ECON4150 - Introductory Econometrics

Seminar 1

Stock and Watson Chapter 2 & 3

Empirical exercise E3.1: Data

- In this exercise we use the data set *CPS92_12.dta*
- Each month the Bureau of Labor Statistics in the U.S. Department of Labor conducts the “Current Population Survey” (CPS).
- The CPS provides data on labor force characteristics of the population, including the level of employment, unemployment, and earnings.
- Approximately 65,000 randomly selected U.S. households are surveyed each month.
- The file CPS92_12 contains the data for 1992 and 2012.
- These data are for full-time workers, defined as workers employed more than 35 hours per week for at least 48 weeks in the previous year.

Empirical exercise E3.1: Data

Series in Data Set:

- FEMALE: 1 if female; 0 if male
- YEAR: Year
- AHE : Average Hourly Earnings
- BACHELOR: 1 if worker has a bachelor's degree; 0 if worker has a high school degree

Empirical exercise E3.1: Data

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
year	15052	2001.886	9.999679	1992	2012
ahe	15052	15.66179	9.44204	1.242788	91.45602
bachelor	15052	.4595403	.4983769	0	1
female	15052	.4252591	.4943987	0	1
age	15052	29.67944	2.822929	25	34

Empirical exercise E3.1: question a)

```
> ttest ahe, by(year) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1992	**7612**	**11.61683**	**.064409**	**5.61948**	**11.49057**	**11.74309**
2012	**7440**	**19.80026**	**.1238916**	**10.68632**	**19.5574**	**20.04312**
combined	**15052**	**15.66179**	**.0769607**	**9.44204**	**15.51094**	**15.81264**
diff		**-8.183424**	**.139634**		**-8.457131**	**-7.909717**

```
    diff = mean( 1992) - mean( 2012)                              t =   -58.6063
Ho: diff = 0                     Satterthwaite's degrees of freedom =   11203.7

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000
```

Empirical exercise E3.1: question b) & c)

b)

```
1 . gen ahe_cpi=ahe if year==2012
  (7612 missing values generated)

2 . replace ahe_cpi=ahe*(229.6/140.3) if year==1992
  (7612 real changes made)

3 .
4 . ttest ahe_cpi, by(year) unequal unpaired

  Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1992	**7612**	**19.01087**	**.1054049**	**9.19624**	**18.80425**	**19.21749**
2012	**7440**	**19.80026**	**.1238916**	**10.68632**	**19.5574**	**20.04312**
combined	**15052**	**19.40105**	**.0812489**	**9.968151**	**19.2418**	**19.56031**
diff		**-.7893878**	**.1626632**		**-1.108228**	**-.4705473**

```
    diff = mean( 1992) - mean( 2012)                              t =    -4.8529
Ho: diff = 0                     Satterthwaite's degrees of freedom =   14619.3

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000
```

c) The results from part (b) adjust for changes in purchasing power. These results should be used.

Empirical exercise E3.1: question d)

```
1 . ttest ahe if year==2012, by(bachelor) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	**3485**	**15.68547**	**.1397007**	**8.247077**	**15.41157**	**15.95938**
1	**3955**	**23.42605**	**.1790882**	**11.26264**	**23.07494**	**23.77717**
combined	**7440**	**19.80026**	**.1238916**	**10.68632**	**19.5574**	**20.04312**
diff		**-7.74058**	**.2271318**		**-8.185825**	**-7.295335**

```
    diff = mean( 0) - mean( 1)                                    t =   -34.0797
Ho: diff = 0                     Satterthwaite's degrees of freedom =   7203.16

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000
```

Empirical exercise E3.1: question e)

```
. ttest ahe_cpi if year==1992, by(bachelor) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	**4650**	**16.31645**	**.1105541**	**7.538785**	**16.09971**	**16.53319**
1	**2962**	**23.2408**	**.1826822**	**9.942341**	**22.8826**	**23.59899**
combined	**7612**	**19.01087**	**.1054049**	**9.19624**	**18.80425**	**19.21749**
diff		**-6.924345**	**.2135298**		**-7.342955**	**-6.505734**

```
    diff = mean( 0) - mean( 1)                                    t =   -32.4280
Ho: diff = 0                     Satterthwaite's degrees of freedom =   5091.99

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000
```

Empirical exercise E3.1: question f) i

```
. ttest ahe_cpi if bachelor==0, by(year) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1992	**4650**	**16.31645**	**.1105541**	**7.538785**	**16.09971**	**16.53319**
2012	**3485**	**15.68547**	**.1397007**	**8.247077**	**15.41157**	**15.95938**
combined	**8135**	**16.04614**	**.0870982**	**7.855756**	**15.87541**	**16.21688**
diff		**.6309787**	**.178153**		**.2817458**	**.9802115**

```
    diff = mean( 1992) - mean( 2012)                              t =     3.5418
Ho: diff = 0                     Satterthwaite's degrees of freedom =  7121.15

