

Solutions for Session 11

Mark Lunt

06/12/2011

```
. set more off  
. clear  
. sysuse uslifeexp, clear  
(U.S. life expectancy, 1900-1999)  
. twoway scatter le year
```

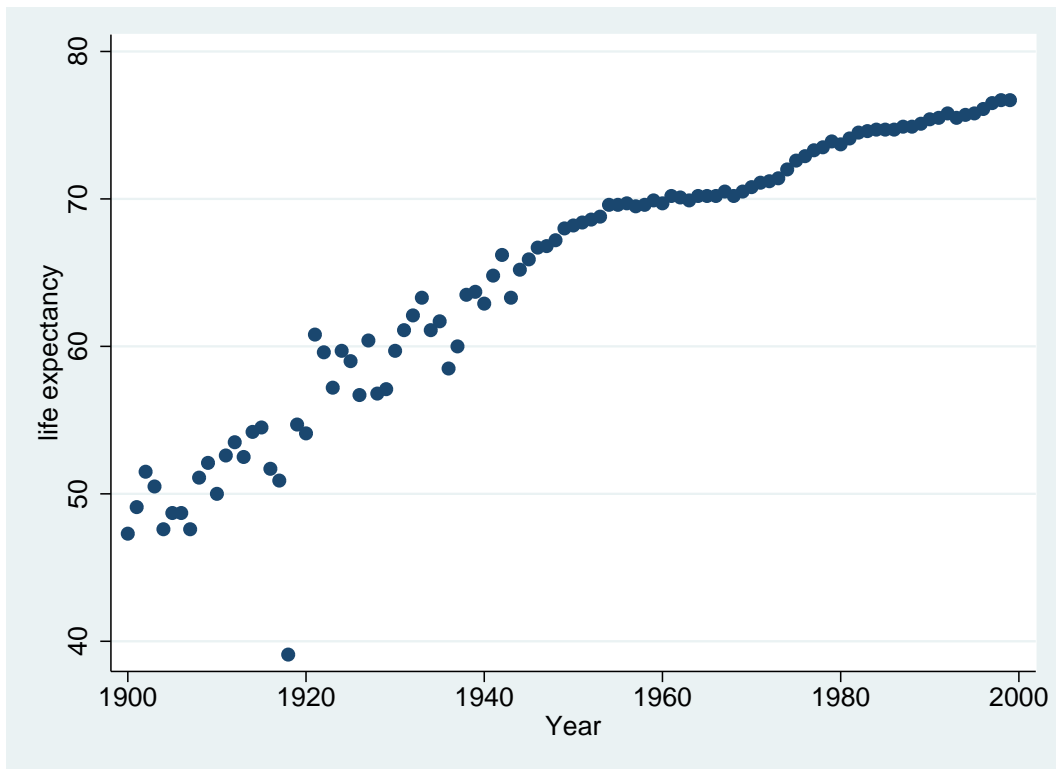


Figure 1: . twoway scatter le year

```

. graph export graph1.eps replace
(file graph1.eps written in EPS format)

. twoway scatter le year, title("U.S. Life Expectancy")

