

Abschlussklausur zur Veranstaltung "**Einführung in die Ökonometrie**"
am 10. Februar 2004, 9.00-10.30 Uhr

Erlaubte Hilfsmittel: Tabelle der statistischen Verteilungen, 4 DIN A4-Seiten eigene Notizen, Taschenrechner, Fremdwörterbuch.

Bearbeitungszeit: 90 Minuten

Die Klausur besteht aus 7 Aufgaben in denen insgesamt 90 Punkte erworben werden können. Die Punktzahl je Aufgabe, bzw. je Teilaufgabe, ist in eckigen Klammern angegeben und entspricht der für die Aufgabe aufzuwendenden Zeit in Minuten.

Das Team der Abteilung Statistik und Ökonometrie wünscht Ihnen viel Erfolg!

Klausurergebnis:

Aufgabe	1a	1b	1c	1d	1e	1f	2a	2b	2c	2d	3	4	5a	5b	5c	5d	6	7	Σ	Note:	
Punkte:																					

Aufgabe 1:

[25]

Im Nachlass eines von Studenten gemeichelten Ökonometrie-Assistenten finden Sie umseitig abgedruckte EViews-Regressionsoutputs. Sie wissen, dass er zuletzt an der Erforschung der studentischen Leistungen mit einer Zufallsstichprobe von 240 Studenten gearbeitet hat. Wie sie aus Tabelle 1 ersehen können, hat er dazu folgendes Model benutzt:

$$L = \beta_1 + \beta_2 SEM + \beta_3 SEM^2 + \beta_4 AGE + \beta_5 AGE * SEM + \beta_6 MNOTE + \beta_7 LETZTES + \varepsilon$$

Als Leistungsindikator (L) verwendete er die Summe der Kreditpunkte aller im Semester erfolgreich abgeschlossenen Veranstaltungen multipliziert mit der jeweiligen Note. Als erklärende Variablen standen ihm die jeweilige Semesterzahl (SEM), das Alter (AGE) und die Matura-Note (MNOTE) zu Verfügung. Zudem wusste er, ob der Student sich im letzten Semester vor Studiumabschluss befand (LETZTES, Dummy, 0: nicht letztes Semester, 1: letztes Semester).

Erweisen Sie dem Assistenten die letzte Ehre, indem Sie sein Werk vollenden und dazu folgende Fragen beantworten:

- a) Interpretieren Sie den Regressionsparameter b_7 statistisch und inhaltlich. [2]
- b) Ergäbe sich ein Unterschied in der Interpretation von b_7 , wenn das Modell nur aus $\beta_1 + \beta_7 LETZTES + \varepsilon$ bestünde? Erläutern Sie. [3]
- c) Testen Sie $H_0: 2\beta_1 + \beta_7 = 46$ auf dem 5%-Signifikanzniveau zweiseitig. Stellen Sie die Alternativhypothese auf. Sie wissen, dass $Cov(b_1, b_7) = 0$. [5]
- d) Berechnen Sie ein 95%-Konfidenzintervall für den marginalen Effekt des Alters auf den Leistungsindikator eines Studenten im 5. Fachsemester und interpretieren Sie es inhaltlich. Sie wissen: $Cov(b_4, b_5) = 0,5$. [8]
- e) Berechnen Sie die Semesterzahl für einen 23jährigen Studenten, in dem seine Leistung am höchsten ist. [3]
- f) Testen Sie, ob die Semesterzahl einen signifikanten Erklärungsbeitrag zum Modell liefert. Geben Sie den kritischen Wert der Teststatistik an, die Sie zum Schliessen heranziehen, oder beschreiben Sie genau, wo man ihn in einer Tabelle statistischer Verteilungen finden würde. [4]

Dependent Variable: L				Tabelle 1
Method: Least Squares				
Included observations:				240
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	168.0000	63.8172	2.6325	0.0090
SEM	12.7500	7.0500	1.8085	0.0718
SEM^2	-27.0420	10.8500	-2.4924	0.0134
AGE	6.5000	5.7000	1.1404	0.2553
AGE * SEM	4.7500	1.5800	3.0063	0.0029
MNOTE	8.2500	26.7600	0.3083	0.7581
LETZTES	-38.0000	20.2500	-1.8765	0.0618
R-squared	0.6185	Mean dependent var		108.5000
Adjusted R-squared	0.6087	S.D. dependent var		14.2750
S.E. of regression	8.9295			
Sum squared resid	18578.2750			

