Klausur: Einführung in die Ökonometrie (Veranstaltungs-Nr. 1232, 4 GHP)

Prüfer: Dr. Ludwig v. Auer

Semester: Sommersemester 2001

Als Hilfsmittel sind zugelassen: nicht-programmierbarer Taschenrechner

Diese Klausur besteht aus drei Aufgaben, die in der zur Verfügung stehenden Zeit (2 Std.) zu bearbeiten sind. Runden Sie Ihre Berechnungen auf drei Stellen nach dem Komma.

Aufgabenstellung:

Aufgabe 1: Die in Magdeburg-Stadtfeld beheimatete linke Protestgruppierung „Fuck the Rich“ ist der Ansicht, daß die Menschen in reicheren Ländern nicht glücklicher sind als Menschen in ärmeren Ländern. Für eine empirische Analyse haben Sie Querschnittsdaten zur Verfügung, welche für 30 Länder jeweils das durchschnittliche Pro-Kopf-Monatseinkommen (x_t, gemessen in \$) und einen numerischen Indikator für das „Glücklichsein“ (y_t, gemessen in „Zufriedenheitspunkten“) angeben. Aus dem Datensatz wurden die folgenden Zwischenergebnisse ermittelt: $S_{xx} = 14.013.830$, $S_{yy} = 290.297$, $S_{xy} = 421.630$ sowie $\overline{x} = 1.047$ und $\overline{y} = 690$.

a) (*16 Punkte*) Schätzen Sie den Niveauparameter und den Steigungsparameter des ökonometrischen Modells. Versuchen Sie durch einen geeigneten einseitigen Test (Signifikanzniveau 10%) die Protestgruppierung davon zu überzeugen, daß „mehr Geld glücklicher macht“.

b) (*3 Punkte*) Wie groß ist das Bestimmtheitsmaß R^2 Ihrer Schätzung? Was sagt dieses Maß hier aus?

c) (*8 Punkte*) Welche Zufriedenheitspunktzahl würden Sie in einem Land erwarten, welches ein Pro–Kopf–Einkommen von 1000 \$ hat? Berechnen Sie auch ein entsprechendes Prognoseintervall (Signifikanzniveau 10%).

d) (*4 Punkte*) Leider scheint bei der Erfassung der *exogenen* Variable ein *konstanter* Meßfehler aufgetreten zu sein. Der Wert der Variable wurde in jedem Land um jeweils 10 \$ zu niedrig gemessen. Um welche Beträge wären Ihre Punktschätzer durch diesen Meßfehler ungefähr verzerrt?

e) (*9 Punkte*) Eine genauere Analyse der Daten zeigt, daß der Meßfehler wohl doch nicht konstant ist, sondern sich gemäß der Annahmen B1 bis B3 eines ökonometrischen Modells verhält. Durch diesen Meßfehler ergibt sich eine negative Korrelation zwischen der Störgröße des ökonometrischen Modells und der exogenen Variable. Erläutern Sie anhand einer geeigneten Grafik, welche Verzerrungen dadurch in Ihren Punktschätzern auftreten.

Aufgabe 2: Ein Politikberater stellt Ihnen eine Studie vor, welche zu dem Schluß kommt, daß die Höhe der Gehälter von höheren Staatsbediensteten (x_{1t}, gemessen als Relation zum Einkommen des höheren Managements in der Privatwirtschaft) die Korruptionsanfälligkeit dieser Bediensteten (y_t, gemessen als Korruptionsfälle pro 100 Bedienstete) erheblich senkt. 20 Beobachtungen (Länder) sind in die Studie eingegangen. Geschätzt wurde das Modell

$$y_t = \alpha + \beta_1 x_{1t} + u_t' \ . \tag{1}$$

Sie selbst vermuten, daß auch das auf Korruption verhängte durchschnittliche Strafmaß (x_{2t}, gemessen in Gehaltseinheiten des höheren Managements) einen Einfluß auf die Korruptionsanfälligkeit ausübt:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \ . \tag{2}$$

a) (*4 Punkte*) Unterstellen Sie, Modell (2) sei das korrekte Modell. Zeigen Sie, daß dann für die Störgröße des Modells (1) Annahme B1 verletzt ist.

