**Klausur:** Einführung in die Ökonometrie (1232) Sommersemester 2008
**Prüfer:** Prof. Dr. Karl-Heinz Paqué, Dr. Bettina Büttner
**Als Hilfsmittel sind zugelassen:** Taschenrechner, Wörterbücher

Die Aufgabenstellung umfasst drei Aufgaben, die alle in der zur Verfügung stehenden Zeit von 120 Minuten zu bearbeiten sind. Sie können maximal 120 Punkte erreichen.

---

**Aufgabe 1 (insgesamt 36 Punkte):** Eine Untersuchung der Löhne von 40 Professoren an einer großen deutschen Universität aus dem Jahr 2004 basierte auf dem folgenden ökonometrischen Modell:

$$\ln L_t = \alpha + \beta_1 A_t + \beta_2 B_t + \beta_3 D_t + u_t \qquad (1)$$

wobei

$L_t$ = jährlicher Lohn des Professors $t$ (in €)
$A_t$ = Anzahl der von Professor $t$ publizierten Artikel
$B_t$ = Anzahl der von Professor $t$ geschriebenen Bücher
$D_t$ = Anzahl der von Professor $t$ betreuten Doktoranden.

Sie selbst vermuten, dass der wahre Zusammenhang besser durch das folgende Modell beschrieben wird:

$$\ln L_t = \alpha + \beta_1 A_t + \beta_4 Y_t + \beta_5 T_t + u_t' \qquad (2)$$

wobei

$Y_t$ = Anzahl der Berufsjahre von Professor $t$
$T_t$ = Anzahl der von Professor $t$ abgelehnten Rufe an andere Universitäten.

a) (*13 Punkte*) Untersuchen Sie anhand des ungenesteten F-Tests, welche der beiden Modellvarianten (1) und (2) die passendere Variante ist. Skizzieren Sie das Vorgehen dieses Testverfahrens und führen Sie das Verfahren durch! Ihnen stehen die folgenden Informationen zur Verfügung: Die KQ-Schätzung des Modells (1) lieferte die Summe der Residuenquadrate $S^I_{\hat{u}\hat{u}} = 8{,}045$, des Modells (2) $S^{II}_{\hat{u}\hat{u}} = 9{,}675$ und des „Mega-Modells" $S^M_{\hat{u}\hat{u}} = 7{,}667$. Welches Modell sollte man laut der Testergebnisse bei einem Signifikanzniveau von 5% vorziehen?

b) (*4 Punkte*) Welche Testergebnisse müssten in dem in Aufgabenteil a) durchgeführten Test auftreten, damit Sie zu der Ansicht kommen, dass man vielleicht doch besser das folgende Modell schätzen sollte:

$$\ln L_t = \alpha + \beta_1 A_t + \beta_2 B_t + \beta_3 D_t + \beta_4 Y_t + \beta_5 T_t + u_t'' \qquad (3)$$

c) (*8 Punkte*) Angenommen, Sie haben sich für Modell (1) entschieden. Nennen Sie die Konsequenzen für Punktschätzer, Intervallschätzer und Hypothesentests, wenn Modell (3) das korrekte Modell ist, aber stattdessen Modell (1) geschätzt wird. Nennen Sie auch die Konsequenzen für den *umgekehrten* Fall, also Modell (1) ist das korrekte Modell, aber stattdessen wird Modell (3) geschätzt.

d) (*4 Punkte*) Unterstellen Sie jetzt, dass Sie sich endgültig für Modell (1) entschieden haben. Neben Modell (1) wurde auch die folgende logarithmische funktionale Spezifikation geschätzt:

$$\ln L_t = \alpha' + \beta_1' \ln A_t + \beta_2' \ln B_t + \beta_3' \ln D_t + u_t''' \qquad (4)$$

Interpretieren Sie die Koeffizienten $\beta_1$ und $\beta_1'$ (in Bezug auf die Veränderungen in den nichtlogarithmierten Variablen)!