Dependent Variable: L				Tabelle 2
Method: Least Squares				
Included observations:				240
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	198.0000	125.0000	1.5840	0.1145
AGE	7.2500	5.3250	1.3615	0.1747
MNOTE	3.7590	12.9750	0.2897	0.7723
LETZTES	-15.7690	18.2750	-0.8629	0.3891
R-squared	0.6040	Mean dependent var		108.5000
Adjusted R-squared	0.5990	S.D. dependent var		14.2750
S.E. of regression	9.0394	F-Test:		
Sum squared resid	19283.7791			

Aufgabe 2:

[17]

- a) Die Schätzung aus Aufgabe 1 wurde separat für 110 Studentinnen (Tab. 3) und 130 Studenten (Tab. 4) durchgeführt. Testen Sie, ob die gefundenen Regressionskoeffizienten für Männer und Frauen gleich sind. [4]
- b) Testen Sie, ob Männer und Frauen die gleiche Fehlertermvarianz haben. Erläutern Sie Ihr Testverfahren im Detail. [4]
- c) Sie wollen überprüfen ob Ihr Ergebnis in a) robust hinsichtlich unterschiedlicher Fehlertermvarianz ist. Dazu schätzen Sie Ihr Modell für beide Teilstichproben erneut mit dem (feasible) GLS-Verfahren und erhalten: $SSE_{\text{Frauen}}^{(fGLS)} = 103$, $SSE_{\text{Männer}}^{(fGLS)} = 123$, $SSE_{\text{beideGeschlechter}}^{(fGLS)} = 250,5$. Testen Sie erneut auf Gleichheit der Regressionskoeffizienten für Männer und Frauen und vergleichen Sie Ihre Ergebnis mit dem aus Teilaufgabe a). [4]
- d) Der Assistent hatte zuletzt folgende Hypothese formuliert: $Var(\varepsilon_i) = \sigma^2 \frac{1}{SEM_i^2}$.

Interpretieren Sie diese Hypothese inhaltlich. Notieren Sie ein neues Schätzmodell für das der KQ-Schätzer effizient ist. Können Sie eine Aussage bezüglich der von Ihnen erwarteten Standardfehler der Regressionskoeffizienten im Vergleich zur Ursprungsschätzung (Tab. 1) machen? [5]

Dependent Variable: L		Tabelle 3
Method: Least Squares		
Included observations:		110
<i>Regressionsergebnisse für 7 Koeffizienten aus Platzgründen gelöscht!</i>		
R-squared	0.7851	Mean dependent var
Adjusted R-squared	0.7675	S.D. dependent var
S.E. of regression	6.8838	
Sum squared resid	3459.2569	

Dependent Variable: L		Tabelle 4
Method: Least Squares		
Included observations:		130
<i>Regressionsergebnisse für 7 Koeffizienten aus Platzgründen gelöscht!</i>		
R-squared	0.5636	Mean dependent var
Adjusted R-squared	0.5524	S.D. dependent var
S.E. of regression	9.6127	
Sum squared resid	14137.9357	

Aufgabe 3:

[15]

Wahr oder Falsch? Tragen Sie für zutreffende Aussagen den Buchstaben w (für wahr), für nicht zutreffende f (für falsch) ein.

(Für jede richtige Antwort gibt es 0,5 Punkte, für jede falsche Antwort werden 0,5 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.)

	Jede F-Statistik ist das Quadrat einer t-Statistik.
	Die Varianz des Vorhersagefehlers im einfachen Regressionsmodell ist am Mittelwert der erklärenden Variablen am geringsten.
	Der Hausman Test testet prinzipiell, ob die durch unterschiedliche Schätzverfahren gewonnenen Schätzwerte für die Parameter gleich sind.
	Beim RESET Test werden zusätzlich zu den erklärenden Variablen ihre vorhergesagten Werte entweder nur im Quadrat oder zusätzlich in dritter Potenz in der Regressionsgleichung berücksichtigt.
	Die KQ Koeffizienten bei multivariaten Modellen werden im Gegensatz zu einfachen Modellen nicht aus der Minimierung der Summe quadrierter Fehlerterme abgeleitet.
	Eine Betrachtung der Residuen bietet dann erste Hinweise auf Autokorrelation in den Fehlertermen, wenn wir eine hohe Korrelation mit einer der erklärenden Variablen feststellen.
	Der "two-stage least squares" Schätzer korrigiert im ersten Schritt alle im Modell genutzten Variablen und führt im zweiten Schritt eine Schätzung mit diesen korrigierten Werten durch.
	Wenn die erklärenden Variablen einer Schätzgleichung mit dem Störterm des Modells korrelieren, ist der KQ Schätzer inkonsistent.
	Monte Carlo Studien untersuchen die Eigenschaften von Schätzverfahren anhand künstlich erzeugter Daten.
	Instrumentvariablenschätzer können das Inkonsistenzproblem im Fall stochastischer erklärender Variablen lösen.
	Auch bei autokorrelierten Fehlertermen sind KQ Schätzer BLUE.
	Eine Betrachtung der Residuen von Schätzern ist hilfreich, um Autokorrelationsmuster zu entdecken.

