Trade-offs? The Impact of WTO Accession on Intimate Partner Violence in Cambodia

By Bilge Erten and Pinar Keskin

I. Access to Data

The analysis reported in the paper is based three sources of individual-level data from Cambodia. First, we use two waves of the Cambodian population census data from IPUMS. The specific waves are:

1. Cambodia 1998 IPUMS census
2. Cambodia 2008 IPUMS census

Both census waves can be accessed through the IPUMS website: https://international.ipums.org/international/. A standard data user registration is required.

Second, we use three waves of the Demographic and Health Surveys (DHS) for Cambodia. These are cross-sections of women aged 15-49. The specific waves are:

3. Cambodia 2000 DHS
4. Cambodia 2005 DHS
5. Cambodia 2014 DHS

The 2000, 2005 and 2014 waves can be accessed through the DHS website: https://dhsprogram.com/. A standard data user registration is required.

Third, we use five waves of the Cambodian Socioeconomic Surveys (CSES). These are annual household surveys conducted by the National Institute of Statistics of Cambodia (NISC). We use the following waves:

6. Cambodia 1996 CSES
7. Cambodia 1999 CSES
8. Cambodia 2003 CSES
9. Cambodia 2009 CSES
10. Cambodia 2014 CSES

The CSES is not publicly available. Researchers interested in accessing the CSES data for purposes of replication should contact NISC through the following website: https://nis.gov.kh/index.php/km/

Finally, we use import tariff data from the WITS–TRAINS database, which can be accessed through the following link:

http://wits.worldbank.org/WITS/WITS/AdvanceQuery/TariffAndTradeAnalysis/AdvancedQueryDefinition.aspx?Page=TariffandTradeAnalysis

Import tariff data can be accessed through the WITS-TRAINS database. A standard data user registration is required.

II. Reconstructing the Final Datasets

The final datasets used in the paper are the following:

i. *census_for_analysis.dta:* This is the cleaned version of the 1999, 2003, 2009 and 2014 CSES datasets. To go from the raw version to the cleaned version, we exclude (i) individuals with missing district information and (ii) individuals with missing information on covariate data. The final dataset contains 1,229,969 individuals aged 15-64.

ii. *DHS_data_for_analysis.dta*: This is the cleaned version of the 2000, 2005, and 2015 DHS datasets. To go from the raw version to the cleaned version, we exclude (i) women with missing district information and (ii) women with missing information on covariate data. The final dataset contains 47,416 women aged 15-49.

iii. *DHS_data_additional.dta:* This is the collapsed version of the dataset, DHS_data_for_analysis.dta, at the district level for additional data analysis. It also includes information on initial sector weights and industry tariff rates over time.

iv. *CSES_data_for_analysis.dta*: This is the cleaned version of the 1999, 2003, 2009 and 2014 CSES datasets. To go from the raw version to the cleaned version, we exclude (i) women with missing district information and (ii) women with missing information on covariate data. The final dataset contains 118,341 individuals aged 15-64.

v. *CSES_data_for_analysis_expenditure:* This is the cleaned version of the expenditure modules from the 1999, 2009 and 2014 CSES datasets at the household level. To go from the raw version to the cleaned version, we exclude (i) households with district information and (ii) households with missing information on covariate data for household heads. The final dataset contains 24,995 households.

vi. *tariffs_3digit.dta:* This is the 3-digit ISIC revision 3 import tariff data obtained from the WITS–TRAINS database. There are 19,845 observations in this dataset.

vii. *sic_code_desc.dta:* This file includes the industry names and codes for the traded industries in the Census data. There are 40 observations in this dataset.

III. Results

The following do files reproduce the tables and figures reported in the paper, and generate the corresponding do files:

i. ***1a Tables using Census data:*** This do file generates all tables using the Census data. These include Table 1 Panel A, Table 2, and Appendix Tables A7, A8, A11, A14, and A17.

ii. ***1b Tables using DHS and CSES data:*** This do file generates all tables using CSES data and all tables using DHS data except Appendix Table A20. These include Table 1 Panel B, Table 3, Table 4, and Appendix Tables A2, A3, A4, A5, A6, A9, A10, A12, A13, A15, A16, A18, A19, and A21.

iii. ***1c Appendix Table A20 and Figure A3:*** This do file generates Appendix Table A20 and Figure A3 using DHS data.

iv. ***2 Figures:*** This do file generates all figures in the paper except Figure A3. These include Figure 1, Figure A1, and Figure A2.

Each do file directly outputs a .tex file for tables and an .eps file for figures, which also reproduce the results used in the paper.