling in of missing data
PW – Parsley Wei cities. This is the convention for results reported in the paper unless otherwise specified.
NSC – No same country (No city pairs are located in the same country).
NT – Non-traded goods (if not indicated, then goods are traded)
LC – Local currency
Indust – Industrial countries
Wide – Data format wide
Price_dispersion – Relative prices of goods across cities.

Other notation is purely for the authors.

All programs were run in STATA version 10 (though they can be run on any subsequent version), with the exception of one program run in MATLAB R2012a, all on a PC computer operating with OS Windows 7.

The file names in quotes below indicate file paths.

STEP (1). Directions for generating intermediate datasets from input datasets:

Input data file path: "BerginGlickWu Replication\data_creation\original"
Output data file path: "BerginGlickWu Replication\data_creation\datasets"
Program file paths: "BerginGlickWu Replication\data_creation\programs\Table1"
"BerginGlickWu Replication\data_creation\programs\Table3"

"BerginGlickWu Replication\data_creation\programs\Table5"
"BerginGlickWu Replication\data_creation\programs\Table6"

All programs specify global and/or local path references at the top. **One will need to change the path references at the top of each program to correspond to the proper directory**. Example of path references below:

```
global programpath "P:\BerginGlickWu Replication\data_creation\programs"
global outpath1 "P:\BerginGlickWu Replication\data_creation\datasets"
global datapath1 "P:\BerginGlickWu Replication\data_creation\original"
```

where programpath has the ado files and serves as the current directory, outpath1 is for the output of intermediate datasets, and datapath1 contains the input dataset(s).

Below we list the programs and the data mapping for generating individual tables:
Input dataset → Intermediate dataset

Programs for Tables 1+2 data:
Path: "data_creation\programs\Table1":
4 programs generate 4 datasets (2 aggregated, 2 disaggregated)
1. cr_aggregate_semiannual_nsc_drop1_f1_wide_indust.do (aggregate, traded)
2. cr_aggregate_semiannual_nt_nsc_drop1_f1_wide_indust.do (aggregate, non-traded)
3. cr_dataset_semiannual_nsc_drop1_f1_wide_indust.do (disaggregate, traded)
4. cr_dataset_semiannual_nt_nsc_drop1_f1_wide_indust.do (disaggregate, non-traded)

Mapping from input to corresponding intermediate dataset:
1. semiannual_aggregate_f1.dta→aggregate_semiannual_nsc_drop1_f1_wide_indust.dta
2. semiannual_aggregate_nt_f1.dta→aggregate_semiannual_nt_nsc_drop1_f1_wide_indust.dta
3. price_dispersion_lc_semiannual_f1.dta→semiannual_nsc_drop1_f1_wide_indust.dta
4. price_dispersion_nt_lc_semiannual_f1.dta→semiannual_nt_nsc_drop1_f1_wide_indust.dta

Programs for Tables 3+4 data:
Path: "data_creation\programs\Table3"
2 programs generate 2 datasets (1 aggregated, 1 disaggregated)
1. cr_aggregate_semiannual_lc _drop1_f1_wide_indust.do (aggregate)
2. cr_dataset_semiannual_lc_drop1_f1_wide_indust.do (disaggregate)

Mapping from input to corresponding intermediate dataset:
1. agg_semiannual_nom_xr_f1.dta, semiannual_aggregate_lc_f1.dta
→aggregate_semiannual_lc_drop1_f1_wide_indust.dta
2. price_dispersion_lc_semiannual_f1.dta→semiannual_lc_drop1_f1_wide_indust.dta

Programs for Table 5 data:
Path: "data_creation\programs\Table5"
2 programs generate 2 datasets
1. cr_imbsdata_agg_lc_wide.do (aggregate)
2. cr_imbsdata_sectoral_lc_wide.do (disaggregate)

Mapping from input to corresponding intermediate dataset:

1. aggprices_lc.csv, xrdata.csv → imbsdata_agg_lc_wide.dta (Imbs et al. data aggregated)
2. prices_lc.csv, US_prices.csv, xrdata.csv→ imbsdata_sectoral_lc_wide.dta (Imbs et al. data disaggregated)