b) (*8 Punkte*) Aus den Daten der in der Studie erfaßten Länder geht hervor, daß die Länder mit relativ besser entlohnten Staatsbediensteten tendentiell auch ein höheres Strafmaß auf Korruption verhängen. Unterstellen Sie, Modell (2) sei das korrekte Modell und $\widehat{\beta}_1$ und $\widehat{\beta}_2$ seien die entsprechenden KQ–Schätzer. Für den KQ–Schätzer des unvollständigen Modells, $\widehat{\beta}_1'$, gilt dann

$$\widehat{\beta}_1' = \widehat{\beta}_1 + \widehat{\beta}_2 \frac{S_{12}}{S_{11}} \ .$$

Erläutern Sie anhand des maßgeblichen Termes dieser Gleichung, inwiefern es zu einer Verzerrung des KQ–Schätzers $\widehat{\beta}_1'$ kommt und ob der Schätzer nach oben oder nach unten verzerrt ist.

c) (*10 Punkte*) Eine dritte Modellvariante ersetzt Variable x_{2t} (Strafmaß) durch die Überwachungsdichte (x_{3t}, gemessen durch Kontrolleure pro 100 höheren Staatsbediensteten):

$$y_t = \alpha + \beta_1 x_{1t} + \beta_3 x_{3t} + u_t'' \ . \tag{3}$$

Beschreiben Sie einen möglichen Test, um zwischen den Modellvarianten (2) und (3) abzuwägen.

d) (*5 Punkte*) Welche Testergebnisse müßten in dem von Ihnen in Aufgabenteil c) vorgeschlagenen Test auftreten, damit Sie zu der Ansicht kommen, daß man vielleicht doch besser das folgende Modell schätzen sollte:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t''' \ . \tag{4}$$

e) (*3 Punkte*) Wenn in Modell (4) eine irrelevante Variable enthalten wäre, welche Auswirkungen hätte dies auf Punktschätzer, Intervallschätzer und Hypothesentest?

f) (*4 Punkte*) Inwiefern können das standardisierte Bestimmtheitsmaß und das Akaike Informationskriterium als Maßzahlen interpretiert werden, welche das verringerte Verzerrungsrisiko und die zusätzliche Schätzvarianz einer zusätzlich in das Modell aufgenommenen Variable in bestimmter Weise gegeneinander abwägen.

g) (*6 Punkte*) Erläutern Sie kurz den Unterschied zwischen der Steinmetz–Methodologie und der Maurer–Methodologie.

Aufgabe 3: Kurzfragen

a) (*4 Punkte*) Erläutern Sie kurz die zwei wesentlichen Schwierigkeiten einer KQ-Schätzung des dynamischen Modells: $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + ... + \beta_K x_{t-K} + v_t$.

b) (*14 Punkte*) Die Volkswirtschaft eines Landes sei durch das makroökonomische Gleichungssystem $c_t = \alpha + \beta y_t + u_t$ und $y_t = c_t + i_t$ beschrieben. Dabei bezeichnet c_t den Konsum, y_t das Sozialprodukt bzw. das Einkommen und i_t die Investitionen einer Volkswirtschaft. Warum ist eine KQ–Schätzung der Konsumgleichung (erste Gleichung) nicht zulässig? Beschreiben Sie ein Schätzverfahren, welches für die Schätzung dieser Gleichung benutzt werden könnte.

c) (*6 Punkte*) Betrachten Sie in der Formelsammlung die Berechungsformeln für den F-Wert im Goldfeld–Quandt Test und den R^2T-Wert im White Test. Erläutern Sie, warum ein großer F-Wert bzw. ein großer R^2T-Wert auf Heteroskedastizität deuten.

d) (*4 Punkte*) Das Ministerium für Wandern und Touristik der Vereinigten Karibischen Urlaubsinseln gibt eine Studie in Auftrag, die untersuchen soll, ob die Sonnenscheindauer pro Jahr (Variablenname *sonne*, gemessen in Stunden) einen signifikanten Einfluß auf die Anzahl der Touristen (Variablenname *touristen*, gemessen in Übernachtungen) ausübt. Die Daten (*sonne.xls*) wurden in ausgewählten Sommerurlaubsregionen erhoben. Folgendes RATS-Programm wurde für die Auswertung der Daten geschrieben:

```
all 34
open data sonne.xls
data(format=xls,org=obs) / sonne touristen
linreg touristen
# constant sonne
```

RATS wirft folgenden verkürzten Output aus:

Variable	Coeff	Std Error	T-Stat	Signif
**********				
1. Constant	547949.049	95448.775	5.741	0.000
2. Sonne	127.854	37.660	3.395	0.001

Interpretieren Sie den von RATS ausgegebenen Wert des Steigungsparameters. Überprüfen Sie anhand des RATS–Outputs, ob der Steigungsparameter signifikant ist (Signifikanzniveau 1%).

