

**Klausur WS 2009/10  
"Angewandte Ökonometrie"**

**22.02.2010  
Dauer 90 Min.  
+ Lesezeit: 10 Min.**

**Termin I**

Name / Surname:

Vorname / First Name:

Immatrikulations-Nr. / Enrolment No.:

**Erlaubte Hilfsmittel: Taschenrechner (nicht programmierbar).  
Allowed aids: pocket calculator (non intelligent)**

**Bitte bestätigen Sie durch Ihre Unterschrift, dass Ihnen bekannt ist, dass jeder Täuschungsversuch zum Abbruch der Arbeit mit der Bewertung "ungenügend" (5.0) führt.**

**Please confirm by signing that you are aware that any attempt to deceive results in a grade "failed" (5.0)**

.....  
Unterschrift / Signature

(nur für die Prüfer / only for examiner)

Punkte/ Points	Problem 1	Problem 2	Problem 3	Gesamt / Total
mögliche / possible	14	15	18	47
erreichte / obtained				

Note / Grade:

.....  
1. Prüfer/1st examiner

.....  
2. Prüfer/2nd examiner

I am not

### Aufgabe 1 (14 Punkte)

Gegeben sei das folgende lineare Regressionsmodell:

$$y = X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 I_N), \quad (1)$$

wobei  $X$  eine deterministische  $(N \times K)$ -Matrix ist mit  $\text{Rang}(X) = K$ .

- a) (2 Punkte) Welche Modellannahme ist im Fall einer exakten Multikollinearität verletzt?
- b) (4 Punkte) Abbildung 1 liefert das Histogramm und einige deskriptive Statistiken für die standardisierten Residuen einer KQ-Schätzung des Modells in Gleichung (1).
- Wie lautet die Null- und Alternativhypothese für den Jarque-Bera-Test und welche Schlussfolgerung würden Sie bei einem Signifikanzniveau von 5% ziehen?
  - Welche Modellannahme in Gleichung (1) ist von Ihrem Ergebnis aus (i) betroffen bzw. muss geändert werden?  
Nennen Sie kurz die Konsequenzen dieser Annahmenänderung hinsichtlich der Effizienz des KQ-Schätzers für  $\beta$  sowie der Gültigkeit der Inferenz ( $t$ - und  $F$ -Test).
- c) (8 Punkte) Abbildung 2 zeigt das Streudiagramm der standardisierten Residuen und angepassten Werte der abhängigen Variable ("Fitted Values") für eine KQ-Schätzung des Modells in Gleichung (1).
- Welche Schlüsse lässt das Streudiagramm hinsichtlich einer möglichen Heteroskedastizität der Fehlerterme zu? Begründen Sie Ihre Antwort.
  - Welche Auswirkungen hätte Heteroskedastizität auf die Erwartungstreue des KQ-Schätzers für  $\beta$  und die Gültigkeit der Inferenz ( $t$ - und  $F$ -Test)?  
Nennen Sie gegebenenfalls Maßnahmen, welche eine erwartungstreue KQ-Schätzung bzw. eine korrekte Inferenz (basierend auf dem KQ-Schätzer) gewährleisten.  
**Hinweis:** *Unterstellen Sie für Ihre Antwort, dass alle übrigen Annahmen in Gleichung (1) weiterhin bestehen.*
  - Welcher Schätzer ist bei Heteroskedastizität der Fehlerterme effizient?  
Geben Sie eine Formel für die Schätzung des Parametervektors  $\beta$  in Matrix-Notation an und benennen Sie die darin auftretenden Matrizen bzw. Vektoren.  
Inwiefern ist die Verwendung dieses Schätzers in der Praxis schwierig?