```

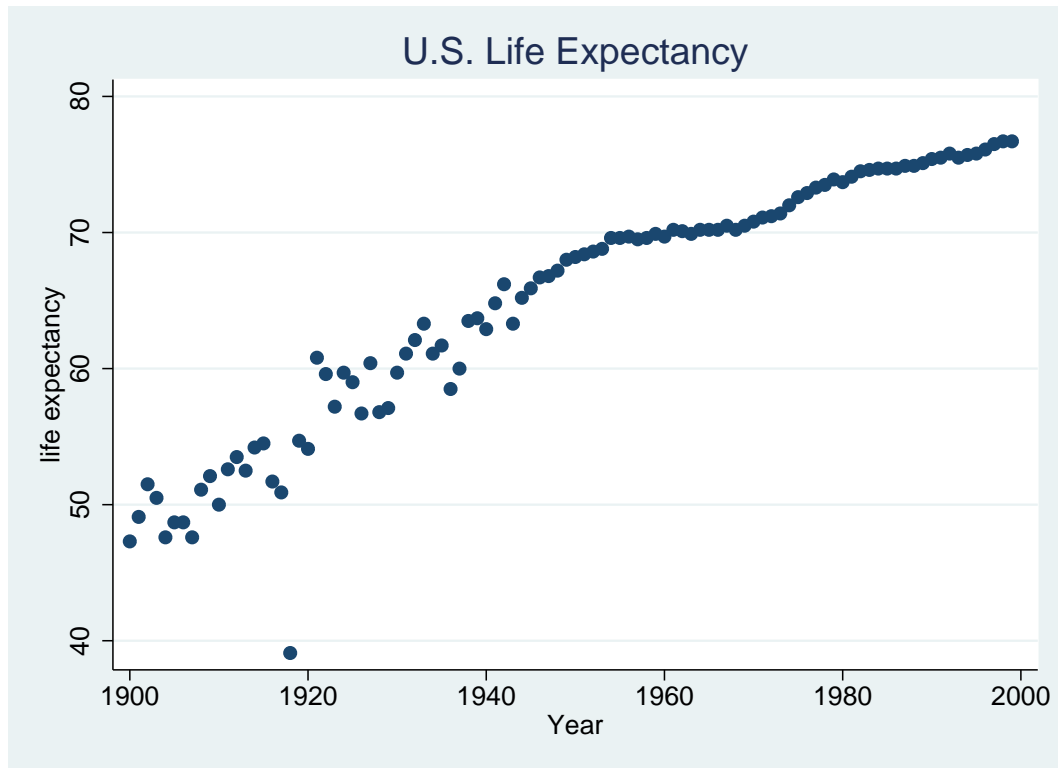


Figure 2: . twoway scatter le year, title("U.S. Life Expectancy")

```

. graph export graph2.eps replace
(file graph2.eps written in EPS format)

. twoway scatter le year, title("U.S. Life Expectancy") ylabel(0(20)80)

. graph export graph3.eps replace
(file graph3.eps written in EPS format)

. twoway scatter le_male year || scatter le_female year

. graph export graph4.eps replace
(file graph4.eps written in EPS format)