e) (*12 Punkte*) Das RATS–Programm aus Aufgabenteil d) wird durch folgenden Programmteil ergänzt:

```
set touristendach = %beta(1) + %beta(2)*sonne
set touristendach2 = touristendach**2
set touristendach3 = touristendach**3
linreg touristen
# constant sonne touristendach2 touristendach3
restrict 2
# 3
# 1 0
# 4
# 1 0
```

Welcher Test wird in diesem Programmteil durchgeführt? Skizzieren Sie kurz die ökonometrische Idee des Tests und beschreiben Sie die notwendigen Arbeitsschritte anhand des Programmteils. Entscheiden Sie den Test anhand des folgenden RATS–Outputs (Signifikanzniveau 5%):

F(2, 30) = 0.97583 with Significance Level 0.38852266.

Lösungs-Skizze

Aufgabe 1

a) Beobachtungsumfang $T = 30$. Berechnung der KQ-Schätzer:

$$\widehat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{421630}{14013830} = 0,030$$
$$\widehat{\alpha} = \overline{y} - \widehat{\beta} \cdot \overline{x} = 690 - 0,030 \cdot 10476,667 = 658,590$$

Schätzer für die Störgrößenvarianz:

$$\begin{aligned} S_{\widehat{u}\widehat{u}} &= S_{yy} - S_{\widehat{y}\widehat{y}} = S_{yy} - \widehat{\beta} S_{xy} \\ &= 290297 - 0,030 \cdot 421630 = 277648,100 \\ \widehat{\sigma}^2 &= S_{\widehat{u}\widehat{u}}/T - 2 = 277648,100/28 = 9916,004 \end{aligned}$$

Einseitiger t-Test:

1. Getestet wird die Nullhypothese: $H_0 : \beta \leq 0$ gegen die Alternativhypothese $H_1 : \beta > 0$ mit Signifikanzniveau $a = 10\%$.
2. Berechnung von $\widehat{se}(\widehat{\beta})$:
$$\widehat{se}(\widehat{\beta}) = \sqrt{\widehat{var}(\widehat{\beta})} = \sqrt{\widehat{\sigma}^2/S_{xx}} = \sqrt{9916,004/14013830} = 0,027$$
3. Berechnung des t-Wertes:
$$t = \frac{\widehat{\beta} - \beta^0}{\widehat{se}(\widehat{\beta})} = \frac{0,030 - 0}{0,027} = 1,111$$
4. Kritischer Wert:
$$t_{0,1} = 1,3125 \text{ (28 Freiheitsgrade)}$$
5. Da $t < t_{0,1}$ kann H_0 nicht abgelehnt werden, d.h. es ist nicht gelungen, durch den Test zu belegen, daß mehr Geld glücklicher macht.

b) Bestimmtheitsmaß:

$$R^2 = \frac{S_{\widehat{y}\widehat{y}}}{S_{yy}} = \frac{S_{yy} - S_{\widehat{u}\widehat{u}}}{S_{yy}} = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{421630^2}{14013830 \cdot 290297} = 0,043698 \approx 4,370\%$$

Nur etwa 4% in der Variation der endogenen Variable ("Zufriedenheitspunktzahl") können durch die Variation im Einkommen erklärt werden.

c) Der Prognosewert beträgt:

$$\widehat{y}_0 = \widehat{\alpha} + \widehat{\beta} x_0 = 658,590 + 0,030 \cdot 1000 = 688,590$$

Berechnung von $\widehat{se}(\widehat{y}_0 - y_0)$

$$\begin{aligned}\widehat{var}(\widehat{y}_0 - y_0) &= \widehat{\sigma}^2 \left[1 + 1/T + (x_0 - \overline{x})^2/S_{xx}\right] \\ &= 9916,004[1 + 1/30 + 47^2/14013830] \\ &= 10248,1\end{aligned}$$

$$\widehat{se}(\widehat{y}_0 - y_0) = \sqrt{\widehat{var}(\widehat{y}_0 - y_0)} = 101,233$$

Das Prognoseintervall mit $a = 10\%$ und $t_{a/2} = 1,7011$ lautet:

$$\begin{aligned}&\left[\widehat{y}_0 - t_{a/2} \cdot \widehat{se}(\widehat{y}_0 - y_0); \widehat{y}_0 + t_{a/2} \cdot \widehat{se}(\widehat{y}_0 - y_0)\right] \\ =&[688,590 - 1,7011 \cdot 101,233; 688,590 + 1,7011 \cdot 101,233] \\ =&[516,383; 860,797]\end{aligned}$$

d) Der Steigungsparameter $\widehat{\beta}$ bleibt unverzerrt, der Niveauparameter $\widehat{\alpha}$ hingegen ist um $\beta \cdot 10$ verzerrt, also um etwa $0,03 \cdot 10 = 0,3$ zu hoch.

e) Negative Korrelation zwischen Störgröße und exogenen Variablen: $cov(x_t, u_t) < 0$. Der Steigungsparameter wird dann unterschätzt, der Niveauparameter überschätzt. Siehe Abbildung 20.2 Teil (b).

