

## Diplomprüfung im Fach Ökonometrie im SS 08 - Aufgabenteil

Name, Vorname	
Matrikelnr.	
Studiengang	
Semester	
Datum	07.08.2008
Raum, Sitzplatz-Nr.	
Unterschrift	

### Vorbemerkungen:

**Anzahl der Aufgaben:** Die Klausur besteht aus 6 Aufgaben, von denen alle bearbeitet werden müssen.

**Bewertung:** Die Prüfung dauert 120 Minuten, es können maximal 120 Punkte erworben werden. Die Punktzahl ist für jede Aufgabe in Klammern angegeben.

**Erlaubte Hilfsmittel:**

- 2 DIN A4-Blätter mit Notizen (Vorder- und Rückseite, also max. 4 DIN A4-Seiten)
- Tabellen der statistischen Verteilungen (sind der Klausur beigefügt)
- Taschenrechner
- Fremdwörterbuch

**Wichtige Hinweise:**

- Sollte es vorkommen, dass die statistischen Tabellen, die dieser Klausur beiliegen, den exakten Wert der Freiheitsgrade nicht ausweisen, machen Sie dies kenntlich und verwenden Sie den nächstgelegenen Wert.
- Sollte es vorkommen, dass bei einer Berechnung eine erforderliche Annahme fehlt, machen Sie dies kenntlich und treffen Sie für den fehlenden Wert eine plausible Annahme.
- Die Aufgaben 5 und 6 sind im Aufgabenteil zu beantworten, die restlichen Aufgaben im Lösungsteil. Verwenden Sie hierbei für jede Aufgabe ein neues Blatt.

**Aufgabe 1:****[27 Punkte]**

Sie interessieren sich für die Determinanten der Preise von Rotweinen. Sie nutzen Querschnittsdaten über 193 verschiedene Weine der Jahrgänge 1989 bis 1991 und schätzen die folgende hedonische Preisfunktion:

$$\log(\text{price}_i) = \beta_0 + \beta_1 \cdot \text{grade}_i + \beta_2 \cdot \text{bord}_i + \beta_3 \cdot \text{an89}_i + \beta_4 \cdot \text{an91}_i + \varepsilon_i$$

- $\text{price}_i$ : Preis des  $i$ -ten Weins  
 $\text{grade}_i$ : Testnote des  $i$ -ten Weins in einem Wein-Ranking  
 $\text{bord}_i$ : Dummy-Variable (=1, wenn Wein  $i$  ein Bordeaux ist, = 0 sonst)  
 $\text{an89}_i$ : Dummy-Variable (=1, wenn Wein  $i$  Jahrgang 1989 ist, = 0 sonst)  
 $\text{an91}_i$ : Dummy-Variable (=1, wenn Wein  $i$  Jahrgang 1991 ist, = 0 sonst)

Die Auswertung ergibt folgenden Output:

```
Call:
lm(formula = log(price) ~ grade + bord + an89 + an91)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.6104     0.3355  10.761 < 2e-16
grade         0.0821     0.0150   5.443 1.62e-07
bord        -0.4192     0.0994   ??? 3.87e-05
an89         0.1212     0.0612   1.980 0.0987
an91         ???       0.0131  -4.493 1.23e-05
---
Residual standard error: 0.4077 on 188 degrees of freedom
Multiple R-squared:  ???, Adjusted R-squared: 0.5067

F-statistic: ??? on 4 and 188 DF, p-value: < 2.2e-16
```

