

Name: _____

Vorname: _____

Matrikel-Nr.: _____

Klausur: 32731 Angewandte Ökonometrie

Termin: 24.09.2019, 14:00–16:00

Prüfer: Prof. Dr. Hans-Jörg Schmerer

Note: _____

Datum: _____

Unterschrift des Prüfers: _____

© 2019

FernUniversität in Hagen

Fakultät für Wirtschaftswissenschaft

Alle Rechte vorbehalten

Hinweise zur Klausur – Bitte unbedingt beachten!

Die ausgeteilten Klausurunterlagen bestehen aus insgesamt 23 Seiten. Bitte kontrollieren Sie sofort, ob Sie ein vollständiges Klausurexemplar erhalten haben und tragen Sie auf dem Deckblatt Ihren Namen und Ihre Matrikelnummer ein.

Die Verwendung eines Taschenrechners ist dann und nur dann erlaubt, wenn dieser einer der folgenden Modellreihen angehört:

Casio fx86 oder Casio fx87,

Texas Instruments TI 30 X II,

Sharp EL 531.

Die Verwendung anderer Taschenrechnermodelle wird als Täuschungsversuch gewertet und mit der Note „nicht ausreichend“ (5,0) sanktioniert. Ob ein Taschenrechner einer der Modellreihen angehört, können Sie selbst überprüfen, indem Sie die vom Hersteller auf dem Rechner angebrachte Modellbezeichnung mit den oben angegebenen Bezeichnungen vergleichen: Bei vollständiger Übereinstimmung ist das Modell erlaubt. Ist die auf dem Rechner angebrachte Modellbezeichnung umfangreicher, enthält aber eine der oben angegebenen Bezeichnungen vollständig, ist das Modell ebenfalls erlaubt. In allen anderen Fällen ist das Modell nicht erlaubt. Eventuelle Vorgänger- oder Nachfolgermodelle, die nicht in der oben aufgeführten Liste enthalten sind, sind ebenfalls nicht erlaubt.

Notieren Sie Ihre Lösungen für Aufgaben 1 bis 3 auf den Lösungsbögen. Sollten Sie zusätzlichen Platz benötigen, können Sie auch die Rückseite der Lösungsbögen verwenden.

Die Klausur besteht aus drei Aufgaben. Es sind alle Aufgaben zu beantworten. Die maximal erreichbare Punktzahl beträgt 100. Die Klausur ist bestanden, wenn Sie mindestens 50 Punkte erzielen.

Aufgabe 1

[40 Punkte]

Sie wollen mögliche Determinanten des Stundenlohns junger Arbeitskräfte empirisch analysieren. In dem zur Verfügung stehenden Datensatz befinden sich Querschnittsdaten zu 2.714 amerikanischen Arbeitskräften, die im Jahr 1979 zwischen 14 und 21 Jahre alt waren. Es handelt sich bei dem Datensatz um eine zufällige Stichprobe. Neben den Stundenlöhnen enthält der Datensatz eine Vielzahl an weiteren Informationen, beispielsweise zur Ausbildung, der ethnischen und religiösen Zugehörigkeit, dem Geschlecht, sowie weiterer Merkmale.

- a) Wie lauten die Annahmen für eine konsistente OLS Schätzung? Erläutern Sie die jeweiligen Annahmen kurz!

b) Sie sind daran interessiert, den Stundenlohn junger Arbeitskräfte zu erklären. Das Ausgangsmodell lautet:

$$\ln(\text{earnings}_i) = \alpha + \beta_1 \text{school}_i + \beta_2 \text{tenure}_i + \beta_3 \text{female}_i + \beta_4 \text{ethnic}_i + \varepsilon_i$$

mit

$\ln(\text{earnings}_i)$	=	logarithmierter Stundenlohn von Arbeitskraft i
school_i	=	Schulzeit in Jahren
tenure_i	=	Anstellung beim aktuellen Arbeitgeber (in Jahren)
female_i	=	Dummy (= 1, wenn das Individuum weiblich ist)
ethnic_i	=	ethnische Herkunft

Welche Vorzeichen erwarten Sie für die jeweiligen Effekte der aufgeführten Variablen? Erläutern Sie Ihre Erwartungen hinsichtlich des geschätzten Vorzeichens kurz!