Aufgabe 2

a) Aus Gleichungen (1) und (2) ergibt sich

$$\begin{aligned}u_t' &= \beta_2 x_{2t} + u_t \\ E(u_t') &= \beta_2 x_{2t} \neq 0\end{aligned}$$

b) Der Verzerrungsterm lautet $\widehat{\beta}_2 \cdot S_{12}/S_{11}$. Es entstünde keine Verzerrung, wenn $S_{12} = 0$. Hier gilt jedoch laut Aufgabenstellung, dass $S_{12} > 0$. $\widehat{\beta}_2$ dürfte negativ sein. Also ergibt sich ein scheinbarer, negativer Einfluß $\widehat{\beta}_2 \cdot S_{12}/S_{11}$, wobei S_{12}/S_{11} den scheinbaren Einfluß des Einkommens auf das Strafmaß bemißt und $\widehat{\beta}_2$ den Einfluß des Strafmaßes auf die Korruptionsanfälligkeit. Der Schätzer β_1' dürfte zu niedrig ausfallen.

c) Ungenesteter F-Test oder J-Test (siehe Lehrbuch Abschnitte 13.2.5 bzw.13.2.6)

d) Falls im ungenesteten F-Test beide Nullhypothesen abgelehnt werden scheint in beiden Modellen etwas relevantes zu fehlen. In Modell (4) ist dann eventuell alles vorhanden. Das gleiche Argument gilt auch für den J-Test.

e) Die Punkt- und Intervallschätzer sind unverzerrt, aber nicht effizient. Hypothesentests führen zu korrekten aber unnötig unscharfen Ergebnissen.

f) Bei $\overline{R}^2$ repräsentiert $S_{\widehat{u}\widehat{u}}$ dass Verzerrungsrisiko und K Schätzvarianz. Das gleiche gilt für das

AIC.

g) siehe Lehrbuch Abschnitt 13.3.1

Aufgabe 3

a) Anzahl der Beobachtungen ist eventuell zu gering, denn es liegen zwar T Beobachtungen vor, aber nur $T - K$ vollständige. Es müssen aus diesen vollständigen Beobachtungen $K + 2$ Parameter geschätzt werden. Ferner weisen die exogenen Variablen oftmals hohe Multikollinearität auf, da sie letztlich aus einer einzigen Variable hervorgehen.

b) Die Störgröße u_t beeinflußt c_t, die Variable c_t wiederum hat einen Einfluß auf y_t, somit beeinflußt die Störgröße u_t auch die Höhe der exogenen Variable y_t; y_t und u_t sind kontemporär korreliert.

Die KQ-Schätzung ist verzerrt und nicht konsistent. IKQ und ZSKQ sind Schätzverfahren, die konsistente Schätzer liefern. (siehe Lehrbuch Abschnitte 23.2 und 23.4)

c) Goldfeldt-Quandt-Test: Der F-Wert wird groß, wenn das durchschnittliche Residuenquadrat der P-Gruppe deutlich größer als das der Z-Gruppe ist. Das bedeutet, daß die Varianz der Störgröße in der P-Gruppe größer als in der Z-Gruppe ist. Es besteht dann Heteroskedastizität.

White-Test: R^2 stammt aus der Regression von u_t^2 auf $x_{1t}, x_{1t}^2, ...$ Ein großes R^2 bedeutet, daß die Störgröße durch die exogenen Variablen gut erklärt wird, also eine Beziehung zwischen Störgrößen und exogenen Variablen besteht.

d) $\beta = 127,854$ bedeutet, das eine zusätzliche Sonnenscheindauer von $1h$ zu 127,854 mehr Übernachtungen führt. Der t-Test der Nullhypothese $H_0 : \beta = 0$ führt zur Ablehnung, da der p-Wert kleiner als 1% ist.

e) RESET-Test (siehe Lehrbuch Abschnitt 14.3.3)

Die Nullhypothese wird nicht verworfen. Es besteht damit kein Hinweis darauf, dass das lineare Modell nicht geeignet ist.