```
                                                                  ___  ____  ____  ____  ____(R)
                                                                 /__    /   ____/   /   ____/
                                                                ___/   /   /___/   /   /___/
                                                                  Statistics/Data analysis

                                                                     User: ali hosseinifar
                                                                  Project: PG5

  ___  ____  ____  ____  ____ ®
 /__    /   ____/   /   ____/      17.0
___/   /   /___/   /   /___/       SE—Standard Edition

 Statistics and Data Science       Copyright 1985-2021 StataCorp LLC
                                   StataCorp
                                   4905 Lakeway Drive
                                   College Station, Texas 77845 USA
                                   800-STATA-PC        https://www.stata.com
                                   979-696-4600        stata@stata.com

Stata license: Unlimited-user network, expiring 22 Dec 2025
Serial number: 401809302839
  Licensed to: seyed ali hosseini far
               University of Genova

Notes:
      1. Unicode is supported; see help unicode advice.
      2. Maximum number of variables is set to 5,000 but can be increased;
          see help set maxvar.

 1 . * Load and clean the dataset

 2 .
 3 . use "PG5.dta", clear

 4 . * Summary of missing data

 5 .
 6 . misstable summarize
                                                                  Obs<.
                                                  +------------------------------------------------------------
                  |                               | Unique
        Variable  |     Obs=.     Obs>.     Obs<. | values        Min         Max
        ----------+-----------------------------------------------------------------
             Wage |          1                    |  9,564       >500        -380    9.632404
         Distance |          1                    |  9,564       >500   .0003438    7.458824
        ----------------------------------------------------------------------------

 7 . * Drop missing values for Wage and Distance

 8 .
 9 . drop if missing(Wage) | missing(Distance)
   (1 observation deleted)

10 . * Drop negative values for Wage

11 .
12 . drop if Wage < 0
   (30 observations deleted)
```

```
13 . * Label categorical variables

14 .
15 . label define rel_label 0 "Catholic" 1 "Other" 2 "Atheist"

16 .
17 . label values Relig rel_label

18 . label variable Wage "Logarithmic Wage"

19 .
20 . label variable Distance "Distance to Counseling (miles)"

21 . * Check for outliers in Wage

22 .
23 . graph box Wage, title("Boxplot of Logarithmic Wage")

24 . * 2) Balancing Tests

25 .
26 . * Categorical variables

27 .
28 . foreach var in Female Degree Fixed Relig {
  2.
29 .     tabulate `var' P, chi2
  3.
30 . }
```

	P		
Female	0	1	Total
Male	**3,922**	**1,229**	**5,151**
Female	**2,958**	**1,425**	**4,383**
Total	**6,880**	**2,654**	**9,534**

Pearson chi2(**1**) = **88.2542** Pr = **0.000**

	P		
Degree	0	1	Total
0	**4,508**	**1,591**	**6,099**
1	**2,372**	**1,063**	**3,435**
Total	**6,880**	**2,654**	**9,534**

Pearson chi2(**1**) = **25.8360** Pr = **0.000**

	P		
Fixed	0	1	Total
0	**1,477**	**1,490**	**2,967**
1	**5,403**	**1,164**	**6,567**
Total	**6,880**	**2,654**	**9,534**

Pearson chi2(**1**) = **1.1e+03** Pr = **0.000**

Relig	P 0	1	Total
Catholic	**3,822**	**1,478**	**5,300**
Other	**2,329**	**884**	**3,213**
Atheist	**729**	**292**	**1,021**
Total	**6,880**	**2,654**	**9,534**

Pearson chi2(**2**) = **0.4696** Pr = **0.791**

```
31 . * Continuous variables

32 .
33 . ttest Wage, by(P)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	**6,880**	**4.802125**	**.0173006**	**1.435015**	**4.76821**	**4.83604**
1	**2,654**	**4.729306**	**.0277597**	**1.430094**	**4.674873**	**4.783739**
Combined	**9,534**	**4.781854**	**.0146857**	**1.433944**	**4.753067**	**4.810641**
diff		**.0728189**	**.0327593**		**.0086036**	**.1370342**

diff = mean(**0**) - mean(**1**) t = **2.2228**
H0: diff = 0 Degrees of freedom = **9532**

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = **0.9869** Pr(|T| > |t|) = **0.0262** Pr(T > t) = **0.0131**