Programs for Tables 6,8,9 data:
Path: "data_creation\programs\Table6"
1 program: cr_dataset_semiannual_lc_newPT3_f1_wide_indust.do

Mapping from input to corresponding intermediate dataset:
price_dispersion_lc_semiannual_f1.dta→semiannual_lc_newPT3_PWT_f1_wide_indust.dta

---

STEP (2). Directions for running programs from intermediate datasets and gathering results:

Each table has a corresponding folder with data, programs, and results. **One will need to change the path references at the top of each program to correspond to the proper directory**. Example of path references below:

```
global programpath "P:\BerginGlickWu Replication \Table 1\programs"
local outpath1 "P:\BerginGlickWu Replication \Table 1\results"
local datapath "P:\BerginGlickWu Replication \Table 1\datasets"
```

Where programpath has the ado files and serves as current directory, outpath1 is for results, and datapath contains the dataset(s). We document below each table's data, programs, and output.

Format (Table #): Table name

Data: "path"
Datasets

Programs: "path"
Program files

Results: "path"
Result files

Directions: Directions for output

Table 1: Stationarity of relative prices

Data: "data_creation\datasets"
aggregate_semiannual_nsc_drop1_f1_wide_indust.dta (Aggregate, traded)
aggregate_semiannual_nt_nsc_drop1_f1_wide_indust.dta (Aggregate, non-traded)
semiannual_nsc_drop1_f1_wide_indust.dta (Disaggregate, traded)
semiannual_nt_nsc_drop1_f1_wide_indust.dta (Disaggregate, non-traded)

Programs: "Table 1\programs"
Aggregate: an_reg1_agg_semiannual_nsc_drop1_f1_indust.do
Disaggregate: an_reg1_semiannual_nsc_drop1_f1_indust.do

Results: "Table 1\results"

Aggregate: reg1_agg_semiannual_nsc_drop1_f1_indust.csv (Traded and non-traded)
Disaggregate: reg1_semiannual_nsc_PW_drop1_f1_indust.csv (Traded),
reg1_semiannual_nt_nsc_PW_drop1_f1_indust.csv (Non-traded)

Directions:
For aggregate results, series #1 corresponds to traded goods, and series #2 corresponds to non-traded goods.

For disaggregate results, take the average of the b_coefs and t-stats (columns C+D) and sum the percent significance levels (columns E,F,G). The average number of observations is the average number of pairs (column B) multiplied by number of time observations (33).

We use t-statistics from Dickey-Fuller tests (see Pesaran 2006, Table IIB), with T= 30 and N = 20.

Table 2: Half-lives in autoregressions of real exchange rates

Data: "data_creation\datasets"
aggregate_semiannual_nsc_drop1_f1_wide_indust.dta (aggregate, traded)
aggregate_semiannual_nt_nsc_drop1_f1_wide_indust.dta (aggregate, non-traded)
semiannual_nsc_drop1_f1_wide_indust.dta (disaggregate, traded)
semiannual_nt_nsc_drop1_f1_wide_indust.dta (disaggregate, non-traded)

Programs:
"Table 2\programs\agg"
an_ccep_agg_semiannual_full_nsc_drop1_f1_indust.do (aggregate)
"Table 2\programs\disagg"
an_ccep_semiannual_full_nsc_drop1_f1_indust.do (disaggregate)

Results: "Table 2\results"
Aggregate: ccep1_agg_semiannual_nsc_drop1_f1_indust.csv (1 lag)
ccep2_agg_semiannual_nsc_drop1_f1_indust.csv (2 lags)

Disaggregate: ccep1_semiannual_full_nsc_PW_drop1_f1_indust.csv (1 lag),
ccep2_semiannual_full_nsc_PW_drop1_f1_indust.csv (2 lags)

Directions:
The aggregate results are evident. For the disaggregates, take the average of the b_coefs across goods, (columns C,E,G,I).

