

Sociology 213, Lab 15

Creating a longitudinal data file from the Current Population Survey.

Most of the lab time for this lab will be devoted to work on your projects. I would like to discuss your project topics and analysis with each of you in lab.

However, I have written a small program to try to match data across consecutive years of the CPS. The article we have read for class by Macpherson and Hirsch demonstrates the use of longitudinal CPS data. The advantage of doing this is the large number of cases that you can generate—400,000 matched cases in their paper—which might allow you to construct a longitudinal data set for a very specific population (i.e. Dominican immigrants, or nurses etc.).

We will run the program together in lab.

As pre-lab homework, I would like you to briefly annotate the ado file to make sure you understand the Stata commands that are used in it.

We will discuss how the program works at the beginning of lab.

Stata ado program for matching CPS data, mergorg2.ado

```
program define mergorg2

global yr1="`1'"
global yr2="`2'"

clear
set more 1

use morg${yr2}

keep if minsamp==8

keep hhid state lineno age marital race sex docc /*
*/ earnwt occ80 uhours earnwke intmonth

rename lineno ylineno
rename age yage
rename marital ymarital
rename uhours yuhours
rename earnwke yearnwke
rename earnwt yearnwt
rename occ80 yocc80
* rename gradeat yed

sort hhid state race sex intmonth

di "number of possible matches " _N
local bign=_N
save merge${yr2}, replace

use morg${yr1}

keep if minsamp==4

keep hhid state lineno age marital race sex docc /*
*/ earnwt occ80 uhours earnwke intmonth

rename lineno xlineno
rename age xage
rename marital xmarital
rename uhours xuhours
rename earnwke xearnwke
rename earnwt xearnwt
rename occ80 xocc80
* rename gradeat xed

sort hhid state race sex intmonth
save merge${yr1}, replace

use merge${yr2}
merge hhid state race sex intmonth using merge${yr1}
```

```

gen double hhid2=real(hhid)
* rename hhid hhida
* rename hhid2 hhid

di _N
tab _merge
keep if _merge==3
gen difage=yage-xage
tab difage
keep if difage>-1 & difage<3

egen n=count(state), by(hhid xlineno)
tab n

keep if n==1

egen n2=count(state), by(hhid ylineno)
tab n2
keep if n2==1

gen b=xmarital==ymarital
di "is marital status the same?"
tab b

sort hhid ylineno

di _N
di _N/`bign'
local mrate=_N/`bign'

gen occ=yocc80
capture drop year
gen year=${yr2}
sort year occ
capture drop _merge

merge year occ using ave_occ8693
rename fem yfem
capture drop n incwage
drop occ
tab _merge
sum yocc80 if _merge==1
drop _merge

gen occ=xocc80
capture drop year
gen year=${yr1}
sort year occ
capture drop _merge

merge year occ using ave_occ8693

tab _merge
sum xocc80 if _merge==1

```

```

rename fem xfem
capture drop n incwage
drop occ
drop _merge

sum xocc yocc

gen ywage=yearnwke/yuhours
gen lnwy=ln(ywage)
gen xwage=xearnwke/xuhours
gen lnwx=ln(xwage)

gen d_lnw=lnwy-lnwx
capture drop d_fem
gen d_fem=yfem-xfem

sum d_ln d_fem
xi: reg d_ln d_fem i.sex [aw=yearnwt]

gen cohort=${yr1}${yr2}

save merge${yr1}${yr2}, replace

di "match rate: " `mrate'

end

```