```
34 .
35 . ttest Distance, by(P)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	**6,880**	**1.982692**	**.0152959**	**1.268731**	**1.952707**	**2.012677**
1	**2,654**	**.7676693**	**.0153322**	**.7898684**	**.7376051**	**.7977336**
Combined	**9,534**	**1.644464**	**.0130825**	**1.277402**	**1.618819**	**1.670108**
diff		**1.215023**	**.0264048**		**1.163264**	**1.266782**

diff = mean(**0**) - mean(**1**) t = **46.0151**
H0: diff = 0 Degrees of freedom = **9532**

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

```
36 .
37 . ttest Sleep, by(P)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]
0	**6,880**	**8.937135**	**.0220986**	**1.832988**	**8.893815**	**8.980455**
1	**2,654**	**8.896251**	**.0374485**	**1.929235**	**8.82282**	**8.969682**
Combined	**9,534**	**8.925754**	**.0190519**	**1.860269**	**8.888408**	**8.9631**
diff		**.0408842**	**.042508**		**-.0424405**	**.1242089**

```
    diff = mean(0) - mean(1)                                      t =   0.9618
H0: diff = 0                                     Degrees of freedom =     9532

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.8319         Pr(|T| > |t|) = 0.3362          Pr(T > t) = 0.1681

38 . * 3) Non-Parametric Approach

39 .
40 . * Propensity score matching

41 .
42 . * Generate treatment indicator

43 .
44 . gen treated = (P == 1)

45 . * Perform Propensity Score Matching

46 .
47 . teffects psmatch (Wage) (treated Degree Exp Female Distance Fixed Sleep Relig), atet

Treatment-effects estimation                   Number of obs      =      9,534
Estimator      : propensity-score matching     Matches: requested =          1
Outcome model  : matching                                  min =          1
Treatment model: logit                                     max =          1
```

Wage	Coefficient	AI robust std. err.	z	P>|z|	[95% conf.	interval]
ATET						
treated (1 vs 0)	**-.5218355**	**.0555129**	**-9.40**	**0.000**	**-.6306388**	**-.4130321**

```
48 . * Check matching balance
```

```
49 .
50 . tebalance summarize
   (refitting the model using the generate() option)

   Covariate balance summary

                           Raw      Matched
   ------------------------------------------
   Number of obs =       9,534        5,308
   Treated obs   =       2,654        2,654
   Control obs   =       6,880        2,654
   ------------------------------------------

   ------------------------------------------------------------------------
                    | Standardized differences          Variance ratio
                    |        Raw     Matched           Raw    Matched
   -----------------+------------------------------------------------------
             Degree |    .1155006    .0061543      1.063115   1.002542
                Exp |   -.1888109   -.0695175      1.361515   1.487623
             Female |    .2152935   -.0083132      1.014697   1.001302
           Distance |    -1.14974   -.0148345      .3875883   .9880614
              Fixed |   -.7612424     .035009      1.460829   1.009965
              Sleep |    -.021727   -.0224942      1.107774   1.141942
              Relig |    .0039534    .0099732      1.017319   1.021593
   ------------------------------------------------------------------------

51 . * Summarize matched data to check balance

52 .
53 . summarize Wage if treated == 1

    Variable |        Obs        Mean    Std. dev.       Min        Max
   ----------+---------------------------------------------------------
        Wage |      2,654    4.729306    1.430094    .127184   9.551585

54 .
55 . summarize Wage if treated == 0

    Variable |        Obs        Mean    Std. dev.       Min        Max
   ----------+---------------------------------------------------------
        Wage |      6,880    4.802125    1.435015   .0394029   9.632404

56 . * 4) Parametric Approach: Classical Wage Equation

57 .
58 . regress Wage P Degree Exp Female Distance Fixed Sleep Relig

      Source |       SS           df       MS      Number of obs   =     9,534
   ----------+----------------------------------   F(8, 9525)      =    711.07
       Model |  7329.36223         8  916.170279   Prob > F        =    0.0000
    Residual |  12272.3362     9,525  1.28843425   R-squared       =    0.3739
   ----------+----------------------------------   Adj R-squared   =    0.3734
       Total |  19601.6985     9,533  2.05619411   Root MSE        =    1.1351
```

Wage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
P	-.3962683	.0313069	-12.66	0.000	-.4576364	-.3349002
Degree	.1810574	.0242617	7.46	0.000	.1334993	.2286155
Exp	-.2750386	.003763	-73.09	0.000	-.2824149	-.2676623
Female	-.0930375	.0234663	-3.96	0.000	-.1390364	-.0470387
Distance	-.081935	.0102364	-8.00	0.000	-.1020005	-.0618695
Fixed	-.1758712	.0271404	-6.48	0.000	-.2290721	-.1226702
Sleep	.0884676	.0062528	14.15	0.000	.0762107	.1007245
Relig	.0053574	.017116	0.31	0.754	-.0281936	.0389083
_cons	7.081344	.078664	90.02	0.000	6.927146	7.235543