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.9998         Pr(|T| > |t|) =  0.0004          Pr(T > t) =  0.0002
```

- Wages of high school graduates fell by an estimated 0.63 dollars per hour from 1992 to 2012 (with a 95% confidence interval of –0.98 to –0.28)

Empirical exercise E3.1: question f) ii

```
. ttest ahe_cpi if bachelor==1, by(year) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
1992	**2962**	**23.2408**	**.1826822**	**9.942341**	**22.8826**	**23.59899**
2012	**3955**	**23.42605**	**.1790882**	**11.26264**	**23.07494**	**23.77717**
combined	**6917**	**23.34672**	**.1288569**	**10.71684**	**23.09412**	**23.59932**
diff		**-.185257**	**.2558229**		**-.6867508**	**.3162368**

```
    diff = mean( 1992) - mean( 2012)                              t =    -0.7242
Ho: diff = 0                     Satterthwaite's degrees of freedom =   6731.34

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.2345         Pr(|T| > |t|) =  0.4690          Pr(T > t) =  0.7655
```

- Wages of college graduates increased by an estimated 0.19 dollars per hour from 1992 to 2012 (with a 95% confidence interval of –0.32 to 0.69).

Empirical exercise E3.1: question f) iii

```
> ttest ahe_cpi if year==1992, by(bachelor) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	**4650**	**16.31645**	**.1105541**	**7.538785**	**16.09971**	**16.53319**
1	**2962**	**23.2408**	**.1826822**	**9.942341**	**22.8826**	**23.59899**
combined	**7612**	**19.01087**	**.1054049**	**9.19624**	**18.80425**	**19.21749**
diff		**-6.924345**	**.2135298**		**-7.342955**	**-6.505734**

```
    diff = mean( 0) - mean( 1)                                    t =   -32.4280
Ho: diff = 0                     Satterthwaite's degrees of freedom =  5091.99

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000

. ttest ahe_cpi if year==2012, by(bachelor) unequal unpaired

Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	**3485**	**15.68547**	**.1397007**	**8.247077**	**15.41157**	**15.95938**
1	**3955**	**23.42605**	**.1790882**	**11.26264**	**23.07494**	**23.77717**
combined	**7440**	**19.80026**	**.1238916**	**10.68632**	**19.5574**	**20.04312**
diff		**-7.74058**	**.2271318**		**-8.185825**	**-7.295335**

```
    diff = mean( 0) - mean( 1)                                    t =   -34.0797
Ho: diff = 0                     Satterthwaite's degrees of freedom =  7203.16

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  0.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  1.0000
```

Empirical exercise E3.1: question g)

```
1 . keep if bachelor==0
  (6917 observations deleted)

2 . ttest ahe_cpi if year == 1992, by(female) unequal unpaired

  Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	2774	17.63061	.1514204	7.975126	17.3337	17.92751
1	1876	14.37324	.1469648	6.365458	14.08501	14.66147
combined	4650	16.31645	.1105541	7.538785	16.09971	16.53319
diff		3.257365	.2110137		2.843675	3.671055

```
    diff = mean( 0) - mean( 1)                                    t =    15.4367
Ho: diff = 0                     Satterthwaite's degrees of freedom =    4522.65

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  1.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  0.0000

3 . ttest ahe_cpi if year == 2012, by(female) unequal unpaired

  Two-sample t test with unequal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	2279	17.04357	.1864512	8.900968	16.67794	17.40921
1	1206	13.11905	.1746657	6.065704	12.77637	13.46173
combined	3485	15.68547	.1397007	8.247077	15.41157	15.95938
diff		3.924525	.2554841		3.4236	4.42545

```
    diff = mean( 0) - mean( 1)                                    t =    15.3611
Ho: diff = 0                     Satterthwaite's degrees of freedom =    3269.91

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) =  1.0000         Pr(|T| > |t|) =  0.0000          Pr(T > t) =  0.0000
```

Empirical exercise E3.1: question g)

Gender Gap in Earnings for High School Graduates

Year	$\bar{Y}_m$	s_m	n_m	$\bar{Y}_w$	s_w	n_w	$\bar{Y}_m - \bar{Y}_w$	$SE(\bar{Y}_m - \bar{Y}_w)$	95% CI
1992	17.63	7.98	2774	14.37	6.37	1876	3.26	0.21	2.84 – 3.67
2012	17.04	8.90	2279	13.12	6.07	1206	3.92	0.26	3.42 – 4.43

- There is a large and statistically significant gender gap in earnings for high school graduates.
- In 2012 the estimated gap was $3.92 per hour; in 1992 the estimated gap was $3.26 per hour (in $2012).
- The estimated gender gap in 2012 is somewhat larger than is the gender gap for college gap for college graduates (which is $3.70 in Table 3.1 in the text).
- Moreover the estimated increase in the gender gap from 1992 to 2012 is also somewhat larger for high school graduates than it was for college graduates ($0.66 for high school graduates versus $0.36 for college graduates).