```

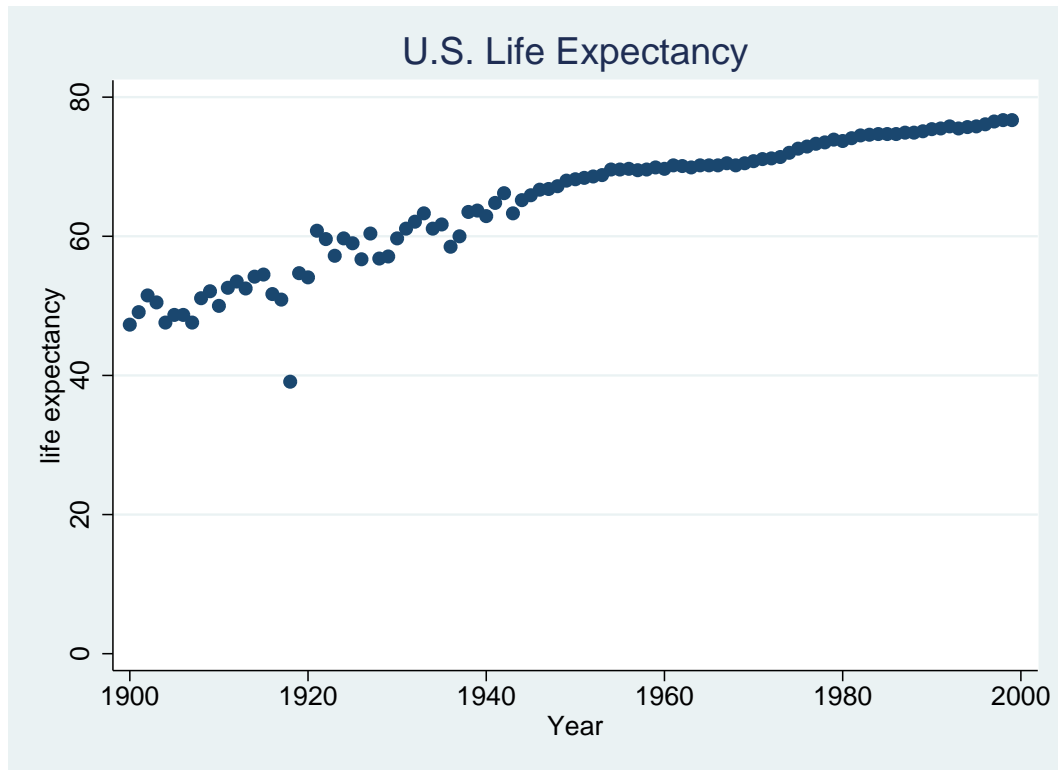


Figure 3: `. twoway scatter le year, title("U.S. Life Expectancy") ylabel(0(20)80)`

```
. twoway scatter le_male year || scatter le_female year || lfitci /* */
le_male year || lfitci le_female year

. graph export graph5.eps replace
(file graph5.eps written in EPS format)

. twoway scatter le_male year || scatter le_female year || lfitci /* */
le_male year || lfitci le_female year, ytitle("Life Expecancy")

. graph export graph6.eps replace
(file graph6.eps written in EPS format)

. sysuse cancer, clear
(Patient Survival in Drug Trial)
```