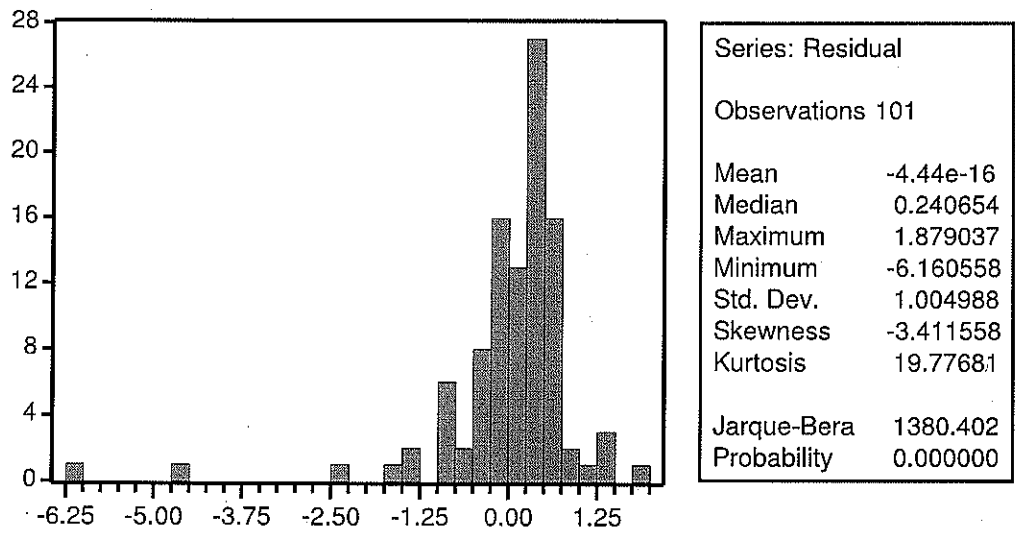


Abbildung 1: Histogramm und deskriptive Statistiken der standardisierten Residuen

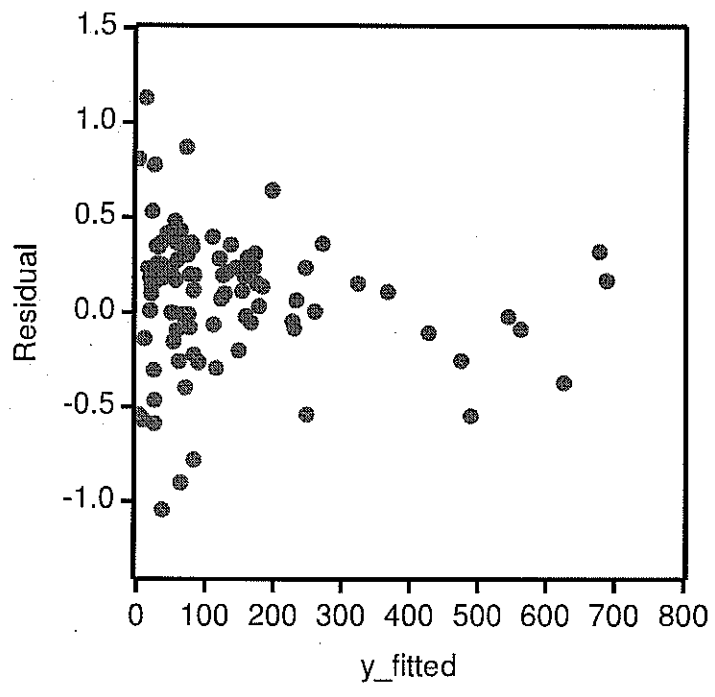


Abbildung 2: Streudiagramm der standardisierten Residuen ("Residual") und angepassten Werte ("y\_fitted")

## Aufgabe 2 (15 Punkte)

- a) (2 Punkte) Beschreiben Sie kurz, wann ein Paneldaten-Modell mit (i) festen oder (ii) zufälligen individuellen Effekten konzeptionell sinnvoll ist. Geben Sie auch jeweils ein kleines Fallbeispiel an.
- b) Mit Hilfe von Paneldaten-Modellen soll untersucht werden, inwieweit die Variablen Berufserfahrung, Gewerkschaftsmitgliedschaft, Eheverhältnis und Anstellung im öffentlichen Sektor den Lohn eines Arbeitnehmers beeinflussen. Dazu liegt ein Paneldatensatz der Dimension  $N=545$  und  $T=8$  mit folgenden Variablen vor:

Variable	Beschreibung
wage	Logarithmierter Lohn
exper	Berufserfahrung
union	Dummy, =1 falls Person Mitglied einer Gewerkschaft ist
mar	Dummy, =1 falls Person verheiratet ist
pub	Dummy, =1 falls Person im öffentlichen Sektor angestellt ist

- (i) (2 Punkte) Abbildung 3 zeigt die Ergebnisse für die “Within”-Schätzung. Welche Parameterschätzwerte fehlen in Abbildung 3? Nennen Sie die Art der Parameter und deren Anzahl.
- (ii) (3 Punkte) Die Tabellen 1 bzw. 2 zeigen die Ergebnisse für den Hausman- bzw. Breusch-Pagan-Test. Geben Sie für beide Tests die Nullhypothese an und interpretieren Sie die Resultate hinsichtlich der Verwendung des verallgemeinerten KQ-Schätzers (GLS).