Calculation of half-lives: Use the coefficients to generate an impulse response that yields the amount of time (semiannual frequency) to reach midpoint between shock and zero. Examples in example_halflife_agg_2lags.xlsx and example_halflife_disagg_2lags.xlsx

---

Table 3: Vector error correction estimates

Data: "data_creation\datasets"
aggregate_semiannual_lc_drop1_f1_wide_indust.dta (aggregate)
semiannual_lc_drop1_f1_wide_indust.dta (disaggregate)

Programs:
"Table 3\3a\programs\agg"

an_ccep_agg_semiannual_ecm_full_drop1_f1_indust.do (aggregate, 3a)
"Table 3\3a\programs\disagg"
an_ccep_semiannual_ecm_full_drop1_f1_indust.do (disaggregate, 3a)
"Table 3\3b\programs\agg"
all_agg.do (aggregate, 3b)
"Table 3\3b\programs\disagg"
all_disagg.do (disaggregate, 3b)

Results:
"Table 3\results\3a" (3a results)
ccep1_agg_semiannual_ecm_full_drop1_f1_indust.csv (e equation, aggregate, 3a)
ccep2_agg_semiannual_ecm_full_drop1_f1_indust.csv (p equation, aggregate, 3a)
ccep1_semiannual_PW_ecm_full_drop1_d1_indust.csv (e equation, disaggregate, 3a)
ccep2_semiannual_PW_ecm_full_drop1_d1_indust.csv (p equation, disaggregate, 3a)

"Table 3\results\3b" (3b results)
ccep_agg_biascorrection_ecm_newPT3_demeanq_semiannual_PW_f1_indust.csv (aggregate, 3b)
ccep_disagg_biascorrection_ecm_newPT3_demeanq_semiannual_PW_f1_indust.csv (disaggregate, 3b)

Directions: Aggregate results are self evident. To calculate the disaggregate results, take averages and standard deviations across goods (columns C-K) for b_coefs and t_stats.

Table 4: Monte Carlo experiment for CCEP estimator

Data: "data_creation\datasets"
aggregate_semiannual_lc_drop1_f1_wide_indust.dta

Programs: "Table 4\programs"
MC1_2var_agg.do

Results: "Table 4\results"
MC_2var_agg_method1.csv

Directions: Columns B-E contain results for the e equation and columns O-R contain results for the p equation.

Table 5: Vector error correction estimates using Imbs etal dataset
Data: "data_creation\datasets"
Imbsdata_agg_lc_wide.dta (aggregate)
Imbsdata_sectoral_lc_wide.dta (disaggregate)

Programs:
"Table 5\agg\programs"
an_ccep_ecm_full_imbsdata_AGG.do (aggregate)
"Table 5\disagg\programs"
an_ccep_ecm_full_imbsdata.do (disaggregate)

Results: "Table 5\results"
ccep1_ecm_full_imbsdata_AGG.csv (e equation results, aggregate)
ccep1_ecm_full_imbsdata.csv (e equation results, disaggregate)

ccep2_ecm_full_imbsdata_AGG.csv (p equation results, aggregate)
ccep2_ecm_full_imbsdata.csv (p equation results, disaggregate)

Directions: Take the average of b_coefs and t-stats (columns C+E) for the disaggregated data. As noted above, 1 indicates the e equation and 2 indicates the p equation.

Table 6: 3-Equation vector error correction estimates

Data: "data_creation\datasets"
semiannual_lc_newPT3_PWT_f1_wide_indust.dta

6a)
Programs 6a: "Table 6\programs\6a"
ccep_semiannual_ecm.do
ccep_semiannual_ecm_averages.do

Results for part a: "Table 6\results\6a"
ecm_newPT3_demeanq_semiannual_PW_f1_indust.csv
ecm_newPT3_demeanq_semiannual_PW_f1_indust_averages.csv

6b)
Programs 6b: "Table 6\programs\6b"
all.do
all_averages.do

Results for part b: "Table 6\results\6b"
ccep_biascorrection_ecm_newPT3_demeanq_semiannual_PW_f1_indust.csv
ccep_biascorrection_ecm_newPT3_demeanq_semiannual_PW_f1_indust_averages.csv

Directions: The averages and standard deviations are calculated in STATA, and results can be found in files with the suffix "averages".

Table 7: Relative contributions of nominal exchange rate and price adjustment to PPP and LOP reversion

Data: Data are already input manually and references coefficient results from Table 6.