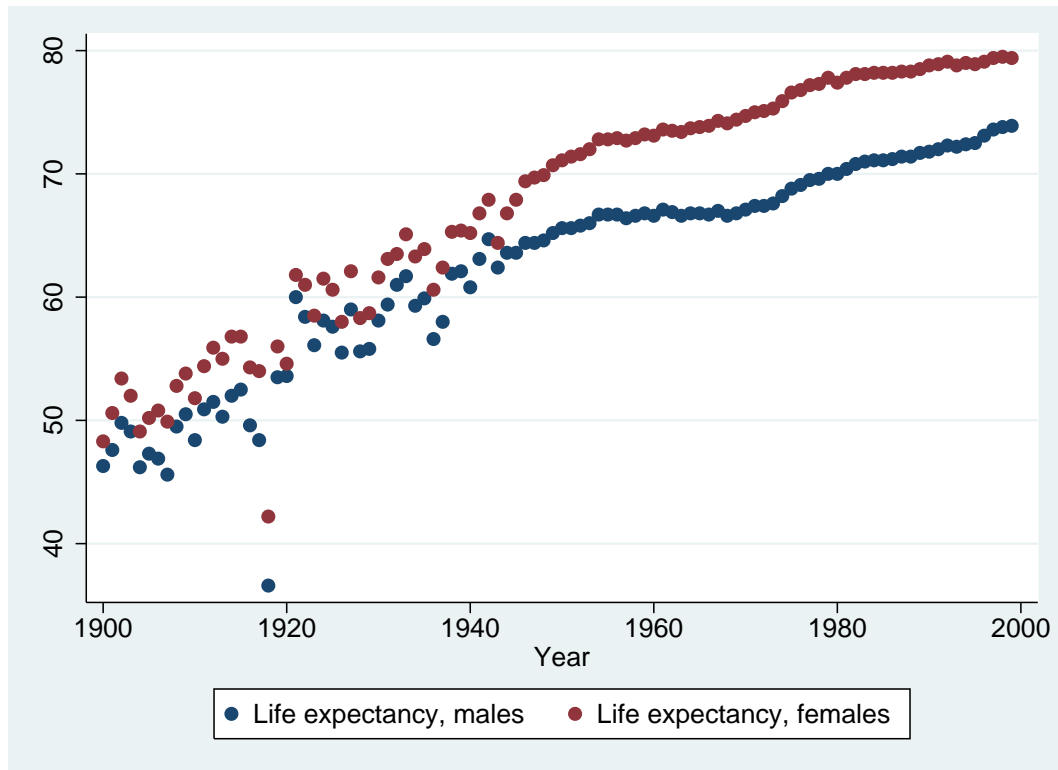


Figure 4: . twoway scatter le'female year — scatter le'male year

```
. describe
Contains data from /usr/local/stata11/ado/base/c/cancer.dta
  obs:      48                Patient Survival in Drug Trial
  vars:      8                3 Mar 2009 16:09
  size:     960 (99.9% of memory free)
```

variable name	storage type	display format	value label	variable label
studytime	int	%8.0g		Months to death or end of exp.
died	int	%8.0g		1 if patient died
drug	int	%8.0g		Drug type (1=placebo)
age	int	%8.0g		Patient's age at start of exp.
_st	byte	%8.0g		
_d	byte	%8.0g		
_t	byte	%10.0g		
_t0	byte	%10.0g		

Sorted by:

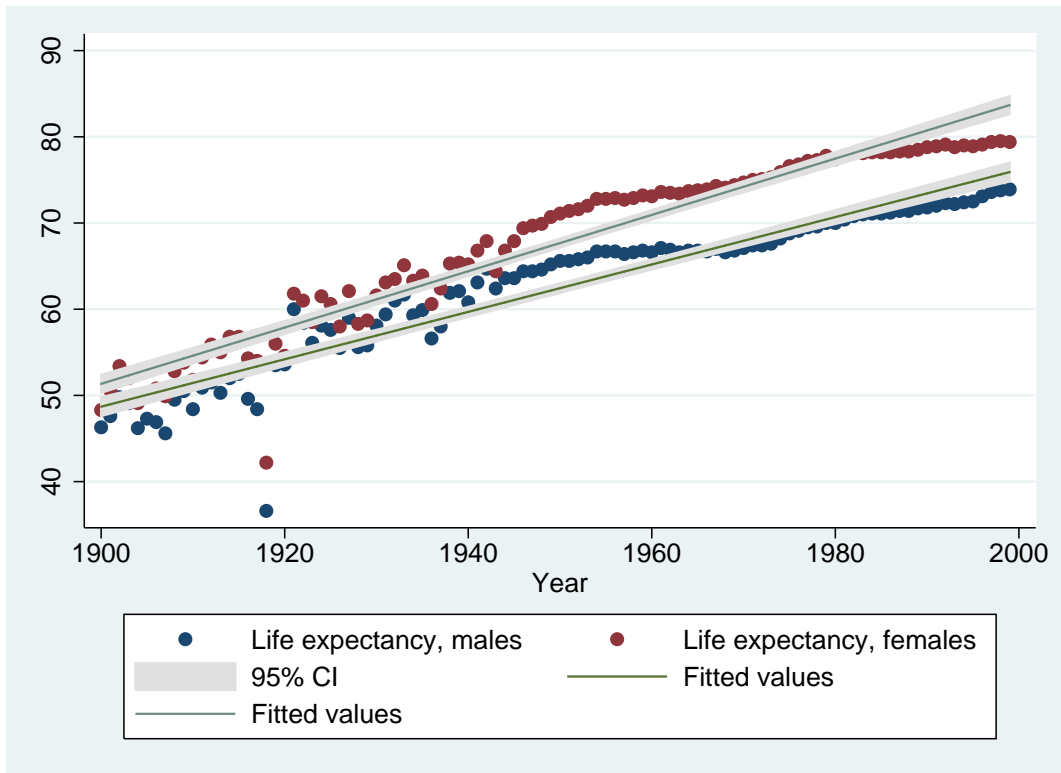


Figure 5: . twoway scatter le`male year — scatter le`female year — lfitci /*

2.1 There are 48 observations

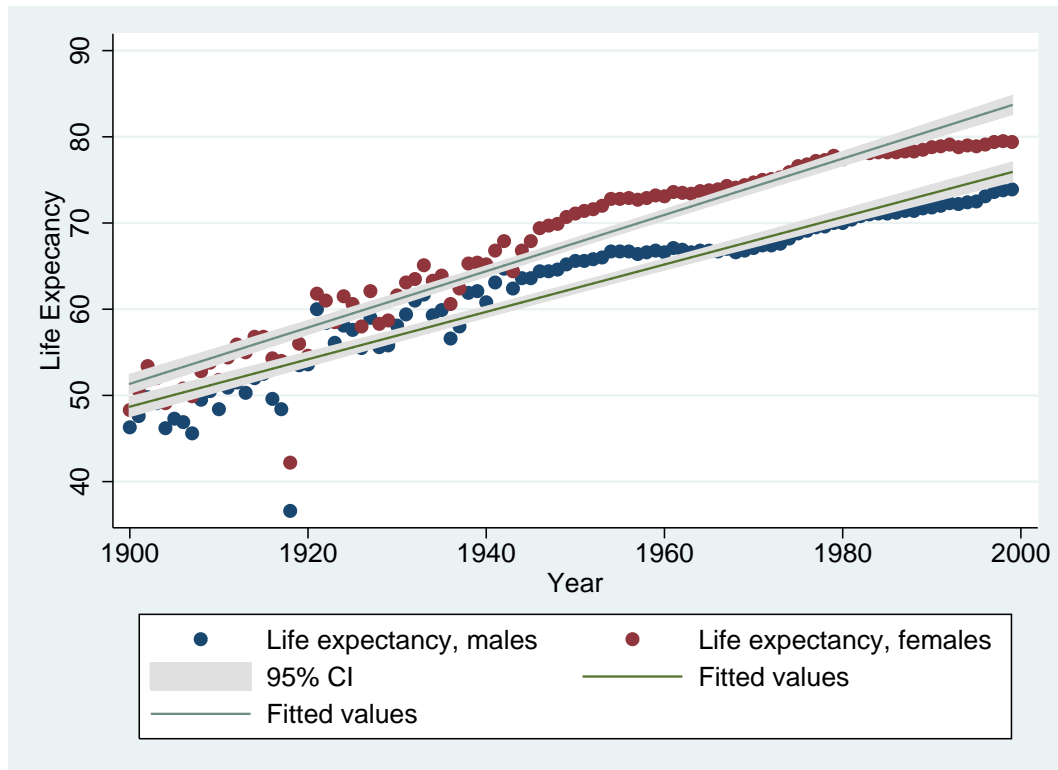


Figure 6: . twoway scatter le`male year — scatter le`female year — lfitted /*

```
. codebook
```

studytime	Months to death or end of exp.				
type:	numeric (int)				
range:	[1,39]		units:	1	
unique values:	28		missing .:	0/48	
mean:	15.5				
std. dev:	10.2563				
percentiles:	10%	25%	50%	75%	90%
	4	7.5	12.5	23	32

died	1 if patient died				
type:	numeric (int)				
range:	[0,1]		units:	1	
unique values:	2		missing .:	0/48	
tabulation:	Freq.	Value			
	17	0			
	31	1			

drug	Drug type (1=placebo)				
type:	numeric (int)	6			
range:	[1,3]		units:	1	
unique values:	3		missing .:	0/48	
tabulation:	Freq.	Value			
	20	1			
	14	2			
	14	3			

age	Patient's age at start of exp.				
-----	--------------------------------	--	--	--	--

