

Data for “Untangling Trade and Technology: Evidence from Local Labor Markets”

David H. Autor
MIT and NBER

David Dorn
University of Zurich and CEPR

Gordon H. Hanson
UCSD and NBER

November 5, 2014

Data files (folder */dta*)

The online data archive contains the following data sets in Stata format:

- *workfile_china.dta* provides data for 722 Commuting Zones times 2 time periods (1990-2000, 2000-2007), and is taken from Autor, Dorn and Hanson (AER 2013). All variables with the prefix “*l_*” refer to start-of-period values while variables with the prefix “*d_*” correspond to 10-year equivalent changes. Variables from this file are used in all Tables.
- *workfile2012.dta* provides data for 722 Commuting Zones times 2 time periods (1980-1990, 1990-2000, 2000-2005), and is taken from Autor and Dorn (AER 2013). All variables with the prefix “*l_*” refer to start-of-period values while variables with the prefix “*d_*” correspond to 10-year equivalent changes. Variables from this file are only used for the results of Table 5.
- *czoneYYYYpor_current_v2.dta* and *czoneYYYYpor_mfg_v3.dta* provide additional Commuting Zone variables that measure unemployment, nonemployment and employment by sector and occupation for different subgroups of the population, and the share of routine employment by Commuting Zone and sector. The files $YYYY=\{1980, 1990, 2000\}$ are based on 5% IPUMS samples of the Decennial Census of Population while the file $YYYY=060708$ is based on the pooled 2006-2008 file of the American Community Survey. Variables from this file are used in all Tables.
- *czone1950por_indstr_v2.dta* provides for each Commuting Zone the share of routine employment, overall and by sector, as predicted by the Commuting Zone’s industrial employment structure in 1950. The data is based on the 1% IPUMS sample of the Decennial Census of Population of 1950. Variables from this file are used in all Tables except Appendix Table 1 as instruments for the start-of-period share of routine employment.
- *cbp_mfgsh.dta* provides for each Commuting Zone the share of manufacturing in total employment according to the County Business Pattern data, and measured separately for the years 1980, 1990 and 2000. This information is used as a control variable in all Tables.

Further documentation of the Autor, Dorn and Hanson (AER 2013) and Autor and Dorn (AER 2013) data, including downloadable file archives for each project and additional information on Commuting Zones, is available from David Dorn’s data page, www.ddorn.net/data.htm.

Computation of results (folder */do*)

The following Stata do-file combines the data files and computes the results of the paper:

- *czone_analysis_untangling_final.do* computes results for Tables 1 to 5 and Appendix Table 1.

Raw results (folder */log*)

The folder */log* contains the regression log file that is produced by the file in the folder */do*.

Tables (folder */tab-fig*)

The Excel file *China-Untangling-Final.xls* provides the full set of Tables that is presented in the paper.