- c) Sie schätzen das Modell durch einen OLS Schätzer mit robusten Standardfehlern und erhalten die folgende Ergebnistabelle:

```

Linear regression           Number of obs   =    2,714
                          F(5, 2708)       =    241.63
                          Prob > F         =    0.0000
                          R-squared        =    0.3337
                          Root MSE      =    .49337
  
```

llearn	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
school	.1064471	.0042417	25.10	0.000	.0981298	.1147643
tenure	.019678	.0014657	13.43	0.000	.016804	.022552
female	-.3033593	.0190095	-15.96	0.000	-.3406338	-.2660848
white	.1942955	.0273745	7.10	0.000	.1406185	.2479726
hispanic	.1412116	.0474086	2.98	0.003	.048251	.2341723
black	0	(omitted)				
_cons	1.162158	.0605476	19.19	0.000	1.043433	1.280882

Interpretieren Sie die Schätzergebnisse kurz!

- d) Sie befürchten, dass die Ergebnisse in c) gegebenenfalls durch das Fehlen der Einflussgröße „Fähigkeiten“ verzerrt sind („omitted variable bias“). Aus diesem Grund fügen Sie dem Modell die Variable **asvabc** hinzu. Diese Variable beinhaltet die Ergebnisse eines Tests, der die Fähigkeiten innerhalb verschiedener Disziplinen (bspw. Allgemeinwissen, Wortschatz, Mathematik, etc.) quantifiziert. Sie erhalten folgende Ergebnistabelle in STATA:

Linear regression		Number of obs	=	2,714		
		F(6, 2707)	=	223.05		
		Prob > F	=	0.0000		
		R-squared	=	0.3502		
		Root MSE	=	.4873		
lnearn	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
school	.0843233	.0049977	16.87	0.000	.0745235	.0941231
tenure	.018331	.0014506	12.64	0.000	.0154866	.0211754
female	-.2914203	.0188299	-15.48	0.000	-.3283428	-.2544978
white	.085538	.0293384	2.92	0.004	.0280101	.143066
hispanic	.0955855	.0468528	2.04	0.041	.0037146	.1874564
black	0	(omitted)				
asvabc	.0109461	.0012838	8.53	0.000	.0084287	.0134635
_cons	1.003085	.0614451	16.32	0.000	.8826012	1.123569

Welche Unterschiede zu den Ergebnissen der vorherigen Schätzung fallen Ihnen auf? Wie lassen sich die Unterschiede zwischen den beiden Schätzungen erklären?

- e) Sie möchten nun untersuchen, ob sowohl die Schulzeit als auch die Vertragslaufzeit bei einem Arbeitgeber gemeinsam einen signifikanten Einfluss auf den Stundenlohn haben. Stellen Sie die zu testende Nullhypothese auf! Welcher Test eignet sich für die Untersuchung? Begründen Sie Ihre Antwort kurz!

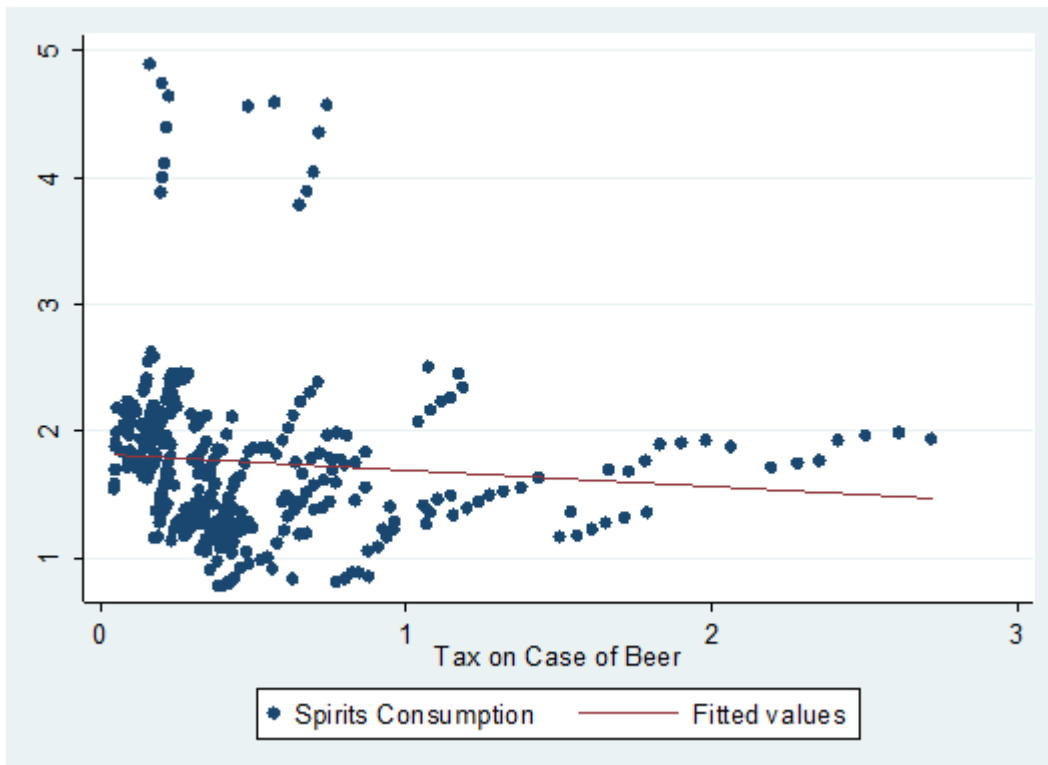
- f) Interpretieren Sie die Aussage, dass der OLS Schätzer „BLUE“ ist. Wofür steht dieses Akronym und welche Eigenschaften des Schätzers werden damit umschrieben?

