Description of data files for
"Looking for Local Labor Market Effects of NAFTA"
Shushanik Hakobyan and John McLaren
Date: September 2015

The following data sets and codes are compiled using Stata/SE 13.1 for Windows (64-bit x86-64). To replicate the results in the paper, first execute "data.do" to create the data files, then run "regressions.do" to obtain the results in Tables 4-10 of the paper.

****

Data sources

****

The raw data (US Census 1990 and 2000, 5% sample) come from IPUMS-USA (usa.ipums.org/usa/), saved as "**usa_00017.dat**". For placebo regression in Table 7, we use US Census 1980 and 1990 (5% sample) from the same source, saved as "**usa_00022.dat**".

To construct industry and conspuma level tariffs, we obtain tariff data from the website of John Romalis. The resulting file is **tariffs1990-2000.dta**. To construct trade weights for aggregating tariffs up to Census industry categories, we use imports from Mexico in 1990 obtained from USITC, saved as **imports1990-2000.dta**. To control for Chinese imports growth, we use US imports from China and total imports in 1990 and 2000, obtained from USITC, saved as **chn_imports1990-2000.dta**. To construct Mexico's RCA index in 1990, we obtain Mexican and world exports from UN Comtrade (accessed via WITS at wits.worldbank.org). The raw data are saved as **worldex19902000.csv**.

concord_hts8_ind1990.dta provides concordance map between HS 8-digit product categories and Census industry categories.

****

Data files

****

main.dta and **main1980.dta** have the same structure.

Variable name	Variable description
conspuma	IPUMS generated identifier for Consistent Public Use Microdata Areas (CONSPUMA). The variable takes values from 1 to 543.
perwt	Person weight
age	Age (restricted between 25 and 64)
age2	Age squared
male	Binary variable that takes the value of 1 if the individual is male
married	Binary variable that takes the value of 1 if the individual is married
white	Binary variable that takes the value of 1 if the individual is white
speakeng	Binary variable that takes the value of 1 if the individual speaks English
empstat	Employment status (1 = employed, 2 = unemployed, 3 = NILF)

ind1990	Census industry code, 1990 basis
incwage	Wage and salary income
yr	Binary variable that takes the value of 1 if the year is 2000
lhs	Binary variable that takes the value of 1 for individuals with less than high school education
hs	Binary variable that takes the value of 1 if the individual is a high school graduate
scol	Binary variable that takes the value of 1 for individuals with some college education (includes associates' degrees)
col	Binary variable that takes the value of 1 if the individual is a college graduate

trade_ind1990_fixed.dta – contains tariff data at the industry level

Variable name	Variable description
mex_t90	US Tariff against Mexico in 1990
mex_t00	US Tariff against Mexico in 2000
mex_dt	Change in tariff
mex_t90_noag	Tariff in 1990 (agric = 0)
mex_dt_noag	Change in tariff (agric = 0)

locvul_mexrca_fixed.dta – contains average local tariff data and local measure of change in share of Chinese imports (at conspuma level)

Variable name	Variable description
conspuma	Consistent Public Use Microdata Area
locvul	Average local tariff in 1990
locdt	Change in average local tariff
locvul_noag	Local tariff in 1990 (agric = 0)
locdt_noag	Change in local tariff (agric = 0)
locdchnmshare	Local employment weighted change in Chinese imports share

mex_rca.dta – contains Mexico's RCA in 1990 for tradable industries
chnmshare.dta – contains change in share of Chinese imports at the industry level
dmshare.dta – contains change in share of Mexican imports at the industry level
locdmshare.dta – contains change in share of Mexican imports at the conspuma level

****

Stata codes

****

data.do – reads the raw data from IPUMS-USA and creates main datasets.

vulnerability.do – creates various measures of vulnerability at the industry- and conspuma level, including initial tariffs and change in tariffs, change in Mexican import share, change in Chinese import share.

regressions.do – runs the regressions reported in Tables 4-10.