- 2.2 the longest followup time is 39 months
- 2.3 There were 3 different treatments in the study (1 being a placebo)
- 2.4 The ages ranged from 47 to 67
- 2.5 The mean age at the start of the study was 55.875
- 2.6 The SD of the followup time was 10.26

. tab drug died, row

Key
frequency row percentage

Drug type (1=placebo)	1 if patient died		Total
	0	1	
1	1 5.00	19 95.00	20 100.00
2	8 57.14	6 42.86	14 100.00
3	8 57.14	6 42.86	14 100.00
Total	17 35.42	31 64.58	48 100.00

- 2.7 19 subjects on placebo died
- 2.8 43% of subjects on treatment 2 died

. summarize age if died == 1

Variable	Obs	Mean	Std. Dev.	Min	Max
age	31	56.80645	5.647533	47	67

3.1 Mean age = 56.8

. summarize age if drug == 1

Variable	Obs	Mean	Std. Dev.	Min	Max
age	20	56.05	5.558067	49	67

3.2 Mean age of those on placebo = 56.1

. summarize age if drug == 1 & died == 1

Variable	Obs	Mean	Std. Dev.	Min	Max
age	19	55.94737	5.690867	49	67

3.3 Mean age of those on placebo who died = 55.9

. bysort died: summ age

-> died = 0

Variable	Obs	Mean	Std. Dev.	Min	Max
age	17	54.17647	5.434097	48	65

-> died = 1

Variable	Obs	Mean	Std. Dev.	Min	Max
age	31	56.80645	5.647533	47	67

. bysort drug: summ age

-> drug = 1

Variable	Obs	Mean	Std. Dev.	Min	Max
age	20	56.05	5.558067	49	67

-> drug = 2

Variable	Obs	Mean	Std. Dev.	Min	Max
age	14	56.92857	6.787594	47	67

-> drug = 3

Variable	Obs	Mean	Std. Dev.	Min	Max
age	14	54.57143	4.636217	48	62

```
. bysort drug died: summ age
```

```
-> drug = 1, died = 0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	1	58	.	58	58

```
-> drug = 1, died = 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	19	55.94737	5.690867	49	67

```
-> drug = 2, died = 0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	8	54.75	5.750776	49	65

```
-> drug = 2, died = 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	6	59.83333	7.467708	47	67

```
-> drug = 3, died = 0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	8	53.125	5.540436	48	62

```
-> drug = 3, died = 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	6	56.5	2.258318	54	60

```
. egen agegrp = cut(age), group(2)
```

```
. bysort agegrp: summ age
```

```
-> agegrp = 0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	23	51	2.486326	47	55

```
-> agegrp = 1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
age	25	60.36	3.650114	56	67

3.4 The group would be split at the median age, which according to codebook was 56

```
. label define agegrp 0 "47-55" 1 "56-67"
```

```
. label values agegrp agegrp
```

```
. tab agegrp died
```

agegrp	1 if patient died		Total
	0	1	
47-55	10	13	23
56-67	7	18	25
Total	17	31	48

```
. bysort agegrp: gen group_size = _N
```

```
. tab group_size
```

group_size	Freq.	Percent	Cum.
23	23	47.92	47.92
25	25	52.08	100.00
Total	48	100.00	

```
. save mycancer, replace  
file mycancer.dta saved
```

```
. foreach x in one two three {  
  2. display " 'x' "  
  3. }  
one  
two  
three
```

```
. foreach x in drug agegrp {
2. tab 'x' died, row
3. }
```

Key
<i>frequency</i> <i>row percentage</i>

Drug type (1=placebo)	1 if patient died		Total
	0	1	
1	1 5.00	19 95.00	20 100.00
2	8 57.14	6 42.86	14 100.00
3	8 57.14	6 42.86	14 100.00
Total	17 35.42	31 64.58	48 100.00