**-Bitte wenden!-**

e) (*5 Punkte*) Es ergaben sich die folgenden Bestimmtheitsmaße:

| Modell | $R^2$ |
|---|---|
| (1) | 90,2 |
| (4) | 88,9 |

Entscheiden Sie auf Basis dieser Informationen, ob es möglich ist, eine Auswahl zwischen Modell (1) und (4) zu treffen. Begründen Sie kurz Ihre Aussage.

f) (*2 Punkte*) Betrachten Sie die folgende funktionale Spezifikation:

$$L_t = \alpha'''' + \beta_1'''' A_t + \beta_2'''' B_t + \beta_3'''' D_t + u_t'''' \qquad (5)$$

Welches einfache Testverfahren würden Sie vorschlagen, um zwischen Modell (1) und Modell (5) die bessere funktionale Form zu ermitteln (keine Beschreibung) und weshalb ist es für diesen Vergleich geeignet?

**Aufgabe 2 (insgesamt 35 Punkte):** Ökonometriker schätzten für die USA basierend auf insgesamt 66 vierteljährlichen Beobachtungen (1961:1 bis 1977:2) mittels KQ–Schätzung die folgende Nachfragefunktion für Kaffee:

$$\widehat{\ln Q_t} = \underset{(3,58)}{1,2787} - \underset{(-2,14)}{0,1647} \ln P_t + \underset{(1,23)}{0,5115} \ln I_t + \underset{(0,55)}{0,1483} \ln W_t - \underset{(-3,36)}{0,0089} T_t - \underset{(-3,74)}{0,0961} D_t \qquad (6)$$

wobei in Klammern die jeweiligen *t*-Statistiken stehen und

$Q_t$ = Pro-Kopf-Konsum von Kaffee (in Pfund (500g)) in Quartal $t$

$P_t$ = Preis von Kaffee (pro Pfund) in Quartal $t$

$I_t$ = Pro-Kopf-Einkommen in Quartal $t$

$W_t$ = Preis von Tee in Quartal $t$

$T_t$ = Trendvariable, $T_1 = 1$ für die erste Beobachtung (1961:1), ..., $T_{66} = 66$ für die letzte Beobachtung (1977:2)

$D_t$ = Dummyvariable; $D_t = \begin{cases} 0 & \text{wenn t im Winterhalbjahr (Herbst und Winter)} \\ 1 & \text{wenn t im Sommerhalbjahr (Frühjahr und Sommer)} \end{cases}$

Hinweis: Der Tabellenwert für die t-Verteilung bei einem einseitigen Signifikanzniveau von 0,025 und 60 Beobachtungen ist 2,00; für 70 Beobachtungen liegt er bei 1,99.

a) (*2 Punkte*) Interpretieren Sie den Koeffizienten von $\ln P_t$!

b) (*4 Punkte*) Ist bei einem Signifikanzniveau von 5% die Kaffeenachfrage preis-unelastisch?

c) (*3 Punkte*) Sind Kaffee und Tee eher Substitute oder eher komplementäre Güter?

d) (*2 Punkte*) Interpretieren Sie den Koeffizienten von $T_t$!

e) (*6 Punkte*) Testen Sie auf einem Signifikanzniveau von 5%, ob der Koeffizient von $\ln I_t$ sich signifikant von 1 unterscheidet!

f) (*2 Punkte*) Begründen Sie kurz, ob es einen saisonalen Einfluss im Kaffeekonsum gibt!

g) (*4 Punkte*) Ein erfahrener Ökonometriker weist Sie darauf hin, dass zwischen $\ln P_t$ und der Störgröße des Modells (6) kontemporäre Korrelation herrschen könnte. Er schlägt vor, die Umsatzsteuer $Z_t$ als Instrumentvariable in einer

Instrumentvariablenschätzung des Modells (6) zu verwenden. Welche zwei Anforderungen muss $Z_t$ erfüllen, um ein geeignetes Instrument darzustellen?

h) (*12 Punkte*) Erläutern Sie die Funktionsweise eines F-Tests anhand der Nullhypothese, welche die gemeinsame Signifikanz der Parameter der logarithmierten exogenen Variablen des Modells (6) überprüft (Nullhypothese, Alternativhypothese, Ablehnungsbereich sowie Intuition)!

