



Übung zur Empirischen Wirtschaftsforschung

Übungsblatt 3

Please examine the following estimations for the log earnings of workers using the SOEP 2012 data and answer the questions below.

Y Bruttomonatseinkommen in €
F Geschlecht, Dummyvariable, 1 für Frauen, 0 für Männer
WEST Westdeutschland, Dummyvariable 1 für West, 0 für Ost

HAUPTORLESS Hauptschulabschluss oder weniger
REAL Realschulabschluss (Referenzgruppe)
ABI (Fach-)Hochschulreife

OHNEAUSBILDUNG ohne Ausbildungsabschluss
BBILDUNG abgeschlossener Berufsausbildung (Referenzgruppe)
UNI (Fach-)Hochschulabschluss

XYR Berufserfahrung (in Jahren)
STUND Tatsächliche Arbeitszeit pro Woche (in Stunden)

Estimation 1

```

=====
Dependent Variable: LOG(Y)
Method: Least Squares
Date: 07/10/17   Time: 15:26
Sample (adjusted): 1 42741
Included observations: 5063 after adjustments
=====

```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.347940	0.126086	34.48402	0.0000
HAUPTORLESS	-0.033728	0.022919	-1.471599	0.1412
ABI	0.190405	0.021236	8.966173	0.0000
KEINAUSBILDUNG	-0.294515	0.028615	-10.29230	0.0000
UNI	0.329552	0.026778	12.30684	0.0000
F	-0.193176	0.018543	-10.41802	0.0000
XYR	0.057108	0.006793	8.407269	0.0000
XYR^2	-0.000515	5.90E-05	-8.726973	0.0000
LOG(STUND)	0.536384	0.079392	6.756116	0.0000

```

=====
R-squared                0.584770    Mean dependent var 7.552216
Adjusted R-squared      0.584113    S.D. dependent var 0.923314
S.E. of regression      0.595439    Akaike info criteri1.802740
Sum squared resid      1791.882    Schwarz criterion  1.814347
Log likelihood          -4554.636    Hannan-Quinn criter1.806805
F-statistic             889.6952    Durbin-Watson stat 1.505352
Prob(F-statistic)      0.000000
=====

```

Interpret the statistical and economic significance of the educational dummy variables.

How is the $\log(\text{STUND})$ variable interpreted? Comment on the statistical significance of the variable .

How are the experience coefficients XYR coefficients interpreted? Why did their influence change relative to the full model estimated in the Einkommensfunktion Übung?

How is the F coefficient interpreted? Please calculate the total earnings of a female with ABI and UNI and compare it to the total income of a corresponding male.

Please interpret the constant term C and calculate the earnings of this group.

Analyze and comment on the estimation quality of the model.

Estimation 2

```

=====
Dependent Variable: LOG(Y)
Method: Least Squares
Date: 07/10/17   Time: 14:32
Sample: 1 42775 IF WEST=1
Included observations: 2397
=====

```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.331193	0.063596	83.82954	0.0000
HAUPTORLESS	-0.123882	0.032522	-3.809229	0.0001
ABI	0.098142	0.029665	3.308295	0.0010
KEINAUSBILDUNG	-0.340302	0.036878	-9.227794	0.0000
UNI	0.302813	0.036347	8.331228	0.0000
F	-0.236466	0.025966	-9.106775	0.0000
XYR	0.098098	0.003098	31.66119	0.0000
XYR^2	-0.000814	4.42E-05	-18.43805	0.0000

```

=====
R-squared          0.634104      Mean dependent var 7.577798
Adjusted R-squared 0.633032      S.D. dependent var 0.936154
S.E. of regression 0.567103      Akaike info criteri1.706780
Sum squared resid  768.3159      Schwarz criterion  1.726077
Log likelihood      -2037.576      Hannan-Quinn criter1.713801
F-statistic         591.4532      Durbin-Watson stat 1.747472
Prob(F-statistic)  0.000000
=====

```

```

=====
Dependent Variable: LOG(Y)
Method: Least Squares
Date: 07/10/17   Time: 14:32
Sample: 1 42775 IF WEST=0
Included observations: 2458
=====

```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.960014	0.070353	70.50160	0.0000
HAUPTORLESS	0.020673	0.033037	0.625771	0.5315
ABI	0.219894	0.032192	6.830677	0.0000
KEINAUSBILDUNG	-0.246725	0.045864	-5.379483	0.0000
UNI	0.384663	0.041403	9.290659	0.0000
F	-0.135146	0.027002	-5.005052	0.0000
XYR	0.103650	0.003430	30.21491	0.0000
XYR^2	-0.000888	4.75E-05	-18.69124	0.0000

```

=====
R-squared          0.554738      Mean dependent var 7.519289
Adjusted R-squared 0.553466      S.D. dependent var 0.916533
S.E. of regression 0.612456      Akaike info criteri1.860572
Sum squared resid  919.0021      Schwarz criterion  1.879472
Log likelihood      -2278.642      Hannan-Quinn criter1.867439
F-statistic         436.0544      Durbin-Watson stat 1.302703
=====

```

What is the difference between the first and second estimation in terms of groups?

Analyze the effect of education and gender on the income for both groups in terms of statistical and economic significance.

What are possible socioeconomic factors that could be behind the differences found between both groups?

Please comment on the estimation quality for both models based on R^2 and S.E. of regression.