**Hinweis:** Legen Sie jeweils ein Signifikanzniveau von 5% zugrunde.

- c) Gegeben sei folgendes Einweg-Fehlerkomponenten-Regressionsmodell:

$$y_{it} = \alpha + x'_{it}\beta + \varepsilon_{it}, \quad \varepsilon_{it} = \mu_i + e_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

wobei  $x_{it} = (x_{1it}, \dots, x_{Kit})'$  (nicht-stochastisch) und  $\beta = (\beta_1, \dots, \beta_K)'$  gilt. Des Weiteren seien alle  $\mu_i$  und  $e_{it}$  unabhängige Zufallsgrößen mit  $\mu_i \sim (0, \sigma_\mu^2)$  und  $e_{it} \sim (0, \sigma_e^2)$ . In Matrix-Notation lässt sich das Modell schreiben als

$$y = \alpha \mathbb{1}_{NT} + X\beta + \epsilon = Z\theta + \epsilon, \quad \epsilon = G\mu + e,$$

wobei  $y$ ,  $\epsilon$  und  $e$  die  $NT$ -Vektoren aller Beobachtungen  $y_{it}$  und Fehler  $\epsilon_{it}$  bzw.  $e_{it}$  sind,  $X$  die  $NT \times K$ -Matrix der erklärenden Variablen mit den Zeilen  $x'_{it}$  ist sowie  $\mu = (\mu_1, \dots, \mu_N)'$  und  $G = I_N \otimes \mathbb{1}_T$  gilt. Darüber hinaus gilt  $Z = [\mathbb{1}_{NT}; X]$  und  $\theta = (\alpha \beta)'$ .

- (i) (3 Punkte) Welche Schätzung für  $\beta$  würden Sie im Falle bekannter Varianzkomponenten  $\sigma_\mu^2$  und  $\sigma_e^2$  verwenden. Erklären Sie kurz wie der Schätzer konzeptionell aufgebaut ist.

**Hinweis:** Die Angabe einer exakten Formel ist nicht erforderlich.

- (ii) (5 Punkte) Zeigen Sie, dass eine “gepoolte” KQ-Schätzung für  $\beta$  erwartungstreu ist.

Dependent Variable: WAGE Method: Panel Least Squares Date: 02/03/10 Time: 17:38 Sample: 1980 1987 Cross-sections included: 545 Total panel (balanced) observations: 4360				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.065698	0.026677	39.94874	0.0000
EXPER	0.116457	0.008431	13.81312	0.0000
EXPER^2	-0.004289	0.000605	-7.083369	0.0000
UNION	0.081203	0.019316	4.203943	0.0000
MAR	0.045106	0.018311	2.463280	0.0138
PUB	0.034927	0.038608	0.904646	0.3657
R-squared	0.619822	Mean dependent var		1.649147
Adjusted R-squared	0.565040	S.D. dependent var		0.532609
S.E. of regression	0.351264	Akaike info criterion		0.862892
Sum squared resid	470.1014	Schwarz criterion		1.667737
Log likelihood	-1331.104	F-statistic		11.31443
Durbin-Watson stat	1.819652	Prob(F-statistic)		0.000000

Abbildung 3: Ergebnisse - "Within"- Schätzung

Tabelle 1: Hausman-Test

Statistik	P-Wert
10.496	0.062

Tabelle 2: Breusch-Pagan-Test

Statistik	P-Wert
3217.140	0.000

### Aufgabe 3 (18 Punkte)

- a) (3 Punkte) Das Grundproblem der Zeitreihenanalyse besteht darin, dass wir Aussagen über die Eigenschaften eines Prozesses  $Y_t$  machen wollen, jedoch nur eine Realisation der Zeitreihe zur Verfügung haben.