**Aufgabe 3 (insgesamt 49 Punkte): Kurzfragen**

a) (*9 Punkte*) Beschreiben Sie den Spezifikationstest von Hausman (Nullhypothese, Alternativhypothese, Ablehnungsbereich und Intuition)!

b) (*10 Punkte*) Zeichnen Sie eine Punktwolke mit gestutzter endogener Variable und erläutern Sie anhand dieser Punktwolke, welche Schätzprobleme auftreten!

c) (*11 Punkte*) Was versteht man unter perfekter Multikollinearität? Welche Probleme beim Schätzen entstehen durch hohe imperfekte Multikollinearität und wie kann man sie mildern? Welche Indizien deuten auf hohe Multikollinearität?

d) (*9 Punkte*) Um eine Gütermarktgleichgewichtsgleichung zu schätzen, wird im nachfolgenden RATS-Programm das logarithmierte reale Einkommen $\ln y_t$ auf den realen Zins $r_t$, den logarithmierten realen Konsum $\ln c_t$ und die logarithmierten Staatsausgaben $\ln g_t$ regressiert. Die Quartalsdaten ($T = 20$) liegen in der Exceltabelle *guetermarkt.xls* vor.

```
cal 2003 1 4
all 2007:4
open data guetermarkt.xls
data(format=xls,org=obs) / y r c g
set ly = log(y)
set lc = log(c)
set lg = log(g)
linreg ly / resid
# constant r lc lg
```

Wie lautet das in den letzten zwei Programmzeilen geschätzte Modell?

Formulieren Sie eine Null- und eine Alternativhypothese, welche die folgende Vermutung testen: „Die Realzins-Semi-Elastizität des realen Einkommen ist kleiner als -1“. Konstruieren Sie die Teststatistik für diesen Test. Gehen Sie von einem Signifikanzniveau von 5% aus und geben Sie den geeigneten kritischen Wert an. Unter welcher Bedingung werden Sie die Nullhypothese ablehnen?

Betrachten Sie den folgenden Befehl in Bezug auf die oben durchgeführte Regression:

```
test
# 3 4
# 0.5 0.5
```

Was für eine Nullhypothese wird mit diesem Befehl getestet und wie würden Sie diese ökonomisch interpretieren? Welcher Verteilung folgt die Teststatistik, die mit diesem Befehl berechnet wurde und wie hoch ist der kritische Wert bei einem Signifikanzniveau von 5%?

**-Bitte wenden!-**

e) (*10 Punkte*) Das RATS-Programm aus Aufgabenteil d) wird durch folgenden Programmteil ergänzt:

```
set resid2 = resid**2
set r2 = r**2
set lc2 = lc**2
set lg2 = lg**2
set rlc = r*lc
set rlg = r*lg
set lclg = lc*lg
set rlclg = r*lc*lg
linreg resid2
# constant r lc lg r2 lc2 lg2 rlc rlg lclg rlclg
compute R2T = %rsquared*%nobs
```

Welches Testverfahren wird in diesem Programmteil durchgeführt? Bei welcher Annahmeverletzung wird dieses Verfahren zur Diagnose eingesetzt? Wie lautet die Nullhypothese dieses Tests? Wie viele Freiheitsgrade hat in diesem Beispiel die für dieses Testverfahren berechnete Teststatistik? Welchen aus der Übung bekannten Befehl werden Sie benutzen, um den *p*-Wert für diesen Test mit RATS zu berechnen? (Geben Sie nur den Namen des Befehls an und nicht die komplette Spezifikation!) Wie hoch ist der kritische Wert bei einem Signifikanzniveau von 5%? Betrachten Sie die folgenden zwei Inputzeilen, die unmittelbar nach den oben angegebenen Befahlen ausgeführt werden und die entsprechende Outputzeile:

```
compute R2 = %rsquared
display 'R2' R2
```

*Output*

*R2 0,89*

Berechnen Sie die Teststatistik für das betrachtete Testverfahren und kommen Sie zu einer Entscheidung bei einem Signifikanzniveau von 5%. Wie hoch ist der *p*-Wert für diesen Test (geben Sie die bestmögliche Approximation an, z.B. 0,95<*p*-Wert<0,975)?