Programs: "Table 7\Programs"
imprs_by_shock.m

Results: "Table 7\Results"
imprs_contributions.xlsx

Directions: As documented in MATLAB, one must select the shockvar # for each corresponding shock (1= e shock, 2 = p shock, 3 = pk shock). The output is listed in MATLAB as well as in results file.

Table 8: Estimates of half-lives conditional on the shock

Data: "BerginGlickWu Replication\data_creation\datasets"
semiannual_lc_newPT3_PWT_f1_wide_indust.dta
irfcoeffs_only_semiannual_PW_f1_indust.dta (from stage 1 in 8a)

8a) This is a 2-stage process. In stage 1 the impulse response function (IRF) coefficients are generated. In stage 2, the half-lives are calculated.

Programs 8a: "Table 8\8a\programs"
Stage 1: an_ccep_semiannual_ecm_newPT3_demeanq_f1_indust.do (saves the IRF coefficients)
Stage 2: an_irf_unrestricted_newPT3_indust.do (generates half-lives):

Results for part a: "Table 8\results\8a"
ccep_halflives_unrestricted_ecm_newPT3_semiannual_PW_f1_indust.csv

8b)
Program 8b: "Table 8\8b\programs"
all.do

Results for part b: "Table 8\results\8b"
ccep_halflives_ecm_newPT3_demeanq_semiannual_PW_f1.csv

Directions: Average each half-life (columns B-G) for results in Table 8a.
Table 8b requires calculations done in excel; see ccep_halflives_ecm_calculation.xlsx for the format.

Table 9: Estimates of speeds of adjustment in expanded autoregression

Data: "BerginGlickWu Replication\data_creation\datasets"
semiannual_lc_newPT3_PWT_f1_wide_indust.dta

Program: "Table 9\programs"
an_ccep_semiannual _Qqk_demeanq_f1_indust.do

Results: "Table 9\results"
Qqk_demeanq_semiannual_PW_f1_indust.csv

Directions: Average columns for ρ coefficients and t-stats for disaggregate (columns L-Q) and for aggregate (columns U-Z). For the standard deviation, take StDev of columns L, O, U, X.

Figures 1-5

Data: "BerginGlickWu Replication\data_creation\datasets"
semiannual_lc_newPT3_PWT_f1_wide_indust.dta
irfcoeffs_newPT3_semiannual_PW_f1_indust.dta (generated from stage 1 below)

Programs: "Figures 1-5\programs"
2-stage process
Stage 1: an_ccep_semiannual_ecm_newPT3_demeanq_f1_indust.do (saves the coefficients and runs IRFs)
Stage 2: an_irf_newPT3_summarize_allsamples.do (graphs figures 1-5. Uses newly created dataset: irfcoeffs_newPT3_semiannual_PW_f1_indust.dta).

Results: "Figures 1-5\Graphs":
Figure 1: CVDsummary_qk_indust.png
Figure 2: CVDsummary_Q_indust.png
Figure 3: irfsummary_pkshock_nobounds_indust.png

Figure 4: irfsummary_eshock_nobounds_indust.png
Figure 5: irfsummary_Pshock_nobounds_indust.png

References:

Crucini, Mario J. and Mototsugu Shintani, 2008. "Persistence in Law-of-One-Price Deviations: Evidence from Micro-data," *Journal of Monetary Economics* 55, 629-644.

Engel, Charles and John H. Rogers, 2004. "European Product Market Integration after the Euro," *Economic Policy* 39, 347-384.

Imbs, Jean M., Haroon Mumtaz, Morten O. Ravn, and Helene Rey, "PPP Strikes Back: Aggregation and the Real Exchange Rate," *Quarterly Journal of Economics* 120:1 (2005), 1-43.

Kilian, Lutz, "Small-Sample Confidence Intervals For Impulse Response Functions," *Review of Economics and Statistics* 80:2 (1998), 218-230.

Parsley, David and Shang-jin Wei, 1996. "Convergence to the Law of One Price Without Trade Barriers or Currency Fluctuations," *Quarterly Journal of Economics* 111, 1211-1236.

Pesaran, M. Hashem, "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence, *Journal of Applied Econometrics* 22:2 (2007), 265-312.