Erklären Sie kurz auf intuitiver Basis, inwiefern uns die beiden Konzepte der Stationarität und Ergodizität dabei helfen, dieses Grundproblem zu "lösen".

Die folgenden Fragen beziehen sich auf die Zeitreihe  $Z_t$  (mit  $t = 1, \dots, 200$ ), welche in Abbildung 4 dargestellt ist.

- b) (5 Punkte) Überprüfen Sie, ob  $Z_t$  eine I(1)-Reihe darstellt, d.h. integriert von der Ordnung 1 ist. Benutzen Sie hierzu die entsprechenden ADF-Tests (Abbildungen 5, 6, 7, 9, 10 und 11). Entscheiden Sie sich unter Verwendung der Zeitreihen in den Abbildungen 4 und 8 jeweils für einen geeigneten Fall des ADF-Tests und begründen Sie ihre Auswahl ((i) ohne Konstante und Trend, (ii) mit Konstante, (iii) mit Konstante und Trend).

Wie lautet die Nullhypothese des ADF-Tests?

- c) (2 Punkte) Abbildung 12 zeigt die empirischen Autokorrelationen und partiellen Autokorrelationen für die ersten Differenzen der Reihe  $Z_t$ .

(i) Begründen Sie kurz, weshalb ein AR(3)-Prozess (und nicht ein anderer AR-Prozess) für die Modellierung herangezogen werden sollte.

(ii) Geben Sie die Gleichung für das Modell aus (i) an.

- d) (4 Punkte) Abbildung 13 zeigt das Korrelogramm der Residuen nach Anpassung eines autoregressiven Modells für  $\Delta Z_t$ .

(i) Erklären Sie kurz auf intuitiver Basis, welche Überlegung der Analyse des Korrelogramms der Residuen eines AR-Modells zugrunde liegt.

(ii) Welche Schlussfolgerungen können Sie aus dem Korrelogramm in Abbildung 13 ziehen?

- e) (4 Punkte) Ist bei einer KQ-Schätzung eines autoregressiven Modells für  $\Delta Z_t$  die Annahme strikter Exogenität der Regressoren erfüllt? Begründen Sie Ihre Antwort kurz.

Ist die KQ-Schätzung der AR-Parameter konsistent? Geben Sie auch hier eine kurze Begründung an.