```
59 . * 5) Instrumental Variables Approach

60 .
61 . * First stage: Distance as an instrument for P

62 .
63 . regress P Distance Degree Exp Female Fixed Sleep Relig
```

Source	SS	df	MS		
Model	600.631262	7	85.804466	Prob > F	0.0000
Residual	1314.56907	9,526	.137998013	R-squared	0.3136
Total	1915.20034	9,533	.200902165	Root MSE	.37148

	Number of obs = 9,534
	F(7, 9526) = 621.78
	Adj R-squared = 0.3131

P	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Distance	-.1492828	.0029805	-50.09	0.000	-.1551252	-.1434404
Degree	.0436872	.0079275	5.51	0.000	.0281477	.0592268
Exp	-.0118004	.0012256	-9.63	0.000	-.0142028	-.009398
Female	.0792942	.0076367	10.38	0.000	.0643247	.0942638
Fixed	-.3287338	.0082188	-40.00	0.000	-.3448445	-.3126231
Sleep	-.0043244	.0020459	-2.11	0.035	-.0083347	-.000314
Relig	.0046409	.0056013	0.83	0.407	-.0063389	.0156207
_cons	.8520571	.024219	35.18	0.000	.8045828	.8995314

```
64 . * Second stage: 2SLS Regression

65 .
66 . ivregress 2sls Wage (P = Distance) Degree Exp Female Fixed Sleep Relig
```

Instrumental variables 2SLS regression — Number of obs = 9,534; Wald chi2(7) = 5360.61; Prob > chi2 = 0.0000; R-squared = 0.3537; Root MSE = 1.1527

Wage	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
P	.1525893	.0619535	2.46	0.014	.0311627	.2740159
Degree	.1570793	.0247934	6.34	0.000	.1084851	.2056736
Exp	-.2685619	.0038829	-69.17	0.000	-.2761722	-.2609516
Female	-.1365588	.0243133	-5.62	0.000	-.1842119	-.0889056
Fixed	.0045568	.0325057	0.14	0.889	-.0591532	.0682669
Sleep	.0908411	.0063499	14.31	0.000	.0783955	.1032866
Relig	.0028102	.0173834	0.16	0.872	-.0312606	.0368809
_cons	6.613686	.0820671	80.59	0.000	6.452838	6.774535

Instrumented: **P**
 Instruments: **Degree Exp Female Fixed Sleep Relig Distance**

```
67 . * Robust standard errors

68 .
69 . ivregress 2sls Wage (P = Distance) Degree Exp Female Fixed Sleep Relig, robust
```

Instrumental variables 2SLS regression

Number of obs	=	**9,534**
Wald chi2(**7**)	=	**326.05**
Prob > chi2	=	**0.0000**
R-squared	=	**0.3537**
Root MSE	=	**1.1527**

Wage	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
P	.1525893	.0621718	2.45	0.014	.0307348	.2744438
Degree	.1570793	.0245257	6.40	0.000	.1090098	.2051488
Exp	-.2685619	.0232665	-11.54	0.000	-.3141634	-.2229604
Female	-.1365588	.0245477	-5.56	0.000	-.1846714	-.0884462
Fixed	.0045568	.0315038	0.14	0.885	-.0571895	.0663031
Sleep	.0908411	.0068589	13.24	0.000	.0773979	.1042843
Relig	.0028102	.0171423	0.16	0.870	-.0307882	.0364085
_cons	6.613686	.2383718	27.75	0.000	6.146486	7.080886

Instrumented: **P**
 Instruments: **Degree Exp Female Fixed Sleep Relig Distance**

```
70 . * Bootstrap for robustness

71 .
72 . bootstrap _b, reps(1000): ivregress 2sls Wage (P = Distance) Degree Exp Female Fixed Sleep Relig
(running ivregress on estimation sample)

Bootstrap replications (1,000)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................   50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................................................   450
..................................................   500
..................................................   550
..................................................   600
..................................................   650
```