Key
<i>frequency</i> <i>row percentage</i>

agegrp	1 if patient died		Total
	0	1	
47-55	10 43.48	13 56.52	23 100.00
56-67	7 28.00	18 72.00	25 100.00
Total	17 35.42	31 64.58	48 100.00

```
. sysuse uslifeexp
(U.S. life expectancy, 1900-1999)

. foreach x of varlist le* {
2. twoway scatter 'x' year, name("'x'")
3. }

. graph export graph7.eps replace
(file graph7.eps written in EPS format)

. sysuse bplong
(fictional blood-pressure data)
```

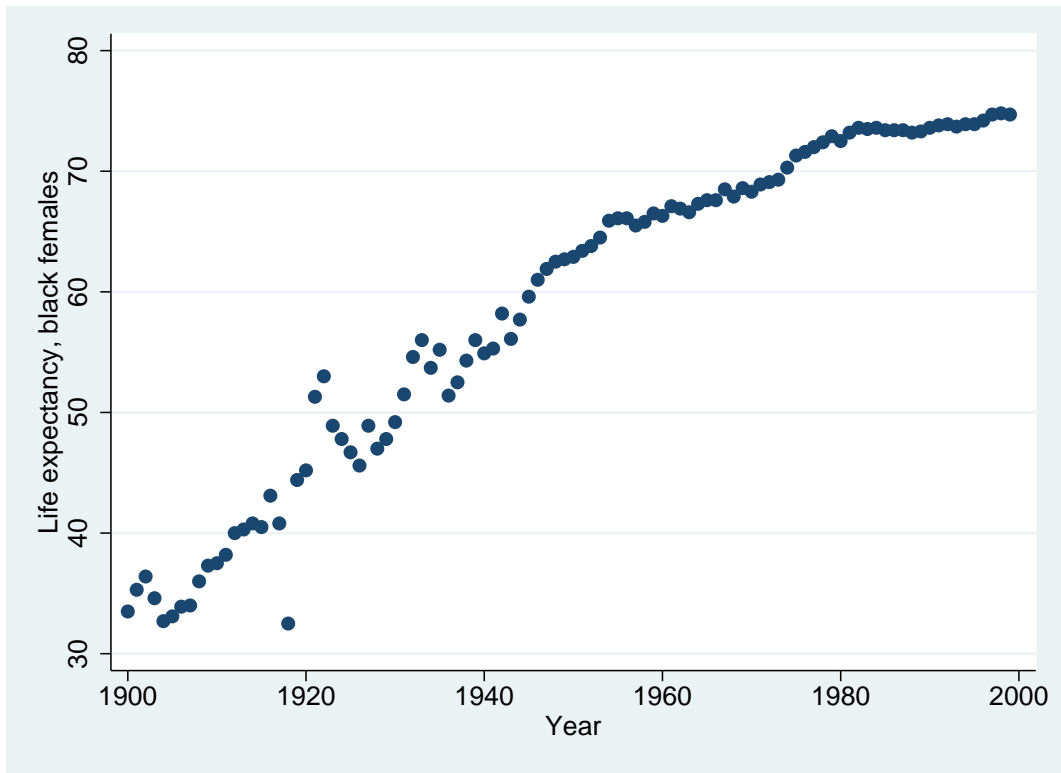


Figure 7: . foreach x of varlist le* {

```
. bysort when: summ bp
```

-> when = Before

Variable	Obs	Mean	Std. Dev.	Min	Max
bp	120	156.45	11.38985	138	185

-> when = After

Variable	Obs	Mean	Std. Dev.	Min	Max
bp	120	151.3583	14.17762	125	185

Mean BP before 156.45, after 151.36

```

. reshape wide bp, i(patient) j(when)
(note: j = 1 2)
Data                long  ->  wide
-----
Number of obs.      240  ->   120
Number of variables  5   ->    5
j variable (2 values) when -> (dropped)
xij variables:
                    bp   ->  bp1 bp2
-----

. summ bp1 bp2