Aufgabe 2

[40 Punkte]

Sie wollen analysieren, wie effektiv bestimmte Gesetze sind, um den Konsum von Alkohol einzuschränken. Für die Analyse steht Ihnen ein Paneldatensatz zur Verfügung. Dieser Datensatz beinhaltet Informationen über verschiedene Gesetze in amerikanischen Bundesstaaten für den Zeitraum 1982 bis 1989. Außerdem beinhaltet der Datensatz weitere Informationen zum Alkoholkonsum, der Arbeitslosenquote, dem pro Kopf-Einkommen, oder der Bevölkerungsgröße im jeweiligen Bundesstaat.

- a) Die nachfolgende Abbildung visualisiert die Korrelation zwischen Alkoholkonsum (Spirits Consumption) und der Höhe der Biersteuer (Tax on Case of Beer).



Entspricht die aufgezeigte Korrelation Ihren Erwartungen? Diskutieren Sie die Abbildung kritisch!

b) Der Zusammenhang zwischen der Biersteuer und dem Alkoholkonsum wird anhand der folgenden Gleichung geschätzt.

$$spircons_{it} = \alpha + \beta_1 beertax_{it} + \beta_2 mlda_{it} + \beta_3 yndrv_{it} + \varepsilon_{it}$$

mit

- spircons_{it}* = Alkoholkonsum
- beertax_{it}* = Reale Steuerarte auf Bier
- mlda_{it}* = Mindestalter für Alkoholkonsum
- yndrv_{it}* = Anteil junger Autofahre (15-24 Jahre alt)

Die folgende Tabelle zeigt die Ergebnisse einer pooled OLS Schätzung.

```

Linear regression                Number of obs   =       336
                                F(3, 47)        =       1.44
                                Prob > F            =       0.2438
                                R-squared          =       0.0201
                                Root MSE       =       .67971
  
```

(Std. Err. adjusted for 48 clusters in state)

<i>spircons</i>	Robust					[95% Conf. Interval]	
	Coef.	Std. Err.	t	P> t			
<i>beertax</i>	-.1127211	.127304	-0.89	0.380	-.3688236	.1433815	
<i>mlda</i>	-.0820245	.0634071	-1.29	0.202	-.2095831	.0455341	
<i>yngdrv</i>	-1.899975	2.199959	-0.86	0.392	-6.325723	2.525772	
<i>_cons</i>	3.842671	1.478234	2.60	0.012	.8688479	6.816493	

Interpretieren Sie die Ergebnisse! Eignet sich diese Methode zur Schätzung des Modells? Begründen Sie Ihre Antwort kurz!

c) Im nächsten Schritt wird das Modell mit Random Effects geschätzt. Die nachfolgende Tabelle veranschaulicht die Ergebnisse dieser Schätzung:

```

Random-effects GLS regression                Number of obs   =       336
Group variable: state                       Number of groups =        48

R-sq:                                       Obs per group:
  within = 0.5977                          min =           7
  between = 0.0493                         avg =          7.0
  overall = 0.0003                         max =           7

corr(u_i, X) = 0 (assumed)                 Wald chi2(3)    =      124.64
                                           Prob > chi2     =       0.0000

```

(Std. Err. adjusted for 48 clusters in state)

spircons	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
beertax	.0040284	.1048856	0.04	0.969	-.2015436	.2096005
mlda	-.0633559	.0175474	-3.61	0.000	-.0977481	-.0289637
yngdrv	5.011034	.9724763	5.15	0.000	3.105016	6.917053
_cons	2.115906	.4835179	4.38	0.000	1.168229	3.063584
sigma_u	.67001066					
sigma_e	.1015866					
rho	.97752816	(fraction of variance due to u_i)				