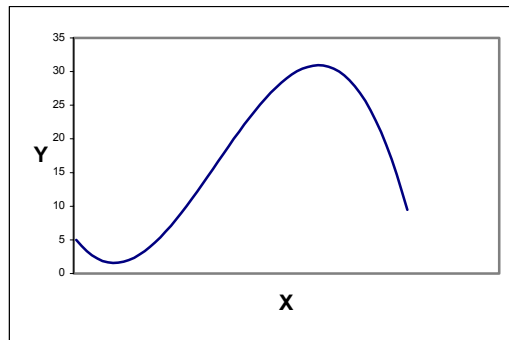
	Bei autokorrelierten Fehlertermen kann die statistische Signifikanz von KQ Parameterschätzern anhand der t-Statistik überprüft werden.
	Der Durbin-Watson Test testet auf Autokorrelation erster oder zweiter Ordnung.
	Das White-Verfahren bietet eine Korrekturmöglichkeit bei in erster Ordnung autokorrelierten Fehlertermen, wenn das Ausmass der Autokorrelation nicht bekannt ist.
	In einem Modell kann jede Variable W als Instrumentvariable genutzt werden, die nicht in der Modellgleichung vorkommt.
	Die verallgemeinerte KQ Schätzung zur Lösung des Problems der Autokorrelation kann nur angewendet werden, wenn Schätzwerte für rho vorliegen.
	Der F-Test auf Gesamtsignifikanz eines Modells prüft ob der R^2 Wert signifikant von 0 verschieden ist.
	Wenn die Varianz des Störterms über alle Beobachtungen der Stichprobe gleich ist, sind die Störterme homoskedastisch.
	Ausgelassenen relevante erklärende Variablen führen notwendigerweise zur Verzerrung des KQ Schätzers.
	Bei heteroskedastischen Störtermen ist der KQ Schätzer nicht mehr unverzerrt.
	Beim RESET Test wird ein F-Test auf die gemeinsame Signifikanz von Steigungsparametern durchgeführt.
	Die Varianz eines Steigungsparameters ist grösser, wenn die Fehlerterme des Modells heteroskedastisch sind, als wenn sie homoskedastisch sind.
	Bei Berücksichtigung irrelevanter erklärender Variablen im Modell ist der KQ Schätzer nur dann unverzerrt, wenn die irrelevanten Variablen mit den relevanten Variablen nicht korreliert sind.
	Beim Lagrange Multiplier Test auf Autokorrelation erster Ordnung kann die Schlussfolgerung anhand eines p-Wertes gezogen werden.
	Das Ausnutzen von aus der Theorie abgeleiteten Restriktionen in Schätzungen führt nur dann zu Verbesserungen der Schätzergebnisse, wenn die Restriktionen zutreffen.
	Der Goldfeld-Quandt Test ist nur für Situationen mit proportionaler Heteroskedastie geeignet.
	Der Jarque-Bera Test prüft ob Schiefe und Varianz der Verteilung einer Zufallsvariablen denen der Normalverteilung entsprechen.
	Bei Messfehlern in den erklärenden Variablen sind die geschätzten Parameter systematisch verzerrt. Positive Parameter werden zu gross, negative zu klein ausgewiesen.
	Schätzer werden dann als konsistent bezeichnet, wenn sich für gegen unendlich wachsende Stichprobengrössen die Wahrscheinlichkeitsverteilung des KQ Schätzers auf den wahren Parameterwert konzentriert.

Aufgabe 4:**[5]**

Sie modellieren die abhängige Variable als Funktion eines Polynom dritten Grades von x

$$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2 + \alpha_3 x_t^3 + \varepsilon_t$$

um den Verlauf der unten dargestellten Kurve abzubilden:



Welche Vorzeichen erwarten Sie für die vier geschätzten Parameter, wenn die Daten der Darstellung entsprechen?

