



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

OHNE kein Schulabschluss (Referenzgruppe)

HAUPT Hauptschulabschluss

REAL Realschulabschluss

FACHSCHU Fachhochschulreife

ABI Abitur

KEINBERUF kein Berufsabschluss (Referenzgruppe)

LEHRE Lehre

MEISTER Meister

UNI Universitätsabschluss

XYR Berufserfahrung (in Years)

STUND Tatsächliche Arbeitszeit pro Woche (in Hours)

PSAMPLE Stichprobenart (1: Deutsche (West), 2: Ausländer (West),
3: Deutsche (Ost), 4: Zuwanderer 1984-93, 5: Ergänzung 1998, 6: Ergänzung
2000, 7: Hocheinkommensbezieher 2002: 8: H Ergänzung 2006, 10: J Auf-
stockung 2011, 11: K Aufstockung 2012)

Estimation 1

```

=====
Dependent Variable: LOG(Y)
Sample: 1 23353 IF PSAMPLE<=3
Included observations: 1508
=====

```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.110953	0.173358	41.01899	0.0000
X _{YR}	0.017095	0.001652	10.34977	0.0000
HAUPT	0.227338	0.175778	1.293322	0.1961
REAL	0.307454	0.174851	1.758374	0.0789
FACHSCHU	0.663023	0.185904	3.566488	0.0004
ABI	0.693863	0.178729	3.882217	0.0001
LEHRE	-0.011533	0.047509	-0.242761	0.8082
MEISTER	0.155557	0.074954	2.075354	0.0381
UNI	0.195439	0.057047	3.425926	0.0006
F	-0.455678	0.038058	-11.97318	0.0000

```

=====
R-squared          0.288393      Mean dependent var 7.671067
Adjusted R-squared 0.284118      S.D. dependent var 0.804808
S.E. of regression 0.680947      Akaike info criteri2.075943
Sum squared resid  694.6050      Schwarz criterion  2.111212
Log likelihood     -1555.261      Hannan-Quinn criter2.089079
F-statistic        67.45501      Durbin-Watson stat 1.852863
Prob(F-statistic) 0.000000
=====

```

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

How is the experience coefficient X_{YR} interpreted?

How is the F coefficient interpreted? Please calculate the total earnings for a female ABI graduate and a male ABI graduate.

Based on the education dummies, what conclusion can you give about the influence of education level on earnings?

What additional variable should be included in the model based on economic theory? What is the expected sign of this variable?

Estimation 2

=====
 Dependent Variable: LOG(Y)
 Sample: 1 23353 IF PSAMPLE=1
 Included observations: 859
 =====

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.273559	0.252263	12.97675	0.0000
LOG(STUND)	1.064255	0.039769	26.76081	0.0000
HAUPT	0.342383	0.219178	1.562122	0.1186
REAL	0.403988	0.218657	1.847589	0.0650
FACHSCHU	0.591498	0.226731	2.608806	0.0092
ABI	0.522211	0.222355	2.348547	0.0191
LEHRE	0.304828	0.046319	6.581118	0.0000
MEISTER	0.526034	0.079115	6.648938	0.0000
UNI	0.575317	0.064700	8.892137	0.0000
F	-0.259492	0.041593	-6.238908	0.0000

=====
 R-squared 0.636947 Mean dependent var 7.703801
 Adjusted R-squared 0.633098 S.D. dependent var 0.867824
 S.E. of regression 0.525662 Akaike info criteri1.563256
 Sum squared resid 234.5959 Schwarz criterion 1.618620
 Log likelihood -661.4183 F-statistic 165.5002
 Durbin-Watson stat 2.041404 Prob(F-statistic) 0.000000
 =====

=====
 Dependent Variable: LOG(Y)
 Sample: 1 23353 IF PSAMPLE=3
 Included observations: 542
 =====

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.471825	0.567586	4.354983	0.0000
LOG(STUND)	1.155669	0.053967	21.41428	0.0000
HAUPT	0.391928	0.543642	0.720931	0.4713
REAL	0.573500	0.537567	1.066844	0.2865
FACHSCHU	0.590584	0.548857	1.076025	0.2824
ABI	0.801848	0.540371	1.483884	0.1384
LEHRE	0.173651	0.064389	2.696896	0.0072
MEISTER	0.305568	0.095057	3.214561	0.0014
UNI	0.409436	0.060586	6.757879	0.0000
F	-0.116087	0.047954	-2.420799	0.0158

=====
 R-squared 0.536560 Mean dependent var 7.508584
 Adjusted R-squared 0.528720 S.D. dependent var 0.779645
 S.E. of regression 0.535225 Akaike info criteri1.606018
 Sum squared resid 152.3997 Schwarz criterion 1.685266
 Log likelihood -425.2308 F-statistic 68.43752
 Durbin-Watson stat 1.465260 Prob(F-statistic) 0.000000
 =====

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

How is the $\log(STUND)$ coefficient interpreted? Please interpret it for both groups.

Analyze the effect of the education dummies for each estimation group and then compare the results between both models. What does the difference mean economically?

Analyze the influence of gender on income for both estimations. What does the difference between both estimations show?

Please comment on the estimation quality for both models. Based on the R^2 results for Estimation 1 and Estimation 2, what conclusion can you give about the coefficients importance?