- a) Berechnen Sie unter Angabe des Rechenwegs (6 Punkte)
- den  $t$ -Wert für  $b_2$ .
  - den geschätzten Koeffizienten für  $\beta_4$ .
  - ein 95% Konfidenzintervall für  $b_3$ .
  - $R^2$ .
  - die F-Statistik des Gesamtmodells (sollten Ihnen der hierfür benötigte  $R^2$ -Wert aus a4) nicht vorliegen, unterstellen Sie  $R^2 = 0.525$ ).
- b) Interpretieren Sie die Koeffizienten  $b_1$  und  $b_3$  inhaltlich und statistisch. (2 Punkte)
- c) Welche Auswirkung hat es auf den Erwartungswert des KQ-Schätzers  $b = \left( \sum_{i=1}^n x_i x_i' \right)^{-1} \sum_{i=1}^n x_i y_i$ , wenn Sie statt des wahren Modells  $y_i = x_i' \beta + z_i' \gamma + \varepsilon_i$  das Modell  $y_i = x_i' \beta + v_i$  schätzen. Leiten Sie Ihre Antwort formal ab. Unterstellen Sie die Gültigkeit der Gauss-Markov-Annahmen. (5 Punkte)
- d) Nehmen Sie umgekehrt an, dass Ihr Modell eine irrelevante Variable enthält. Welche Konsequenz hätte dies für die Eigenschaften Ihrer Schätzergebnisse? (1 Punkt)
- e) Sie vermuten nun, dass Sie das Modell fehlspezifiziert haben. Um dies zu überprüfen, wollen Sie einen RESET-Test durchführen. Stellen Sie die Hilfsregression für eine Variante dieses Tests auf und erläutern Sie die Idee und Schlusslogik des Tests. (4 Punkte)
- f) Sie schätzen ein weiteres Modell, in welchem Sie Aroma-Intensität als erklärende Größe aufnehmen. Sie fragen sich anschließend, welches der beiden Modelle Sie bevorzugen sollten. Berechnen Sie hierzu die AIC- und BIC-Werte und entscheiden Sie (Hinweis: Im Modell ohne Aroma-Intensität ist  $\log \frac{1}{N} \sum e_i^2 = 23.0358$ , im Modell mit Aroma-Intensität ist  $\log \frac{1}{N} \sum e_i^2 = 21.0244$ ). (4 Punkte)

- g) Sie vermuten schließlich, dass die Preisdeterminanten aus dem Modell in a) unterschiedlichen Einfluss auf Bordeaux-Weine bzw. auf die anderen Weine haben. (5 Punkte)
- g1) Skizzieren Sie kurz verbal die Intuition und die Vorgehensweise eines geeigneten Tests, der untersucht, ob ein Strukturunterschied vorliegt.
- g2) Geben Sie zusätzlich Null- und Alternativhypothese, die im Modell vorliegenden Freiheitsgrade und den kritischen Wert der Teststatistik an.
- g3) Wie lautet Ihre Schlussfolgerung bei einem empirischen Wert der Teststatistik von 4.55?

**Aufgabe 2:**

**[10 Punkte]**

Gegeben sei ein Störtermprozess mit folgender Struktur:

$$\varepsilon_t = u_t + u_{t-1} + u_{t-2},$$

wobei  $u_t \sim N(0, \sigma_u^2)$  und  $cov(u_t, u_{t-s}) = 0 \quad \forall s \neq 0, t \neq 0$ .

- a) Bestimmen Sie die Varianz von  $\varepsilon$  und die Kovarianz benachbarter Störterme  $\varepsilon_t$  und  $\varepsilon_{t-1}$ . (4 Punkte)
- b) Stellen Sie die Varianz-Kovarianz-Matrix von  $\varepsilon_t$  dar. (4 Punkte)
- c) Beschreiben Sie einen empirischen Zusammenhang, bei dem eine solche Fehlertermstruktur vorkommen kann. (2 Punkte)

**Aufgabe 3:**

**[9 Punkte]**

Sie wissen um die Überfischung verschiedener Fischarten und interessieren sich daher für den Markt für Kabeljau. Ein einfaches Modell dieses Marktes lässt sich in folgender Form darstellen:

$$\text{Nachfrage: } q_d = \alpha_1 p + \alpha_2 y + e_d; \quad (1)$$

$$\text{Angebot: } q_s = \beta_1 p + e_s, \quad (2)$$

wobei die nachgefragte Menge ( $q_d$ ) eine Funktion des Preises ( $p$ ) und des Einkommens ( $y$ ) sei, die angebotene Menge ( $q_s$ ) sei eine Funktion des Preises ( $p$ );  $e_s$  und  $e_d$  seien die Störterme des Modells.