Welche Unterschiede zu den Ergebnissen der OLS Schätzung fallen Ihnen auf? Diskutieren Sie kurz die Validität des Schätzers!

d) Im nächsten Schritt wird das Modell mit Fixed Effects geschätzt. Die nachfolgende Tabelle fasst die Ergebnisse zusammen:

```

Fixed-effects (within) regression      Number of obs   =      336
Group variable: state                 Number of groups =      48

R-sq:                                 Obs per group:
  within = 0.7524                      min =          7
  between = 0.0025                     avg =         7.0
  overall = 0.0359                      max =          7

corr(u_i, Xb) = -0.0203                F(9,47)         =      30.76
                                           Prob > F         =      0.0000

```

(Std. Err. adjusted for 48 clusters in state)

spircons	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
beertax	-.1402529	.1592867	-0.88	0.383	-.4606965	.1801907
mla	-.0122508	.0156014	-0.79	0.436	-.0436368	.0191353
yngrdv	.7707772	.6880737	1.12	0.268	-.6134486	2.155003
year						
1983	-.0388253	.0087212	-4.45	0.000	-.0563701	-.0212805
1984	-.0749342	.014133	-5.30	0.000	-.1033663	-.0465022
1985	-.1273377	.0193694	-6.57	0.000	-.166304	-.0883714
1986	-.2412993	.0288518	-8.36	0.000	-.2993416	-.1832569
1987	-.274035	.0346852	-7.90	0.000	-.3438125	-.2042574
1988	-.3126208	.0404394	-7.73	0.000	-.3939743	-.2312673
_cons	2.085685	.3679397	5.67	0.000	1.345485	2.825884
sigma_u	.67333269					
sigma_e	.08057853					
rho	.98588101	(fraction of variance due to u_i)				

Welche Unterschiede fallen Ihnen zum Ergebnis zur Schätzung mit Random Effects auf? Sind die beiden Schätzungen miteinander vergleichbar?

e) Diskutieren Sie kurz die Vor- und Nachteile des Fixed Effects Schätzers!

- f) Inwiefern unterscheiden sich die Annahmen zwischen Fixed Effects und Random Effects Schätzern? Vergleichen Sie die Annahmen und erläutern Sie kurz!

Aufgabe 3**[20 Punkte]**

Sie möchten untersuchen, welchen Einfluss die Größe eines Unternehmens auf dessen Exportwahrscheinlichkeit hat. Für diese Untersuchung liegt Ihnen ein (imaginärer) Datensatz mit zwanzig Unternehmen vor. Die Unternehmen haben angegeben, ob Sie ins Ausland exportieren oder nicht und wie viele Mitarbeiter Sie beschäftigen. Außerdem sei der Gewinn der jeweiligen Unternehmung bekannt.

Exportstatus (ja=1)	Anzahl der Mitarbeiter	Gewinn
1	52	400.569
1	63	789.123
0	15	125.456
1	45	350.410
0	8	80.125
0	22	108.417
0	31	250.321
1	102	1.264.753
1	56	1.000.254
0	32	169.145
1	61	897.125
0	12	105.145
0	23	250.789
1	35	650.189
1	41	780.405
1	56	569.951
1	89	1.987.257
0	12	80.456
0	37	350.982
0	18	120.247

- a) Diskutieren Sie die vorliegenden Daten! Welche Besonderheiten in der Datenstruktur und möglichen Zusammenhängen zwischen der Exportvariable und den Variablen Anzahl an Mitarbeitern und Gewinn fallen Ihnen auf?

b) Sie wollen den in a) skizzierten Zusammenhang mit einem ökonometrischen Modell schätzen. Ist ein OLS Schätzer für die Analyse geeignet? Begründen Sie Ihre Antwort und nennen Sie gegebenenfalls alternative Ansätze!

c) Stellen Sie die Likelihood-Funktion für ein binär kodiertes Regressionsmodell (Probit/Logit) ohne Kontrollvariablen auf!

- d) Lösen Sie das Maximierungsproblem allgemein, um die Wahrscheinlichkeit P zu ermitteln. Zeigen Sie, dass die Wahrscheinlichkeit genau dem Durchschnitt der Variable Export entspricht.

- e) Wie kann der Einfluss der Variable Größe ("Anzahl der Mitarbeiter") im Modell berücksichtigt werden?

ZUSATZPAPIER