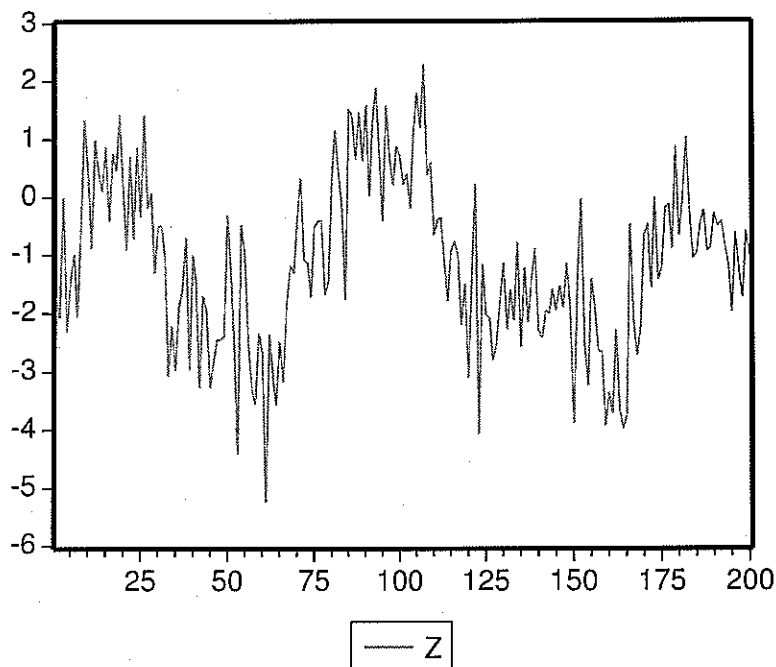


Abbildung 4:  $Z_t$

Exogenous: Constant				
Lag Length: 3 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.345785	0.1588
Test critical values:				
	1% level		-3.463749	
	5% level		-2.876123	
	10% level		-2.574622	
*Mackinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(Z)				
Method: Least Squares				
Date: 01/26/10 Time: 12:52				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Z(-1)	-0.135578	0.057796	-2.345785	0.0200
D(Z(-1))	-0.566864	0.081029	-6.995831	0.0000
D(Z(-2))	-0.448014	0.080291	-5.579898	0.0000
D(Z(-3))	-0.220477	0.069854	-3.156251	0.0019
C	-0.141944	0.094773	-1.497731	0.1359
R-squared	0.353935	Mean dependent var		0.006661
Adjusted R-squared	0.340405	S.D. dependent var		1.217694
S.E. of regression	0.988956	Akaike info criterion		2.840845
Sum squared resid	186.8045	Schwarz criterion		2.924471
Log likelihood	-273.4028	F-statistic		26.15897
Durbin-Watson stat	2.031433	Prob(F-statistic)		0.000000

Abbildung 5: ADF-Test für  $Z_t$  (mit Konstante)



Exogenous: None Lag Length: 3 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-1.801765	0.0682
Test critical values:				
	1% level		-2.576814	
	5% level		-1.942456	
	10% level		-1.615622	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(Z)				
Method: Least Squares				
Date: 01/26/10 Time: 12:56				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Z(-1)	-0.077869	0.043218	-1.801765	0.0732
D(Z(-1))	-0.609716	0.076054	-8.016829	0.0000
D(Z(-2))	-0.476918	0.078189	-6.099573	0.0000
D(Z(-3))	-0.235363	0.069367	-3.393026	0.0008
R-squared	0.346347	Mean dependent var		0.006661
Adjusted R-squared	0.336134	S.D. dependent var		1.217694
S.E. of regression	0.992153	Akaike info criterion		2.842317
Sum squared resid	188.9984	Schwarz criterion		2.909218
Log likelihood	-274.5471	Durbin-Watson stat		2.039881

Abbildung 6: ADF-Test für  $Z_t$  (ohne Konstante)

Exogenous: Constant, Linear Trend Lag Length: 3 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.382869	0.3873
Test critical values:				
	1% level		-4.005662	
	5% level		-3.432917	
	10% level		-3.140265	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(Z)				
Method: Least Squares				
Date: 01/26/10 Time: 12:57				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Z(-1)	-0.139320	0.058467	-2.382869	0.0182
D(Z(-1))	-0.563987	0.081429	-6.926142	0.0000
D(Z(-2))	-0.446567	0.080515	-5.546364	0.0000
D(Z(-3))	-0.219860	0.070010	-3.140400	0.0020
C	-0.086184	0.152605	-0.564757	0.5729
@TREND(1)	-0.000590	0.001263	-0.466790	0.6412
R-squared	0.354675	Mean dependent var		0.006661
Adjusted R-squared	0.337693	S.D. dependent var		1.217694
S.E. of regression	0.990987	Akaike info criterion		2.849903
Sum squared resid	186.5905	Schwarz criterion		2.950254
Log likelihood	-273.2905	F-statistic		20.88506
Durbin-Watson stat	2.031931	Prob(F-statistic)		0.000000

Abbildung 7: ADF-Test für  $Z_t$  (mit Konstante und Trend)