- a) Welches Problem tritt hierbei für die Schätzung des Steigungsparameters  $\beta_1$  auf? Erläutern Sie kurz und benennen Sie eine Konsequenz für den KQ-Schätzer. (2 Punkte)
- b) Wie kann das Problem ohne zusätzliche Information auf Modellgleichungsebene umgangen werden? (1 Punkt)
- c) Sie interessieren sich zudem für die zukünftige Entwicklung des Kabeljaumarktes und möchten Vorhersagen generieren. Leiten Sie allgemein für das lineare Modell  $y_i = x_i' \beta + \varepsilon_i$  die Varianz des Vorhersagefehlers ab und zeigen Sie formal, in welcher Form diese von  $\sigma^2$  abhängt. (6 Punkte)

**Aufgabe 4:**

**[14 Punkte]**

Sie interessieren sich für die Ersparnisbildung in Haushalten. Ihnen liegen Paneldaten der Haushalte  $i$  zum Zeitpunkt  $t$  für die Ersparnisse ( $y_{it}$ ) und das verfügbare Haushaltseinkommen ( $w_{it}$ ) vor.

- a) Sie wollen zunächst das gepoolte KQ-Modell  $y_{it} = \beta_1 + \beta_2 w_{it} + v_{it}$  schätzen. Sie vermuten jedoch, dass die Einkommen mit Fehler gemessen wurden,  $x_{it} = w_{it} + u_{it}$ . Erläutern Sie kurz verbal, welche Auswirkung dies auf den KQ-Schätzer für die Konstante und den Steigungsparameter hat. (2 Punkte)

- b) Sie wollen Panelschätzmethoden verwenden. Erläutern Sie, warum sich die Schätzergebnisse für einerseits between und fixed effects Schätzer und andererseits random effects und fixed effects Schätzer teils erheblich unterscheiden können (Hinweis: Für den Unterschied zwischen random effects und fixed effects lassen sich mehrere Gründe anführen). (6 Punkte)
- c) Ein Test für die Wahl zwischen dem random effects und dem fixed effects Schätzer ist der Hausman Test. (6 Punkte)
- c1) Skizzieren Sie kurz die Idee des Tests.
- c2) Geben Sie Null- und Alternativhypothese an und erläutern Sie diese kurz.
- c3) Die empirische Teststatistik beträgt 11.05 bei 5 Freiheitsgraden. Geben Sie den kritischen Wert der Teststatistik und Ihre Schlusslogik an.

**Aufgabe 5:**

**[45 Punkte]**

Wahr oder falsch? Tragen Sie für jede der folgenden Aussagen ein „w“ für „wahr“ oder ein „f“ für „falsch“ ein. Für jede richtige Antwort gibt es 0.75 Punkte, für jede falsche Antwort werden 0.75 Punkte abgezogen. Die Gesamtpunktzahl kann nicht negativ werden.