```
.................................................   700
.................................................   750
.................................................   800
.................................................   850
.................................................   900
.................................................   950
................................................. 1,000
```

Instrumental variables 2SLS regression — Number of obs = **9,534**; Wald chi2(**7**) = **369.96**; Prob > chi2 = **0.0000**; R-squared = **0.3537**; Root MSE = **1.1527**

Wage	Observed coefficient	Bootstrap std. err.	z	P>|z|	Normal-based [95% conf.	interval]
P	**.1525893**	**.0629318**	**2.42**	**0.015**	**.0292452**	**.2759334**
Degree	**.1570793**	**.023809**	**6.60**	**0.000**	**.1104145**	**.2037442**
Exp	**-.2685619**	**.0215532**	**-12.46**	**0.000**	**-.3108053**	**-.2263185**
Female	**-.1365588**	**.0239964**	**-5.69**	**0.000**	**-.1835909**	**-.0895266**
Fixed	**.0045568**	**.0314737**	**0.14**	**0.885**	**-.0571305**	**.0662442**
Sleep	**.0908411**	**.0067295**	**13.50**	**0.000**	**.0776516**	**.1040305**
Relig	**.0028102**	**.0164666**	**0.17**	**0.864**	**-.0294638**	**.0350841**
_cons	**6.613686**	**.2219736**	**29.79**	**0.000**	**6.178626**	**7.048747**

Instrumented: **P**
Instruments: **Degree Exp Female Fixed Sleep Relig Distance**

```
73 . * Test strength of the instrument

74 .
75 . estat firststage
```

First-stage regression summary statistics

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	Bootstrap F(1,9526)	Prob > F
P	**0.3136**	**0.3131**	**0.2085**	**2508.65**	**0.0000**

```
76 . * 6) Robustness Checks

77 .
78 . * Alternative model specifications

79 .
80 . regress Wage P Degree Exp Female Fixed Sleep Relig if Distance < 3
```

Source	SS	df	MS		
Model	**6653.42952**	**7**	**950.489932**	Number of obs	= **8,017**
Residual	**9662.44666**	**8,009**	**1.20644858**	F(7, 8009)	= **787.84**
				Prob > F	= **0.0000**
				R-squared	= **0.4078**
				Adj R-squared	= **0.4073**
Total	**16315.8762**	**8,016**	**2.0354137**	Root MSE	= **1.0984**

Wage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
P	-.3647111	.028633	-12.74	0.000	-.4208392	-.3085829
Degree	.1936748	.0256054	7.56	0.000	.1434815	.2438682
Exp	-.3019745	.0041892	-72.08	0.000	-.3101864	-.2937627
Female	-.0977873	.0247584	-3.95	0.000	-.1463203	-.0492544
Fixed	-.2154289	.0286574	-7.52	0.000	-.2716048	-.159253
Sleep	.0937405	.0067222	13.94	0.000	.0805633	.1069177
Relig	.0070061	.0180065	0.39	0.697	-.0282913	.0423036
_cons	7.212169	.0811621	88.86	0.000	7.05307	7.371268

```
81 . * Save results for presentation

82 .
83 . eststo clear

84 .
85 . eststo: regress Wage P Degree Exp Female Distance Fixed Sleep Relig
```

Source	SS	df	MS		
				Number of obs	= 9,534
				F(8, 9525)	= 711.07
Model	7329.36223	8	916.170279	Prob > F	= 0.0000
Residual	12272.3362	9,525	1.28843425	R-squared	= 0.3739
				Adj R-squared	= 0.3734
Total	19601.6985	9,533	2.05619411	Root MSE	= 1.1351

Wage	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
P	-.3962683	.0313069	-12.66	0.000	-.4576364	-.3349002
Degree	.1810574	.0242617	7.46	0.000	.1334993	.2286155
Exp	-.2750386	.003763	-73.09	0.000	-.2824149	-.2676623
Female	-.0930375	.0234663	-3.96	0.000	-.1390364	-.0470387
Distance	-.081935	.0102364	-8.00	0.000	-.1020005	-.0618695
Fixed	-.1758712	.0271404	-6.48	0.000	-.2290721	-.1226702
Sleep	.0884676	.0062528	14.15	0.000	.0762107	.1007245
Relig	.0053574	.017116	0.31	0.754	-.0281936	.0389083
_cons	7.081344	.078664	90.02	0.000	6.927146	7.235543

```
(est1 stored)

86 .
87 . esttab using "results.rtf", replace title("Regression Results")
(file results.rtf not found)
(output written to results.rtf)

88 .
```