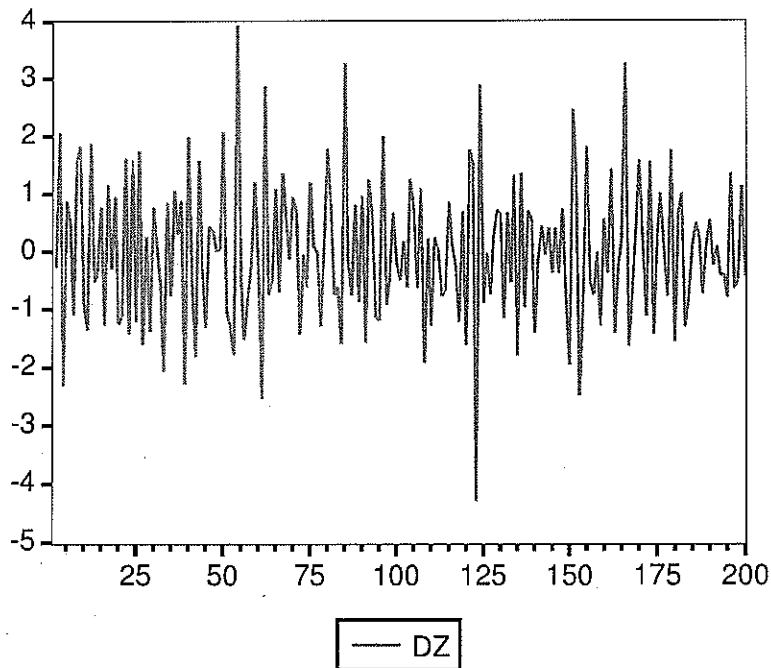


Abbildung 8: Erste Differenzen von  $Z_t$  ( $\Delta Z_t$ )

Exogenous: Constant				
Lag Length: 2 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-14.04905	0.0000
Test critical values:				
	1% level		-3.463749	
	5% level		-2.876123	
	10% level		-2.574622	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DZ)				
Method: Least Squares				
Date: 01/26/10 Time: 12:58				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DZ(-1)	-2.438277	0.173555	-14.04905	0.0000
D(DZ(-1))	0.770571	0.126880	6.073214	0.0000
D(DZ(-2))	0.254988	0.069083	3.691013	0.0003
C	0.006266	0.071464	0.087674	0.9302
R-squared	0.767808	Mean dependent var		0.009652
Adjusted R-squared	0.764180	S.D. dependent var		2.060253
S.E. of regression	1.000485	Akaike info criterion		2.859044
Sum squared resid	192.1863	Schwarz criterion		2.925944
Log likelihood	-276.1863	F-statistic		211.6343
Durbin-Watson stat	2.049244	Prob(F-statistic)		0.000000

Abbildung 9: ADF-Test für erste Differenzen von  $Z_t$  (mit Konstante)

Exogenous: None				
Lag Length: 2 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-14.08519	0.0000
Test critical values:				
	1% level		-2.576814	
	5% level		-1.942456	
	10% level		-1.615622	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DZ)				
Method: Least Squares				
Date: 01/26/10 Time: 12:59				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DZ(-1)	-2.438255	0.173108	-14.08519	0.0000
D(DZ(-1))	0.770544	0.126553	6.088692	0.0000
D(DZ(-2))	0.254972	0.068905	3.700324	0.0003
R-squared	0.767799	Mean dependent var		0.009652
Adjusted R-squared	0.765393	S.D. dependent var		2.060253
S.E. of regression	0.997910	Akaike info criterion		2.848880
Sum squared resid	192.1940	Schwarz criterion		2.899055
Log likelihood	-276.1902	Durbin-Watson stat		2.049147

Abbildung 10: ADF-Test für erste Differenzen von  $Z_t$  (ohne Konstante)

Exogenous: Constant, Linear Trend				
Lag Length: 2 (Automatic based on SIC, MAXLAG=14)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-14.01380	0.0000
Test critical values:				
	1% level		-4.005562	
	5% level		-3.432917	
	10% level		-3.140265	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(DZ)				
Method: Least Squares				
Date: 01/26/10 Time: 13:00				
Sample (adjusted): 5 200				
Included observations: 196 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DZ(-1)	-2.438475	0.174005	-14.01380	0.0000
D(DZ(-1))	0.770797	0.127216	6.058974	0.0000
D(DZ(-2))	0.255088	0.069264	3.682830	0.0003
C	0.024229	0.147170	0.164632	0.8694
@TREND(1)	-0.000177	0.001266	-0.139736	0.8890
R-squared	0.767832	Mean dependent var		0.009652
Adjusted R-squared	0.762970	S.D. dependent var		2.060253
S.E. of regression	1.003049	Akaike info criterion		2.869146
Sum squared resid	192.1667	Schwarz criterion		2.952771
Log likelihood	-276.1763	F-statistic		157.9201
Durbin-Watson stat	2.049528	Prob(F-statistic)		0.000000

Abbildung 11: ADF-Test für erste Differenzen von  $Z_t$  (mit Konstante und Trend)