```

Variable	Obs	Mean	Std. Dev.	Min	Max
bp1	120	156.45	11.38985	138	185
bp2	120	151.3583	14.17762	125	185

```

. sysuse uslifeexp, clear
(U.S. life expectancy, 1900-1999)

. rename le le_total

. reshape long le, i(year) j(group) string
(note: j = _b _bfemale _bmale _female _male _total _w _wfemale _wmale)
Data                wide  ->  long
-----
Number of obs.      100  ->   900
Number of variables  10  ->    3
j variable (9 values) ->  group
xij variables:
    le_b le_bfemale ... le_wmale ->  le
-----

. tab group

```

group	Freq.	Percent	Cum.
_b	100	11.11	11.11
_bfemale	100	11.11	22.22
_bmale	100	11.11	33.33
_female	100	11.11	44.44
_male	100	11.11	55.56
_total	100	11.11	66.67
_w	100	11.11	77.78
_wfemale	100	11.11	88.89
_wmale	100	11.11	100.00
Total	900	100.00	

```

. twoway scatter le year if group == "_male" || scatter le year if group == "_female"

. graph export graph8.eps replace
(file graph8.eps written in EPS format)

```

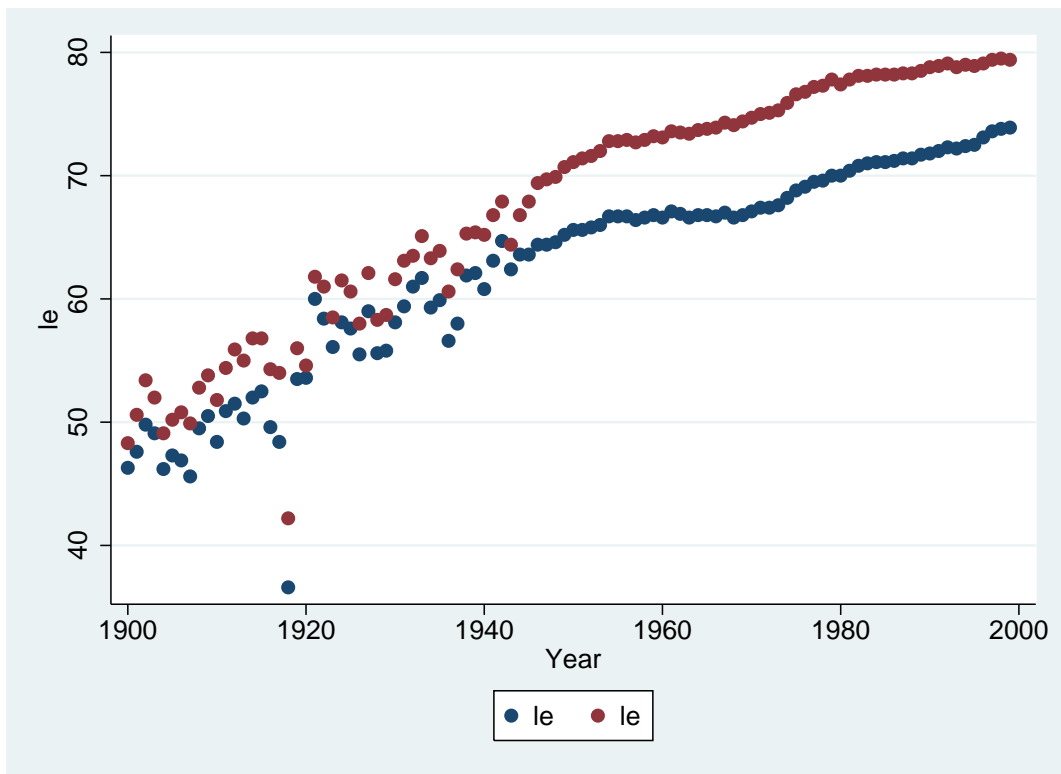


Figure 8: . twoway scatter le year if group == "male" — scatter le year if group == "f