	Die t-Verteilung ist eine symmetrische Verteilungsfunktion.
	Der kritische Wert der Durbin-Watson Teststatistik variiert mit der Anzahl der Regressoren im Modell.
	Das angepasste $R^2$ einer Schätzung kann bei Hinzufügen von erklärenden Variablen sinken.
	Signifikante Koeffizienten können kleiner als 1 sein.
	Die Annahme $V\{\varepsilon\} = \sigma^2 I$ gilt unter Autokorrelation, aber nicht unter Heteroskedastie.
	Bei einem gegebenen Schätzverfahren und Modell können sich für unterschiedliche Stichproben verschiedene Schätzwerte ergeben.
	Im linearen Modell gibt die Regressionskonstante den Mittelwert der abhängigen Variable an.
	Wenn gilt $E\{x_{i2}\varepsilon_i\} = 0$ , sagen wir, dass $x_{i2}$ eine endogene erklärende Variable ist.
	In der Paneldatenanalyse wird standardmäßig unterstellt, dass sich die Steigungsparameter über die Zeit ändern.
	AR(1) Fehlerterme sind heteroskedastisch.
	Viele Multikollinearitätsprobleme können durch Vergrößerung der Stichprobe reduziert werden.
	Der within Schätzer kann als Regressoren die beobachtungsspezifischen (i,t) Abweichungen vom Mittelwert der Variable pro Beobachtungseinheit (i) verwenden.
	Die Standardnormalverteilung hat einen Erwartungswert von Eins.
	Der p-Wert gibt das kleinstmögliche Signifikanzniveau an, auf dem die Nullhypothese verworfen werden kann.
	Zu den Gauss-Markov-Annahmen gehört, dass der Erwartungswert der Regressionskonstante Null ist.
	Der Wert des $R^2$ kann nicht kleiner als Null sein.
	Die Nullhypothese $\beta \leq \beta_k$ wird am 5 Prozent Signifikanzniveau bei 500 Freiheitsgraden verworfen, wenn der empirische t-Wert größer als 1,645 ist.

	Der KQ-Schätzer erzeugt für Steigungsparameter, aber nicht für die Regressionskonstante Zufallsvariablen.
	Je größer der Typ II Fehler eines Tests, umso größer muss der Typ I Fehler sein.
	Negative Autokorrelation kann nicht mit dem LM-Test getestet werden.
	Schätzgleichungen mit quadratischen erklärenden Variablen erfordern nichtlineare Schätzverfahren.
	Newey-West Standardfehler stellen eine Erweiterung von White-Standardfehlern auf den Fall der Autokorrelation dar.
	Das Auslassen einer irrelevanten erklärenden Variablen führt zu überhöhten Standardfehlern.
	Bei strikter Exogenität der erklärenden Variablen ist der fixed effects-Schätzer konsistent.
	Im Falle von Heteroskedastie sind die Störterme autokorreliert.
	Quadratische Matrizen haben mehrere Inversen.
	Der "Encompassing Test" wird genutzt, um alternative Spezifikationen der abhängigen Variablen gegeneinander zu testen.
	Bei Autokorrelation sind die mit den KQ-Schätzern ausgewiesenen p-Werte ungültig.
	Der White-Schätzer benutzt gruppenspezifische Mittelwerte, um heteroskedastische Koeffizienten zu berechnen.
	Ein Parameterschätzer ist effizient, wenn er gegen seinen wahren Wert konvergiert.
	Der Breusch-Pagan Test basiert auf der Regression der geschätzten KQ-Residuen auf die erklärenden Größen $x$ .
	Der RESET Test nutzt Potenzen der Residuen, um ein Modell auf Fehlspezifikation zu überprüfen.
	Bei Paneldaten lassen sich Probleme ausgelassener zeitkonstanter Variablen durch den fixed effects Schätzer lösen.
	Die Breite des Vorhersageintervalls für die abhängige Variable variiert mit der Streuung des Störterms.
	Autokorrelation führt zu verzerrten KQ-Schätzern.
	Das $R^2$ ist nur gültig, wenn das Modell mit Konstante geschätzt wird.
	Der gepoolte KQ-Schätzer nutzt im Panel sowohl die Information der Within-Variation als auch der Between-Variation.
	Die Teststatistik des Breusch-Pagan Tests ist asymptotisch Chi-quadrat verteilt.
	Der Durbin-Watson Test verallgemeinert den White-Test.
	In Querschnittsdaten kann es nicht zu zeitlich bedingter Autokorrelation kommen.
	Die Parameter eines FGLS-Modells müssen vor der Interpretation transformiert werden.
	Die Qualität von Instrumentvariablenschätzern hängt von der Qualität der Instrumente ab.