Correlogram of DZ

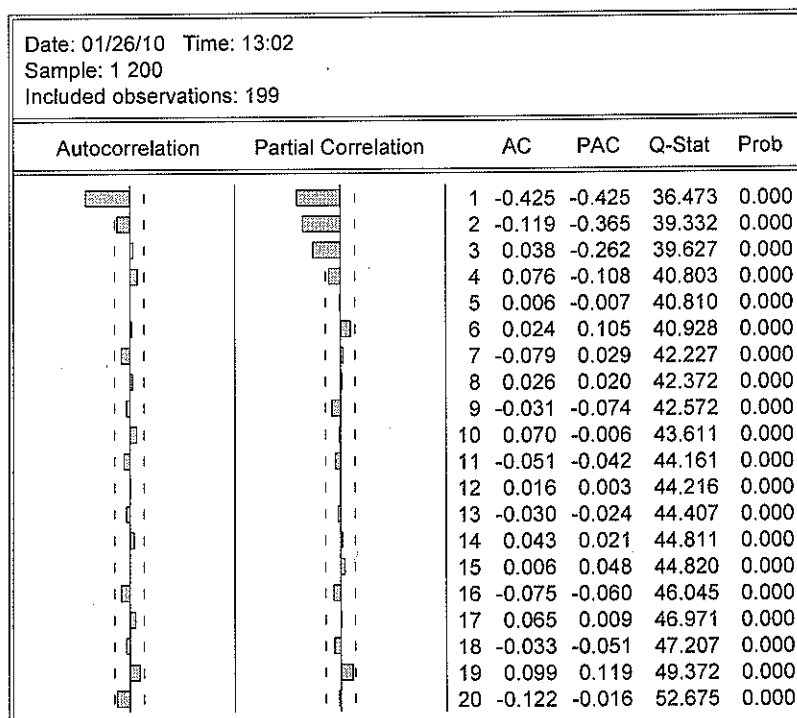


Abbildung 12: Empirische Autokorrelationen und partielle Autokorrelationen der ersten Differenzen von  $Z_t$

Correlogram of Residuals

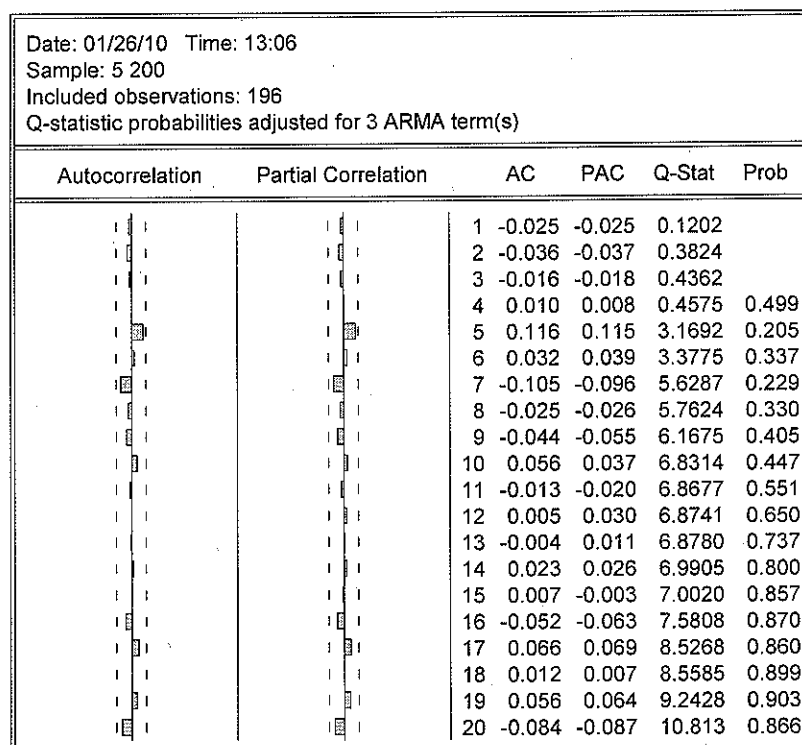


Abbildung 13: Empirische Autokorrelationen und partielle Autokorrelationen der Residuen