Soc213, Lab 11: Historical Trends in Gender Occupational Segregation and the Male/Female Wage Gap
(updated spring 2003)

Goal: In lab11 we will look at trends in gender occupational segregation using IPUMS data. While the topic we are studying is gender inequality, this same general approach will work for time-trends on any demographic characteristic. Because the IPUMS data is so large if we keep all the cases, we will collapse (aggregate) the information to the occupation level, where it will be easier to work with (notice mf.do, below).

Basic overview:
This lab takes individual data for each census from 1900-2000 and collapses the data by occupation. Then it uses the data sets of occupation to analyze trends in occupational segregation from 1900-2000.

Lab assignment: Turn in answers for questions 1-5.

Tip: if you want to stop a do file partway through to see what the data looks like, type a nonsense command. I.e., asdf

Directions:
The file lab11.zip includes the collapsed data for 1900-1990. Your goal is to add the data for 2000 and complete the lab.
1) Go to the IPUMS web page, www.ipums.org, to see the extract process for the 2000 data (you don't actually need to make the extract, since I have done it for you…this is a "virtual" step.
2) Uncompress the extract for the 2000 data. If your computer doesn't have winzip (which will decompress a .Z archive), use gzip.
Directions on using gzip: open up a dos prompt. Change the dos prompt to your working directory. Example:
h:
cd lab11
then, type: gzip –d tedmo001.dat.Z
3) load the data into Stata and collapse it using mf.do. Make the necessary changes to mf.do to load tedmo001.dat in and save it as mf2000

Question 1: *After the infix command in mf.do, what is a case? After the collapse command in mf.do, what is a case? Why would we want to do this?*

4) Read over occseg.do so that you understand it. Then, make the necessary changes to occseg.do so that it also uses the mf2000.dta that you just made. As currently written, it just uses the 1900-1990 data. (Hint: follow the existing pattern)

Question 2: *How does the reshape command*
```
reshape wide @incwage @n, i(occ1950 year) j(sex2) string
```
change the data? What is a case after this command? Why would I want to do this?

Question 3: The challenging part of occseg.do is calculating the index of dissimilarity for male/female occupational segregation. The formula is the same we used to calculate racial segregation. *Write down the mathematical formula for the index of dissimilarity and then explain, step by step, how occseg.do uses Stata to calculate this index.* Hint: you will need to use "help command" to understand how the egen commands work.

5) Run occseg.do and print the graph.

Question 4: *Briefly describe (2-4 sentences) the trends in gender occupational segregation and women's labor force participation.*

6) Look at the 3 list commands at the end of the occseg.do file. This list the most segregated male and female occupations and the most even (i.e. closest to 50%) occupations.

Question 5: *Change occseg.do to repeat these three list commands for a specific year (you choose the year), and print the results and turn it in. Are you surprised by the results or do they correspond to your intuition about historical patterns of m/f occupational segregation?*

```
------------mf.do-----------------


* cd to the working directory
pwd

clear
set mem 65m

infix using tedmo041.dct
* change this to your data dictionary

tab empstatg sex if empstatg>0, col
keep if empstatg==1

* sex: 1=male, 2=female

* note: if you do not have incwage in your data (before 1940, I think)
* then, remove the asterisk from the next line:

* gen incwage=-9

recode incwage 999999=.

collapse (mean) incwage (rawsum) n=perwt [fw=perwt], by(occ1950 sex)

save mf1980, replace
* change this to your year
```

```
----------------occseg.do-------------------

capture log close
log using occseg, text replace
clear

use mf1900
gen year=1900

append using mf1910
recode year .=1910

append using mf1920
recode year .=1920

append using mf1940
recode year .=1940


append using mf1950
recode year .=1950


append using mf1960
recode year .=1960

append using mf1970
recode year .=1970

append using mf1980
recode year .=1980


append using mf1990
recode year .=1990




tab year

gen occ2=occ1950

do occ1950


drop if occ2>=980
recode incwage -9=.

gen str2 sex2="m"
replace sex2="f" if sex==2
drop sex

sort year occ2 sex
des
list year occ2 n sex in 1/10
reshape wide @incwage @n, i(occ1950 year) j(sex2) string
```

```
recode mn .=0
recode fn .=0
recode mincwage .=0
recode fincwage .=0
sort year occ2
list year occ2 mn fn in 1/10
des

* calculate segregation index (index of dissimilarity)

replace fn=fn/1000
replace mn=mn/1000

egen ftot=sum(fn), by(year)
egen mtot=sum(mn), by(year)

gen fpct=fn/ftot
gen mpct=mn/mtot

gen dif=abs(fpct-mpct)



egen dif2=sum(dif), by(year)

replace dif2=dif2/2

rename dif2 segindex

lab var segindex "occupational segregation index"
tab year, sum(segindex)

gen totpf=ftot/(mtot+ftot)

lab var totpf "overall pct female in labor force"
tab year, sum(totpf)

*calculate m/f wage ratio
gen wm=mincwage*mn
gen wf=fincwage*fn
egen wmtot=sum(wm), by(year)
egen wftot=sum(wf), by(year)
replace wmtot=wmtot/mtot
replace wftot=wftot/ftot
gen wratio=wftot/wmtot
lab var wratio "overall m/f wage ratio"

gen owratio=fincwage/mincwage
lab var owratio "occ m/f wage ratio"

gen pctf=fn/(mn+fn)
lab var pctf "occ % female"

tab year, sum(wratio)


save temp, replace
```

```
sort year
by year: keep if _n==1

graph segindex wratio totpf year, c(mmm) xlab ylab(0, .25, .5, .75, 1)
saving(lab11, replace)

clear
use temp

lab var mn "number of men in occ, in 1,000s"
lab var fn "number of women in occ, in 1,000s"

gen occ=occ1950
lab val occ occ1950
format occ %12.0g


format %4.3f pctf owratio wratio totpf segindex  mpct fpct
format %5.1f fn mn
format %4.0f year occ2

gsort -pctf
list occ occ2 year pctf owratio fn mn mpct fpct in 1/20
gsort pctf
list occ occ2 year pctf owratio fn mn  mpct fpct in 1/20

gen equal=abs(pctf-.5)
gsort equal
list occ occ2 year pctf owratio fn mn  mpct fpct in 1/20
```