Aufgabe 5:**[15]**

Sie nutzen Daten über Mitarbeitern eines Unternehmens, um die nachfolgende Lohngleichung mit einem KQ Schätzer zu schätzen. Dabei ist der Lohn in 10'000 CHF pro Jahr gemessen. Alle erklärenden Variablen sind Indikatoren, die Ihnen angeben, ob die Person weiblich ist (W), eine Matur hat (Mat), eine abgeschlossene Berufsausbildung hat (Ber), einen Tertiärabschluss (Ter) erworben hat, in der Abteilung 1 des Unternehmens arbeitet ($Abt1$) und in der Abteilung 2 des Unternehmens arbeitet ($Abt2$).

$$Lohn = \alpha + \beta_1 W + \beta_2 Mat + \beta_3 Ber + \beta_4 Ter + \beta_5 Abt1 + \beta_6 Abt2 + \beta_7 (Abt2 * W) + \varepsilon$$

- Beschreiben Sie die Charakteristika der Referenzperson. [3]
- Sie haben nur eine kleine Stichprobe. Wie beeinflusst dies Ihre Schätzergebnisse? [2]
- Um das Problem zu reduzieren, konsultieren Sie Fachleute, die Ihnen 2 Aussagen hinterlassen:
 - Der mittlere Lohnvorteil derjenigen mit Matur, Berufs- oder Tertiärbildung gegenüber denjenigen ohne diese Ausbildungsabschlüsse beträgt 35'000 CHF.
 - Abteilung 2 ist frauenfeindlich. Wenn Sie dort als Frau beschäftigt sind, verdienen Sie 10'000 CHF weniger als sonst gleiche Männer in dieser Abteilung.Wie können Ihnen diese Aussagen zu besseren Schätzergebnissen verhelfen? Erläutern Sie, wie Sie diese Informationen nutzen können und präsentieren Sie eine neue Schätzgleichung. [8]
- Nachdem Sie Ihre Schätzung durchgeführt haben, ruft einer der beiden Fachleute an und teilt mit, dass er sich hinsichtlich der Lohnunterschiede für Frauen in der Abteilungsnummer geirrt hat. Was bedeutet das für Ihre gerade abgeschlossene Untersuchung? [2]

Aufgabe 6:**[3]**

Erläutern Sie, warum folgende Aussage korrekt ist: Bei einem Fehlerterm, der einem Autokorrelationsprozess erster Ordnung mit einem Wert von $\rho = 0.5$ folgt, beträgt die Kovarianz der Fehler in Periode t und $t+3$ nur dann 0.125, wenn dieser Fehler standardnormalverteilt ist.

Aufgabe 7:**[10]**

Sind folgende Aussagen richtig? Erläutern Sie **stichwortartig** Ihre Auffassung (Bsp.: "Stimmt, weil..." bzw. "Stimmt nicht, weil..."). Nur bei korrekter Begründung wird die Antwort mit einem Punkt pro Frage honoriert.

	Die Störtermvarianz bei Anwendung von White's Schätzer ist grösser als bei Anwendung eines generalisierten KQ Schätzers.
	Bei Fehlern, die in erster Ordnung autokorreliert sind, sind die Störterme von Beobachtungen die mindestens 2 Perioden auseinanderliegen nicht miteinander korreliert.
	Die Verzerrung des KQ Schätzers bei ausgelassenen, relevanten erklärenden Variablen ist geringer wenn die ausgelassenen mit den berücksichtigten Variablen hoch korreliert sind.
	Bei in erster Ordnung autokorrelierten Störprozessen wirkt ein Zufallsschock in Periode t umso stärker auf den Störterm in Periode t , je höher der Betrag des Parameters ρ ist.
	Im Gegensatz zu Punktschätzern sind Intervallschätzer von Autokorrelation nicht beeinflusst.

	Bei heteroskedastischen Varianzen müssen die Intervallschätzer für Steigungsparameter angepasst werden.
	Die Varianz autokorrelierter Störterme ist homoskedastisch.
	Hohe Korrelation unter erklärenden Variablen führt zu fehlerhaften Schätzungen der Parametervarianz.
	Es ist möglich, den Parameter rho bei autokorrelierten Störtermen erster Ordnung als Steigungsparameter in einer KQ Schätzung zu schätzen.
	Zur Berechnung eines effizienten Schätzers muss im Fall autokorrelierter Störterme auf die jeweils erste Beobachtung einer Zeitreihe verzichtet werden.