	Die Wahl der Gewichtungsmatrix $W$ ändert nichts an der Konsistenz des GMM Schätzers.
	Um $K$ Parameter zu identifizieren, benötigt man $K - 1$ Momentenbedingungen.
	Wird beim GMM Schätzer nicht die optimale Gewichtungsmatrix $W$ gewählt, so sind die Parameterschätzer verzerrt.
	Bei Modellen in struktureller Form ist es möglich, dass erklärende Variablen mit dem Störterm korreliert sind.
	Der nichtlineare Kleinstquadrateschätzer bestimmt diejenigen Parameter, die die quadrierte Summe der Störterme minimieren.
	HAC Standardfehler unterstellen auf eine feste Anzahl von Beobachtungseinheiten begrenzte Störtermkorrelationen.
	Das GMM-Verfahren schätzt Parameter direkt auf Basis von nichtlinearen Momentenbedingungen.
	Homoskedastische Störterme haben eine nicht konstante Varianz.
	Im Rahmen einer GIVE-Schätzung mit mehreren endogenen Variablen und mehreren Instrumenten ist es egal, welcher endogenen Variablen welches Instrument zugeordnet wird.
	Auf Basis linearer Modelle geschätzte Koeffizienten sind kausale Effekte.
	GMM Verfahren benötigen keine Annahmen bezüglich der Verteilungsfunktion.
	Die Annahme $\varepsilon_i \sim i.i.d.(0, \sigma^2)$ schließt sowohl Heteroskedastie als auch Autokorrelation aus.
	Die Durbin-Watson Teststatistik kann approximiert werden als $dw \approx 2 - 2\hat{\rho}$ .
	Bei Autokorrelation erster Ordnung wird der Wert eines Störterms unmittelbar vom Wert des Störterms der vorausgehenden Perioden beeinflusst.
	Mithilfe eines linearen Regressionsmodells lassen sich Elastizitäten schätzen.
	Ohne Kenntnis der Kovarianz zwischen den beiden Schätzern kann der Hausman Test zur Entscheidung zwischen fixed und random effects Schätzung nicht durchgeführt werden.
	Ein schwaches Instrument liegt dann vor, wenn die Instrumentvariable eine Varianz unter 10 hat.
	Der GLS Schätzer im random effects Modell stellt einen Mittelwert zwischen random effects und fixed effects Schätzer dar.

### Aufgabe 6:

[15 Punkte]

Wahr oder falsch? Begründen Sie Ihre Auffassung (Bsp.: "Stimmt, weil..." bzw. "Stimmt nicht, weil..."). Nur bei korrekter Begründung erhält jede richtige Antwort 1.5 Punkte; Angaben **ohne Begründung** werden **nicht gewertet**.

	Der Wert des $R^2$ kann nicht größer sein als 1.
--	--

	Bei Autokorrelation spielt die Art der erklärenden Variablen für die Konsistenz der KQ-Schätzung eine Rolle.
	Bei Vorliegen von Autokorrelation sind die Störterme aller Beobachtungen miteinander korreliert.
	Die F-Statistik kann nicht negativ werden.
	Das random effects Modell kann mittels erster Differenzen geschätzt werden.
	Wenn im einfachen linearen Modell Messfehler in der erklärenden Variablen vorliegen, ist abschätzbar, in welche Richtung der Steigungsparameter verzerrt ist.
	Das Signifikanzniveau einseitiger Hypothesentests ist kleiner als das zweiseitiger Tests.
	Der between-Schätzer ist effizient, der random effects-Schätzer nicht.
	Der Durbin-Watson Test auf Autokorrelation ist nur bei bestimmten Modellspezifikationen anwendbar.
	Unter den Gauss-Markov-Annahmen sind KQ Schätzer nicht exakt normalverteilt.