Readme file . Anna Aizer. July 16, 2009.

1) Creating the measures of wages
   a. aer_census1990.do: from 1990 ipums sample(census.dta), generate **censusraceyr.dta** which includes the breakdown of employment by industry for males and females in 1990. For a handful of counties not represented in the ipums data due to small population size, nearest county was used.
   b. aer_bartik.do: merges the files censusraceyr.dta with NAICS data that includes the annual employment and wages by industry for all counties in California. The NAICS files are numbered naics90.dta – naics103.dta corresponding to the year (1990-2003). Naicsca03 contains data on employment and wages by industry for the entire state by year. The final dataset is called **bartik03_aer.dta** and it contains the measures of female and male wages used in the analysis.
   c. Bartik03_industry.do :redefining the wage gap based on changes in industrial composition of the county. This file merges censusraceyr.dta with census2000.dta and generates linear interpolations. These data are then merged with 1990 data on wages by industry. The resulting dataset is **bartik03_industry.dta**.
   d. Bartik03_aeriv.do creates the wage ratio and instrument for the IV regressions in table 4 column 2**. Bartik03_aeriv.dta**

2) Additional data to merge together:
   a. Data on immigration by county and year **immigration.dta**. These data are available from the California Department of Finance, Research Unit. See http://www.dof.ca.gov/research/
   b. Data on unemployment and per capita income by county and year **unempinc03.dta.** These data are available from the California Department of Finance, Economic and Financial Research Unit. See http://www.dof.ca.gov/research/
   c. Data on domestic violence and non-intimate homicide by county, race and year **cadv_2003.dta.** These data come from the State of California, Department of Justice, Criminal Justice Statistics Center. See http://ag.ca.gov/crime.php.
   d. Population counts by gender county race and year **demogb03.dta** . These data are available from the California Department of Finance, Demographic Research Unit. See http://www.dof.ca.gov/research/
   e. College enrollment by gender county race and year **ccenroll.dta**. These data are available at the California Postsecondary Education commission. See http://www.cpec.ca.gov/OnLineData/SelectFirstOptions.ASP?ReportType=Enroll
   f. Primary care clinic data by county and year **clinicall.dta.** These data came from OSPHD and can be accessed: http://www.oshpd.ca.gov/hid/Products/Hospitals/Utilization/PC_SC_Util_Info.html. Earlier years must be requested from OSHPD.

g. Drug admissions by county and year **drugs1990-2003.dta**

h. Male arrests for domestic violence by county race and year **malearrests1990-2003.dta.** These data come from the State of California, Department of Justice, Criminal Justice Statistics Center. See http://ag.ca.gov/crime.php.

i. Male releases/detentions from prison by county race and year **caincar1990-2003.dta.** These data are available from ICPSR as part of the National Corrections Reporting program series.

3) California hospital data is not publicly available, but interested researchers can apply for these data from the state of California OSHPD.

OSHPD
Healthcare Information Resource Center (HIRC), Suite 250
400 R Street
Sacramento, CA 95811-6213
Phone: (916) 326-3802
Fax: (916) 324-9242

   a. The following programs extract the hospital discharge data: disch90.do disch91.do disch92.do disch93.do disch94.do… Combine2.do combines all the data for 1990-2000 into **hosp2.dta**. Finally, combine2_2003.do which extracts data 2001-2003 and combines with data from the 1990s (hosp2.dta). Resulting dataset is **hosp2_2003.dta** (this dataset is <u>not</u> included). Ecode.do is a subroutine.
   b. Makeall03_aer.do is the program that combines all the above datasets and generates the final dataset for analysis**: dataall03_aer.dta.**

4) Analysis files
   a. Aer_analysis.do is the main program that calls in the following sub-routines: aerind_oct15.do and weekend.do as well as oct9globals.do